

Article

A Decentralized Receiver in Gaussian Interference

Christian D. Chapman ^{1,*}, Hans Mittelmann ², Adam R. Margetts ³ and Daniel W. Bliss ¹

¹ School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85281, USA; d.w.bliss@asu.edu

² School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85281, USA; mittelmann@asu.edu

³ MIT Lincoln Laboratory, Lexington, MA 02421, USA; margetts@ll.mit.edu

* Correspondence: cdchapm2@asu.edu; Tel.: +1-541-619-8726

Received: 1 February 2018; Accepted: 9 April 2018; Published: 11 April 2018



Abstract: Bounds are developed on the maximum communications rate between a transmitter and a fusion node aided by a cluster of distributed receivers with limited resources for cooperation, all in the presence of an additive Gaussian interferer. The receivers cannot communicate with one another and can only convey processed versions of their observations to the fusion center through a Local Array Network (LAN) with limited total throughput. The effectiveness of each bound's approach for mitigating a strong interferer is assessed over a wide range of channels. It is seen that, if resources are shared effectively, even a simple quantize-and-forward strategy can mitigate an interferer 20 dB stronger than the signal in a diverse range of spatially Ricean channels. Monte-Carlo experiments for the bounds reveal that, while achievable rates are stable when varying the receiver's observed scattered-path to line-of-sight signal power, the receivers must adapt how they share resources in response to this change. The bounds analyzed are proven to be achievable and are seen to be tight with capacity when LAN resources are either ample or limited.

Keywords: distributed reception; communications networks; channel capacity; relay channels; interference mitigation

1. Introduction

We are at a remarkable 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. This has two effects. First, it increases the likelihood of interference between systems and necessitates that systems be designed to maintain good performance in these conditions. Second, it creates an opportunity for groups of independent radios to be used as a distributed array to improve performance. Considering this new setting, we study the problem of receiving a signal in the presence of interference with the help of a distributed array, as illustrated in Figure 1.

We develop bounds on the maximum communications rate from a single transmitter to a base node that is provided side information from a distributed receive array in the presence of interference. The nodes that provide side information to the base are identified as *helpers*, and we assert they do not communicate with one another, but only forward information to the base through a previously established reliable Local Array Network (LAN) link, and that this link only supports a maximum total throughput which all the helpers must share.

This receive system can mitigate a strong interferer in a wide variety of environments. 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].

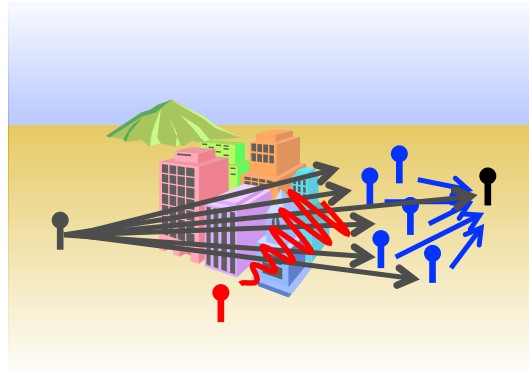


Figure 1. A broadcast node speaks to many helper nodes, which forward limited amounts of information to a base node. Additive interference from neighboring systems is present in the link from broadcaster to helpers.

1.1. Results

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 3). The strongest lower bound is given in Theorem 7.
- Performance characteristics of these rates in the presence of an interferer (Section 4.1).
- Behavior of the strategies in various scattering environments (Section 4.2). The scattering environment is seen to not affect average performance if LAN resource sharing adapts to the channel.
- A strengthening of an existing achievability proof for the system, where the same rate is achieved using less cooperation between nodes (Remark 7).

1.2. Background

The problem of communicating with the help of an array of distributed receivers has been studied in a variety of contexts, most generally as an instance of communications over a single-input multiple-output (SIMO) channel. A variety of SIMO channels have been extensively analyzed, but most analyses are subtly different than the problem considered here. Our work is done in the context of information theory and bounds achievable communications rates, while in contrast most existing studies work towards minimizing distortion metrics or bit error rates.

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 tighter inner bound than all those in [3] is included in Theorem 7, and full proofs are provided for all the results.

Achievable communications rates for this network topology were studied in References [1,2], although their bounds do not directly apply to channels with a Gaussian interferer. The studies provide an example demonstrating the sub-optimality of a Gaussian broadcaster when no interference is present, and suggest that non-Gaussian signaling techniques are needed for this sort of network. In contrast, we demonstrate in Section 4 that an achievable rate using Gaussian signaling typically comes quite close to the system's upper bound in many practical regimes.

The helpers in the system we have described seem like relay nodes but there is an important distinction. Here, the ensemble of links from each helper to the base are like those in a graphical network, available for design within distributional and rate constraints. This is not usual in the context of relay networks, including those studied in very general setting such as Reference [4]. Studies such as References [5–7] detailed 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 the others, so beam forming studies are not directly applicable.

Performance of specific coding schemes for this system were studied in References [8–10]. In particular, [10] was able to identify a scheme that can perform to within 1.5 dB of allowing the base and receivers to communicate without constraints. Results in our study further work towards characterizing achievable rates of the system rather than designing and analyzing the performance of specific modulation schemes.

The topology we consider is similar to many-help-one problems such as the “Chief Executive Officer (CEO) problem” posed by Berger in Reference [11]. In this scenario, some node called the CEO seeks to estimate a source by listening to “agent” nodes which communicate to the CEO at fixed total rates. Many variations of this problem have been studied, for instance in Reference [12] in limit with number of agents. The focus of the CEO problem and most of its derivatives are to estimate a source to within some distortion (often mean squared error), whereas the focus here is on finding achievable rates for lossless communications.

The noisy Slepian–Wolf problem [13] may 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.

2. Problem Setup

Throughout the paper, we use the notation in Table 1. A broadcaster seeks to convey a message to a base at some rate $R > 0$ bits per time period. The message M is uniformly distributed along $[1 : 2^{TR}]$ over some $T \in \mathbb{N}$ time periods. Our goal is to determine the greatest rate R at which the base can recover M with low probability of error.

Table 1. Notation and terminology.

$\mathcal{CN}(0, \Sigma)$	Complex circularly symmetric normal distribution with zero mean and Hermitian covariance matrix Σ
$I(\cdot; \cdot), H(\cdot), h(\cdot)$	Mutual information, Shannon entropy and differential entropy
$\mathbf{A}^\dagger, \mathbf{a}^\dagger$	Conjugate transpose of matrix \mathbf{A} , vector \mathbf{a}
$\mathbf{a}_S; S \subseteq T$	Vector with components indexed by S
$ \mathbf{A} $	Determinant of a matrix \mathbf{A}
$[a : b]$	Integers from a to b , inclusive
$\mathbf{a} \succeq 0$	Vector \mathbf{a} with non-negative elements
$\mathbf{A} \succeq 0$	Positive semidefinite matrix \mathbf{A}
$\mathbf{1}_N$	N -row vector of 1 s
\mathbf{A}_{jk}	Element on row j , column k of a matrix \mathbf{A}
$\text{diag}(\mathbf{a})$	Diagonal matrix where the (i, i) th diagonal element is the i th element of vector \mathbf{a}
$\mathbf{I}_{n \times n}$	$n \times n$ identity matrix
$\ \mathbf{x}\ $	Vector magnitude/2-norm
\mathbb{R}_+	Nonnegative real numbers

Before communicating, the broadcaster and base agree on a modulation scheme $\phi : [1 : 2^{TR}] \rightarrow \mathbb{C}^T$, where each possible message $m \in [1 : 2^{TR}]$ is mapped to a T -length signal through ϕ . Each of these signals is denoted $\mathbf{x}_m = (x_{m,1}, \dots, x_{m,T}) = \phi(m)$, and it is chosen within the broadcaster’s power constraint: $\frac{1}{T} \sum_{\ell=1}^T |x_{m,\ell}|^2 \leq 1$. To send message M , the broadcaster transmits \mathbf{x}_M , which we denote $\mathbf{X} = (X_1, \dots, X_T) = \mathbf{x}_M$.

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 $0-N$, we identify the base as the “0th” receiver, and call the rest *helpers*. For $n \in [0 : N]$, at each time t , the n th receiver observes

$$Y_{n,t} = h_n X_t + W_{n,t}, \tag{1}$$

where

- the channel $\mathbf{h} \triangleq \begin{bmatrix} h_0 \\ \vdots \\ h_N \end{bmatrix} \in \mathbb{C}^{(N+1) \times 1}$ is constant over time t , and
- the noise $(W_{0,t}, \dots, W_{N,t}) \sim \mathcal{CN}(0, \mathbf{\Sigma})$ has covariance $\mathbf{\Sigma}$ across receivers, and is independent over t .

The receiver observations are written as vectors:

$$\mathbf{Y}_n \triangleq (Y_{n,1}, \dots, Y_{n,T}) \in \mathbb{C}^T, \tag{2}$$

$$\mathbf{Y} \triangleq (\mathbf{Y}_0, \mathbf{Y}_1, \dots, \mathbf{Y}_N) \in (\mathbb{C}^T)^{N+1}. \tag{3}$$

Helper n for $n \in [1 : N]$ employs a (possibly randomly chosen) vector quantizer Q_n on its received sequence \mathbf{Y}_n to produce a coarse r_n -bitrate summary of its observations, $U_n \triangleq Q_n(\mathbf{Y}_n)$. The vector of helper messages is denoted:

$$\mathbf{U} \triangleq (U_1, \dots, U_N) \tag{4}$$

$$= (Q_1(\mathbf{Y}_1), \dots, Q_N(\mathbf{Y}_N)) \tag{5}$$

and the quantizers $(Q_1, \dots, Q_N) = \mathbf{Q}$ satisfy two properties:

- Every node in the system is informed of each quantizer’s behavior, and quantizers produce their compression using only local information. Formally, the probability distribution of $(\mathbf{Q}, \mathbf{X}, \mathbf{Y}, \mathbf{U})$ factors:

$$P_{(\mathbf{Q}, \mathbf{X}, \mathbf{Y}, \mathbf{U})} = P_{\mathbf{Q}} \cdot P_{\mathbf{X}|\mathbf{Q}} \cdot P_{\mathbf{Y}|\mathbf{X}} \cdot \prod_{n=1}^N P_{U_n|\mathbf{Y}_n, \mathbf{Q}}. \tag{6}$$

- Quantizations are within the LAN constraint:

$$\frac{1}{T} H(U_n | \mathbf{Q}) \leq r_n, \quad n \in [1 : N]. \tag{7}$$

The vector of the helper’s quantization rates is denoted:

$$\mathbf{r} \triangleq (r_1, \dots, r_N) \in \mathbb{C}^{(N+1) \times 1}. \tag{8}$$

Each U_n is conveyed precisely to the base over a Local Array Network (LAN). We assume the LAN only supports a limited amount of communication between helpers and the base. For instance, the helpers must all share a small frequency band. We model this by asserting that feasible rate vectors \mathbf{r} belong to a bounded set $\mathcal{R}_{\text{LAN}}(L)$ with a sum-capacity $L > 0$:

$$\mathcal{R}_{\text{LAN}}(L) \triangleq \left\{ \mathbf{r} \mid \sum_{n=1}^N r_n \leq L, r_n \geq 0 \right\}. \tag{9}$$

We refer to the condition that $\mathbf{r} \in \mathcal{R}_{\text{LAN}}(L)$ as the *LAN constraint*. The base reconstructs an estimate $\hat{\mathbf{X}}$ of \mathbf{X} using its own full-precision observation \mathbf{Y}_0 and side information \mathbf{U} from the helpers, then recovers an estimate \hat{M} of message M from $\hat{\mathbf{X}}$. A block diagram of the system is shown in Figure 2.

