Achievability Performance Bounds for Integer-Forcing Source Coding

Elad Domanovitz and Uri Erez

Abstract

Integer-forcing source coding has been proposed as a low-complexity method for compression of distributed correlated Gaussian sources. In this scheme, each encoder quantizes its observation using the same fine lattice and reduces the result modulo a coarse lattice. Rather than directly recovering the individual quantized signals, the decoder first recovers a full-rank set of judiciously chosen integer linear combinations of the quantized signals, and then inverts it. It has been observed that the method works very well for "most" but not all source covariance matrices. The present work quantifies the measure of bad covariance matrices by studying the probability that integer-forcing source coding fails as a function of the allocated rate, where the probability is with respect to a random orthonormal transformation that is applied to the sources prior to quantization. For the important case where the signals to be compressed correspond to the antenna inputs of relays in an i.i.d. Rayleigh fading environment, this orthonormal transformation can be viewed as being performed by nature. Hence, the results provide performance guarantees for distributed source coding via integer forcing in this scenario.

I. Introduction

Integer-forcing (IF) source coding, proposed in [1], is a scheme for distributed lossy compression of correlated Gaussian sources under a minimum mean squared error distortion measure. Similar to its channel coding counterpart, in this scheme, all encoders use the same nested lattice codebook. Each encoder quantizes its observation using the fine lattice as a quantizer and reduces the result modulo the coarse lattice, which plays the role of binning. Rather than directly recovering the individual quantized signals, the decoder first recovers a full-rank set of judiciously chosen integer linear combinations of the quantized signals, and then inverts it. An appealing feature of integer-forcing source coding, not shared by previously proposed practical methods (e.g., Wyner-Ziv coding) for the distributed source coding problem, is its inherent symmetry, supporting equal distortion and quantization rates. A potential application of IF source coding is to distributed compression of signals received at several relays as suggested in [1] and further explored in [2].

Similar to IF channel coding, IF source coding works well for "most" but not all Gaussian vector sources. Following the approach of [3], in the present work we quantify the measure of bad source covariance matrices by considering a randomized version of IF source coding where a random unitary transformation is applied to the sources prior to quantization. While in general such a transformation implies joint processing at the encoders, we note that in some natural scenarios, including that of distributed compression of signals received at relays in an i.i.d. Rayleigh fading environment, the random transformation is actually performed by nature.[1] In fact, it was already empirically observed in [1] that IF source coding performs very well in the latter scenario.

The rest of this paper is organized as follows. Section II formulates the problem of distributed compression of Gaussian sources in a compound vector source setting and provides some relevant background on IF source coding. Section III describes randomly precoded IF source coding and its empirical performance. Section IV derives upper bounds on the probability of failure of randomly-precoded IF as a function of the excess rate. In Section V, deterministic linear precoding is considered. A bound on the worst-case excess rate needed is derived for any number of sources with any correlation matrix when space-time precoding derived from non-vanishing determinant codes is used. Further, we show that this bound can be significantly tightened for the case of uncorrelated sources. Concluding remarks appear in Section V.

II. Problem Formulation and Background

In this section we provide the problem formulation and briefly recall the achievable rates of IF source coding as developed in [1].

The work of E. Domanovitz and U. Erez was supported in part by the Israel Science Foundation under Grant No. 1956/16 and by the Heron consortium via the Israel Ministry of Economy and Industry.

E. Domanovitz and U. Erez are with the Department of Electrical Engineering – Systems, Tel Aviv University, Tel Aviv, Israel (email: domanovi,uri@eng.tau.ac.il).

[1] This follows since the left and right singular vector matrices of an i.i.d. Gaussian matrix $\mathbf{K}$ are equal to the eigenvector matrices of the Wishart ensembles $\mathbf{K}\mathbf{K}^T$ and $\mathbf{K}^T\mathbf{K}$, respectively. The latter are known to be uniformly (Haar) distributed. See, e.g., Chapter 4.6 in [4].

A. Distributed Compression of Gaussian Sources

We start by recalling the classical problem of distributed lossy compression of jointly Gaussian real random variables under a quadratic distortion measure. Specifically, we consider a distributed source coding setting with K encoding terminals and one decoder. Each of the K encoders has access to a vector of n i.i.d. realizations of the random variable x_k, $k = 1, ..., K$.[2] The random vector $\mathbf{x} = (x_1, \ldots, x_K)^T$ (corresponding to the different sources) is assumed to be Gaussian with zero mean and covariance matrix $\mathbf{K_{xx}} \triangleq \mathbb{E}\left(\mathbf{xx^T}\right)$.

Each encoder maps its observation $\mathbf{x_k}$ to an index using the encoding function

$$\mathcal{E}_k \ : \ \mathbb{R}^n \to \left\{1, ..., 2^{nR_k}\right\}, \tag{1}$$

and sends the index to the decoder.

The decoder is equipped with K decoding functions

$$\mathcal{D}_k \ : \ \left\{1, ..., 2^{nR_1}\right\} \times \cdots \times \left\{1, ..., 2^{nR_K}\right\} \to \mathbb{R}^n, \tag{2}$$

where $k = 1, ..., K$. Upon receiving K indices, one from each terminal, it generates the estimates

$$\hat{\mathbf{x}}_k = \mathcal{D}_k\left(\mathcal{E}_1(\mathbf{x}_1), ..., \mathcal{E}_K(\mathbf{x}_K)\right), \quad k = 1, ..., K. \tag{3}$$

A rate-distortion vector $(R_1, ..., R_K; d_1, ..., d_K)$ is achievable if there exist encoding functions, $\mathcal{E}_1, ..., \mathcal{E}_K$, and decoding functions, $\mathcal{D}_1, ..., \mathcal{D}_K$, such that $\frac{1}{n}\mathbb{E}\left(\|\mathbf{x}_k - \hat{\mathbf{x}}_k\|^2\right) \leq d_k$, for all $k = 1, ..., K$.

We focus on the symmetric case where $d_1 = \cdots = d_K = d$ and $R_1 = \cdots = R_K = R/K$ where we denote the sum rate by R. The best known achievable scheme (for this symmetric setting) is that of Berger and Tung [5], for which the following (in general, suboptimal) sum rate is achievable

$$\begin{aligned}\sum_{k=1}^{K} R_k &\geq \frac{1}{2}\log\det\left(\mathbf{I} + \frac{1}{d}\mathbf{K_{xx}}\right) \\ &\triangleq R_{\mathrm{BT}}.\end{aligned} \tag{4}$$

As shown in [1], R_{BT} is a lower bound on the achievable rate of IF source coding. We will refer to R_{BT} as the Berger-Tung benchmark. To simplify notation, we note that d can be "absorbed" into $\mathbf{K_{xx}}$. Hence, without loss of generality, we assume throughout that $d = 1$.

