

Distributed Recovery of a Gaussian Source in Interference with Successive Lattice Processing

Christian Chapman ^{1,*}, Matthew Kinsinger ¹, Ameya Agaskar ² and Daniel W. Bliss ¹

¹ School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85281, USA

² MIT Lincoln Laboratory, Lexington, MT 02420, USA

* Correspondence: cdchapm2@asu.edu

Received: 6 August 2019; Accepted: 27 August 2019; Published: 30 August 2019



Abstract: A scheme for recovery of a signal by distributed listeners in the presence of Gaussian interference is constructed by exhausting an “iterative power reduction” property. An upper bound for the scheme’s achieved mean-squared-error distortion is derived. The strategy exposes a parameter search problem, which, when solved, causes the scheme to outperform others of its kind. Performance of a blocklength-one scheme is simulated and is seen to improve over plain source coding without compression in the presence of many interferers, and experiences less outages over ensembles of channels.

Keywords: network information theory; distributed source coding; lattice codes

1. Introduction

Consider a situation where multiple users have the common goal of recovering a broadcast in the presence of much stronger interfering signals. If the listening nodes are able to form a local area network (LAN) among one another, the nodes can mitigate interference by sharing information through the LAN and processing the messages they share. This scenario could arise, for instance, when a network of co-located nodes must aid a neighbor’s reception, when a group of cellular nodes have the common goal of receiving a code-domain-multiple-access transmission in a crowded environment, or when an adversary jams a broadcast node meant to serve multiple users. A high level model of this scenario is shown in Figure 1. When the amount of information transmission among listeners is limited (as it well may be if they must conserve power or bandwidth), they must take care to only forward novel information to their neighbors. This paper presents a lattice-algebra-based structured coding scheme for the observers in this scenario in an additive Gaussian noise channel with additive Gaussian interferers.

The strategy presented allows for strong recovery of the signal of interest in terms of mean-squared-error, even when the amount of information traded over the LAN is limited. It compares favorably to previous structured coding strategies of its kind, but still has a performance gap to random coding bounds in some regimes.

The scheme is stated in terms of a reduced version of the system where instead of communicating to one another, all listeners forward digital messages over a reliable LAN link to a base node which processes the messages into an estimate of the signal of interest. The LAN is modeled such that, for each channel use, each receiver forwards a fixed number of bits to the base. The LAN may be such that each receiver forwards to the base at the same rate, or the LAN may supports any receiver-to-base rates such that their sum-rate is below some maximum total throughput. This may arise if the LAN link is, e.g., an out-of-band frequency-domain-multiple-access channel where redistribution of receiver-to-base bitrates corresponds to redistribution of bandwidth resources. Both these situations are tested in

Section 5. A more technical version of the model in Figure 1 is shown in Figure 2. The strategy is referenced in this paper as “Successive Integer-Forcing Many-help-one” (SIFM).

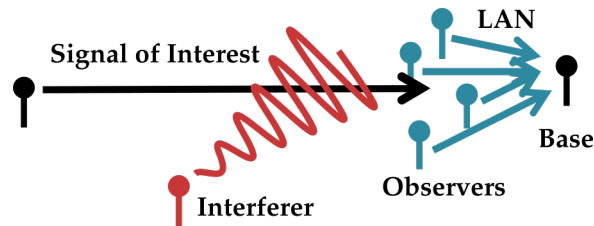


Figure 1. High-level view of the system. A base node seeks to recover a signal of interest to low mean-squared-error distortion in the presence of interference. The base receives information about the signal from several observer nodes. Observers encode their observations separately and forward a limited amount of information to the base over a LAN. The base uses this information to estimate the signal of interest. Note that in some scenarios the base could actually be a virtual node, representing all the messages any single observer has received over the LAN.

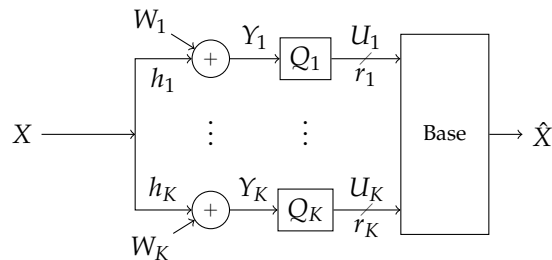


Figure 2. Block diagram of the system. A signal X is broadcast through an additive-white-Gaussian-noise channel in the presence of Gaussian interference, creating correlated additive noise (W_1, \dots, W_K) . The signal is observed at K receivers labeled Q_1, \dots, Q_K . The k th receiver observes Y_k and processes its observation into a rate- r_k message U_k . The messages are forwarded losslessly over a LAN to a base which processes the messages into an estimate of the signal, \hat{X} .

1.1. Relation to Other Literature

This is a strategy for the *Gaussian many-help-one source coding problem* [1], allowing for the case where the “receiver being helped” cannot provide side information to the decoder. The problem is also similar in structure to the CEO problem [2], but most studies on the CEO problem do not model correlated noise across observers. The rate-distortion region of this problem is achieved by random coding for some covariance structures [1], but is unknown in general. The rate-distortion region for the case of two helpers is known within bounds [3]. Unstructured coding schemes for the problem have long been known, although, given some results Wagner [4], Nazer and Gastpar [5] and Krithivasan and Pradhan [6], it seems likely that more structured coding strategies can outperform them in general.

The asymptotic version of the scheme is a direct application of general results and ideas from [7] to the many-help-one problem, although derivation differs, yielding a different characterization of its achieved rate region in terms of Algorithm 1. There is a lot of existing literature on lattice signal processing for this scenario, and similar techniques have been applied to closely related problems. The procedures which comprise SIFM have also been described notionally [8] and are contained within more general multiple-input-multiple-output (MIMO) communication studies [9] (Chapter 12). The authors are not aware of any direct presentation of this strategy for this problem until this work.

Integer Forcing Source Coding (IFSC) [8] is closely related to the present scheme. Further analytic properties of IFSC, simulations and robust configuration strategies for such codes are investigated in [10,11]. IFSC is extended to use the exhaustive power-reduction property in [12]. The studies just listed examine lattice techniques for a slightly different problem: design a coding strategy for the decoder to recover an estimate of all the observers’ signals, not just a single component of interest. Relaxing this constraint allows for a larger rate region for the many-help-one problem investigated

here. Distributed-source-coding schemes such as IFSC and its related strategies provide weak solutions for the present problem as they intrinsically involve recovery of superfluous information. Likewise, the present scheme performs poorly for the IFSC distributed-source-coding problem as it produces a rank-deficient estimate of the sources.

SIFM is a strict generalization of a scheme by Krithivasan and Pradhan (KP) [13]. Although conceptually very similar, the generalizations of SIFM over KP necessitate a near ground-up re-description of the strategy. This parameterization of the KP scheme in terms of SIFM is shown in Appendix D. Unfortunately, SIFM replaces KP's formulaic choice of certain parameters with a difficult continuous-domain nonlinear parameter search problem described in Section 5.2. This comes with the benefit of SIFM reliably outperforming KP, as shown in Section 5.

The lattice messages involved in such schemes are still non-trivially jointly distributed and could be further compressed to yield performance improvements Wagner [4], Yang and Xiong [14], Yang and Xiong [15] and Cheng et al. [16]. Most results on this correlation [4,14] focus on the case of two observers, and the compression described in [16] is inexhaustive. SIFM suffers from the same problem, and the amount of redundancy still present in observer messages has not yet been totally characterized. Analysis of the network scheme in the method presented here enables further study of these encoding redundancies. This will be shown in a forthcoming document.

1.2. Contributions

- A strategy for recovery of a Gaussian source using side information from many receivers in the presence of interference is presented in Section 3. This scheme outperforms others of its kind by exhausting a power-reduction property.
- An asymptotic version of the scheme is presented in Section 4, yielding Theorem 1.
- An algorithm for calculating the asymptotic scheme's performance in Theorem 1 is presented as Algorithm 1.

1.3. Outline

The problem is set up analytically in Section 2. A table of symbols and notation is included in Table 1. Some algebraic properties of lattices are required to describe the strategy, and are reviewed briefly in Section 3. The scheme is described in terms of successful decoding events, which are constructed for arbitrary lattices in Definition 3. The definition yields a performance upper bound in Lemma 2. In limit with blocklength and certain choice of lattices, Lemma 2 yields an asymptotic performance in Corollary 1. This bound, along with several others, is plotted over various regimes in Section 5.

Table 1. Symbols and notation.

$[n]$	Whole numbers from 1 to n
A, a, \vec{a}, \vec{A}	Matrix, column vector, vector, random vector
A^\dagger, a^\dagger	Hermitian conjugate
$\text{diag } \vec{a}$	Square diagonal matrix with diagonals \vec{a}
$\text{pinv}(\cdot)$	Moore-Penrose pseudoinverse
$\mathbb{P}(\cdot)$	Current context's unconditioned probability measure
$\mathcal{N}(0, \Sigma)$	Normal distribution with zero mean, covariance Σ
$X \sim f$	X is a random variable distributed like f
$\text{var}(X)$	Variance (or covariance matrix) of X
$X^n, f(x^n)$	Vector of n independent trials of a random variable distributed like X , a function whose input is intended to be such a variable
$\text{var}(a b)$	Conditional variance (or covariance matrix) of a given observation b
$\mathbb{E}[a b]$	Conditional expectation of a given observations b
$\text{round}_L(\cdot), \text{mod}_L(\cdot)$	Lattice round, modulo to a lattice L (when it is clear what base region is associated with L) (see Section 3.1).
$\text{mod}(\cdot)$	Lattice modulo to the current context's coarse lattice L_c 's base region B_c .