Channel h and noise covariance Σ are assumed to be static throughout the transmission of each message and are known to all the receive nodes. Both the transmitter and the receivers are assumed to have knowledge of the set of feasible rates, $\mathcal{R}_{\text{LAN}}(L)$. The environment determines:

- channel fades h ,
- noise covariance Σ and
- maximum LAN throughput L .

Available for design are:

- quantizers Q ,
- message modulation $\{x_m\}_{m \in [1:2^{TR}]}$,
- rates each helper should send to the base r and
- fusion and decoding methods at the base to produce \hat{M} .

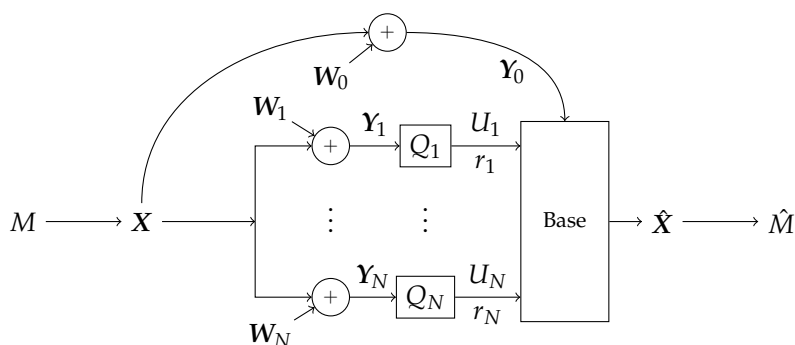


Figure 2. The system in consideration. A message M is broadcast as X with average power 1. The signal is received by a base and N helper nodes, occluded by correlated AWGN W . Helper n for $n \in [1 : N]$ quantizes Y_n through Q_n to produce an r_n -bit summary U_n . The quantized observations U and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

We say a communication rate R is *achievable* if for any $\epsilon > 0$ then for large enough T there is some distribution of quantizers Q satisfying $\frac{1}{T}H(Q_k(Y_k)|Q) \leq r_k$ for each k , encoders $\phi_Q : [1 : 2^{TR}] \rightarrow \mathcal{X}^T$ and decoders f_Q where:

$$P(f_Q(\mathbf{U}) \neq M) < \epsilon. \tag{10}$$

Theorem 1. The maximum communications rate achievable with arbitrarily low probability of error is:

$$C = \sup_T \max_{\substack{(\mathbf{X}, \mathbf{Y}, \mathbf{Q}, \mathbf{U}) \sim P \in \mathcal{S}_T \\ \mathbb{E}|\mathbf{X}|^2 \leq 1}} \frac{1}{T} I(\mathbf{X}; \mathbf{Y}_0, \mathbf{U} | \mathbf{Q}) \tag{11}$$

where \mathcal{S}_T is the collection of joint distributions with block length T that satisfy conditions in Equations (6) and (7) for some $\mathbf{r} \in \mathcal{R}_{\text{LAN}}(L)$.

Proof. C is achievable through the usual noisy channel coding argument [14]. For any achievable rate R , applying Fano’s inequality to $H(M|Q)$ gives $R \leq C$. \square

In the sequel, we bound C from Theorem 1 with computable expressions.

3. Bounds on Communications Rates

The system’s capacity C can be upper bounded by considering a stronger channel where all helpers are informed of each other’s observations. Compute:

$$I(\mathbf{X}; \mathbf{Y}_0, \mathbf{U} | \mathbf{Q}) = I(\mathbf{X}; \mathbf{U} | \mathbf{Y}_0, \mathbf{Q}) + I(\mathbf{X}; \mathbf{Y}_0 | \mathbf{Q}) \tag{12}$$

$$\leq H(\mathbf{U} | \mathbf{Q}) + I(\mathbf{X}; \mathbf{Y}_0) \tag{13}$$

$$\leq T \cdot \left(L + \log_2 \left[1 + \|h_0\|^2 / \Sigma_{0,0} \right] \right), \tag{14}$$

where the first inequality follows since $\mathbf{Q} \rightarrow \mathbf{X} \rightarrow \mathbf{Y}_0$ is a Markov chain (Equation (6)). 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 [15], the capacity of this stronger channel is $\max_{P_X: E[|X|^2] < 1} I(\mathbf{X}; \mathbf{Y})$ (here, (\mathbf{X}, \mathbf{Y}) are variables at a single timepoint) which, by a derivation given in [16], simplifies to the familiar upper bound:

$$\log_2 \left[\frac{|\Sigma + \mathbf{h}\mathbf{h}^\dagger|}{|\Sigma|} \right]. \tag{15}$$

Taking the minimum of this bound and Equation (14) yields an upper bound for the original system. This is an instance of a general *cut-set upper bound* from [17].

Remark 1. Any achievable rate R must satisfy

$$R \leq \min \left\{ \log_2 \left[\frac{|\Sigma + \mathbf{h}\mathbf{h}^\dagger|}{|\Sigma|} \right], L + \log_2 \left[1 + \|h_0\|^2 / \Sigma_{1,1} \right] \right\}. \tag{16}$$

Proof. Justified by the preceding discussion. \square

The rest of the section is dedicated to four communication strategies for the system in order of increasing complexity and diversity utilization. The strongest bound is presented in Theorem 7. It is described using the context of the preceding two bounds.

3.1. 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 was characterized in Reference [18], and in particular its sum-rate was given in Reference [19]. In the scalar case, this sum rate reduces to:

$$\max_n \log_2 \left[\frac{\Sigma_{n,n} + |h_n|^2}{\Sigma_{n,n}} \right], \tag{17}$$

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

Remark 2. The following rate is achievable

$$R_{BC}(L) = \max \left\{ \log_2 \left[\frac{\Sigma_{0,0} + |h_0|^2}{\Sigma_{0,0}} \right], \min \left\{ L, \max_n \log_2 \left[\frac{\Sigma_{n,n} + |h_n|^2}{\Sigma_{n,n}} \right] \right\} \right\}. \tag{18}$$

The rate is achievable with the following facilities:

- Helpers are informed of which helper is the best (best is derived from \mathbf{h}, Σ, L).
- The transmitter and best helper coordinate a codebook of appropriate rate and distribution.

Proof. By having the broadcaster send a message to the n th receiver, a rate

$$\log_2 \left[\frac{\Sigma_{n,n} + |h_n|^2}{\Sigma_{n,n}} \right] \tag{19}$$

is achievable from transmitter-to-receiver. If $n = 0$ (i.e., the transmitter broadcasts to the base), then Equation (19) is achievable from transmitter-to-base. Otherwise, Equation (19) is achievable from transmitter-to-base if it is less than L by forwarding the receiver’s decoding over the LAN. \square

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

$$L \leq \max_n \log_2 \left[\frac{\Sigma_{n,n} + |h_n|^2}{\Sigma_{n,n}} \right] \tag{20}$$

then the system’s capacity is L .

Proof. By assumption, the upper bound in Equation (16) equals L . This is achievable by Remark 2. \square

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.

3.2. Achievable Rate through Gaussian Distortion

Rate-distortion theory shows that 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 $\mathcal{D} > 0$, the rate must be at least

$$\mathcal{R}(\mathcal{D}) = \log_2(\sigma^2) - \log_2(\mathcal{D}). \tag{21}$$

To achieve this rate, the encoding operation must emulate the following test channel (Reference [14]):

$$Y \xrightarrow{\alpha/\beta} \oplus \xrightarrow{\beta} Z = \alpha Y + \beta W \tag{22}$$

$W \sim \mathcal{CN}(0, 1)$

where $\alpha = \sqrt{1 - \mathcal{D}/\sigma^2}$ and $\beta = \sqrt{\mathcal{D}}$. Dithered lattice quantization [20] is one method that can be used to realize such a quantizer in limit with block length in addition to having other nice analytic properties. Using this method, in limit with block length, the base’s estimates for a helper’s observation at any single timepoint can be made to approximate

$$Z_n(r_n) \triangleq Y_n + W_{Q,n}(r_n) \tag{23}$$

where $W_{Q,n}(r_n)$ is independent Gaussian distortion with variance some function of the helper’s rate r_n . We will now derive this function. Define

$$\mathbf{W}_Q(\mathbf{r}) \triangleq (W_{Q,1}(r_1), \dots, W_{Q,N}(r_N)) \in \mathbb{C}^{N \times 1}, \tag{24}$$

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

$$\mathbf{Z}(\mathbf{r}) \triangleq \mathbf{Y} + \mathbf{W}_Q(\mathbf{r}). \tag{25}$$

The subscript Q here is a decoration to distinguish quantization distortion terms $W_Q, W_{Q,n}$ from environment noise, W, W_n .

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 Equation (21) equal to r_n and

solving for \mathcal{D} , with σ^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 (22)) 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:

$$\text{Var}(W_{Q,n}(r_n)) = \frac{\|h_n\|^2 + \Sigma_{n,n}}{2^{r_n} - 1}. \tag{26}$$

By Equations (1) and (26), 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

$$\mathbf{D}(\mathbf{r}) \triangleq \Sigma + \text{diag} \left(0, \frac{\|h_1\|^2 + \Sigma_{1,1}}{2^{r_1} - 1}, \dots, \frac{\|h_N\|^2 + \Sigma_{N,N}}{2^{r_N} - 1} \right). \tag{27}$$

We refer to this system as a *Gaussian distortion system*. A block diagram of this system is shown in Figure 3. This system is equivalent to a multi-antenna Gaussian receiver and its achievable rate has a closed form:

$$R_G(\mathbf{r}) \triangleq I(X; \mathbf{Z}(\mathbf{r})) = \log_2 \left[\frac{|\mathbf{D}(\mathbf{r}) + \mathbf{h}\mathbf{h}^\dagger|}{|\mathbf{D}(\mathbf{r})|} \right]. \tag{28}$$

Theorem 2. For a distributed receive system as described in Section 2 with noise covariance matrix Σ , LAN constraint L and fixed average helper quantization rates $\mathbf{r} \in \mathcal{R}_{\text{LAN}}(L)$, then a rate $R_G(\mathbf{r})$ is achievable with the following facilities:

- The transmitter and base coordinate a codebook of appropriate rate and distribution.
- Each n -th helper is informed of an encoding rate r_n (optimal choice derived from \mathbf{h}, Σ, L).
- Each helper coordinates a random dither with the base.
- The base is informed of $\mathbf{r}, \mathbf{h}, \Sigma, L$.

All the details of this derivation are formalized in Appendix B.