B. Compound Source Model And Scheme Outage Formulation

Consider distributed lossy compression of a vector of Gaussian sources

$$\mathbf{x} \sim \mathcal{N}(0, \mathbf{K_{xx}}). \tag{5}$$

We define the following compound class of Gaussian sources, having the same value of R_{BT}, via their covariance matrix:

$$\mathcal{K}(R_{\mathrm{BT}}) = \left\{\mathbf{K_{xx}} \in \mathbb{R}^{K\times K} : \frac{1}{2}\log\det\left(\mathbf{I} + \mathbf{K_{xx}}\right) = R_{\mathrm{BT}}\right\}. \tag{6}$$

We quantify the measure of the set of source covariance matrices by considering outage events, i.e., those events (sources) where integer forcing fails to achieve the desired level of distortion even though the rate exceeds R_{BT}. More broadly, for a given quantization scheme, denote the necessary rate to achieve $d = 1$ for a given covariance matrix $\mathbf{K_{xx}}$ as $R_{\mathrm{scheme}}(\mathbf{K_{xx}})$. Then, given a target rate $R > R_{\mathrm{BT}}$ and a covariance matrix $\mathbf{K_{xx}} \in \mathcal{K}(R_{\mathrm{BT}})$, a scheme outage occurs when $R_{\mathrm{scheme}}(\mathbf{K_{xx}}) > R$.

To quantify the measure of "bad" covariance matrices, we follow [3] and apply a random orthonormal precoding matrix to the (vector of) source samples prior to encoding. As mentioned above, this amounts to joint processing of the samples and hence the problem is no longer distributed in general. Nonetheless, as in the scenario described in Section IV-A, in certain statistical settings, this precoding operation is redundant as it can be viewed as being performed by nature.

Applying a precoding matrix to the source vector, we obtain a transformed source vector

$$\tilde{\mathbf{x}} = \mathbf{P}\mathbf{x}, \tag{7}$$

with covariance matrix

$$\mathbf{K}_{\tilde{x}\tilde{x}} = \mathbf{P}\mathbf{K_{xx}}\mathbf{P}^T. \tag{8}$$

It follows that the achievable rate of a quantization scheme for the precoded source is $R_{\mathrm{scheme}}(\mathbf{K}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}})$. When $\mathbf{P}$ is drawn at random, the latter rate is also random. The worst-case (WC) scheme outage probability is defined in turn as

$$P^{\mathrm{WC}}_{\mathrm{out,scheme}}\left(R_{\mathrm{BT}}, \Delta R\right)$$

[2]The time axis will be suppressed in the sequel and vector notation will be reserved to describe samples taken from different sources.

$$= \sup_{\mathbf{K_{xx}} \in \mathcal{K}(R_{\mathrm{BT}})} \Pr\left(R_{\mathrm{scheme}}(\mathbf{K}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}) > R_{\mathrm{BT}} + \Delta R\right), \tag{9}$$

where the probability is over the ensemble of unitary precoding matrices considered and ΔR is the gap to the Berger-Tung benchmark.

In the sequel, we quantify the tradeoff between the quantization rate R (or equivalently, between the excess rate $\Delta R = R - R_{\mathrm{BT}}$) and the outage probability $P^{\mathrm{WC}}_{\mathrm{out,IF}}\left(R_{\mathrm{BT}}, \Delta R\right)$ as defined in (9).

C. Integer-Forcing Source Coding

In a manner similar to IF equalization for channel coding, IF can be applied to the problem of distributed lossy compression. The approach is based on standard quantization followed by lattice-based binning. However, in the IF framework, the decoder first uses the bin indices for recovering linear combinations with integer coefficients of the quantized signals, and only then recovers the quantized signals themselves.

For our purposes, it suffices to state only the achievable rates of IF source coding. We refer the reader to [1] for the derivation and proofs.

We recall Theorem 1 from [1], stating that for any covariance matrix $\mathbf{K_{xx}}$, IF source coding can achieve any (sum) rate satisfying

$$\begin{aligned} R &> R_{\mathrm{IF}}(\mathbf{K_{xx}}) \\ &\triangleq \frac{K}{2}\log\left(\min_{\substack{\mathbf{A}\in\mathbb{Z}^{K\times K}\\ \det\mathbf{A}\neq 0}} \max_{k=1,\ldots,K} \mathbf{a}_k^T\left(\mathbf{I}+\mathbf{K_{xx}}\right)\mathbf{a}_k\right), \end{aligned} \tag{10}$$

where $\mathbf{a}_k^T$ is the kth row of the integer matrix $\mathbf{A}$. We denote by

$$R_{\mathrm{IF,m}}(\mathbf{K}_{xx}, \mathbf{A}) \triangleq \mathbf{a}_k^T\left(\mathbf{I}+\mathbf{K}_{xx}\right)\mathbf{a}_k \tag{11}$$

the effective rate that can be achieved at the m'th equation.

The matrix $\mathbf{I}+\mathbf{K_{xx}}$ is symmetric and positive definite, and therefore it admits a Cholesky decomposition

$$\mathbf{I}+\mathbf{K_{xx}} = \mathbf{F}\mathbf{F}^T. \tag{12}$$

With this notation, we have

$$R_{\mathrm{IF}}(\mathbf{K_{xx}}) = \frac{K}{2}\log\left(\min_{\substack{\mathbf{A}\in\mathbb{Z}^{K\times K}\\ \det\mathbf{A}\neq 0}} \max_{k=1,\ldots,K} \|\mathbf{F}^T\mathbf{a_k}\|^{\mathbf{2}}\right). \tag{13}$$

Denote by $\Lambda(\mathbf{F}^T)$ the K-dimensional lattice spanned by the matrix $\mathbf{F}^T$, i.e.,

$$\Lambda(\mathbf{F}^T) \triangleq \left\{\mathbf{F}^T\mathbf{a}:\ \mathbf{a}\in\mathbb{Z}^{\mathbf{K}}\right\}. \tag{14}$$

Then the problem of finding the optimal matrix $\mathbf{A}$ is equivalent to finding the shortest set of K linearly independent vectors in $\Lambda(\mathbf{F}^T)$. Denoting the kth successive minimum of the lattice by $\lambda_k(\mathbf{F}^T)$, we have

$$R_{\mathrm{IF}}(\mathbf{K_{xx}}) = \frac{K}{2}\log\left(\lambda_K^2(\mathbf{F}^T)\right). \tag{15}$$