2. Problem Setup

A Gaussian source X is observed through a static channel by K receivers in the presence of additive Gaussian interference and noise, represented by a conglomerate term W_k . At each discrete time, receiver $k \in [K]$ observes a real sample:

$$Y_k := h_k X + W_k \quad (1)$$

where $(h_1, \dots, h_K) = \vec{h} \in \mathbb{R}^k$, $(W_1, \dots, W_K) = \vec{W} \sim \mathcal{N}(0, \Sigma_W)$, $X \sim \mathcal{N}(0, 1)$ and $X \perp \vec{W}$. The above trial is repeated identically and independently n times to form time-expanded variables $X^n \in \mathbb{R}^n$, $\tilde{Y}^n, \tilde{W}^n \in (\mathbb{R}^n)^K$. A base node seeks to recover X^n to low mean-squared-error.

2.1. Encoder

Receiver rates $r_1, \dots, r_K > 0$ are fixed. Receiver $k \in [K]$ has an *encoder* $\text{enc}_k : \mathbb{R}^n \rightarrow [2^{nr_k}]$, which forms an *encoding* $U_k = \text{enc}_k(Y_k^n)$. Each receiver forwards its encoding to the base perfectly.

2.2. Decoder

The base has a *decoder* $\text{dec} : \prod_k [2^{nr_k}] \rightarrow \mathbb{R}^n$ which processes the encodings $(U_1, \dots, U_K) = \vec{U}$ into an estimate $\hat{X}^n := \text{dec}(\vec{U})$.

2.3. Achievable Distortion

For rates $(r_k)_{k \in [K]}$ and noise covariance Σ_W , a distortion d^2 is *achievable* if there is a collection of encoders-and-decoder pairs $((\text{enc}_k)_{k \in [K]}, \text{dec})$, the collection indexed by, say, ℓ , where $\inf_{\ell} \frac{1}{n} \mathbb{E} [\|X^n - \hat{X}^n\|^2] \leq d^2$.

3. Successive Integer Forcing Many-Help-One Scheme

A block diagram of the scheme is shown in Figure 4. Broadly, the strategy operates as follows. *Design:* First, a blocklength n is chosen and a “coarse lattice” $L_c \subset \mathbb{R}^n$ is chosen according to some design to be specified. One assumes the LAN is established and allows for reliable communication from receivers to the base node, each k th receiver at some rate r_k . Given these rates, a “fine lattice” $L_k \supset L_c$ is chosen for each k th receiver according to some design to be specified. At each k th receiver, a quantization dither $W_{d,k}$ is selected randomly uniformly over the base region of the receiver’s fine lattice, L_k .

Quantization in this scheme involves a dither term. The dither causes quantization noise to manifest as additive, independent of the input signal, and uniform over the base of L_k . These properties are demonstrated via the *crypto lemma* [9] (Theorem 4.1.1). Use of dither is described in detail in the Appendix proofs and must be included here for a complete description of the scheme’s operation but is inessential to an initial broad understanding of the scheme.

One assumes that the covariance between the transmitter’s signal and all the receivers’ observations is known at all the receivers and stable over n observations. As a function of rates $r_1, \dots, r_K > 0$ and the channel covariance matrix, some scale parameters $\alpha_1, \dots, \alpha_K > 0$ are designed.

Operation:

1. Receiver k , labeled Q_k in Figure 4, normalizes all its observed sequence of n samples by $\alpha_k / \sqrt{\text{var } Y_k}$ so that expected-power-per-sample is 1.
2. Receiver k quantizes the result from the previous step by adding dither $W_{d,k}$, then rounding the result onto a nearby point on its fine lattice L_k .
3. Receiver k takes the modulo of the rounded result onto L_c , producing a point in L_c ’s “modulo-space”. L_k must be designed with respect to L_c such that the result of this step has entropy-rate less than r_k . Receiver k forwards this result to the base. See Figure 3 for an example design of L_c and L_k for blocklength $n = 2$ and receiver rate $r_k = 2$ bits per sample.

4. The decoder, labeled “base” in Figure 4, receives all the receiver messages, each message being some point in L_c -modulo-space, and removes the dither by subtracting the chosen dithers from each message, and taking the L_c -modulo of the result.
5. The decoder recombines them in a way that the result is no longer in modulo-space, but some linear combination of the quantizations.
6. The decoder recombines this recovered component from all the original modulo-space messages to produce new modulo-space points with the just-recovered component removed. The removal process is illustrated in Figure 6.
7. This removal allows for new recombinations to allow recovery of different components. This process is repeated until no more components can be recovered. All repetitions of Steps 5–7 are represented as the block labeled “Dec” in Figure 4.
8. All recovered components are used to estimate the source.

This is described in full precision below. To simplify exposition, from here on, assume $\text{var } Y_1 = \dots = \text{var } Y_K = 1$ so that normalization is not necessary.

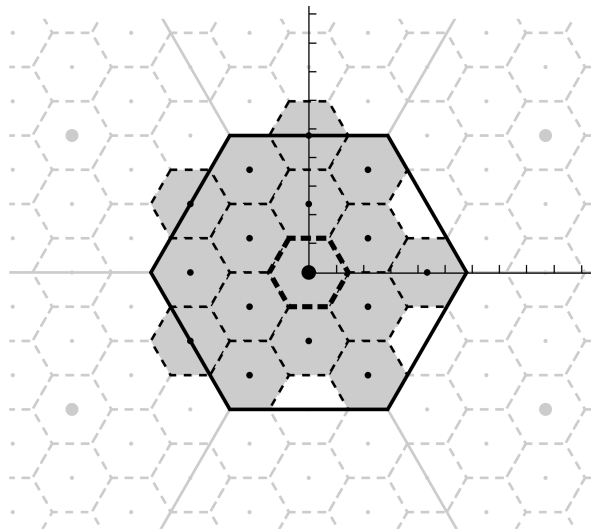


Figure 3. An example of nested lattices in \mathbb{R}^2 . Points from a coarse lattice L_c are drawn as large dots. Points from a fine lattice $L_k \supset L_c$ besides those in L_c are drawn as small dots. A base region B_c for L_c and its translates $B_c + L_c$ are outlined by solid lines, and a base region B_k for L_k and its translates $B_k + L_k$ are outlined by dashed lines. The base regions for points in $L_k \cap B_c$ are shaded. Notice that $L_k \cap B_c$ has 2^4 points, so 4 bits are required to describe a general point in $L_k \cap B_c$, or 2 bits per sample.

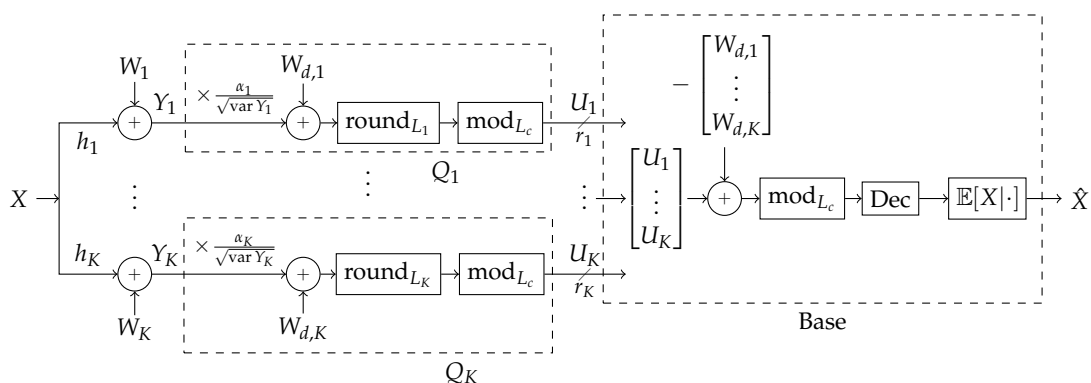


Figure 4. Block diagram of strategy, expanded from general block diagram in Figure 2. See the beginning of Section 3 for details.

3.1. Background on Lattices

Before the scheme can be described, some preliminary definitions and statements are needed.

Definition 1. A lattice $L \subset \mathbb{R}^n$ is an infinite set of discrete points closed under addition and subtraction.

Definition 2. A region $s \subset \mathbb{R}^n$ is a base region for a lattice L if the sets $(l + s)_{l \in L}$ are all disjoint, their union forms \mathbb{R}^n , and if s has its moment at the origin.

One can define modulo and rounding operations relative to a base region s for a lattice L :

$$\text{round}_s : \mathbb{R}^n \rightarrow L, \quad (2)$$

$$\text{round}_s(x) := l \in L \text{ where } (x - l) \in s, \quad (3)$$

$$\text{mod}_s : \mathbb{R}^n \rightarrow s, \quad (4)$$

$$\text{mod}_s(x) := x - \text{round}_s(x). \quad (5)$$

Note that round_s is well defined since s being a *base region* implies there is one and only one lattice point l that satisfies its prescription. These operations have some useful properties. A graphical example of the lattice round and modulo operations for a lattice in \mathbb{R}^2 and a hexagonal base region s is shown in Figure 5.