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

$$R_{G,\max}(L) \triangleq \max_{\mathbf{r} \in \mathcal{R}_{\text{LAN}}(L)} R_G(\mathbf{r}). \tag{29}$$

Unfortunately, this expression cannot be simplified much in general. Maximization of R_G over $\mathcal{R}_{\text{LAN}}(\cdot)$ can be performed efficiently using quasi-convex optimization algorithms:

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

Proof. The maximum of R_G follows the maximum of $-\mathbf{h}^\dagger \mathbf{D}(\mathbf{r})^{-1} \mathbf{h}$ 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 [21]). Further, a one-dimensional function is quasi-concave if anywhere its derivative is 0 and its second derivative is below 0.

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

$$\frac{d}{dt} f(t) = \mathbf{A}_t \mathbf{D}_t^{-1} \mathbf{h} \tag{30}$$

$$\frac{d^2}{dt^2} f(t) = (\mathbf{D}_t^{-1} \mathbf{h})^\dagger \left[\frac{d}{dt} \mathbf{A}_t - 2\mathbf{A}_t \mathbf{D}_t^{-1} \mathbf{A}_t \right] (\mathbf{D}_t^{-1} \mathbf{h}). \tag{31}$$

If $\frac{d}{dt} f(t) = 0$, then $\mathbf{D}_t^{-1} \mathbf{h}$ is in the null space of \mathbf{A}_t so the $-2\mathbf{A}_t \mathbf{D}_t^{-1} \mathbf{A}_t$ term in Equation (31) vanishes. $\frac{d}{dt} \mathbf{A}_t$ is negative definite in $\mathbb{R}_+^{N \times 1}$ so $\frac{d^2}{dt^2} f(t) < 0$ there, establishing f 's quasi-concavity and thus that of $-\mathbf{h}^\dagger \mathbf{D}_t^{-1} \mathbf{h}$. Since $-\mathbf{h}^\dagger \mathbf{D}_t^{-1} \mathbf{h}$ is quasi-concave along all of $\mathbb{R}_+^{N \times 1}$, then it is also quasi-concave in any convex restriction of that domain. $\mathcal{R}_{\text{LAN}}(L)$ is a simplex, which is convex. \square

$R_{G,\max}(L)$ is tight with the upper bound in Equation (16), in limit with LAN throughput L .

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

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

$$R_G(L/N \cdot \mathbf{1}_N) \rightarrow \log_2 \left[\frac{|\Sigma + hh^\dagger|}{|\Sigma|} \right] \tag{32}$$

by continuity of \log_2 , sums, and the matrix determinant (in terms of matrix coordinates). The limit in Equation (32) is greater than or equal to Equation (16), establishing the remark. \square

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.

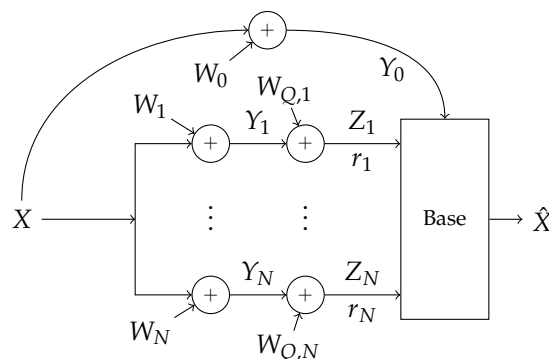


Figure 3. Gaussian distortion System: Quantizer distortion is modeled as additive white Gaussian noise, where enough noise is added so that the capacity across quantizers is some given rate vector \mathbf{r} .

3.3. Achievable Rate through Distributed Compression

Helper n for $n \in [1 : N]$ in a realization of the Gaussian distortion system described in Section 3.2, 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 can be compressed before forwarding to allow for less noise to be introduced in quantization.

The Slepian–Wolf theorem [22] shows that if the LAN is such that helpers can encode at rates $\boldsymbol{\rho} \triangleq (\rho_1, \dots, \rho_N)$, then the helpers can losslessly convey the encodings to the base at lower rates $\mathbf{r} \preceq \boldsymbol{\rho}$, as long as \mathbf{r} and the encodings \mathbf{U} satisfy the conditions that for all subsets $S \subseteq [1 : N]$, then:

$$H(\mathbf{U}_S | \mathbf{U}_{S^c}, Y_0, \mathbf{Q}) < \sum_{n \in S} r_n. \tag{33}$$

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 right before Equation (23), $H(\mathbf{U}_S | \mathbf{U}_{S^c}, Y_0, \mathbf{Q})$ can be made close to $I(Y_S; \mathbf{Z}_S(\boldsymbol{\rho}) | \mathbf{Z}_{S^c}(\boldsymbol{\rho}), Y_0)$ with time expansion. The details of this are given in Lemma 2. Maximizing over the LAN constraint in Equation (9), the following rate is achievable:

$$R_{DC}(L) \triangleq \max_{\mathbf{r} \in \mathcal{R}_{LAN}(L)} \max_{\boldsymbol{\rho} \in \mathcal{R}_{DC}(\mathbf{r})} R_G(\boldsymbol{\rho}) \tag{34}$$

where

$$\mathcal{R}_{DC}(\mathbf{r}) = \left\{ \boldsymbol{\rho} : \forall S \subseteq [1 : N], I(\mathbf{Y}_S; \mathbf{Z}_S(\boldsymbol{\rho}) | \mathbf{Z}_{S^c}(\boldsymbol{\rho}), Y_0) < \sum_{n \in S} r_n \right\}. \tag{35}$$

$\mathbf{Z}(\cdot)$ in Equation (35) is defined as in Equation (25). We can then state the following:

Theorem 3. For a distributed receive system as described in Section 2 with noise covariance matrix $\boldsymbol{\Sigma}$ and LAN constraint L , $R_{DC}(L)$ is achievable with the following facilities:

- The transmitter and base coordinate a codebook of appropriate rate and distribution.
- Each n -th helper is informed of a quantization rate ρ_n and binning rate r_n , (optimal choice derived from $\mathbf{h}, \boldsymbol{\Sigma}, L$).
- Each helper coordinates a random dither with the base.
- The base is informed of $\boldsymbol{\rho}, \mathbf{r}, \mathbf{h}, \boldsymbol{\Sigma}, L$.

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 Equation (33) is satisfied. However, in analog to Corollary 1 from Reference [1], the rate in 6 can be improved by expanding $\mathcal{R}_{DC}(\mathbf{r})$ 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 $\mathcal{R}_{DC}(\mathbf{r})$ so that helpers cannot convey \mathbf{U} to the base unambiguously, some extra correlation with X is retained through this distortion.

Theorem 4. For a distributed receive system as described in Section 2 with noise covariance matrix $\boldsymbol{\Sigma}$ and LAN constraint L , the following rate is achievable:

$$R_{\overline{DC}}(L) \triangleq \max_{\substack{\mathbf{r} \in \mathcal{R}_{LAN}(L), \\ \lambda > 0}} \max_{\boldsymbol{\rho} \in \mathcal{R}_{DC}^\lambda(\mathbf{r})} R_G(\boldsymbol{\rho}) - \lambda \tag{36}$$

where

$$\mathcal{R}_{DC}^\lambda(\mathbf{r}) \triangleq \left\{ \boldsymbol{\rho} : \forall S \subseteq [1 : N], I(\mathbf{Y}_S; \mathbf{Z}_S(\boldsymbol{\rho}) | \mathbf{Z}_{S^c}(\boldsymbol{\rho}), Y_0) < \sum_{n \in S} r_n + \lambda \right\}. \tag{37}$$

The rate requires the following facilities:

- The transmitter and base coordinate a codebook of appropriate rate and distribution.
- Each n -th helper is informed of a quantization rate ρ_n , binning rate r_n and hull parameter λ (optimal choice derived from $\mathbf{h}, \boldsymbol{\Sigma}, L$).
- Each helper coordinates a random dither with the base.
- The base is informed of $\boldsymbol{\rho}, \mathbf{r}, \lambda, \mathbf{h}, \boldsymbol{\Sigma}, L$.

All the details of Theorems 6 and 7 are shown in Appendix B.

Remark 6. 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.
- Construction of helper messages is independent of the transmitter’s codebook \mathcal{X} .

This is demonstrated in Appendix C. The last three assumptions are necessary:

- Since the base has codebook 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 $R_{DC}(L)$ is strictly less than the upper bound in Equation (16), but, by Remark 3, this upper bound is achieved in some regimes.

Theorem 7 and Remark 7 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. Sanderovich et al. [1] used a “nomadicity” assumption which asserts that the mapping from transmitter messages to codewords is not present at the helpers, whereas Theorem 7 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.

Similar to 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.

4. 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 Σ , and
- channel h (assumed to be static and precisely estimated a priori per each channel use).

In all graphs, the rate from Equation (16) is called “Upper Bound”, the rate from Equation (28) is “Gaussian Distortion”, the rate from Theorem 7 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.

4.1. Performance in the Presence of an Interferer

One motivation for increasing receive diversity is to better mitigate the effect of interference, so it is important to study the extent to which the strategies from Section 3 can do so.

We model interference as a zero-mean Gaussian broadcast independent of our system. Our system’s nodes are informed of nothing more than the scales at which the interferer appears in our receivers’ observations. This assumption prohibits using dirty paper coding [23] and other transmit-side interference mitigation strategies. With this model, the covariance matrix of noise terms associated with a particular interferer seen by the receivers is a matrix $\mathbf{A} = \mathbf{a}\mathbf{a}^\dagger$, where $\mathbf{a} \in \mathbb{C}^{N+1}$. In all trials run, a single interferer was assumed to be present at all nodes with power α and random phase so that

$$\Sigma = \mathbf{I}_{N+1 \times N+1} + \alpha \cdot \mathbf{a}\mathbf{a}^\dagger. \tag{38}$$

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

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

Sanderovich et al. [1] provided an example of a system such as the one in the present work where a Gaussian transmitter is significantly 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 signaling 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.

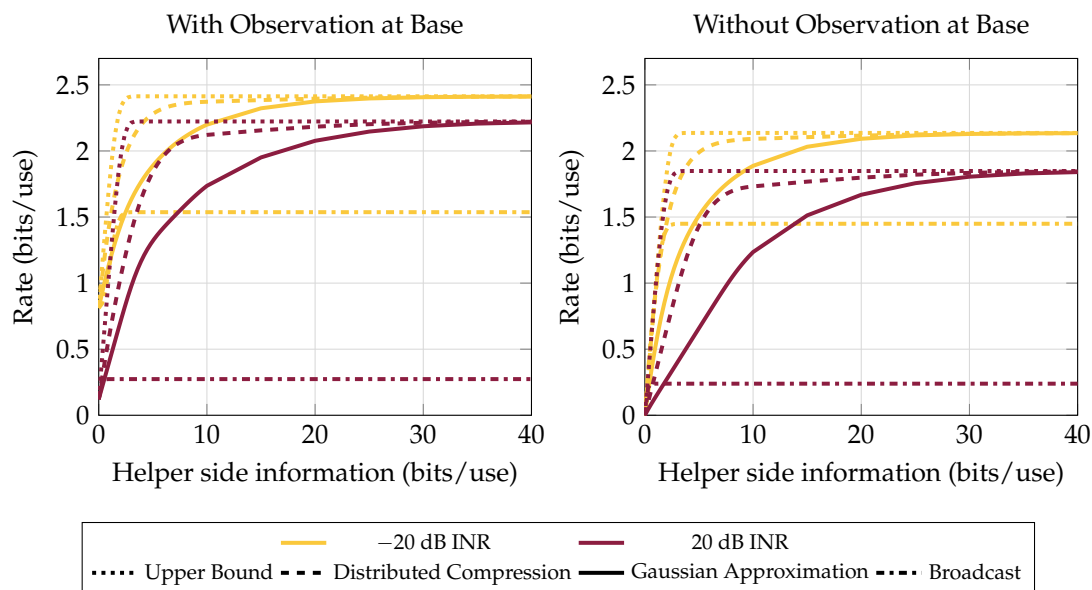


Figure 4. Bound performance versus LAN constraint (that is, the total rate available to hub from all the helpers combined). Four helpers with 0 dB average receiver SNR, averaged over 1000 channels with gains $\sim \mathcal{CN}(0, I)$. Interferer present at each receiver with a uniform random phase and specified INR. Single-user decode-and-forward is greatly affected by interference (Compare the Broadcast curves from -20 dB to 20 dB INR). Distributed compression only offers significant benefits over Gaussian compress-and-forward in strong interference and when LAN resources are scarce.