Just as successive interference cancellation significantly improves the achievable rate of IF equalization in channel coding, an analogous scheme can be implemented in the case of IF source coding. Specifically, for a given full-rand integer matrix $\mathbf{A}$, let $\mathbf{L}$ be defined by the Cholesky decomposition

$$\mathbf{A}\left(\mathbf{I}+\mathbf{K_{xx}}\right)\mathbf{A}^T = \mathbf{L}\mathbf{L}^T \tag{16}$$

and denote the mth element of the diagonal of $\mathbf{L}$ by $\ell_{m,m}$. Then, as shown in [6] and [7], the achievable rate of successive IF source coding (which we denote as IF-SUC) for this choice of $\mathbf{A}$ is given by

$$R_{\mathrm{IF-SUC}}(\mathbf{K_{xx}};\mathbf{A}) = K\cdot\max_{m=1,\ldots,K} R_{\mathrm{IF-SUC},m}(\mathbf{K_{xx}}), \tag{17}$$

where

$$R_{\mathrm{IF-SUC},m}(\mathbf{K_{xx}};\mathbf{A}) = \frac{1}{2}\log\left(\ell^2_{m,m}\right). \tag{18}$$

Finally, by optimizing over the choice of $\mathbf{A}$, we obtain

$$R_{\mathrm{IF-SUC}}(\mathbf{K_{xx}}) = \min_{\substack{\mathbf{A}\in\mathbb{Z}^{K\times K}\\ \det\mathbf{A}\neq 0}} R_{\mathrm{IF-SUC}}(\mathbf{K_{xx}};\mathbf{A}). \tag{19}$$

While the gap between R_{IF} (and even more so $R_{\text{IF-SUC}}$) and R_{BT} is quite small for most covariance matrices, it can nevertheless be arbitrarily large. We next quantify the measure of bad covariance matrices by considering randomly-precoded IF source coding.

III. CRE-Precoded IF Source Coding and its Empirical Performance

Recalling (10), and with a slight abuse of notation, the rate of IF source coding for a given precoding matrix $\mathbf{P}$ is denoted by

$$\begin{aligned} R_{\text{IF}}(\mathbf{K_{xx}}, \mathbf{P}) &\triangleq R_{\text{IF}}(\mathbf{P}\mathbf{K_{xx}}\mathbf{P}^T) \\ &= \frac{K}{2}\log\left(\min_{\substack{\mathbf{A}\in\mathbb{Z}^{K\times K} \\ \det \mathbf{A}\neq 0}} \max_{k=1,\ldots,K} \mathbf{a}_k^T\left(\mathbf{I}+\mathbf{P}\mathbf{K_{xx}}\mathbf{P}^T\right)\mathbf{a}_k\right). \end{aligned} \tag{20}$$

Since $\mathbf{K_{xx}}$ is symmetric, it allows orthonormal diagonalization

$$\mathbf{I}+\mathbf{K_{xx}} = \mathbf{U}\mathbf{D}\mathbf{U}^T. \tag{21}$$

When unitary precoding is applied, we have

$$\mathbf{I}+\mathbf{P}\mathbf{K_{xx}}\mathbf{P}^T = \mathbf{P}\mathbf{U}\mathbf{D}\mathbf{U}^T\mathbf{P}^T. \tag{22}$$

To quantify the measure of "bad" sources, we consider precoding matrices that are uniformly (Haar) distributed over the group of orthonormal matrices. Such a matrix ensemble is referred to as the circular real ensemble (CRE) and is defined by the unique distribution on orthonormal matrices that is invariant under left and right orthonornal transformations [8]. That is, given a random matrix $\mathbf{P}$ drawn from the CRE, for any orthonormal matrix $\grave{\mathbf{U}}$, both $\mathbf{P}\grave{\mathbf{U}}$ and $\grave{\mathbf{U}}\mathbf{P}$ are equal in distribution to $\mathbf{P}$. Since $\mathbf{P}\mathbf{U}^T$ is equal in distribution to $\mathbf{P}$ for CRE precoding, for the sake of computing outage probabilities, we may simply assume that $\mathbf{U}^T$ (and also $\mathbf{U}$) is drawn from the CRE.

For a specific choice of integer vector $\mathbf{a_k}$, we define (again, with a slight abuse of notation)

$$\begin{aligned} R_{\text{IF}}(\mathbf{D}, \mathbf{U}; \mathbf{a}_k) &\triangleq \frac{1}{2}\log\left(\mathbf{a}_k^T\mathbf{U}\mathbf{D}\mathbf{U}^T\mathbf{a}_k\right) \\ &= \frac{1}{2}\log\left(\|\mathbf{D}^{1/2}\mathbf{U}^T\mathbf{a}_k\|^2\right), \end{aligned} \tag{23}$$

and correspondingly

$$\begin{aligned} &R_{\text{IF}}(\mathbf{D}, \mathbf{U}) \triangleq \\ &\frac{K}{2}\log\left(\min_{\substack{\mathbf{A}\in\mathbb{Z}^{K\times K} \\ \det \mathbf{A}\neq 0}} \max_{k=1,\ldots,K} \|\mathbf{D}^{1/2}\mathbf{U}^T\mathbf{a}_k\|^2\right). \end{aligned} \tag{24}$$

Let Λ be the lattice spanned by $\mathbf{G} = \mathbf{D}^{1/2}\mathbf{U}^T$. Then (24) may be rewritten as

$$R_{\text{IF}}(\mathbf{D}, \mathbf{U}) = \frac{K}{2}\log\left(\lambda_K^2(\Lambda)\right). \tag{25}$$

Let us denote the set of all diagonal matrices having the same value of R_{BT}, i.e.,

$$\mathbb{D}(R_{\text{BT}}, K) = \{\mathbf{D} \ : \ \det(\mathbf{D}) = 2^{2R_{\text{BT}}}\}. \tag{26}$$

We may thus rewrite the worst-case outage probability of IF source coding, defined in (9), as

$$\begin{aligned} &P_{\text{out,IF}}^{\text{WC}}(R_{\text{BT}}, \Delta R) \\ &= \sup_{\mathbf{D}\in\mathbb{D}(R_{\text{BT}},K)} \Pr\left(R_{\text{IF}}(\mathbf{D}, \mathbf{U}) > R_{\text{BT}} + \Delta R\right), \end{aligned} \tag{27}$$

where the probability is with respect to the random selection of $\mathbf{U}$ that is drawn from the CRE.

To illustrate the worst-case performance of CRE-precoded IF, we present its empirical performance for the case of a two-dimensional compound Gaussian source vector, where the outage probability (27) is computed via Monte-Carlo simulation.