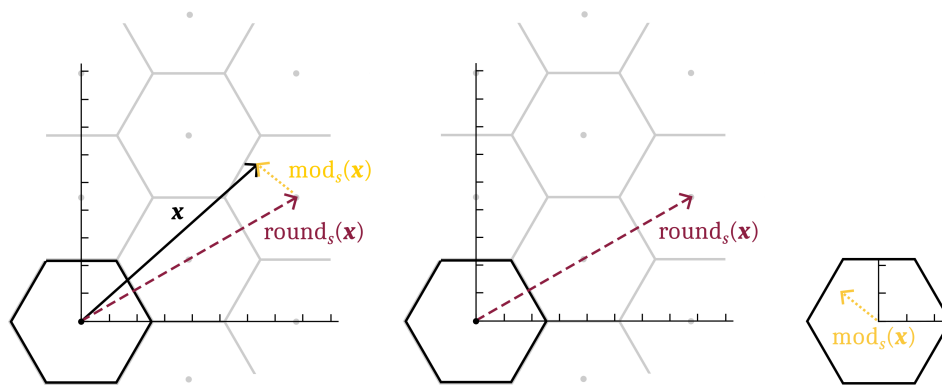


Figure 5. A graphical example of lattice modulo and rounding operations on a vector $x \in \mathbb{R}^2$ for a particular lattice (drawn as dots) and hexagonal base region s (outlined in black, translates outlined by solid gray lines). Notice $x = \text{round}_s(x) + \text{mod}_s(x)$, that the image of $\text{round}_s(\cdot)$ is the lattice, and that the image of $\text{mod}_s(\cdot)$ is s .

Property 1. (*mod is the identity within s*) $x \in s$ has $\text{mod}_s(x) = x$.

Proof. $\text{mod}_s(x) = x - \text{round}_s(x) = x$. \square

Property 2. $x \in \mathbb{R}^n, l \in L$ have $\text{round}_s(x + l) = \text{round}_s(x) + l$.

Proof. $(x + l) - (\text{round}_s(x) + l) = (x - \text{round}_s(x)) \in s$. Thus, by definition, $\text{round}_s(x + l) = \text{round}_s(x) + l$. \square

Property 3. $x \in \mathbb{R}^n, l \in L$ have $\text{mod}_s(x + l) = \text{mod}_s(x)$.

Proof. $\text{mod}_s(x + l) = (x + l) - \text{round}_s(x + l) = (x + l) - \text{round}_s(x) - l = x - \text{round}_s(x) = \text{mod}_s(x)$. The second equality is by Property 2. \square

Property 4. (Lattice modulo is distributive) $x, y \in \mathbb{R}^n$ have:

$$\text{mod}_s(\text{mod}_s(x) + y) = \text{mod}_s(x + y). \quad (6)$$

Proof. $\text{mod}_s(\text{mod}_s(x) + y) = \text{mod}_s(x + y - \text{round}_s(x)) = \text{mod}_s(x + y)$. The last equality is by Property 3. \square

An example of how Property 4 can be used is shown in Figure 6. The efficacy of the lattice strategy presented in this paper is in exhaustive use of this technique.

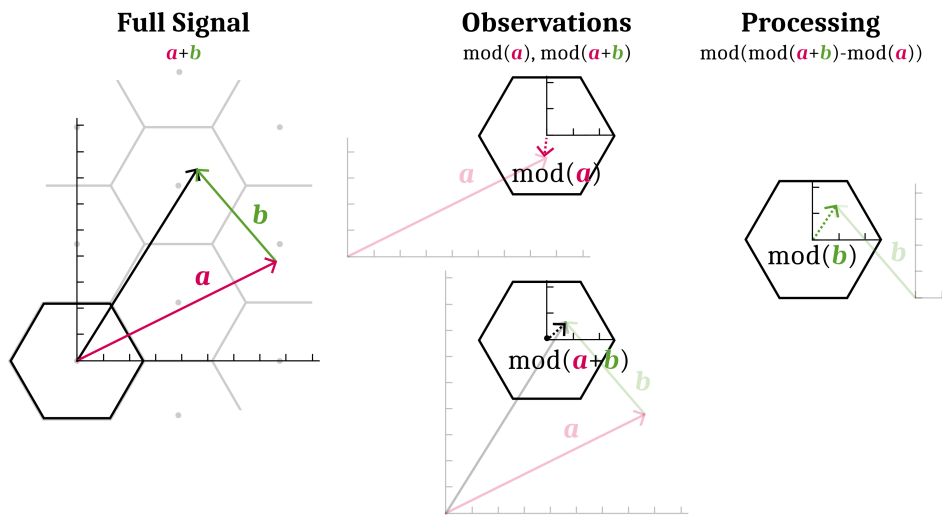


Figure 6. Removal of a vector component a from modulo $\text{mod}(a+b)$. Drawn in the same context as Figure 5. The full signal, $a+b$ without any modulo operations, is plotted on the left. Say that available for processing are the vectors $\text{mod}(a)$ and $\text{mod}(a+b)$, illustrated in the middle. A particular processing of $\text{mod}(a)$ and $\text{mod}(a+b)$ produces $\text{mod}(b)$, shown on the right. Non-modulo components are drawn lightly under each modulo.

3.2. Lattice Scheme Description

Fix the following parameters:

- Blocklength n
- Receiver scales $\vec{\alpha} = (\alpha_1, \dots, \alpha_K) \in \mathbb{R}^K$
- “Fine” lattices $L_k \subset \mathbb{R}^n$ each with a base region $B_k \subset \mathbb{R}^n$, $k \in [K]$
- A “coarse” lattice $L_c \subset \mathbb{R}^n$ with a base region B_c , where $L_k \supset L_c$ and $\frac{1}{n} \log |B_c \cap L_k| \leq r_k$ for each $k \in [K]$.
- Functions $\phi_k : (B_c \cap L_k) \rightarrow [2^{nr_k}]$ which enumerate their domain’s points
- Dither variables $\tilde{W}_d = (W_{d,1}, \dots, W_{d,K})$, with $W_{d,k} \sim \text{unif } B_k$ independent over k

For brevity we neglect to denote application of the enumeration ϕ_k and its inverse when it is clear from context where it should be applied.

3.2.1. Encoders

Each encoder first scales its observation then quantizes the result by rounding with dither onto its fine lattice L_k . This discretizes the source’s observation onto a countable collection of points, but it may still be too high-rate to forward to the base directly. The encoder wraps the discretization onto the coarse lattice’s base region B_c by applying mod_{B_c} . The domain reduction from all of L_k to only points within $B_c \cap L_k$ reduces the discretization’s entropy enough to forward it to the decoder.

Construct the encoder for receiver k as $\text{enc}_k : \mathbb{R}^n \rightarrow [2^{nr_k}]$:

$$\text{enc}_k(y_k^n) := \phi_k(\text{mod}_{L_c}(\text{round}_{L_k}(\alpha_k y_k^n + W_{d,k}))) \quad (7)$$

3.2.2. Decoders

The decoder produces estimates of particular integer linear combinations $\mathbf{a}_1, \dots, \mathbf{a}_K$ of the source observations by processing the encodings in stages. In stage k , the decoder recovers combination \mathbf{a}_k . In all future stages, the \mathbf{a}_k component is used to aid recovery. This is described in more detail below.

To construct each stage, it is necessary to describe the covariance between receiver quantizations. By [9] (Theorem 4.1.1),

$$\text{round}_{L_k}(\alpha_k Y_k^n + W_{d,k}) - W_{d,k} = \alpha_k Y_k^n - \tilde{W}_{d,k} \quad (8)$$

and $\tilde{W}_{d,k} \sim \text{unif } B_k$ is independent of Y_k^n . Denote $\tilde{W}_d = (\tilde{W}_{d,1}, \dots, \tilde{W}_{d,K})$. In addition, denote the scaled receiver observations as $\vec{\alpha Y} = (\alpha_1 Y_1, \dots, \alpha_K Y_K)$, and similarly for time-expanded \vec{Y}^n . Then, on average over time, the receivers' dithered rounding to their fine lattices effectively adds noise of the following covariance to the observations:

$$\Sigma_Q := \text{diag} \left(\left(\mathbb{E} \left[\frac{1}{n} \|\tilde{W}_{d,k}\|^2 \right] \right)_{k \in [K]} \right). \quad (9)$$

A covariance matrix \mathbf{C} between the sources after dithered rounding to the fine lattices can be written as the time-averaged covariance of vectors in Equation (8):

$$\mathbf{C} := \mathbb{E} \left[\frac{1}{n} (\vec{\alpha Y}^n - \tilde{W}_d)^\dagger (\vec{\alpha Y}^n - \tilde{W}_d) \right]. \quad (10)$$

$$= (\text{diag } \vec{\alpha}) \Sigma (\text{diag } \vec{\alpha})^\dagger + \Sigma_Q. \quad (11)$$

Now, some events over outcomes of $(\vec{\alpha Y}^n - \tilde{W}_d)$ are constructed.

Definition 3. For some $K' \in [K]$, fix an integer matrix $\mathbf{A} \in \mathbb{Z}^{K \times K'}$, call its columns $\mathbf{a}_1, \dots, \mathbf{a}_{K'} \in \mathbb{Z}^K$, and take \mathbf{A}_k to be the first k columns of \mathbf{A} .

In terms of \mathbf{A} , define matrices for each $k = 1, \dots, K'$:

$$\mathbf{S}_k(\mathbf{v}) := \arg \min_{\mathbf{u} \in \mathbb{R}^k} \text{var} \left((\mathbf{v} - \mathbf{A}_k \mathbf{u})^\dagger (\vec{\alpha Y}^n - \tilde{W}_d) \right) \quad (12)$$

$$\mathbf{R}_k(\mathbf{v}) := [\mathbf{I}_{K \times K} - \mathbf{A}_k \mathbf{S}_k] \mathbf{v}. \quad (13)$$

The arg min above derived in closed form in Appendix C.