4.2. 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 be modeled with a Rician distribution [24]. 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 $K \rightarrow \infty$, then the distribution approaches a point mass at Ω (Reference [25]).

The type of scattering environment does not greatly affect the average rate achievable by any bound other than the Broadcast bound (Figure A1). 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 the other parameters, in high scattering, it is most helpful for the base to draw most 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 interference 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.

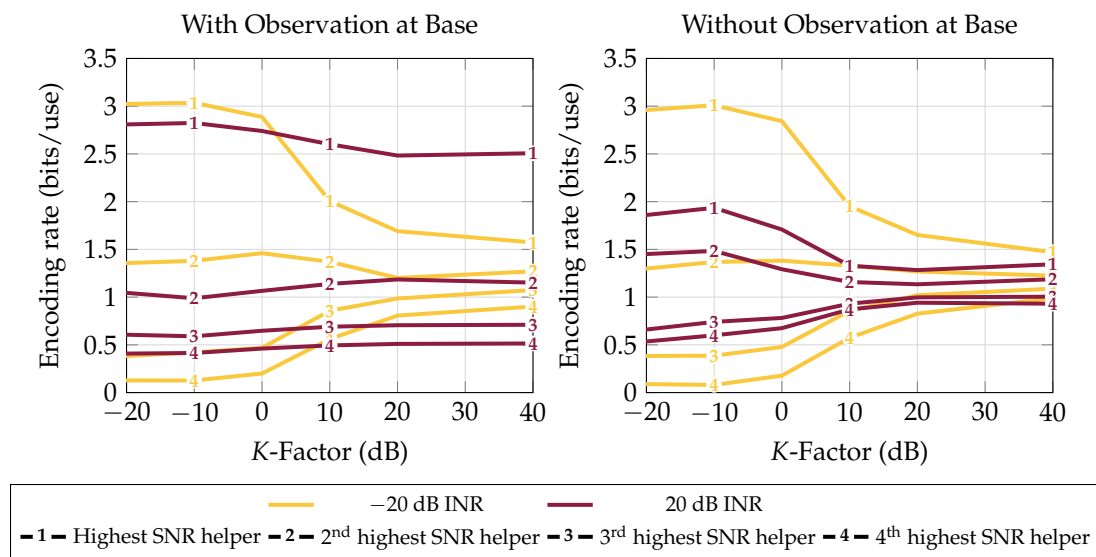


Figure 5. Average rate at which each helper forwards to base in distributed compression (Theorem 7) versus K -factor (scattered-to-direct-path received signal power ratio). Four helpers with 0 dB average receive SNR and a LAN constraint of $L = 5$ bits per channel use of side information the helpers share to inform the base of their observations. Averaged over 1000 channels with gains $\sim \mathcal{CN}(0, I)$. Interferer present at each receiver with a uniform random phase and specified INR. 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.

5. Conclusions

One upper bound and three lower bounds on communication rate were developed. A simple upper bound (Remark 1) was derived, achievable when the LAN constraint is stringent (small L), and achievable in limit as the LAN constraint is relaxed ($L \rightarrow \infty$). An achievable rate (Theorem 5) considering quantizers which add Gaussian distortion was seen to provide large gains in interference over a rate that does not fully use the distributed receive array (Remark 2). The rate (Theorem 7) achieved when a distributed compression stage is added to the previous technique provides gains over Theorem 5 when interference is strong and helper-to-base communication is limited, but may be more difficult to implement.

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 helper observations are ignored by the base, while, in line-of-sight environments, each helper informs the base equally. The presence of a strong interferer dampens this effect, since in this regime the base needs more spatial diversity to mitigate the interferer.

An immediate goal for future work is to devise practical implementations of the presented strategies. This may include refining the models of the LAN and helpers to more precisely reflect system capabilities.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1099-4300/20/4/269/s1>.

Acknowledgments: We would like to thank Shankar Ragi for aid in producing datasets for this paper. We would also like to thank Bryan Paul for his comments and corrections. Support by AFOSR under grant FA 9550-15-1-0351. DISTRIBUTION STATEMENT A: Approved for Public Release, Distribution Unlimited. 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.

Author Contributions: Daneil W. Bliss, Adam Margetts and Christian Chapman developed the theoretical work in the paper. Christian Chapman and Hans Mittelmann performed the simulations. Christian Chapman wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

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 $\mathcal{R}_{\overline{DC}}(r) \subseteq \mathbb{R}_+^K$, hence the domain's non-convexity.

To overcome this, an iterative interior point method was used to find the maximum: each constraint $f(x) < 0$ is replaced with a stricter constraint, $f(x) + \beta < 0$ for some $\beta > 0$ and a minimization of the objective is performed under the new constraint. The optimal point is passed as the initial guess for the next iteration, where the objective function is minimized with updated constraints $f(x) + \beta' < 0$ with $0 < \beta' < \beta$. The method is iterated until the constraint is practically equivalent to $f(x) < 0$ or the optimal point converged. Each individual minimization was performed using sequential least-squares programming (SLSQP) through SciPy [26]. The code for this is included in the Supplementary Materials.

Appendix B. Achievability of Bounds

Several lemmas are needed to prove bound achievability. The following lemma demonstrates that lattice encodings become close in joint information content to the equivalent random variables in a Gaussian distortion system from Section 3.2. Lemma 1 demonstrates the joint differential entropies between source estimates from the encodings are close to the equivalent differential entropies in the Gaussian distortion system. Lemma 2 demonstrates the encodings themselves have joint Shannon entropies that are close to mutual informations in the Gaussian distortion system.

Lemma 1. *Take helper observations \mathbf{Y} and independent dithers $(\mathbf{W}_{Q,n})_{n \in [0:N]}$ where, for each n , the dither vector $\mathbf{W}_{Q,n}$ is distributed along the base cell of a regular lattice in \mathbb{C}^t whose normalized second moment is $\sigma_{Q,n}^2$. For any $A \subset [0 : N]$, $\varepsilon > 0$ and t large enough,*

$$\left| \frac{1}{t} h(\mathbf{Y}_A - (\mathbf{W}_{Q,n})_{n \in A}) - h(\mathbf{Y}_A + \mathbf{W}_A) \right| < \varepsilon \quad (\text{A1})$$

Proof. The proof is similar to Theorem 6 from [20]. Notice that, for any additive noise channel, $x \rightarrow x + n$ then $I(x; x + n) = h(x + n) - h(n)$ so that $h(x + n) = h(n) + I(x; x + n)$. Then,

$$\frac{1}{t}h(\mathbf{Y}_A - (\mathbf{W}_{Q,n})_{n \in A}) = \frac{1}{t}[h((\mathbf{W}_{Q,n})_{n \in A}) + I(\mathbf{Y}_A; \mathbf{Y}_A - (\mathbf{W}_{Q,n})_{n \in A})] \tag{A2}$$

$$= \frac{1}{t}[h((\mathbf{W}_{Q,n})_{n \in A}) - th(W_A) + th(W_A) + I(\mathbf{Y}_A; \mathbf{Y}_A - (\mathbf{W}_{Q,n})_{n \in A})] \tag{A3}$$

$$= \frac{1}{t}[\sum_{n \in A} h(\mathbf{W}_{Q,n}) - th(W_n) + th(W_A) + I(\mathbf{Y}_A; \mathbf{Y}_A - (\mathbf{W}_{Q,n})_{n \in A})] \tag{A4}$$

$$\geq \frac{1}{t}[\sum_{n \in A} h(\mathbf{W}_{Q,n}) - th(W_n) + th(W_A) + tI(\mathbf{Y}_A; \mathbf{Y}_A + W_A)] \tag{A5}$$

$$= \frac{1}{t}[\sum_{n \in A} h(\mathbf{W}_{Q,n}) - th(W_n) + th(\mathbf{Y}_A + W_A)] \tag{A6}$$

$$= h(\mathbf{Y}_A + W_A) + \frac{1}{t} \sum_{n \in A} D(P_{\mathbf{W}_{Q,n}} \| \prod_{\ell=1}^t P_{W_n}) \xrightarrow{t \uparrow \infty} h(\mathbf{Y}_A + W_A). \tag{A7}$$

where Equation (A4) follows since the terms are independent, Equation (A5) follows by the worst additive noise lemma [27], and the substitution and convergence in Equation (A7) follow by Section III of [20]. In addition:

$$\frac{1}{t}h(\mathbf{Y}_A - (\mathbf{W}_{Q,n})_{n \in A}) = \frac{1}{t}[h((\mathbf{W}_{Q,n})_{n \in A}) + I(\mathbf{Y}_A; \mathbf{Y}_A - (\mathbf{W}_{Q,n})_{n \in A})] \tag{A8}$$

$$\leq \frac{1}{t}h((\mathbf{W}_{Q,n})_{n \in A}) + I(\mathbf{Y}_A; \mathbf{Y}_A + W_A) + \sum_{n \in A} D(P_{\mathbf{W}_{Q,n}} \| \prod_{\ell=1}^t P_{W_n}) \tag{A9}$$

$$< I(\mathbf{Y}_A; \mathbf{Y}_A + W_A) + \sum_{n \in A} \frac{1}{t}h(\mathbf{W}_{Q,n}) + \varepsilon/2 \tag{A10}$$

$$\leq I(\mathbf{Y}_A; \mathbf{Y}_A + W_A) + \sum_{n \in A} h(W_n) + \varepsilon \tag{A11}$$

$$= I(\mathbf{Y}_A; \mathbf{Y}_A + W_A) + h(W) + \varepsilon = h(\mathbf{Y}_A + W_A) + \varepsilon \tag{A12}$$

where Equation (A4) follows by Lemma 1 from [28] and Equations (A10) and (A11) follow since the terms in $(\mathbf{W}_{Q,n})_{n \in A}$ are independent and by Section III of [20]. □