Figure 1 depicts the results for different values of R_{BT}. Rather than plotting the worst-case outage probability, its complement is depicted, i.e., we plot the probability that the rate of IF falls below $R_{\text{BT}} + \Delta R$. As can be seen from the figure, the WC outage probability (as a function of ΔR) converges to a limiting curve as R_{BT} increases.

Figure 2 depicts the results for the single (high) rate $R_{\text{BT}} = 16$. The required compression rate required to support a given worst-case outage probability constraint is marked, for several outage probabilities. We observe that:

- For 10% worst-case outage probability, a gap of $\Delta R = 3.292$ bits (or 1.646 bits per source) is required.

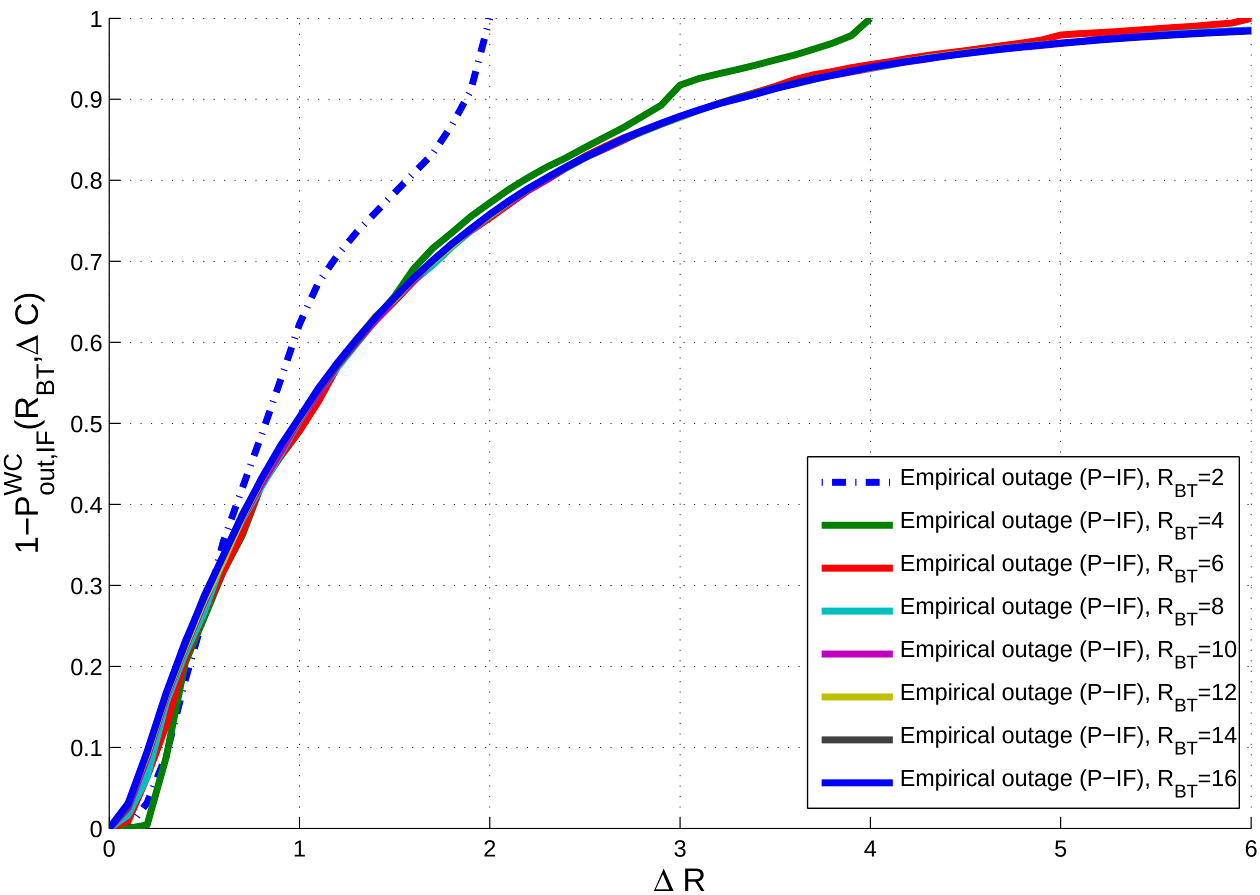

Fig. 1. Empirical results for (the complement of) the worst-case outage probability of IF source coding when applied to a two-dimensional compound Gaussian source vector as a function of ΔR, for various values of $R_{\rm BT}$.

- For 5% worst-case outage probability, a gap of $\Delta R = 4.293$ bits (or 2.1465 bits per source) is required.
- For 1% worst-case outage probability, a gap of $\Delta R = 6.665$ bits (or 3.3325 bits per source) is required.

IV. Upper Bounds on the Outage Probability for CRE-Precoded Integer-Forcing Source Coding

In this section we develop achievability bounds for randomly-precoded IF source coding. As the derivation is very much along the lines of the results for the analogous problem in channel coding as developed in [3], we refer to results from the latter in many points.

The next lemma provides an upper bound on the outage probability of precoded IF source coding as a function of $R_{\rm BT}$ and the rate gap ΔR (as well as the number of sources and $d_{\max}$ defined below). Denote

$$\mathbb{A}(\beta,\delta,K) = \left\{ \mathbf{a} \in \mathbb{Z}^K : 0 < \|\mathbf{a}\| < \sqrt{\frac{\beta}{\delta}} \right\}. \tag{28}$$

Lemma 1. *For any K Gaussian sources such that $\mathbf{D} \in \mathbb{D}(R_{\rm BT}, K)$, and for $\mathbf{U}$ drawn from the CRE, we have*

$$\begin{aligned} &\Pr\left(R_{\rm IF}(\mathbf{D},\mathbf{U}) < R_{\rm BT} + \Delta R\right) \\ &\leq \sum_{\mathbf{a}\in\mathbb{A}(\beta,1/d_{\max},K)} \frac{K\alpha(K)^{\frac{K-1}{2}} 2^{-\frac{K-1}{K}(R_{\rm BT}+\Delta R)}}{\frac{\|\mathbf{a}\|^{K-1}}{2^{R_{\rm BT}}}\sqrt{d_{\max}}} \end{aligned} \tag{29}$$

where $d_{\max} = \max_i \mathbf{D}_{i,i}$, the set $\mathbb{A}(\beta, 1/d_{\max}, K)$ is defined in (28)*, and where $\alpha(K)$ is defined in* (34) *below.*