In addition, in terms of \mathbf{A} , define events $\mathcal{M}(\mathbf{A}_1), \dots, \mathcal{M}(\mathbf{A}_{K'}) \subset (\mathbb{R}^n)^K$:

$$\mathcal{M}(\mathbf{A}_1) := \{\mathbf{z} \in \mathbb{R}^{K \times n} : \mathbf{a}_1^\dagger \mathbf{z} \in B_c\}, \quad (14)$$

$$\mathcal{M}(\mathbf{A}_k) := \mathcal{M}(\mathbf{A}_{k-1}) \cap \left\{ \mathbf{z} \in \mathbb{R}^{K \times n} : (\mathbf{R}_{k-1} \mathbf{a}_k)^\dagger \mathbf{z} \in B_c \right\}. \quad (15)$$

The events designate when a particular processing of the encodings successfully produces an estimate of the observations *without modulo*:

Lemma 1. Fix $\mathbf{A} \in \mathbb{Z}^{K \times K'}$ as in Definition 3. Take $U_k = \text{enc}_k(Y_k^n)$ for $k \in [K]$. There is a function f where $f(\vec{U}) = \mathbf{A}^\dagger (\vec{\alpha Y}^n - \tilde{W}_d)$ whenever $(\vec{\alpha Y}^n - \tilde{W}_d) \in \mathcal{M}(\mathbf{A})$.

The functions $f_1, \dots, f_{K'}$ for each output of f , $f_{[k]}$ denoting the first k outputs, are given as follows, all mod taken with respect to coarse base region B_c .

$$f_1(\vec{U}) := \text{mod} \left(\mathbf{a}_1^\dagger (\vec{U} - \tilde{W}_d) \right), \quad (16)$$

$$\begin{aligned} f_m(\vec{U}) := & \text{mod} \left\{ \mathbf{a}_m^\dagger (\vec{U} - \tilde{W}_d) - \dots \right. \\ & \left. \text{mod} \left([\mathbf{S}_{m-1} \mathbf{a}_m]^\dagger f_{[m-1]}(\vec{U}) \right) \right\} + \dots \\ & [\mathbf{S}_{m-1} \mathbf{a}_m]^\dagger f_{[m-1]}(\vec{U}). \end{aligned} \quad (17)$$

The proof is delayed to Appendix A.

This aspect of general reuse of all previously recovered components for the recovery of a new one is the source of benefit of the present scheme over [13]. Comments in [9] among other places describe such a strategy.

A decoder can be realized from each event from Definition 3 by using a linear estimator on the output of f gotten from that event in Lemma 1. The likelihood of each event is quite sensitive to channel covariance Σ , receiver scalings $\vec{\alpha}$ and integer vector \mathbf{a}_ℓ . As layers of component recovery are added, the events become increasingly unlikely. Thus, only a few decoders perform reliably.

3.2.3. Decoder Performance

We now bound the worst-case performance of such decoders.

Definition 4. In the context of Lemma 1, take $\mathbf{e}_{X|\mathbf{A}}$ to be the coefficients of the best linear unbiased estimator for X given $\mathbf{A}^\dagger(\vec{\alpha}\vec{Y} - \vec{W}_Q)$, $\vec{W}_Q \sim \mathcal{N}(0, \Sigma_Q)$. Then, for $\Delta > \text{var}(X|\mathbf{A}^\dagger(\vec{\alpha}\vec{Y} - \vec{W}_Q))$, define a decoder:

$$\text{dec}(\vec{U}; \mathbf{A}) := \mathbf{e}_{X|\mathbf{A}}^\dagger f(\vec{U}), \quad (18)$$

or 0 if the observed average power of Equation (18) is greater than Δ .

The mean-squared-error distortion each such decoder achieves can be upper bounded in terms of the probability of the event $\mathcal{M}(\mathbf{A})$:

Lemma 2. Take a decoder from Definition 4 and define

$$d^2 := 1 - \text{var}(X|\mathbf{e}_{X|\mathbf{A}}^\dagger(\vec{\alpha}\vec{Y} + \vec{W}_Q)). \quad (19)$$

Take \mathcal{E} to be the event where the decoding is zero. Then, no worse than the following mean-squared-error is achieved in estimating X^n :

$$d^2 + (\Delta + 1)^2 \cdot \sqrt{3\mathbb{P}(\mathcal{M}^C)} + \sqrt{3\mathbb{P}(\mathcal{E})} \quad (20)$$

Proof. The goal is to approximate the integral:

$$\begin{aligned} \int \frac{1}{n} \|\text{dec}(\vec{U}; \mathbf{A}) - X^n\|^2 d\mathbb{P} &= \int_{\mathcal{M}} \frac{1}{n} \|\text{dec}(\vec{U}; \mathbf{A}) - X^n\|^2 d\mathbb{P} + \dots \\ &\quad \int_{\mathcal{M}^C} \frac{1}{n} \|\text{dec}(\vec{U}; \mathbf{A}) - X^n\|^2 d\mathbb{P}. \end{aligned} \quad (21)$$

Bound the first summand of Equation (21):

$$\int_{\mathcal{M}} \frac{1}{n} \|\text{dec}(\vec{U}; \mathbf{A}) - X^n\|^2 d\mathbb{P} = \frac{1}{n} \int_{\mathcal{M}} \mathbf{1}_{\mathcal{E}^C} \|E[X^n|\mathbf{A}(\vec{\alpha}\vec{Y} - \vec{W}_Q)] - X^n\|^2 + \mathbf{1}_{\mathcal{E}} \|X^n\|^2 d\mathbb{P} \quad (22)$$

$$< \frac{1}{n} \int \|E[X^n|\mathbf{A}(\vec{\alpha}\vec{Y} - \vec{W}_Q)] - X^n\|^2 d\mathbb{P} + \frac{1}{n} \int_{\mathcal{E}} \|X^n\|^2 d\mathbb{P} \quad (23)$$

$$= d^2 + \frac{1}{n} \int_{\mathcal{E}} \|X^n\|^2 d\mathbb{P}. \quad (24)$$

where the first equality follows by Lemma 1 and choice of decoder. Bound the second summand of Equation (21):

$$\int_{\mathcal{M}^C} \frac{1}{n} \|\text{dec}(\vec{U}; \mathbf{A}) - X^n\|^2 d\mathbb{P} \leq \frac{1}{n} \int_{\mathcal{M}^C} \|(1 + \Delta)X^n\|^2 d\mathbb{P} \quad (25)$$

$$= (1 + \Delta)^2 \frac{1}{n} \int_{\mathcal{M}^C} \|X^n\|^2 d\mathbb{P}. \quad (26)$$

Applying Hölder's inequality as follows to the two bounds yields the result:

$$\int_S \|X^n\| d\mathbb{P} \leq \left[\int \mathbf{1}_S^2 d\mathbb{P} \right]^{1/2} \cdot \left[\int \|X\|^4 d\mathbb{P} \right]^{1/2} \quad (27)$$

$$= \sqrt{3\mathbb{P}(S)}. \quad (28)$$

□

The bound on distortion suffices for the asymptotic analysis in Section 4, but is quite coarse in low dimension.

4. Asymptotic Scheme

Analysis of the scheme in limit with blocklength over particular choice of $L_c, (L_k)_{k \in [K]}$ yields a nice characterization of its performance. Theorem 1 demonstrates that, in limit with blocklength and certain lattice design, there is essentially one decoder which performs at least as well as any others. The subspace of receiver observations it reliably recovers is characterized. First, a matrix definition is needed to state this result.

Definition 5. For $\varepsilon > 0$ define covariance matrices $\Sigma_{Q\infty}, \mathbf{C}_\infty \in \mathbb{R}^{K \times K}$

$$\Sigma_{Q\infty} := \text{diag} \left(2^{-2r_1+\varepsilon}, \dots, 2^{-2r_K+\varepsilon} \right). \quad (29)$$

$$\mathbf{C}_\infty := (\text{diag } \vec{\alpha})^\dagger \Sigma (\text{diag } \vec{\alpha}) + \Sigma_{Q\infty}. \quad (30)$$

\mathbf{C}_∞ represents the covariance between quantized observations achieved in limit with blocklength when the lattices $(L_k)_{k \in [K]}$ are chosen well.

Theorem 1. Take $\varepsilon > 0$ small and \mathbf{C}_∞ from Equation (29). Define S to be the smallest subspace in \mathbb{R}^K with the property that all integer vectors $\mathbf{v} \in \mathbb{Z}^K$ have either $\min_{\mathbf{s} \in S} (\mathbf{v} - \mathbf{s})^\dagger \mathbf{C}_\infty (\mathbf{v} - \mathbf{s}) \geq 1$ or $\mathbf{v} \in S$. Fix \mathbf{P}_∞ as the projection onto S . Then, for large enough blocklength n and certain encoders, some processing f of the encodings has with high probability

$$f(\vec{U}) = \mathbf{P}_\infty (\vec{\alpha} \vec{Y}^n + \vec{W}'_Q), \quad (31)$$

where $\vec{W}'_Q = (W'_{Q,1}, \dots, W'_{Q,K})$ has independent components and $\mathbb{E}(\frac{1}{n} \|W'_{Q,k}\|^2) < 2^{-2r_k+2\varepsilon}$.