Lemma 2. For length- t helper observations \mathbf{Y} , dithered lattice encodings \mathbf{U} of \mathbf{Y} using dithers $(\tilde{\mathbf{W}}_{Q,n})_{n \in [0:N]}$ where each \mathbf{U}_n encodes \mathbf{Y}_n onto a regular lattice in \mathbb{C}^t whose cells have normalized second moment $\sigma_{Q,n}^2$. Then, for any $\varepsilon > 0$ and if t is large enough, any $S \subseteq [0 : N]$ has

$$\left| \frac{1}{t}H(\mathbf{U}_S | (\tilde{\mathbf{W}}_{Q,n})_{n \in S}) - I(\mathbf{Y}_S; \mathbf{Z}_S(\mathbf{r})) \right| < \varepsilon. \tag{A13}$$

In Equations (A13), $\mathbf{Z}(\mathbf{r}) = \mathbf{Y} + \mathbf{W}$, with $\mathbf{Y} = (Y_0, \dots, Y_N) \in \mathbb{C}^{N+1}$ distributed the same as helper observations at a single time point, and

$$\mathbf{W} = (W_0, \dots, W_N) \sim \mathcal{CN} \left(0, \text{diag} \left(\sigma_{Q,1}^2, \dots, \sigma_{Q,N}^2 \right) \right). \tag{A14}$$

Proof. We prove this by induction. By Theorem 1 in Reference [29], for any $n \in [0 : N]$

$$H(\mathbf{U}_n | \tilde{\mathbf{W}}_{Q,n}) = I(\mathbf{Y}_n; \mathbf{Y}_n - \tilde{\mathbf{W}}_{Q,n}). \tag{A15}$$

Theorem 6 in Reference [20] has that

$$\frac{1}{t}I(\mathbf{Y}_n; \mathbf{Y}_n - \tilde{\mathbf{W}}_{Q,n}) \xrightarrow{t \uparrow \infty} I(\mathbf{Y}_n; \mathbf{Y}_n + W_{Q,n}) = I(\mathbf{Y}_n; \mathbf{Z}_n(r_n)). \tag{A16}$$

Thus, the statement holds for any singleton $\{n\}$. Take any disjoint $A, B \subset [0 : N]$ for which the claim holds for $\epsilon' > 0$. By construction, given $(\tilde{W}_{Q,n})_{n \in A \cup B}$, there is a function $f_{A \cup B}$ where $f_{A \cup B}(\mathbf{U}_{A \cup B}) = \mathbf{Y}_{A \cup B} - (\tilde{W}_{Q,n})_{n \in A \cup B}^*$ with $(\tilde{W}_{Q,n})_{n \in A \cup B}^*$ independent of but distributed like $(\tilde{W}_{Q,n})_{n \in A \cup B}$ [20]. Compute:

$$\frac{1}{t} H(\mathbf{U}_{A \cup B} | (\tilde{W}_{Q,n})_{n \in A \cup B}) \geq \frac{1}{t} H(f_{A \cup B}(\mathbf{U}_{A \cup B}) | (\tilde{W}_{Q,n})_{n \in A \cup B}) \tag{A17}$$

$$\geq \frac{1}{t} I(f_{A \cup B}(\mathbf{U}_{A \cup B}); \mathbf{Y}_{A \cup B} | (\tilde{W}_{Q,n})_{n \in A \cup B}) \tag{A18}$$

$$= \frac{1}{t} I(f_{A \cup B}(\mathbf{U}_{A \cup B}); \mathbf{Y}_{A \cup B}) \tag{A19}$$

$$\geq I(\mathbf{Y}_{A \cup B}; \mathbf{Z}_{A \cup B}(\mathbf{r})). \tag{A20}$$

The first inequality follows by the data processing inequality. Equation (A19) follows since $(\tilde{W}_{Q,n})_{n \in A \cup B}$ is independent of the other variables in the term. The last inequality follows by the worst additive noise lemma [27]. For t large enough,

$$\frac{1}{t} H(\mathbf{U}_{A \cup B} | (\tilde{W}_{Q,n})_{n \in A \cup B}) = \frac{1}{t} [H(\mathbf{U}_A | (\tilde{W}_{Q,n})_{n \in A \cup B}) + \tag{A21}$$

$$H(\mathbf{U}_B | (\tilde{W}_{Q,n})_{n \in A \cup B}) - I(\mathbf{U}_A; \mathbf{U}_B | (\tilde{W}_{Q,n})_{n \in A \cup B})] \\ \leq I(\mathbf{Y}_A; \mathbf{Z}_A(\mathbf{r})) + I(\mathbf{Y}_B; \mathbf{Z}_B(\mathbf{r})) - \frac{1}{t} I(\mathbf{U}_A; \mathbf{U}_B | (\tilde{W}_{Q,n})_{n \in A \cup B}) + 2\epsilon. \tag{A22}$$

Similar to before, given $(\tilde{W}_{Q,n})_{n \in S}, S = A, B$ there are deterministic functions f_A, f_B where $f_A(\mathbf{U}_A) = \mathbf{Y}_A - (\tilde{W}_{Q,n})_{n \in A}^*$ and $f_B(\mathbf{U}_B) = \mathbf{Y}_B - (\tilde{W}_{Q,n})_{n \in B}^*$ with $(\tilde{W}_{Q,n})_{n \in S}^*$ iid to $(\tilde{W}_{Q,n})_{n \in S}, S = A, B$. Compute:

$$\frac{1}{t} I(\mathbf{U}_A; \mathbf{U}_B | (\tilde{W}_{Q,n})_{n \in A \cup B}) \geq \frac{1}{t} I(f_A(\mathbf{U}_A); f_B(\mathbf{U}_B) | (\tilde{W}_{Q,n})_{n \in A \cup B}) \tag{A23}$$

$$= \frac{1}{t} I(f_A(\mathbf{U}_A); f_B(\mathbf{U}_B)) \tag{A24}$$

$$= \frac{1}{t} [h(\mathbf{Y}_A - (\tilde{W}_{Q,n})_{n \in A}^*) + h(\mathbf{Y}_B - (\tilde{W}_{Q,n})_{n \in B}^*) - \dots \tag{A25}$$

$$h(\mathbf{Y}_A - (\tilde{W}_{Q,n})_{n \in A}^*, \mathbf{Y}_B - (\tilde{W}_{Q,n})_{n \in B}^*)] \\ \geq h(\mathbf{Y}_A + W_A) + h(\mathbf{Y}_B + W_B) - h(\mathbf{Y}_A + W_A, \mathbf{Y}_B + W_B) - 3\epsilon' \tag{A26}$$

$$= I(\mathbf{Z}_A(\mathbf{r}); \mathbf{Z}_B(\mathbf{r})) - 3\epsilon' \tag{A27}$$

where, again, the first line follows by the data processing inequality and the second line since the conditional is independent of the rest of the mutual information's terms. For large enough t , the second inequality follows by Lemma 1. Combining Equations (A20), (A22) and (A27),

$$0 \leq H(\mathbf{U}_{A \cup B}^t) - I(\mathbf{Y}_{A \cup B}^t; \mathbf{Z}_{A \cup B}^t(\mathbf{r})) \leq 5\epsilon'. \tag{A28}$$

Since $S \subseteq [1 : N]$, the inductive step will only need to be used up to N times. Letting $0 < \epsilon' < \epsilon \cdot 5^{-N}$, the initial statement holds. \square

A small result is also needed to show that the base having its own full precision observation is approximately equivalent to receiving a high-bitrate helper message. Recall that the base is associated with helper index 0.

Lemma 3. *If $\forall S \subseteq [1 : N]$, then*

$$I(\mathbf{Y}_S, \mathbf{Z}_S(\mathbf{r}) | \mathbf{Z}_{S^c}(\mathbf{r}), Y_0) < \sum_{m \in S} r_m, \tag{A29}$$

then, some value r_0 can be chosen sufficiently large so that for any $S \subseteq [0 : N]$

$$I(\mathbf{Y}_S, \mathbf{Z}_S(\mathbf{r}) | \mathbf{Z}_{S^c}(\mathbf{r})) < \sum_{m \in S} r_m. \tag{A30}$$

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

$$I(\mathbf{Y}_S, \mathbf{Z}_S(\mathbf{r}) | \mathbf{Z}_{S^c}(\mathbf{r}), Y_0) < \sum_{m \in S} r_m - \delta. \tag{A31}$$

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

$$\|I(\mathbf{Y}_S; \mathbf{Z}_S(\mathbf{r}) | \mathbf{Z}_{S^c}(\mathbf{r}), Y_0) - I(\mathbf{Y}_S; \mathbf{Z}_S(\mathbf{r}) | \mathbf{Z}_{S^c}(\mathbf{r}), Z_0(r_0))\| \leq \delta. \tag{A32}$$

There is guaranteed to be such an r_0 because the above expression is a continuous function of the covariance matrix of $(Y_0, \mathbf{Y}, Z_0(r_0), \mathbf{Z}(\mathbf{r}))$ 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 $\frac{\|h_0\|^2 + \Sigma_{1,1}}{2^{r_0} - 1} \rightarrow 0$ as $r_0 \rightarrow \infty$. By Equations (A31) and (A32), 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 $S \subseteq [1 : N]$,

$$I(\mathbf{Y}_S^t, \mathbf{Y}_0^t; \mathbf{Z}_S^t(\mathbf{r}), \mathbf{Z}_0^t(\mathbf{r}) | \mathbf{Z}_{(S \cup \{0\})^c}^t(\mathbf{r})) \tag{A33}$$

$$= I(\mathbf{Y}_S^t; \mathbf{Z}_S^t(\mathbf{r}) | \mathbf{Z}_{(S \cup \{0\})^c}^t(\mathbf{r})) + I(\mathbf{Y}_0^t; \mathbf{Z}_0^t(\mathbf{r}) | \mathbf{Z}_{[1:N]}^t(\mathbf{r})) \tag{A34}$$

$$\leq I(\mathbf{Y}_S^t; \mathbf{Z}_S^t(\mathbf{r}) | \mathbf{Z}_{(S \cup \{0\})^c}^t(\mathbf{r})) + r_0 \tag{A35}$$

$$\leq \sum_{m \in S} r_m + r_0 \tag{A36}$$

where Equation (A35) follows because conditioning reduces mutual information, and Equation (A36) follows from what was shown at the start. \square

Theorem 5. For a distributed receive system as described in Section 2 with noise covariance matrix Σ , LAN constraint L and fixed average helper quantization rates $\mathbf{r} \in \mathcal{R}_{\text{LAN}}(L)$, a rate $R_G(\mathbf{r})$ is achievable.

Proof. If some $n \in [0 : N]$ has $r_n = 0$, the system is equivalent to the case where the n th receiver’s observation is not present, so, without loss of generality, we assert that $r_n > \epsilon$. Fix some rate $R < R_G(\mathbf{r})$ and a block length $T = t^2 \in \mathbb{N}$. Form a message M uniform on $[1 : 2^{tR}]$.

- **Outline**

Form a rate- R codebook with length- T codewords. At receiver $n \in [0 : N]$, form a length- t regular lattice encoder with lattice chosen coarse enough that its encodings have rate r_n . Combine encodings at base and find the typical codeword. Observe that probability of error is low by recognizing that each t -length encoding from a helper approximately contains the information content of the helper’s observation plus Gaussian distortion.