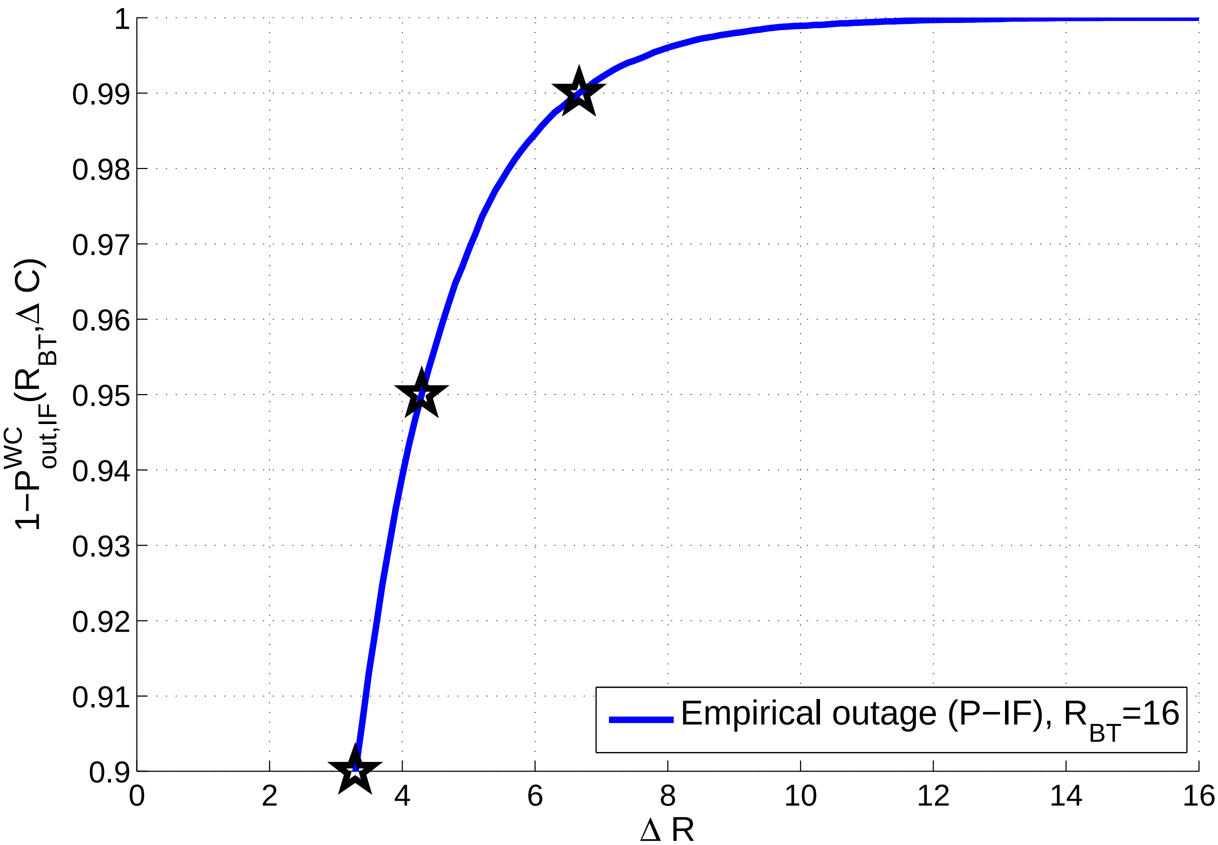

Fig. 2. Zoom in on the empirical worst-case outage probability for a two-dimensional compound Gaussian source vector with $R_{\mathrm{BT}} = 16$.

Proof. Let Λ^* denote the dual lattice of Λ and note that it is spanned by the matrix

$$(\mathbf{G}^T)^{-1} = \mathbf{D}^{-1/2}\mathbf{U}^T. \tag{30}$$

The successive minima of Λ and Λ^* are related by (Theorem 2.4 in [9])

$$\lambda_1(\Lambda^*)^2\lambda_K(\Lambda)^2 \leq \frac{K+3}{4}\bar{\gamma}_K^2, \tag{31}$$

where

$$\bar{\gamma}_K = \max\{\gamma_i : 1 \leq i \leq K\} \tag{32}$$

with γ_K denoting Hermite's constant.

The tightest known bound for Hermite's constant, as derived in [10], is

$$\gamma_K \leq \left(\frac{2}{\pi}\right)\Gamma\left(2+\frac{K}{2}\right)^{2/K}. \tag{33}$$

Since this is an increasing function of K, it follows that $\bar{\gamma}_K$ is smaller than the r.h.s. of (33). Combining the latter with the exact values of the Hermite constant for dimensions for which it is known, we define

$$\alpha(K) = \begin{cases} \frac{K+3}{4}\gamma_K^2, & K = 1-8, 24 \\ \frac{K+3}{4}\frac{2}{\pi}\Gamma\left(2+\frac{K}{2}\right)^{2/K}, & \text{otherwise} \end{cases}. \tag{34}$$

Therefore, we may bound the achievable rates of IF via the dual lattice as follows

$$R_{\mathrm{IF}}(\mathbf{D},\mathbf{U}) \leq \frac{K}{2}\log\left(\alpha(K)\frac{1}{\lambda_1(\Lambda^*)^2}\right). \tag{35}$$

Hence, we have

$$\begin{aligned}&\Pr\left(R_{\mathrm{IF}}(\mathbf{D},\mathbf{U}) > R_{\mathrm{BT}}+\Delta R\right)\\&\leq \Pr\left(\frac{K}{2}\log\left(\alpha(K)\frac{1}{\lambda_1(\Lambda^*)^2}\right) > R_{\mathrm{BT}}+\Delta R\right)\\&= \Pr\left(\lambda_1(\Lambda^*)^2 < \alpha(K)2^{-\frac{2}{K}(R_{\mathrm{BT}}+\Delta R)}\right).\end{aligned} \tag{36}$$

Denote

$$\beta = \alpha(K)2^{-\frac{2}{K}(R_{\mathrm{BT}}+\Delta R)}. \tag{37}$$

We wish to bound (36), or equivalently, we wish to bound

$$\Pr\left(\lambda_1^2(\Lambda^*) < \beta\right) = \Pr\left(\lambda_1(\Lambda^*) < \sqrt{\beta}\right) \tag{38}$$

for a given matrix $\mathbf{D} \in \mathbb{D}(R_{\mathrm{BT}}, K)$. Note that the event $\lambda_1(\Lambda^*) < \sqrt{\beta}$ is equivalent to the event

$$\bigcup_{\mathbf{a}\in\mathbb{Z}^K\setminus\{\mathbf{0}\}} ||\mathbf{D}^{-1/2}\mathbf{U}^T\mathbf{a}|| < \sqrt{\beta}. \tag{39}$$