Corollary 1. A decoder provided side information as in Theorem 1 can achieve the following MSE distortion in estimating X :

$$d_L^2 := 1 - \vec{\alpha}^\dagger \text{pinv}(\mathbf{P}_\infty \mathbf{C}_\infty \mathbf{P}_\infty) \vec{\alpha}. \quad (32)$$

Proof. Apply a linear estimation for the source on f 's output. □

The proof is given in Appendix B. The theorem is demonstrated by observing that, if lattices are chosen well, then events from Definition 3 approach probability zero or one, and that \mathbf{P}_∞ 's image coincides with the span of the high-probability-events' vectors. Computation of the projection \mathbf{P}_∞ can be done via repeated reduction of $\mathbf{C}_\infty^{1/2}$:

Algorithm 1 Compute projection \mathbf{P}_∞ and processing stages \mathbf{A} from \mathbf{C}_∞ .

```

 $\mathbf{A} \leftarrow [], \mathbf{a} \leftarrow \text{SLVC}(\mathbf{C}_\infty^{1/2}), \mathbf{R} \leftarrow \mathbf{I}_{K \times K},$ 
while  $0 < (\mathbf{R}\mathbf{a})^\dagger \mathbf{C}_\infty \mathbf{R}\mathbf{a} < 1$  do
   $\mathbf{A} \leftarrow [\mathbf{A}, \mathbf{a}]$ 
   $\mathbf{R} \leftarrow \mathbf{I}_{K \times K} - \mathbf{A} \text{pinv}(\mathbf{A}^\dagger \mathbf{C}_\infty \mathbf{A}) \mathbf{A}^\dagger \mathbf{C}_\infty$ 
   $\mathbf{a} \leftarrow \text{SLVC}(\mathbf{C}_\infty^{1/2} \mathbf{R})$ 
end while
 $\mathbf{P}_\infty \leftarrow \mathbf{A}(\mathbf{A}^\dagger \mathbf{A})^{-1} \mathbf{A}^\dagger$ 
return  $\mathbf{P}_\infty, \mathbf{A}$ 

```

In Algorithm 1, the subroutine $\text{SLVC}(\mathbf{B})$, “Shortest Lattice Vector Coordinates” returns the nonzero integer vector \mathbf{a} that minimizes the norm of $\mathbf{B}\mathbf{a}$ while $\mathbf{B}\mathbf{a} \neq 0$. $\text{SLVC}(\cdot)$ can be implemented using a lattice enumeration algorithm like one in [17] together with the Lenstra-Lenstra-Lovász (LLL) algorithm to convert a set of spanning lattice vectors into a basis [18]. Algorithm 1 indeed returns \mathbf{P}_∞ since it is lifted from \mathbf{P}_∞ ’s construction in the proof of Theorem 1 given in Appendix B.

Complexity of Scheme

The complexity of the scheme in operation is identical to that of codes with similar structure. Examples include IFSC and KP, as discussed in the Introduction. In particular, if $g(n)$ grows with the amount of operations used to evaluate round_L (L being the involved lattice for which round_L is hardest to compute), then it follows from the scheme description that the time complexity of each encoder is $O(n + g(n))$ and the time complexity of the base’s decoder is $O(nK + K^2 \cdot g(n))$. The practicality of such a scheme is then primarily dependent on the existence of lattices which both satisfy the “goodness” properties described here and have rounding, modulo operations of tractable complexity. Several propositions for such lattice structures exist, for instance *LDPC lattices* ($O(n)$) [19] and *polar lattices* ($O(n \log n)$) [20].

Configuration of the scheme is also a difficult computational problem. Determination of the optimal configuration is a non-smooth continuous-domain search for the objective given in Corollary 1 over choice of encoder scalings $\vec{\alpha} \in \mathbb{R}^+$. Each objective evaluation involves computing Algorithm 1. Due to this algorithm involving several shortest-lattice-vector-problems in dimension up to K , one implementation of this algorithm has time complexity $O(K2^K)$. This complexity reduces considerably if the strict shortest-lattice-vector problem is relaxed to allow approximate solutions such as those provided by the LLL algorithm. No methods for global optimization of the objective are known, but computationally tractable approximations can provide strong performance as demonstrated in Section 5.

5. Performance

Here, the performance of the scheme is compared to existing results. All numerical results deal with complex circularly-symmetric Gaussian channels. Although the SIFM derivations shown are in terms of real channels for easier exposition, they apply just as well to the complex case by considering each complex observation as two appropriately correlated real observations. Each data point was computed on average over 200 randomly generated channel covariances selected with the prescribed statistics. An outer bound, representing the performance if the base had access to the receivers’ observations in full precision, is also shown.

5.1. Quantize and Forward

One strategy much simpler than SIFM is for each observer forward its own quantized representation of its observations to the base. A “saturating uniform quantizer” was simulated, where each real observation was clipped to some interval, and then rounded onto a quantization step using the prescribed data rate. This curve is labeled “1D Quantize and Forward”.

Performance when observers used rate-distortion quantizers for Gaussian sources was also computed and is labeled “HD Quantize and Forward”. The performance for the rate-distortion quantizer strategy when observer bitrates were allowed to vary within a sum-rate was also plotted as “HD Quantize and Forward (Variable-Bitrate)”.

These techniques are strong when the LAN rate is severely limited or the interference is low, as shown in Figures 7 and 9. The function of joint compression in these regimes is more subtle since there are less superfluous correlated components among receiver observations to eliminate. Quantize and Forward schemes are weak relative to the others in the presence of many strong interferers, as shown in Figures 7–10.

5.2. Asymptotic Scheme

The distortion SIFM achieves is dependent on choice of observer scales $\vec{\alpha}$ and is highly nonconvex in terms of them due to the discontinuity of Algorithm 1's outputs. For this reason, some form of search for good scale parameters is required for each fixed channel covariance and observer rate vector \vec{r} . For performance evaluation, the search problem was solved approximately.

Recall that the asymptotic scheme described here includes one by Krithivasan and Pradhan (KP, see Introduction) as a special case. Details of the parameterization of the KP scheme in terms of SIFM are shown in Appendix D. It was observed empirically that scaling all receiver's observation by the same constant can yield strong performance. The strongest between uniform and KP scaling was chosen as an initial guess and was improved by taking random steps.

5.3. One-Shot Scheme

If the involved lattices are all one-dimensional (i.e., nested intervals), then the likelihood of each event in Definition 3 is straightforward to approximate via Monte-Carlo simulation. This enables a slight modification of Algorithm 1 to be used for identifying a strong decoding strategy for given receiver scalings. Although the upper bound in Lemma 2 applies to the one-shot strategy, it is often too weak in low dimension to be informative of a scheme's true distortion. Distortion was estimated by simulation instead. In plots, this scheme is labeled "1D SIFM".

Unfortunately, the best identified configurations for the one-shot scheme did not consistently provide significant improvement over uniform quantization as the high dimensional scheme had. This is thought to be because, even in extreme observations, the uniform quantizer saturates and still provides a reasonable representation of the source. In contrast, the same observation in SIFM creates a lattice wrapping distortion, which significantly affects the final decoding result. Reducing the likelihood of such errors in low dimension requires conservative choice of scales, which further degrades performance. However, such problems diminish as higher dimensional lattices are used.

Performance for the asymptotic SIFM strategy when observer bitrates were allowed to vary within a sum-rate was also plotted as "HD SIFM (Variable-Bitrate)".

5.4. Versus Increasing Receiver Rates

The performance of the various schemes considered is shown in Figure 7 for five receivers, each messaging to the base at rate varying from 2 bits-per-complex-sample to 16 bits-per-complex-sample, in the presence of three interferers each appearing at each observer at an average of 20 dB above a 0 dB signal of interest.

Coincidentally, the performance of the KP scheme with equal bitrates was observed in all plots to closely match the performance of the variable-bitrate Quantize and Forward scheme.

As shown in Figure 8, there is still a significant gap at low bitrates between the performance of SIFM and the bound given in [21]. This bound is based on non-structured joint-compression of quantizations (like Berger-Tung source coding [22] but binned for recovery of a source rather than all the observers' quantizations). The gap in performance could be due to a combination of factors. First, the scheme from [21] is designed so that receiver messages are independent of one another, while in SIFM some inter-message dependences are still present after lattice processing. This means SIFM messages could be jointly compressed to improve performance. Another contributing factor could be poor solution of the search problem mentioned in Section 5.2.

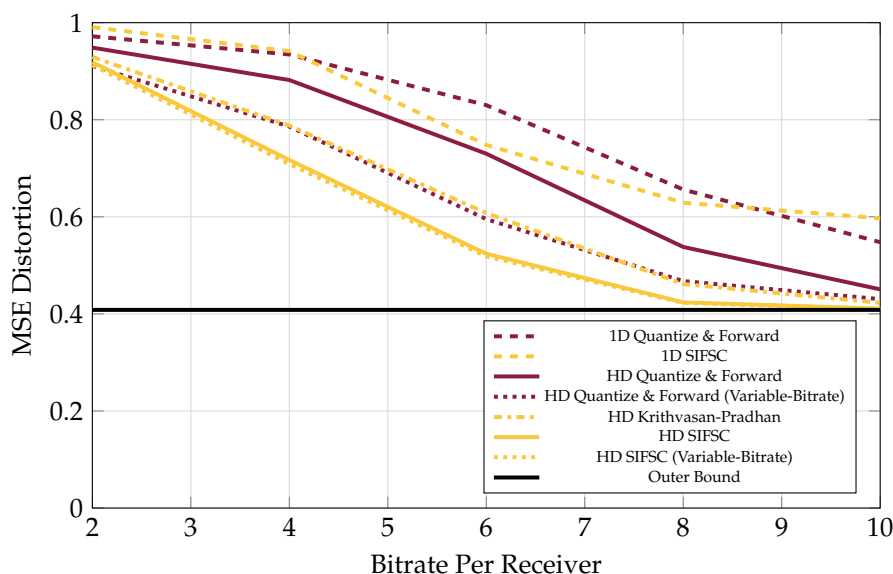


Figure 7. MSE performance versus receiver bitrate for five receivers the presence of three 20 dB interferers, all observing a 0 dB source. SIFM consistently significantly outperforms the rest of the schemes, and does not appear to benefit much from redistribution of observer bitrates.