- **Transmitter setup**

Generate a codebook mapping $\phi : [1 : 2^{tR}] \rightarrow \mathbb{C}^{T \times 1}$ where:

$$\phi(m) = (x_{m,1}, \dots, x_{m,T}) \in \mathbb{C}^{T \times 1} \tag{A37}$$

with each component $x_{m,\ell}$ drawn iid from $\mathcal{CN}(0, 1)$. Reveal ϕ to the transmitter and base.

- **Helper encoder setup**

For each helper n , $n \in [1 : N]$, generate a dither vector $\tilde{\mathbf{W}}_{Q,n} = (\tilde{\mathbf{W}}_{Q,n}^\ell)_\ell \in (\mathbb{C}^{t \times 1})^t$, where each successive t -segment is uniform in the base region of the Voroni partition \mathcal{P}_n^t of a regular, white t -dimensional lattice \mathcal{L}_n^t which has the following normalized-second-moment:

$$G^t(\mathcal{L}_n^t, \mathcal{P}_n^t) = \frac{\|h_n\|^2 + \Sigma_{n,n}}{2^{r_n - \varepsilon} - 1}. \tag{A38}$$

These terms are detailed at the beginning of Reference [20]. Reveal $\tilde{\mathbf{W}}_{Q,n}$ to the base and helper n . Similarly, at the base, generate a dither vector $\tilde{\mathbf{W}}_{Q,0} = (\tilde{\mathbf{W}}_{Q,0}^\ell)_\ell \in (\mathbb{C}^{t \times 1})^t$ with r_0 chosen large enough for Lemma 3 to hold.

- **Transmission**

To send message $M \in [1 : 2^{TR}]$, have the transmitter broadcast $\mathbf{X} = (X_1, \dots, X_T) = \phi(M)$.

- **Helper encoding and forwarding**

Helper n ($n \in [0 : N]$) observes a sequence of length T , $\mathbf{Y}_n = (\mathbf{Y}_n^\ell)_\ell \in (\mathbb{C}^{t \times 1})^t$. Now $\mathbf{Y}_n - \tilde{\mathbf{W}}_{Q,n} = (\mathbf{Y}_n^\ell - \tilde{\mathbf{W}}_{Q,n}^\ell)_\ell \in \mathbb{C}^T = (\mathbb{C}^t)^t$ is composed of t consecutive length- t segments. For the ℓ -th length- t segment ($\ell \in [1 : t]$), form a quantization $U_n^\ell \in \mathcal{L}_n^t$ by finding the point in \mathcal{L}_n^t associated with the region in \mathcal{P}_n^t in which that segment resides. The properties of such U_n are the subject of References [20,29]. By the extension of Theorem 1 in Appendix A of Reference [29],

$$H(U_n^\ell | \tilde{\mathbf{W}}_{Q,n}) = \frac{1}{t} I(\mathbf{Y}_n; \mathbf{Y}_n - \tilde{\mathbf{W}}_{Q,n}). \tag{A39}$$

Further, by Theorem 6 in Reference [20],

$$\frac{1}{t} I(\mathbf{Y}_n; \mathbf{Y}_n - \tilde{\mathbf{W}}_{Q,n}) \xrightarrow{t \uparrow \infty} I(\mathbf{Y}_n; \mathbf{Y}_n + \mathbf{W}_{Q,n}) = r_n - \varepsilon \tag{A40}$$

where Y_n has the distribution of any individual component of \mathbf{Y}_n and $W_{Q,n} \sim \mathcal{CN}(0, G^t(\mathcal{L}_n^t, \mathcal{P}_n^t))$ (in agreement with notation in Section 3.2). Thus, the t encoded messages produced by helper n , $(U_n^\ell)_\ell$, are within the LAN constraint for large enough block length T . Helper $n \in [1 : N]$ forwards $(U_n^\ell)_\ell$ to the base.

- **Decoding**

At the base, receive all of the lattice quantizations,

$$\mathbf{U} = ((U_n^\ell)_{\ell \in [1:t]})_{n \in [0:N]}. \tag{A41}$$

Take $A_\varepsilon^T(\mathbf{X}, \mathbf{U} | \mathbf{W}_Q)$ to be the collection of

$$(\mathbf{x}_m, \mathbf{u}) \in \text{Range}(\phi) \times \left(\prod_{n=0}^N \mathcal{L}_n^t \right)^t \tag{A42}$$

which are jointly ε -weakly-typical with respect to the joint distribution of (\mathbf{X}, \mathbf{U}) , conditional on all the dithers $(\tilde{\mathbf{W}}_{Q,n})_{n \in [0:N]}$. Weak- and joint-typicality are defined in Reference [14].

At the base, find a message estimate $\hat{m} \in [1 : 2^{TR}]$ where $(\phi(\hat{m}), \mathbf{U}) \in A_\varepsilon^T(\mathbf{X}, \mathbf{U} | \mathbf{W}_Q)$. Declare error events \mathcal{E}_0 if M is not found to be typical with \mathbf{U} , and \mathcal{E}_1 if there is some $\hat{m} \neq M$ where $(\phi(\hat{m}), \mathbf{U}) \in A_\varepsilon^T(\mathbf{X}, \mathbf{U} | \mathbf{W}_Q)$.

• **Error analysis**

By typicality and the law of large numbers, $P(\mathcal{E}_0) \rightarrow 0$ as $t \rightarrow \infty$. In addition,

$$P(\mathcal{E}_1) \leq \sum_{\hat{m} \neq M} P\left(\{\phi(\hat{m}) : (\phi(\hat{m}), \mathbf{U}) \in A_\epsilon^T(\mathbf{X}, \mathbf{U} | \mathbf{W}_Q)\}\right) \tag{A43}$$

$$\leq \sum_{\hat{m} \neq M} 2^{-t \cdot (I(X_1, \dots, X_t; \mathbf{U}_n^1)_{n \in [0:N]} | (\tilde{\mathbf{W}}_{Q,n}^1)_{n \in [0:N]}) - 3\epsilon} \tag{A44}$$

$$< 2^{-t \cdot (I(X_1, \dots, X_t; \mathbf{U}_n^1)_{n \in [0:N]} | (\tilde{\mathbf{W}}_{Q,n}^1)_{n \in [0:N]}) - 3\epsilon} + tR \tag{A45}$$

$$= 2^{-t \cdot (h(X_1, \dots, X_t) - h(X_1, \dots, X_t | (\mathbf{U}_n^1)_{n \in [0:N]}, (\tilde{\mathbf{W}}_{Q,n}^1)_{n \in [0:N]})) - 3\epsilon - tR} \tag{A46}$$

$$\leq 2^{-t \cdot (h(X_1, \dots, X_t) - h(X_1, \dots, X_t | (\mathbf{Y}_n^\ell - \tilde{\mathbf{W}}_{Q,n}^1)_n)) - 3\epsilon - tR} \tag{A47}$$

$$\leq 2^{-t \cdot (t \cdot I(X; \mathbf{Z}(\mathbf{r} - \mathbf{1} \cdot \epsilon))) - 4\epsilon - tR} \tag{A48}$$

$\mathbf{Z}(\cdot)$ is defined as in Equation (25). Equation (A47) follows by the data processing inequality. Equation (A48) follows by combining the entropies into a mutual information, then using the worst additive noise lemma [27]. Thus, if R is chosen less than $I(X; \mathbf{Z}(\mathbf{r} - \mathbf{1} \cdot \epsilon)) - 4\epsilon$ then $P(\mathcal{E}_0 \cup \mathcal{E}_1) \rightarrow 0$ as $T \rightarrow \infty$. Mutual information is lower semi-continuous so $I(X; \mathbf{Z}(\mathbf{r} - \mathbf{1} \cdot \epsilon)) - 4\epsilon$ can be made arbitrarily close to $R_G(\mathbf{r})$. \square

Theorem 6. For a distributed receive system as described in Section 2 with noise covariance matrix Σ and LAN constraint L , a rate $R_{DC}(L)$ is achievable.

Proof. Apply Theorem 7 with $\lambda = 0$ (proof shown below). \square

Theorem 7. For a distributed receive system as described in Section 2 with noise covariance matrix Σ and LAN constraint L then $R_{DC}(L)$ is achievable.

Proof. It is enough to show that any component in the maximization from Equation (36) is achievable. Fix $\lambda > 0$, a helper rate vector $\mathbf{r} \in \mathcal{R}_{LAN}(L)$, and a compression rate vector $\boldsymbol{\rho} \in \mathcal{R}_{DC}^\lambda(\mathbf{r})$. If some $n \in [0 : N]$ has $r_n = 0$ or $\rho_n = 0$, the system is equivalent to the case where the n th helper is not present, so, without loss of generality, we assert that each $r_n, \rho_n > \epsilon$. Fix some rate $R < R_G(\boldsymbol{\rho} - \mathbf{1} \cdot \epsilon) - \lambda$ and a block length $T = t^2 \in \mathbb{N}$. Form a message M uniform on $[1 : 2^{tR}]$.

• **Outline**

Form a rate- R codebook with length- T codewords. At receiver $n \in [0 : N]$, form a length- t regular lattice encoder with lattice chosen coarse enough that its encodings have rate r_n . Randomly bin the encodings down to rate ρ_n . At the base find the codeword jointly typical with the binned encodings. Observe that probability of error is low by recognizing that the unbinned t -length encodings from helpers approximately contain the joint-information content of the helper’s observations plus Gaussian distortion.

• **Transmitter setup**

Generate a codebook mapping $\phi : [1 : 2^{tR}] \rightarrow \mathbb{C}^{T \times 1}$ where

$$\phi(m) = (x_{m,1}, \dots, x_{m,T}) \in \mathbb{C}^{T \times 1} \tag{A49}$$

with each component $x_{m,\ell}$ drawn iid from $\mathcal{CN}(0, 1)$. Reveal ϕ to the transmitter and base.

• **Helper encoder setup**

For each helper n , $n \in [1 : N]$, generate a dither vector $\tilde{W}_{Q,n} = (\tilde{W}_{Q,n}^\ell)_\ell \in (\mathbb{C}^{t \times 1})^t$, where each successive t -segment is uniform in the base region of the Voroni partition \mathcal{P}_n^t of a regular, white t -dimensional lattice \mathcal{L}_n^t which has the following normalized-second-moment:

$$G^t(\mathcal{L}_n^t, \mathcal{P}_n^t) = \frac{\|h_n\|^2 + \Sigma_{n,n}}{2^{\rho_n - \epsilon} - 1}. \tag{A50}$$

These terms are detailed at the beginning of Reference [20]. Reveal $\tilde{W}_{Q,n}$ to the base and helper n . Similarly at the base generate a dither vector $\tilde{W}_{Q,0} = (\tilde{W}_{Q,0}^\ell)_\ell \in (\mathbb{C}^{t \times 1})^t$ with ρ_0 chosen large enough for Lemma 3 to hold.

For each receiver $n \in [0 : N]$, form a random mapping $\text{Index}_n : \mathcal{L}_n^t \rightarrow [1 : 2^{tr_n}]$ which takes on iid values at each $u_n \in \mathcal{L}_n^t$:

$$\text{Index}_n(u_n) \sim \text{Uniform}([1 : 2^{tr_n}]) \tag{A51}$$

Index_n is the binning scheme used by receiver n . Distribute each Index_n to helper n and the base.