Applying the union bound yields

$$\Pr\left(\lambda_1(\Lambda^*) < \sqrt{\beta}\right) \leq \sum_{\mathbf{a}\in\mathbb{Z}^K\setminus\{\mathbf{0}\}} \Pr\left(||\mathbf{D}^{-1/2}\mathbf{U}^T\mathbf{a}|| < \sqrt{\beta}\right). \tag{40}$$

Note that whenever $\frac{||\mathbf{a}||}{\sqrt{d_{\max}}} \geq \sqrt{\beta}$, we have

$$\Pr\left(||\mathbf{D}^{-1/2}\mathbf{U}^T\mathbf{a}|| < \sqrt{\beta}\right) = 0. \tag{41}$$

Therefore, substituting $1/d_{\max}$ in (28), the set of relevant vectors $\mathbf{a}$ is

$$\mathbb{A}(\beta, 1/d_{\max}, K) = \left\{\mathbf{a}\in\mathbb{Z}^K : 0 < ||\mathbf{a}|| < \sqrt{\beta d_{\max}}\right\}. \tag{42}$$

It follows from (40) and (41) that

$$\Pr\left(\lambda_1(\Lambda^*) < \sqrt{\beta}\right) \leq \sum_{\mathbf{a}\in\mathbb{A}(\beta,1/d_{\max},K)} \Pr\left(\|\mathbf{D}^{-1/2}\mathbf{U}^T\mathbf{a}\| < \sqrt{\beta}\right). \tag{43}$$

The rest of the proof follows the footsteps of Lemma 2 in [3] and is given in Appendix A. □

While Lemma 1 provides an explicit bound on the outage probability, in order to calculate it, one needs to go over all diagonal matrices in $\mathbb{D}(R_{\mathrm{BT}}, K)$ and for each such diagonal matrix, sum over all the relevant integer vectors in $\mathbb{A}(\beta, 1/d_{\max}, K)$. Hence, the bound can be evaluated only for moderate compression rates and for a small number of sources. The following theorem, that may be viewed as the counterpart of Theorem 1 in [3], provides a looser, yet very simple closed-form bound. Another advantage for this bound is that it does not depend on the Berger-Tung achievable rate.

Theorem 1. *For any K sources such that $\mathbf{D} \in \mathbb{D}(R_{\mathrm{BT}}, K)$, and for $\mathbf{U}$ drawn from the CRE we have*

$$\Pr\left(R_{\mathrm{IF}}(\mathbf{D},\mathbf{U}) > C+\Delta R\right) \leq c(K)2^{-\Delta R}, \tag{44}$$

where

$$c(K) = K\alpha(K)^{\frac{K}{2}}(K+c_{\max})\frac{\pi^{K/2}}{\Gamma(K/2+1)}, \tag{45}$$

$$\alpha(K) = \frac{K+3}{4}\frac{2}{\pi}\Gamma\left(2+\frac{K}{2}\right)^{2/K} \tag{46}$$

and

$$c_{\max} = \begin{cases}\left(2+\frac{\sqrt{K}}{2}\right)^K - \left(1-\frac{\sqrt{K}}{2}\right)^K & 1\leq K<4\\ \left(1+\sqrt{K}\right)^K & K\geq 4\end{cases}. \tag{47}$$

Note that $c(K)$ is a constant that depends only on the number of sources K.

Proof. See Appendix B. □

Similarly to the case of IF channel coding (cf., Section IV-C in [3]), analyzing Theorem 1 reveals that there are two main sources for looseness that may be further tightened:

- *Union bound* - While there is an inherent loss in the union bound, in fact, some terms in the summation (43) may be completely dropped.[3] Specifically, using Corollary 1 in [3], the set $\mathbb{A}(\beta, 1/d_{\max}, K)$ appearing in the summation in (1) may be replaced by the smaller set $\mathbb{B}(\beta, 1/d_{\max}, K)$ where

$$\mathbb{B}(\beta, d, K) = \left\{ \mathbf{a} \in \mathbb{Z}^K : 0 < \|\mathbf{a}\| < \sqrt{\frac{\beta}{d}} \text{ and } \nexists 0 < c < 1 \text{ s.t. } c\mathbf{a} \in \mathbb{Z}^K \right\}. \tag{48}$$

- *Dual Lattice* - Bounding via the dual lattice induces a loss reflected in (31). This may be circumvented for the case of a two-dimensional source vector by using IF-SUC, as accomplished in Lemma 2 and Theorem 2 which we present next.

Lemma 2. *For a two-dimensional Gaussian source vector such that* $\mathbf{D} \in \mathbb{D}(R_{\mathrm{BT}}, K)$*, and for* $\mathbf{U}$ *drawn from the CRE, we have*

$$\Pr\left(R_{\mathrm{IF-SUC}}(\mathbf{D}, \mathbf{U}) > R_{\mathrm{BT}} + \Delta R\right) \leq \sum_{\mathbf{a} \in \mathbb{B}(\beta, d_{\min}, K)} \frac{2\sqrt{\beta}}{\|\mathbf{a}\| 2^{R_{\mathrm{BT}}} \frac{1}{\sqrt{d_{\min}}}} \tag{49}$$

where $d_{\min} = \min_i \mathbf{D}_{i,i}$ *and* $\mathbb{A}(\beta, d_{\min}, K)$ *is defined in* (28)*,*

Theorem 2. *For a two-dimensional Gaussian source vector such that* $\mathbf{D} \in \mathbb{D}(R_{\mathrm{BT}}, K)$*, and for* $\mathbf{U}$ *drawn from the CRE, we have*

$$\Pr\left(R_{\mathrm{IF-SUC}}(\mathbf{D}, \mathbf{U}) > C + \Delta R\right) \leq c'(K) 2^{-\Delta R}, \tag{50}$$

where

$$c'(K) = 2\pi \left(5 + 3\sqrt{2}\right). \tag{51}$$

Proof. See Appendix C. □

Figure 3 depicts the bounds derived as well as results of a Monte Carlo evaluation of (9) for the case of a two-dimensional Gaussian and CRE precoding.[4] When calculating the empirical curves and the lemmas, we assumed high quantization rates ($R_{\mathrm{BT}} = 14$). The lemmas were calculated by going over a grid of values of d_1 and d_2 satisfying $d_1 d_2 = 2^{2 \cdot 14}$.

A. Application: Distributed Compression for Cloud Radio Access Networks

Since we described IF source coding as well as the precoding over the reals, we outline the application of IF source coding for the cloud radio access network (C-RAN) scenario assuming a real channel model. We then comment on the adaptation of the scheme to the more realistic scenario of a complex channel.