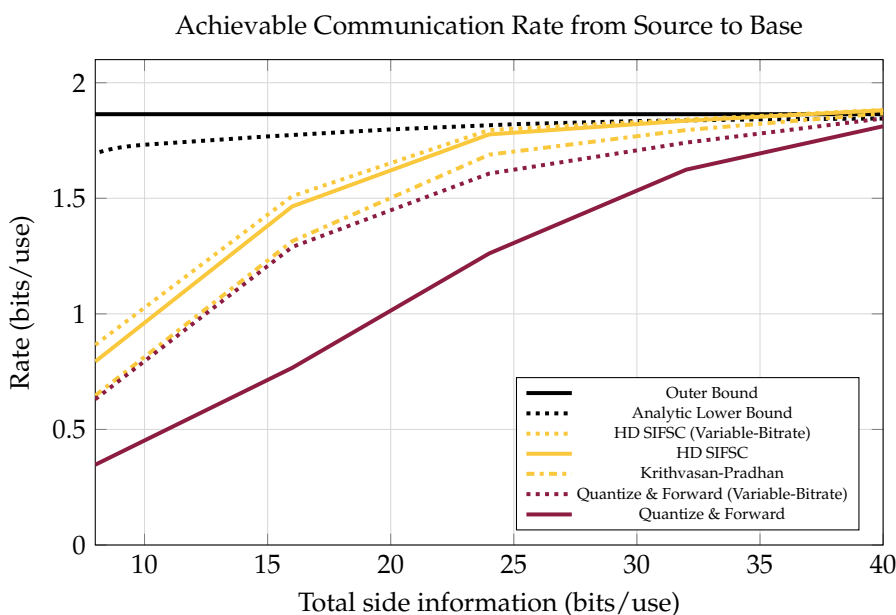


Figure 8. Achievable end-to-end communication rate between source and base for various sources, including an achievable rate using Berger-Tung-like lossy distributed source coding. Four receivers each observing a source at 0 dB in the presence of one 20 dB interferer. Notice that at low bitrates there is a large gap between this bound and the achieved SIFM rate. This could either be due to the fact that SIFM messages are not independent of each other and could be further jointly compressed, or because the SIFM messages are sub-optimally configured.

5.5. Versus Adding Interferers

Performance for the various schemes is shown in Figure 9 for a system of five receivers, each messaging to the base at a rate of 6 bits-per-complex-sample, in the presence of 0–5 independent interferers, each appearing at each observer an average of 20 dB above a 0 dB signal of interest.

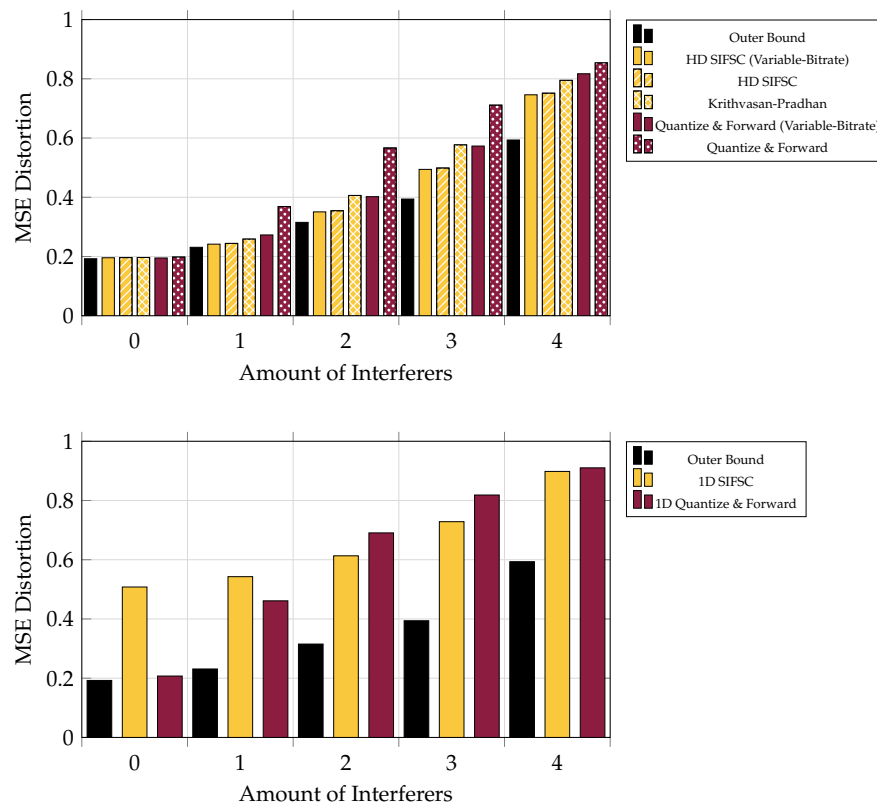


Figure 9. MSE performance versus amount of 20 dB interferers, when each of five observer encodes at a rate of 6 bits per observation (or 30 bits total shared among observers for Variable-Rate strategies). SIFM consistently outperforms the rest of the strategies and was not observed to improve much by reconfiguring observer bitrates. One-shot SIFM did not consistently outperform much simpler one-shot quantizers.

5.6. One-Shot Outage Probability

One may be interested in the likelihood of being able to recover a signal to above some acceptable noise threshold over an ensemble of channels. Although in low-interference regimes the one-shot SIFM scheme is reliably outperformed by the much simpler one-shot Quantize and Forward scheme (Figure 9), one-shot SIFM is much less likely to perform exceptionally poorly when averaged over the current channel model (see Figure 10).

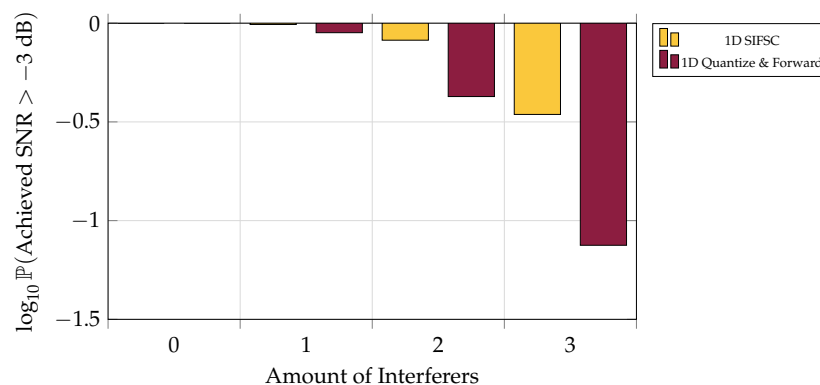


Figure 10. Likelihood of scheme providing recovery of X better than -3 dB SNR versus amount of 20 dB interferers, when each of five observer encodes at a rate of 6 bits per observation. For this threshold and ensemble of channels, one-shot SIFM performs roughly as reliably as one-shot Quantize and Forward with one less 20 dB interferer present.

6. Conclusions

Successive Integer-Forcing Many-Help-One (SIFM) is a lattice-algebra-based strategy for distributed coding of a Gaussian source in correlated noise. For good choice of parameters, SIFM consistently outperforms many of the strategies it generalizes. Finding good parameters for SIFM is a difficult non-convex search problem but reasonably strong solutions can be found through well-initialized random search.

A one-shot implementation of the scheme usually outperforms plain uncompressed quantization when multiple interferers are present, but often performs worse when there are few. This is probably due to SIFM's heavy dependence on the absence of tail events that are somewhat common in low blocklength versions of the scheme. Despite this, the one-shot scheme is more typically above low-SNR thresholds in certain ensembles of channels. It is expected that performance would improve greatly if higher dimensional nested lattices were used.

Many results for related lattice schemes remain to be ported to the present case. Configurations for IFSC encoders robust to incomplete channel state knowledge are designed in Reference [10] (Chapter 4). Robustness of such schemes of IFSC to Doppler and delay spread in a certain system model are investigated in Griffin [11].

There is still some gap between the best achievable rate and SIFM, as seen in Figure 8. This is at least partially due to redundancies in SIFM messages. Some redundancies still exist between SIFM messages Wagner [4] Yang and Xiong [14]. This indicates that further compression of the messages is possible, and a study on these redundancies is forthcoming.

Author Contributions: Conceptualization, C.C., A.A. and D.W.B.; methodology, C.C., A.A. and D.W.B.; software, C.C. and M.K.; validation, C.C., A.A. and D.W.B.; formal analysis, C.C.; investigation, C.C.; resources, D.W.B.; data curation, C.C. and D.W.B.; writing-original draft preparation, C.C.; writing-review and editing, C.C.; visualization, C.C.; supervision, D.W.B.; project administration, D.W.B. and A.A.; funding acquisition, D.W.B.