• **Transmission**

To send a message $M \in [1 : 2^{TR}]$, have the transmitter broadcast $\mathbf{X} = (X_1, \dots, X_T) = \phi(M)$.

• **Helper encoding**

Helper n ($n \in [0 : N]$) observes a sequence of length T , $\mathbf{Y}_n = (\mathbf{Y}_n^\ell)_\ell \in (\mathbb{C}^{t \times 1})^t$. Now $\mathbf{Y}_n - \tilde{W}_{Q,n} = (\mathbf{Y}_n^\ell - \tilde{W}_{Q,n}^\ell)_\ell \in \mathbb{C}^T = (\mathbb{C}^t)^t$ is composed of t consecutive length- t segments. For the ℓ -th length- t segment, $\ell \in [1 : t]$, form a quantization $U_n^\ell \in \mathcal{L}_n^t$ by finding the point in \mathcal{L}_n^t associated with the region in \mathcal{P}_n^t in which that segment resides. The properties of such U_n are the subject of References [20,29].

Applying Lemma 2, then for large enough t , any $S \subseteq [0 : N]$ has

$$\frac{1}{t} H(U_S^\ell | U_{S^c}^\ell, (\tilde{W}_{Q,n})_{n \in [0:N]}) \xrightarrow{t \uparrow \infty} I(\mathbf{Z}_S(\rho); \mathbf{Y}_S | \mathbf{Z}_{S^c}(\rho)) < \sum_{n \in S} r_n \tag{A52}$$

where the right-hand inequality comes from choice of ρ .

• **Helper joint-compression and forwarding**

At receiver $n \in [0 : N]$, form binned encodings $(V_n^\ell)_{\ell \in [1:t]}$ where for each ℓ , $V_n^\ell = \text{Index}_n(U_n^\ell)$. Note that $|\text{Range}(\text{Index}_n)| \leq 2^{tr_n}$ so $H(V_n^\ell) \leq t \cdot r_n$ and $H((V_n^\ell)_\ell) \leq T \cdot r_n$ so that the LAN constraint is satisfied.

Each receiver forwards $(V_n^\ell)_\ell$ to the base.

• **Decoding**

At the base, receive all the binned helper encodings $\mathbf{V} = ((V_n^\ell)_\ell)_n$.

Take $A_\epsilon^T(\mathbf{X}, \mathbf{U} | \mathbf{W}_Q)$ to be the set of

$$\left(\mathbf{x}_m, (\mathbf{u}^1, \dots, \mathbf{u}^t) \right) \in \text{Range}(\phi) \times \left(\prod_{n=0}^N \mathcal{L}_n^t \right)^t \tag{A53}$$

which are jointly ϵ -weakly-typical with respect to the joint distribution of (\mathbf{X}, \mathbf{U}) , conditional on all the dithers $(\tilde{W}_{Q,n})_{n \in [0:N]}$. Weak- and joint-typicality are defined in Reference [14].

For an ensemble of bin indices from all the receivers, $\mathbf{v} = ((v_n^\ell)_\ell)_n \in \prod_{n=0}^N [1 : 2^{tr_n}]^t$, define:

$$B_v \triangleq \left\{ \mathbf{u} \in \left(\prod_{n=0}^N \mathcal{L}_n^t \right)^t \mid \text{Index}_k(u_k^\ell) = v_k^\ell, k \in [0 : N], \ell \in [1 : t] \right\}. \tag{A54}$$

Each B_v represents the set of helper lattice quantizations represented by the ensemble of helper bin indices v . At the base, find \hat{m} for which there is some $\hat{\mathbf{u}} \in B_V$ where $(\phi(\hat{m}), \hat{\mathbf{u}}) \in A_\epsilon^T(\mathbf{X}, \mathbf{U}|\mathbf{W}_Q)$. Declare the message associated with \hat{x} to be the broadcast.

• **Error analysis**

We have the following error events:

- \mathcal{E}_{JT} : \mathbf{X} is not typical with anything in $B_{V,[0:N]}$.
- $\mathcal{E}_{\hat{m},S}$: For $S \subseteq [0 : N]$, there is some $\hat{m} \neq M$ and $\hat{\mathbf{u}} = (\hat{\mathbf{u}}_S, \hat{\mathbf{u}}_{S^c}) \in B_V$ where $(\phi(\hat{m}), \hat{\mathbf{u}}) \in A_\epsilon^T(\mathbf{X}, \mathbf{U}|\mathbf{W}_Q)$ and $\hat{\mathbf{u}}_{S^c} = \mathbf{U}_{S^c}$, but any $n \in S$ has $\hat{\mathbf{u}}_n \neq \mathbf{U}_n$. $\mathcal{E}_{\hat{m},S}$ denotes the situation where the base identifies the wrong quantizations for all the receivers in S .

By the law of large numbers and properties of typical sets, as t becomes large, eventually $P(\mathcal{E}_{JT}) < \epsilon$. For each $S \subseteq [0 : N]$, take $A_\epsilon^T(\mathbf{U}|\mathbf{W}_Q)$ to be the collection of jointly-typical sequences in $(\prod_n \mathcal{L}_n^t)^t$ up to the joint distribution on $(\mathbf{U}_n^\ell)_{n \in S}$ (any $\ell \in [1 : t]$), conditioned on $(\tilde{\mathbf{W}}_{Q,n})_n$. Now compute:

$$\sum_{\hat{m} \neq M} P(\mathcal{E}_{\hat{m},S}) \leq 2^{TR} \cdot \sum_{\substack{\tilde{\mathbf{u}} \in A_\epsilon^T(\mathbf{U}|\mathbf{W}_Q): \\ \tilde{\mathbf{u}}_{S^c} = \mathbf{U}_{S^c}}} \mathbf{1}_{B_V}(\tilde{\mathbf{u}}) \cdot P\left((\tilde{\mathbf{X}}, \tilde{\mathbf{u}}) \in A_\epsilon^T(\mathbf{X}, \mathbf{U}|\mathbf{W}_Q)\right) \tag{A55}$$

where $\tilde{\mathbf{X}}$ is a random variable independent of and distributed identically to \mathbf{X} . Then, with high probability for large enough t ,

$$\sum_{\hat{m} \neq M} P(\mathcal{E}_{\hat{m},S}) \leq 2^{TR} \cdot \sum_{\substack{\tilde{\mathbf{u}} \in A_\epsilon^T(\mathbf{U}|\mathbf{W}_Q): \\ \tilde{\mathbf{u}}_{S^c} = \mathbf{U}_{S^c}}} \mathbf{1}_{B_V}(\tilde{\mathbf{u}}) \cdot 2^{-t \cdot (I(\mathbf{X}; \mathbf{U} | (\mathbf{W}_{Q,n})_n) - 3\epsilon)} \tag{A56}$$

$$\leq 2^{TR} \cdot \left| \left\{ \tilde{\mathbf{u}} \in A_\epsilon^T(\mathbf{U}|\mathbf{W}_Q) \mid \tilde{\mathbf{u}}_{S^c} = \mathbf{U}_{S^c} \right\} \right| \cdot (1 + \epsilon) \cdot P(\mathbf{U} \in B_V) \cdot 2^{-t \cdot (I(\mathbf{X}; \mathbf{U} | (\mathbf{W}_{Q,n})_n) - 3\epsilon)} \tag{A57}$$

$$\leq 2^{TR} \cdot 2^{t \cdot (H(\mathbf{U}_S | \mathbf{U}_{S^c}, (\mathbf{W}_{Q,n})_n) + \epsilon)} \cdot 2^{-t \cdot (t \cdot \sum_{n \in S} r_n - \epsilon)} \cdot 2^{-t \cdot (I(\mathbf{X}; \mathbf{U} | (\mathbf{W}_{Q,n})_n) - 3\epsilon)} \tag{A58}$$

where Equation (A58) follows by construction of B_V . Taking the log,

$$\log \left(\sum_{\hat{m} \neq M} P(\mathcal{E}_{\hat{m},S}) \right) \tag{A59}$$

$$\leq TR + t \cdot (H(\mathbf{U}_S | \mathbf{U}_{S^c}, (\mathbf{W}_{Q,n})_n) + \epsilon) - t \cdot (t \cdot \sum_{n \in S} r_n - \epsilon) - t \cdot (I(\mathbf{X}; \mathbf{U} | (\mathbf{W}_{Q,n})_n) - 3\epsilon) + \epsilon \tag{A60}$$

$$\leq t \cdot (H(\mathbf{U}_S | \mathbf{U}_{S^c}, (\mathbf{W}_{Q,n})_n) - t \cdot \sum_{n \in S} r_n + tR - I(\mathbf{X}; \mathbf{U} | (\mathbf{W}_{Q,n})_n) + 6\epsilon) \tag{A61}$$

$$\leq t^2 \cdot (I(\mathbf{Y}_S; \mathbf{Z}_S(\rho - \mathbf{1} \cdot \epsilon) | \mathbf{Z}_{S^c}(\rho - \mathbf{1} \cdot \epsilon)) - \sum_{n \in S} r_n + R - I(\mathbf{X}; \mathbf{Z}(\rho - \mathbf{1} \cdot \epsilon))) + 8\epsilon/t. \tag{A62}$$

Equation (A62) follows by using Lemma 2 on $H(\mathbf{U}_S | \mathbf{U}_{S^c})/t$ and using reasoning identical to Equations (A45)–(A48) on $I(\mathbf{X}; \mathbf{U} | (\mathbf{W}_{Q,n})_n)$. Recall that R was chosen so that $R \leq R_G(\rho - \mathbf{1} \cdot \epsilon) - \lambda = I(\mathbf{X}; \mathbf{Z}(\rho - \mathbf{1} \cdot \epsilon)) - \lambda$. Then, Equation (A62) will approach $-\infty$ as $t \rightarrow \infty$ (thereby $P(\cup_{\hat{m},S} \mathcal{E}_{\hat{m},S})$ approaches 0) if all $S \subseteq [0 : N]$ satisfy:

$$I(\mathbf{Y}_S; \mathbf{Z}_S(\rho - \mathbf{1} \cdot \epsilon) | \mathbf{Z}_{S^c}(\rho - \mathbf{1} \cdot \epsilon)) < \lambda + \sum_{m \in S} r_m - 8\epsilon. \tag{A63}$$

By assumption that $\rho \in \mathcal{R}_{DC}^\lambda(\mathbf{r})$ and for small enough ϵ , then Equation (A63) holds for each S . Since all error events approach 0, then a rate of $R_G(\rho - \mathbf{1} \cdot \epsilon) - \lambda$ is achievable. For small enough ϵ , by lower semi-continuity of mutual information $R_G(\rho - \mathbf{1} \cdot \epsilon) - \lambda$ can be made arbitrarily close to $R_G(\rho) - \lambda$. \square

Appendix C. Proof of Conditional Capacity

Remark 7. The capacity of the system is $R_{\overline{DC}}(L)$ under the following restrictions:

- Σ 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.
- Construction of helper messages is independent of the transmitter’s codebook \mathcal{X} .