Consider the C-RAN scenario depicted in Figure 4 where M transmitters send their data (that is modeled as an i.i.d. Gaussian source vector) over a $K \times M$ MIMO broadcast channel $\mathbf{H} \in \mathbb{R}^{K \times M}$. The data is received at K receivers (relays) that wish to compress and forward it for processing (decoding) at a central node via rate-constrained noiseless bit pipes.

As we wish to minimize the distortion at the central node subject to the rate constraints, this is a distributed lossy source coding problem. See depiction in Figure 4.

Here, the covariance matrix of the received signals at the relays is given by

$$\mathbf{K}_{\mathbf{xx}} = \mathsf{SNR}\mathbf{H}\mathbf{H}^T + \mathbf{I}. \tag{52}$$

We note that we can "absorb" the SNR into the channel and hence we set $\mathsf{SNR} = 1$, so that

$$\mathbf{K}_{\mathbf{xx}} = \mathbf{H}\mathbf{H}^T + \mathbf{I}. \tag{53}$$

We further assume that the entries of the channel matrix $\mathbf{H}$ are Gaussian i.i.d., i.e. $\forall_{i,j} \mathbf{H}_{i,j} \sim \mathcal{N}(0, \sigma^2)$. As mentioned in the introduction, the SVD of this matrix

$$\mathbf{H} = \tilde{\mathbf{U}} \tilde{\mathbf{\Sigma}} \tilde{\mathbf{V}}^T \tag{54}$$

[3] Similar to the derivation in Section IV-C in [3], a simple factor of 2 can be deduced (regardless of the rate and number of sources) by noting that $\mathbf{a}$ and $-\mathbf{a}$ result in the same outcome and hence there is no need to account for both cases.

[4] The bounds are computed after applying a factor of 2 to the lemmas in accordance to footnote 3. Again, rather than plotting the worst-case outage probability, we plot its complement.

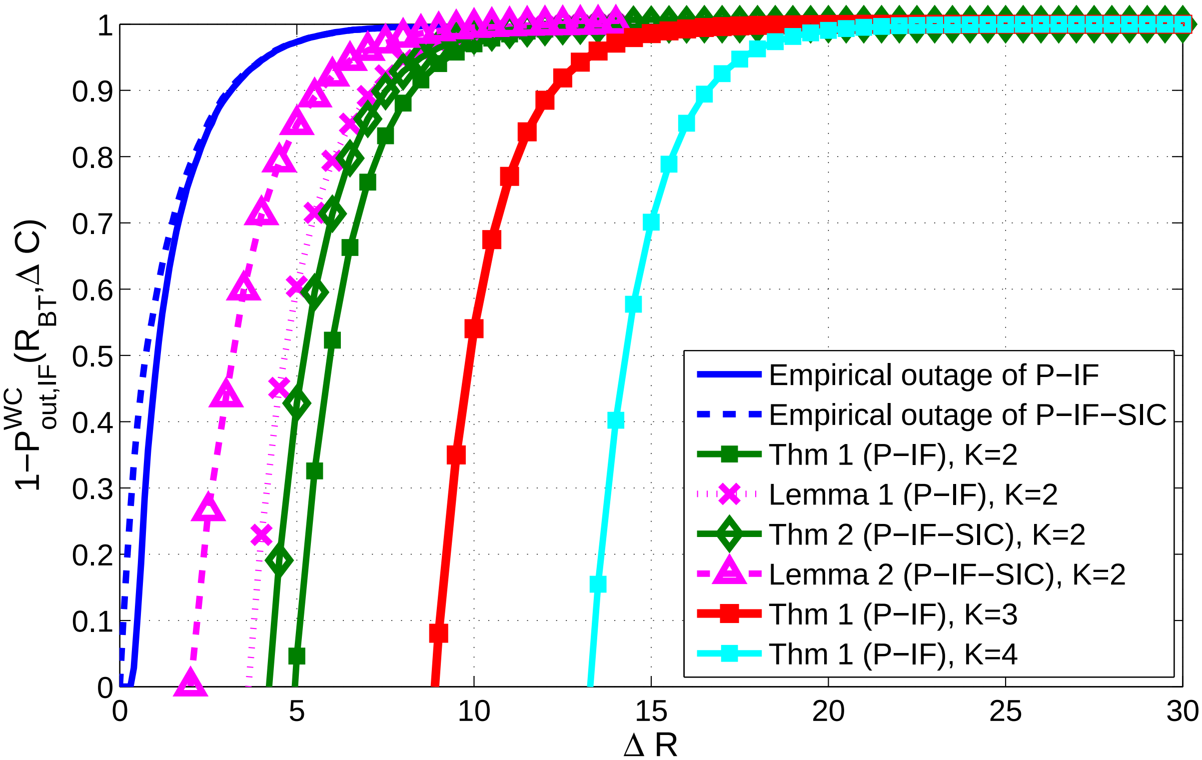

Fig. 3. Upper bounds on the outage probability of IF source coding for various values of source dimension K.