Funding: This research was funded by the United States Department of the Air Force under Air Force Contract No. FA8702-15-D-0001. Distribution Statement A. Approved for public release. Distribution is unlimited. This material is based upon work supported by the Department of the Air Force under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Department of the Air Force.

Acknowledgments: We thank Alex Chiriyath for his valuable comments in composing this document.

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

Appendix A. Proof of Lemma 1, Construction of Mod-Elementary Decoders

Proof. Assume $\mathcal{M}(\mathbf{A}_j)$ occurs. Then, $\mathcal{M}(\mathbf{A}_1) \supset \dots \supset \mathcal{M}(\mathbf{A}_{j-1})$ have occurred as well. Compute:

$$f_1(\vec{U}) = \text{mod} \left(\mathbf{a}_1^\dagger (\text{mod}(\alpha_k Y_k^n) - \tilde{W}_{d,k})_k \right) \quad (\text{A1})$$

$$= \text{mod} \left(\mathbf{a}_1^\dagger (\text{mod}(\alpha_k Y_k^n) - \text{mod}(\tilde{W}_{d,k}))_k \right) \quad (\text{A2})$$

$$= \text{mod} \left(\mathbf{a}_1^\dagger (\vec{\alpha Y}^n - \tilde{W}_d) \right) \quad (\text{A3})$$

$$= \mathbf{a}_1^\dagger (\vec{\alpha Y}^n - \tilde{W}_d). \quad (\text{A4})$$

where Equation (A1) is due to the crypto lemma [9] (Theorem 4.1.1), Equation (A2) is since $W'_{d,k} \in B_c$, Equation (A3) is by Property 4 and Equation (A4) is since $\mathcal{M}(\mathbf{A}_1)$ occurred.

Now, for any $m > 1$, say $f_{[m-1]}$ has:

$$f_{[m-1]}(\vec{U}) = \mathbf{A}_{[m-1]}^\dagger (\vec{\alpha Y}^n - \tilde{W}_{d,k}). \quad (\text{A5})$$

Then, compute (similarly to the base case):

$$f_m(\vec{U}) = \text{mod} \left\{ \text{mod} \left(\mathbf{a}_m^\dagger (\vec{\alpha Y}^n - \tilde{W}_d) \right) - \dots \right. \quad (\text{A6})$$

$$\left. \text{mod} \left([\mathbf{S}_{m-1} \mathbf{a}_m]^\dagger \mathbf{A}_{[m-1]}^\dagger (\vec{\alpha Y}^n - \tilde{W}_d) \right) \right\} + \dots$$

$$[\mathbf{S}_{m-1} \mathbf{a}_m]^\dagger \mathbf{A}_{[m-1]}^\dagger (\vec{\alpha Y}^n - \tilde{W}_d)$$

$$= \text{mod} \left\{ \mathbf{a}_m^\dagger (\vec{\alpha Y}^n - \tilde{W}_d) - \dots \right. \quad (\text{A7})$$

$$\left. [\mathbf{S}_{m-1} \mathbf{a}_m]^\dagger \mathbf{A}_{[m-1]}^\dagger (\vec{\alpha Y}^n - \tilde{W}_d) \right\} + \dots$$

$$[\mathbf{S}_{m-1} \mathbf{a}_m]^\dagger \mathbf{A}_{[m-1]}^\dagger (\vec{\alpha Y}^n - \tilde{W}_d)$$

$$= \text{mod} \left\{ [\mathbf{R}_{m-1} \mathbf{a}_m]^\dagger (\vec{\alpha Y}^n + \tilde{W}_d) \right\} + \dots \quad (\text{A8})$$

$$[\mathbf{S}_{m-1} \mathbf{a}_m]^\dagger \mathbf{A}_{[m-1]}^\dagger (\vec{\alpha Y}^n - \tilde{W}_d).$$

Since $\mathcal{M}(\mathbf{A}_m)$ has occurred, then $[\mathbf{R}_{m-1} \mathbf{a}_m]^\dagger (\vec{\alpha Y}^n + \tilde{W}_d) \in B_c$, hence by Property 1, Equation (A8) reduces to:

$$f_m(\vec{U}) = [\mathbf{R}_{m-1} \mathbf{a}_m]^\dagger (\vec{\alpha Y}^n + \tilde{W}_d) + \dots \quad (\text{A9})$$

$$[\mathbf{S}_{m-1} \mathbf{a}_m]^\dagger \mathbf{A}_{[m-1]}^\dagger (\vec{\alpha Y}^n - \tilde{W}_d)$$

$$= \mathbf{A}_m^\dagger (\vec{\alpha Y}^n - \tilde{W}_d). \quad (\text{A10})$$

Thus, by induction, the described mod-elementary functions behave as desired. \square

Appendix B. Proof of Theorem 1, Asymptotic Lattice Scheme Performance

Proof. Using the construction in [23] (Theorem 2), take sequences of lattices $L_c^{(n)}, (L_k^{(n)})_{k \in [K]}$ with base regions $B_c^{(n)}, (B_k^{(n)})_k$ where:

- $L_c^{(n)}$ with $B_c^{(n)}$ is good for channel coding in the presence of semi norm-ergodic noise, [23] (Definition 4) scaled to handle semi norm-ergodic [23] (Definition 2) noise of power less than 1.
- Each $L_k^{(n)} \subset L_c^{(n)}$ is good for mean squared error quantization [23] (Definition 5), with $\lim_n \frac{1}{n} \mathbb{E}[\|u_k^{(n)}\|^2] < 2^{-2r_k + \varepsilon}$, taking each $u_k^{(n)} \sim \text{unif } B_k^{(n)}$ independent.
- Eventually, in n , $\frac{1}{n} \log(|L_c^{(n)} \cap L_k^{(n)}|) < r_k$.

For each $k \in [K]$ take $Z_k^n := \alpha_k Y_k^n - \tilde{W}_{d,k}$, where $(\tilde{W}_{d,k})_{k \in [K]}$ is distributed as $(u_k^{(n)})_{k \in [K]}$ and is independent of $(Y_k^n)_{k \in [K]}$. Each event $\mathcal{M}(\mathbf{A}_j)$ has associated with it some linear combination $\mathbf{R}_{j-1} = I$ if $j = 1$ or as in Definition 3 otherwise. Then, by [13] (Appendix V), the linear combination $(\mathbf{R}_{j-1} \mathbf{a}_j)^\dagger (Z_k^n)_{k \in [K]}$ probably lands in the base of a lattice cell as long as that lattice is good for coding for powers greater than $p = \frac{1}{n} \mathbb{E}[\|(\mathbf{R}_{j-1} \mathbf{a}_j)^\dagger (Z_k^n)_{k \in [K]}\|^2] = (\mathbf{R}_{j-1} \mathbf{a}_j)^\dagger \mathbf{C}_\infty (\mathbf{R}_{j-1} \mathbf{a}_j)$.

Since $L_c^{(n)}$ is good for channel coding in the presence of semi norm-ergodic noise with power less than 1 [23] then (by definition) Z_k^n occurs in $\mathcal{M}(\mathbf{A}_j)$ with arbitrarily high probability eventually in n if $p < 1$, and with arbitrarily low probability if $p > 1$. The case where $p = 1$ never occurs when using $\varepsilon > 0$ that affects \mathbf{C}_∞ so that every $(\mathbf{R}_{j-1} \mathbf{a}_j)^\dagger \mathbf{C}_\infty (\mathbf{R}_{j-1} \mathbf{a}_j)$ is irrational (only countable $\mathbf{R}_{j-1}, \mathbf{a}_j$ are possible, thus ε small enough always exist).

We now demonstrate that, in this situation, among the high-probability events from Definition 3, there are some whose conditions cannot be strengthened. Any event $\mathcal{M}(\mathbf{A}_j)$ with probability eventually high and associated projection \mathbf{R}_{j-1} either has some vector $\mathbf{a} \in \mathbb{Z}^K$ with \mathbf{a} not in \mathbf{R}_{j-1} 's null space and $\frac{1}{n} \mathbb{E}[\|(\mathbf{R}_{j-1} \mathbf{a})^\dagger (Z_k^n)_{k \in [K]}\|^2] < 1$, or no such vector \mathbf{a} exists. If there is such an \mathbf{a} , then, taking $\mathbf{a}_{j+1} = \mathbf{a}$, the event $\mathcal{M}([\mathbf{A}_j, \mathbf{a}_{j+1}])$ will also have eventually high probability. Repeating the argument,

such \mathbf{a} s can only be found up to $m < K$ times: by then, the matrix \mathbf{A}_j has column basis for all \mathbb{R}^K , or all choice of $\mathbf{a} \in \mathbb{Z}^K$ yields $\frac{1}{n} \mathbb{E}[\|(\mathbf{R}_{j-1}\mathbf{a})^\dagger (Z_k^n)_{k \in [K]}\|^2] > 1$.

The result of this maximal strengthening of $\mathcal{M}(\mathbf{A}_j)$'s conditions to $\mathcal{M}(\mathbf{A}_{j+m})$ has associated with it a projection \mathbf{R}_- with the property that any integer vector $\mathbf{a} \in \mathbb{Z}^K$ has either \mathbf{a} in \mathbf{R}_- 's null space or $\frac{1}{n} \mathbb{E}[\|(\mathbf{R}_-\mathbf{a})^\dagger (Z_k^n)_{k \in [K]}\|^2] = (\mathbf{R}_-\mathbf{a})^\dagger \mathbf{C}_\infty (\mathbf{R}_-\mathbf{a}) > 1$. Any projections created in such a way from two different events must be equal, since if the first strengthened event's projection's null space had a vector the second's did not, then the second strengthened event could be further strengthened using the vectors in the first event's sequence.