Proof. By Theorem 5 in Reference [1] (variable names altered, constants adapted for complex variables), the capacity of the system under the assumed restrictions is

$$\max_{\substack{r \in \mathcal{R}_{LAN}(L), \\ \mathbf{V} \in \mathcal{V}}} C_{r,\mathbf{V}} \tag{A64}$$

$$C_{r,\mathbf{V}} \triangleq \min_{S \subseteq [1:N]} \left\{ \sum_{m \in S} [r_m - I(V_m; Y_m | X)] + I(\mathbf{V}_S; X) \right\}. \tag{A65}$$

In Equation (A65), \mathcal{V} is the collection of random vectors $\mathbf{V} = (V_1, \dots, V_N)$ whose components are of the form $V_n = Y_n + \tilde{W}_n$ with independently distributed \tilde{W}_n :

$$\tilde{W}_n \sim \mathcal{CN} \left(0, \sigma_n^2 \cdot \frac{2^{-2v_n}}{1 - 2^{-2v_n}} \right) \tag{A66}$$

for any $v_n \geq 0$. Since both the variance in Equation (A66) as a function of v_n and Equation (26) as a function of r_n are injective on $(0, \infty)$, then for any $\mathbf{V} \in \mathcal{V}$ there is some ρ (in the context of Equation (37); either inside or outside $\cup_{\lambda \in \mathbb{R}} \mathcal{R}_{DC}^\lambda(L)$) which will yield a variable $\mathbf{Z}(\rho)$ with identical distribution to \mathbf{V} .

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

$$C_{r,\mathbf{V}} = \max \left\{ c : \forall S \subseteq [1 : N], \quad c \leq I(\mathbf{V}_S; X) + \sum_{m \in S} r_m - I(V_m; Y_m | X) \right\} \tag{A67}$$

$$\leq \max \left\{ c : \forall S \subseteq [1 : N], \quad c - I(\mathbf{V}_S; X) + I(\mathbf{V}_S; \mathbf{Y}_S | X) \leq \sum_{m \in S} r_m \right\} \tag{A68}$$

$$\leq \max \left\{ c : \forall S \subseteq [1 : N], \quad [c - I(X; \mathbf{V})] + I(\mathbf{Y}_S; \mathbf{V}_S | \mathbf{V}_S) \leq \sum_{m \in S} r_m \right\} \tag{A69}$$

$$= \max_{\lambda \in \mathbb{R}} \max \left\{ I(X; \mathbf{V}) - \lambda : \forall S \subseteq [1 : N] I(\mathbf{Y}_S; \mathbf{V}_S | \mathbf{V}_S) \leq \sum_{m \in S} r_m + \lambda \right\} \tag{A70}$$

where $I(\mathbf{V}_S; \mathbf{Y}_S | X) \leq \sum_{m \in S} I(V_m; Y_m | X)$ follows because, for each n , $V_n \leftrightarrow (X, Y_n) \leftrightarrow (\mathbf{V}_{\{j \neq n\}}, \mathbf{Y}_{\{j \neq n\}})$ is a Markov chain. Maximizing over r , Equation (A70) is identical to the set of rates shown in Theorem 7. \square

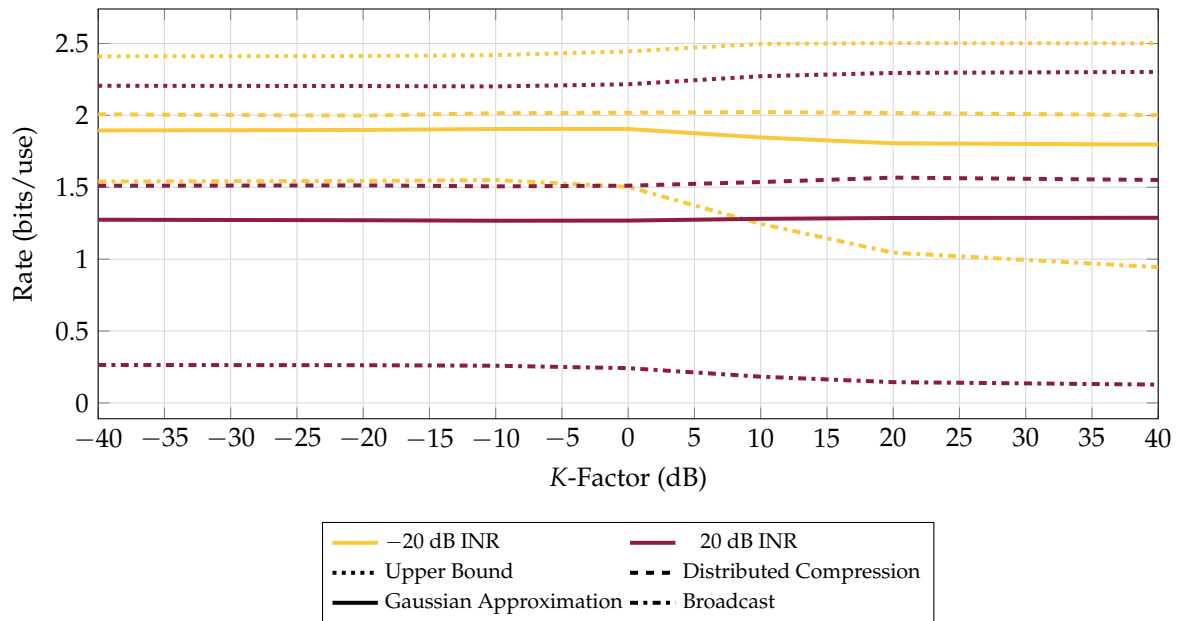


Figure A1. Achievable rates versus K . Four helpers with an observation available at the base and LAN constraint $L = 5$ bits shared among helpers. Averaged over 500 channel realizations with uniform independent phase and average receiver SNR fixed at 0 dB. Scattering environment has negligible impact on achievable rates for these bounds, except in the Broadcast bound where receive diversity is not well utilized. Similar results are obtained when the base does not have its own observation.

References

1. Sanderovich, A.; Shamai, S.; Steinberg, Y.; Kramer, G. Communication via Decentralized Processing. *IEEE Trans. Inf. Theory* **2008**, *54*, 3008–3023.
2. Sanderovich, A.; Shamai, S.; Steinberg, Y. On Upper bounds for Decentralized MIMO Receiver. In Proceedings of the 2007 IEEE International Symposium on Information Theory, Nice, France, 24–29 June 2007; pp. 1056–1060.
3. Chapman, C.; Margetts, A.R.; Bliss, D.W. Bounds on the Information Capacity of a Broadcast Channel with Quantizing Receivers. In Proceedings of the IEEE 2015 49th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 8–11 November 2015; pp. 281–285.
4. Kramer, G.; Gastpar, M.; Gupta, P. Cooperative Strategies and Capacity Theorems for Relay Networks. *IEEE Trans. Inf. Theory* **2005**, *51*, 3037–3063.
5. Bliss, D.W.; Kraut, S.; Agaskar, A. Transmit and Receive Space-Time-Frequency Adaptive Processing for Cooperative Distributed MIMO Communications. In Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, 25–30 March 2012; pp. 5221–5224.
6. Quitin, F.; Irish, A.; Madhow, U. Distributed Receive Beamforming: A Scalable Architecture and Its Proof of Concept. In Proceedings of the IEEE 77th Vehicular Technology Conference (VTC Spring), Dresden, Germany, 2–5 June 2013; pp. 1–5.
7. Vagenas, E.D.; Karadimas, P.; Kotsopoulos, S.A. Ergodic Capacity for the SIMO Nakagami-M Channel. *EURASIP J. Wirel. Commun. Netw.* **2009**, *2009*, 802067.
8. Love, D.J.; Choi, J.; Bidigare, P. Receive Spatial Coding for Distributed Diversity. In Proceedings of the 2013 Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 3–6 November 2013; pp. 602–606.
9. Choi, J.; Love, D.J.; Bidigare, P. Coded Distributed Diversity: A Novel Distributed Reception Technique for Wireless Communication Systems. *IEEE Trans. Signal Process.* **2015**, *63*, 1310–1321.
10. Brown, D.R.; Ni, M.; Madhow, U.; Bidigare, P. Distributed Reception with Coarsely-Quantized Observation Exchanges. In Proceedings of the 2013 47th Annual Conference on Information Sciences and Systems (CISS), Altmore, MD, USA, 20–22 March 2013; pp. 1–6.

11. Berger, T.; Zhang, Z.; Viswanathan, H. The CEO Problem [Multiterminal Source Coding]. *IEEE Trans. Inf. Theory* **1996**, *42*, 887–902.
12. Tavildar, S.; Viswanath, P.; Wagner, A.B. The Gaussian Many-Help-One Distributed Source Coding Problem. *IEEE Trans. Inf. Theory* **2010**, *56*, 564–581.
13. Yedla, A.; Pfister, H.D.; Narayanan, K.R. Code Design for the Noisy Slepian-Wolf Problem. *IEEE Trans. Commun.* **2013**, *61*, 2535–2545.
14. Cover, T.M.; Thomas, J.A. *Elements Information Theory*; John Wiley & Sons: Hoboken, NJ, USA, 2012.
15. Shannon, C.E. A Mathematical Theory of Communication. *ACM SIGMOBILE Mob. Comput. Commun. Rev.* **2001**, *5*, 3–55.
16. Telatar, I.E. Capacity of Multi-Antenna Gaussian Channels. *Eur. Trans. Telecommun.* **1999**, *10*, 585–595.
17. Gamal, A.E.; Kim, Y. *Network Information Theory*; Cambridge University Press: Cambridge, UK, 2011.
18. Weingarten, H.; Steinberg, S.; Shamai, S. The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel. *IEEE Trans. Inf. Theory* **2006**, *52*, 3936–3964.
19. Wei, Y.; Cioffi, J.M. Sum Capacity of Gaussian Vector Broadcast Channels. *IEEE Trans. Inf. Theory* **2004**, *50*, 1875–1892.
20. Zamir, R.; Feder, M. On lattice Quantization Noise. *IEEE Trans. Inf. Theory* **1996**, *42*, 1152–1159.
21. Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: Cambridge, UK, 2004.
22. Slepian, D.; Wolf, J.K. Noiseless Coding of Correlated Information Sources. *IEEE Trans. Inf. Theory* **1973**, *19*, 471–480.
23. Costa, M. Writing on Dirty Paper (corresp.). *IEEE Trans. Inf. Theory* **1983**, *29*, 439–441.
24. Goldsmith, A. *Wireless Communications*; Cambridge University Press: Cambridge, UK, 2005.
25. Bliss, D.W.; Govindasamy, S. *Adaptive Wireless Communications: MIMO Channels and Networks*; Cambridge University Press: Cambridge, UK, 2013.
26. Jones, E.; Oliphant, T.; Peterson, P. SciPy: Open Source Scientific Tools for Python. 2001. Available online: <http://www.scipy.org/> (accessed on 10 September 2017).
27. Diggavi, S.N.; Cover, T.M. The Worst Additive Noise Under a Covariance Constraint. *IEEE Trans. Inf. Theory* **2001**, *47*, 3072–3081.
28. Ihara, S. On the Capacity of Channels with Additive Non-Gaussian Noise. *Inf. Control* **1978**, *37*, 34–39.
29. Zamir, R.; Feder, M. On Universal Quantization by Randomized Uniform/Lattice Quantizers. *IEEE Trans. Inf. Theory* **1992**, *38*, 428–436.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).