By above, if \mathbf{R}_- has nonzero range, and if a nonzero subspace U does not contain ($\text{image } \mathbf{R}_-$), then some nonzero integer vector $\mathbf{a} \in \mathbb{Z}^K$ away from U has $\min_{\mathbf{s} \in U} (\mathbf{a} - \mathbf{s})^\dagger \mathbf{C}_\infty (\mathbf{a} - \mathbf{s}) < 1$. Thus, $S = \text{image}(\mathbf{R}_-)$ is the smallest of all subspaces with the property that any $\mathbf{a} \in \mathbb{Z}^K$ has either $\min_{\mathbf{s} \in S} (\mathbf{a} - \mathbf{s})^\dagger \mathbf{C}_\infty (\mathbf{a} - \mathbf{s}) \geq 1$ or $\mathbf{a} \in S$ otherwise.

Finally apply Lemma 1 and Corollary 2 to any high-probability event associated with projection \mathbf{R}_- . Since the event eventually has arbitrarily high probability, choosing $\Delta > 1$ eventually gives expected distortion arbitrarily close to d_L^2 . \square

Appendix C. Closed Form for $\mathbf{S}_k(\mathbf{v})$

For exposition, the definition of $\mathbf{S}_k(\mathbf{v})$ is shown semantically, as a minimizer, as opposed to in closed form. Its closed form is readily derived:

$$\arg \min_{\mathbf{u} \in \mathbb{R}^k} (\mathbf{v} - \mathbf{A}^{(k)}\mathbf{u})^\dagger \mathbf{C} (\mathbf{v} - \mathbf{A}^{(k)}\mathbf{u}) \quad (\text{A11})$$

$$= \arg \min_{\mathbf{u} \in \mathbb{R}^k} -2 \text{Re} \left\{ \mathbf{v}^\dagger \mathbf{C} \mathbf{A}^{(k)} \mathbf{u} \right\} + \mathbf{u}^\dagger \mathbf{A}^{(k)\dagger} \mathbf{C} \mathbf{A}^{(k)} \mathbf{u} \quad (\text{A12})$$

$$= \left[\text{pinv}(\mathbf{A}^{(k)\dagger} \mathbf{C} \mathbf{A}^{(k)}) \mathbf{A}^{(k)\dagger} \mathbf{C} \right] \mathbf{v}. \quad (\text{A13})$$

Appendix D. KP Parameterization

Here, the parameterization of the KP scheme [13], provided in terms of the asymptotic scheme and Lemma 1, is shown. First, fix some $\mathbf{c} \in \mathbb{R}^K$ and an ordered partition $(\Theta_1, \dots, \Theta_p)$ of $[K]$. With these variable choices, the KP scheme specifies use of \mathbf{A} with columns $\mathbf{a}_m = \mathbf{1}_{\Theta_m}$. A matrix with this structure is denoted with subscript as \mathbf{A}_{KP} .

It also specifies receiver scaling coefficients $\vec{\alpha}$ defined in parts over the partition sets. Starting at $m = 1$ and up through $m = p$, then the components $(\alpha_k)_{k \in \Theta_m}$ are specified by a minimization problem dependent on the result of previous steps. Take $\vec{W}_{Q\infty} \sim \mathcal{N}(0, \Sigma_{Q\infty})$. In addition, take \mathbf{c}_{Θ_k} to be the $|\Theta_k|$ -vector with coefficients taken from \mathbf{c} at indices in Θ_k . Now, starting at $m = 1$ and up through $m = p$, define, if possible:

$$(\alpha_k)_{\Theta_m} := \sqrt{\frac{1 - \text{var}(\mathbf{a}_k^\dagger \vec{W}_{Q\infty})}{\text{var}(\mathbf{c}_{\Theta_m} \vec{Y} | \mathbf{A}_{KP, m-1}(\vec{\alpha} \vec{Y} + \vec{W}_{Q\infty}))}} \mathbf{c}_{\Theta_m}. \quad (\text{A14})$$

If the square-root in Equation (A14) does not exist (i.e., if the numerator is not positive), then a KP scheme with the partition $(\Theta_1, \dots, \Theta_p)$ cannot be designed for specified \mathbf{c} and encoder rates (R_1, \dots, R_K) . For any \mathbf{c} and nonzero rates, there is always some partition that works. For example, a singleton partition works. A well-formed KP scheme guarantees achievable average distortion:

$$d_{KP}^2 = \text{var} \left(X | \mathbf{A}_{KP}^\dagger (\vec{\alpha} \vec{Y} + \vec{W}_{Q\infty}) \right). \quad (\text{A15})$$

References

1. Tavildar, S.; Viswanath, P.; Wagner, A.B. The Gaussian many-help-one distributed source coding problem. *IEEE Trans. Inf. Theory* **2009**, *56*, 564–581. [\[CrossRef\]](#)
2. Berger, T.; Zhang, Z.; Viswanathan, H. The CEO Problem [multiterminal source coding]. *IEEE Trans. Inf. Theory* **1996**, *42*, 887–902. [\[CrossRef\]](#)
3. Maddah-Ali, M.A.; Tse, D. Interference neutralization in distributed lossy source coding. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), Austin, TX, USA, 12–18 June 2010; pp. 166–170.
4. Wagner, A.B. On Distributed Compression of Linear Functions. *IEEE Trans. Inf. Theory* **2011**, *57*, 79–94. [\[CrossRef\]](#)
5. Nazer, B.; Gastpar, M. The Case for Structured Random Codes in Network Capacity Theorems. *Eur. Trans. Telecommun.* **2008**, *19*, 455–474. [\[CrossRef\]](#)
6. Krithivasan, D.; Pradhan, S.S. Distributed Source Coding Using Abelian Group Codes: A New Achievable Rate-Distortion Region. *IEEE Trans. Inf. Theory* **2011**, *57*, 1495–1519. [\[CrossRef\]](#)
7. Nazer, B.; Cadambe, V.R.; Ntranos, V.; Caire, G. Expanding the Compute-and-Forward Framework: Unequal Powers, Signal Levels, and Multiple Linear Combinations. *IEEE Trans. Inf. Theory* **2016**, *62*, 4879–4909. [\[CrossRef\]](#)
8. Ordentlich, O.; Erez, U. Integer-forcing source coding. *IEEE Trans. Inf. Theory* **2017**, *63*, 1253–1269. [\[CrossRef\]](#)
9. Zamir, R. *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation, and Multiuser Information Theory*; Cambridge University Press: Cambridge, UK, 2014.
10. Teixeira, P.J.M. Distributed Source Coding Based on Integer-Forcing. Master's Thesis, Técnico Lisboa, Lisboa, Portugal, June 2018.
11. Griffin, J.C. Exploring data compression for a distributed aerial relay application. Master's Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, June 2017.
12. He, W.; Nazer, B. Integer-Forcing Source Coding: Successive Cancellation and Source-Channel Duality. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, 10–15 Junly 2016; pp. 155–159.
13. Krithivasan, D.; Pradhan, S.S. Lattices for Distributed Source Coding: Jointly Gaussian Sources and Reconstruction of a Linear Function. In *International Symposium on Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 178–187.
14. Yang, Y.; Xiong, Z. An improved lattice-based scheme for lossy distributed compression of linear functions. In Proceedings of the Information Theory and Applications Workshop, San Diego, CA, USA, 6–11 February 2011; pp. 1–5.
15. Yang, Y.; Xiong, Z. Distributed Compression of Linear Functions: Partial Sum-Rate Tightness and Gap to Optimal Sum-Rate. *IEEE Trans. Inf. Theory* **2014**, *60*, 2835–2855. [\[CrossRef\]](#)
16. Cheng, H.; Yuan, X.; Tan, Y. Generalized Compute-Compress-and-Forward. *IEEE Trans. Inf. Theory* **2018**, *65*, 462–481. [\[CrossRef\]](#)
17. Schnorr, C.P.; Euchner, M. Lattice Basis Reduction: Improved Practical Algorithms and Solving Subset Sum Problems. *Math. Program.* **1994**, *66*, 181–199.
18. Buchmann, J.; Pohst, M. Computing a Lattice Basis from a System of Generating Vectors. In *European Conference on Computer Algebra*; Springer: Berlin/Heidelberg, Germany, 1987; pp. 54–63.
19. da Silva, P.R.B.; Silva, D. Multilevel LDPC Lattices With Efficient Encoding and Decoding and a Generalization of Construction D' . *IEEE Trans. Inf. Theory* **2018**, *65*, 3246–3260. [\[CrossRef\]](#)
20. Liu, L.; Yan, Y.; Ling, C.; Wu, X. Construction of capacity-achieving lattice codes: Polar lattices. *IEEE Trans. Commun.* **2018**, *67*, 915–928. [\[CrossRef\]](#)
21. Chapman, C.D.; Mittelmann, H.; Margetts, A.R.; Bliss, D.W. A Decentralized Receiver in Gaussian Interference. *Entropy* **2018**, *20*, 269. [\[CrossRef\]](#)
22. El Gamal, A.; Kim, Y.H. *Network Information Theory*; Cambridge University Press: Cambridge, UK, 2011.
23. Ordentlich, O.; Erez, U. A simple proof for the existence of “good” pairs of nested lattices. *IEEE Trans. Inf. Theory* **2016**, *62*, 4439–4453. [\[CrossRef\]](#)