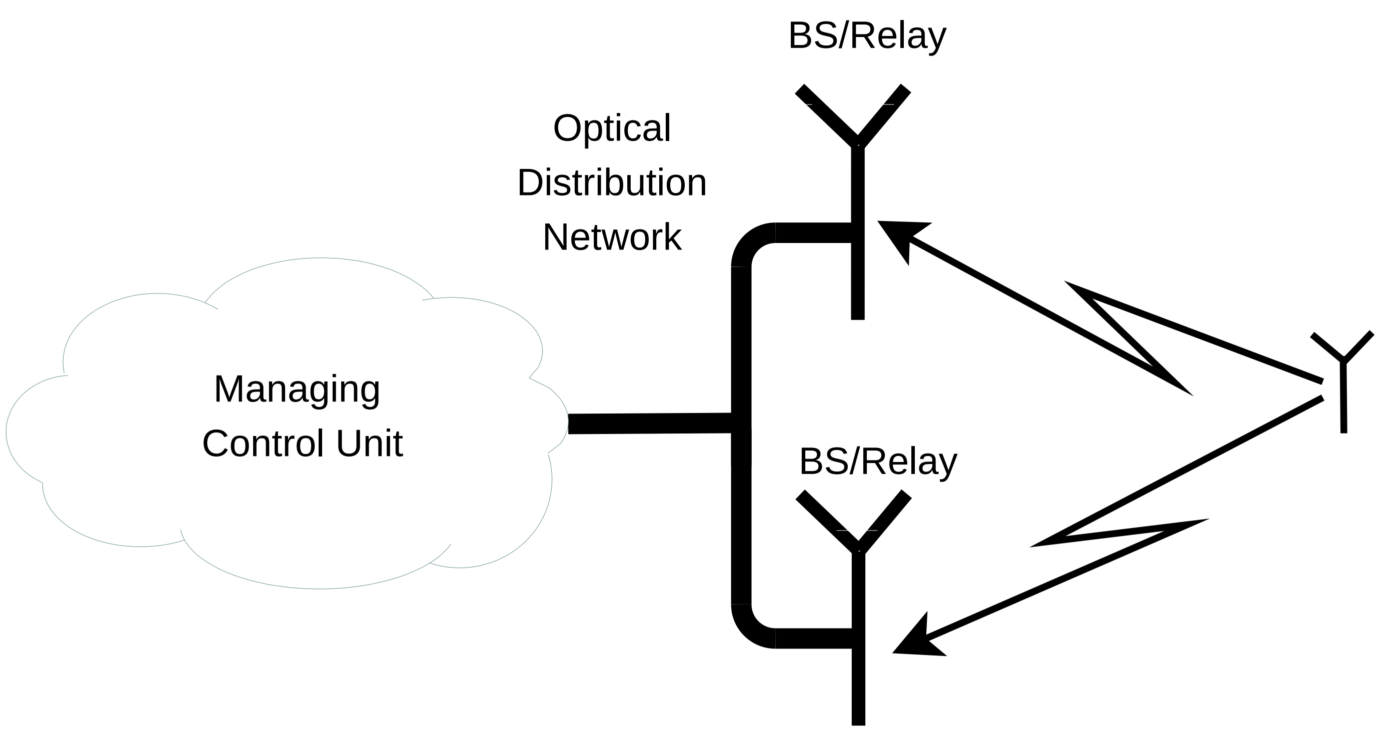

Fig. 4. Cloud radio access network communication scenario. Two relays compress and forward the correlated signals they receive from several users.

satisfies that $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{V}}$ belong to the CRE. We may therefore express the (random) covariance matrix as

$$\mathbf{K}_{\mathbf{xx}} = \tilde{\mathbf{U}}(\tilde{\mathbf{\Sigma}} + \mathbf{I})(\tilde{\mathbf{\Sigma}} + \mathbf{I})^T \tilde{\mathbf{U}}^T, \tag{55}$$

where $\tilde{\mathbf{U}}$ is drawn from the CRE. It follows that the precoding matrix $\mathbf{P}$ is redundant (as we assumed that $\mathbf{P}$ is also drawn from the CRE). Thus, the analysis above holds also for the considered scenario.

Specifically, assuming the encoders are subject to an equal rate constraint, then for a given distortion level, the relation between the compression rate of IF source coding and the guaranteed outage probability (for meeting the prescribed distortion) is bounded using Theorem 1 above.

We note that just as precoded IF channel coding can be applied to complex channels as described in [3], so can precoded IF source coding be extended to complex Gaussian sources. In describing an outage event in this case we assume that the precoding matrix is drawn from the circular unitary ensemble (CUE). The bounds derived above (replacing K with $2K$ in all derivations) for the relation between the compression rate of IF source coding and the worst-case outage probability hold for the C-RAN scenario over complex Gaussian channels $\mathbf{H} \in \mathbb{C}^{K \times M}$, where the CUE precoding can be viewed as been performed by nature.

V. Performance Guarantees for Integer-Forcing Source Coding with Deterministic Precoding

In this section, we consider the performance of IF source coding when used in conjunction with judiciously chosen (deterministic) precoding. Worst-case performance will be measured in a stricter sense than in previous sections; namely, no outage is allowed.

Achieving the goal of no outage for the case of general Gaussian sources requires performing space-time precoding, whereas for the case of parallel (independent) sources, space-only precoding still suffices.

We note that while doing away with outage events is desirable, it does come at a price. Namely, the precoding assumed in this section requires joint processing of the different sources prior to quantization and is thus precluded in a distributed setting.

On the other hand, the precoded IF scheme considered in this section has several advantages with respect to traditional source coding of correlated sources. Namely, the traditional approach requires utilizing the a statistical characterization of the source at the encoder side, e.g., via a prediction filter or via applying a transformation (such as DCT or DFT) along with *bit loading*.

In contrast, IF as considered in this section is applied after applying a *universal* transformation, i.e., a transformation that is independent of the source statistics. In addition, all samples are quantized at the same bit rate. Knowledge of the statistics of the source needs of course be utilized, but only at the decoder side, a propertry that may be advatageous in certain applications. This is similar to the case of Wyner-Ziv compression but whereas the latter requires in general the use of bit allocation (unless the source is stationary), IF source coding does not.

A. Additive Bound for General Sources

Similar to the case of channel coding, we can derive a worst-case additive bound for the gap to the Berger-Tung benchmark. Achieving this guaranteed performance requires joint algebraic number theoretic based space-time precoding at the encoders. The following theorem is due to Or Ordentlich [11].

Theorem 3 (Ordentlich)**.** *For any K sources with covariance matrix $\mathbf{K}_{\mathbf{xx}}$ and Berger-Tung benchmark R_{BT}, the excess rate with respect to the Berger-Tung benchmark (normalized per the number of time-extensions used) of space-time IF source coding with an NVD precoding matrix with minimum determinant $\delta_{\min}$ is bounded by*

$$R_{\mathrm{IF}} - R_{\mathrm{BT}} \le 2K^3 \log(2K^2) + K^2 \log \frac{1}{\delta_{\min}}. \tag{56}$$

Proof. See Appendix D. □

We note that similarly to the case of IF for channel coding, the gap to the Berger-Tung benchmark is large and thus it has limited applicability.

B. Uncorrelated Sources

For the special case of uncorrelated Gaussian sources, much tighter bounds (in comparison to Theorem 3) on the worst-case quantization rate of IF for a given distortion level may be obtained. First, space-time precoding may be replaced with precoding over space only. This allows to obtain a tighter counterpart to Theorem 3, as derived in Section 2.4.2 of [6].

We next derive yet tighter performance guarantees, also following ideas developed in [6], by numerically evaluating the performance of IF source coding over a “densely” quantized set of source (diagonal) covariance matrices belonging to the compound class, and then bounding the excess rate w.r.t. to the evaluated ones for any possible source vector in the compound class.

In the case of uncorrelated sources, the covariance matrix $\mathbf{K_{xx}}$ is diagonal. Hence, (55) becomes

$$\begin{aligned}\mathbf{K_{xx}} &= \tilde{\mathbf{\Sigma}}\tilde{\mathbf{\Sigma}}^T \\ &\triangleq \mathbf{S}, \end{aligned} \tag{57}$$

where

$$\mathbf{S} = \begin{bmatrix} s_1^2 & 0 & \cdots & 0 \\ 0 & s_2^2 & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & s_K^2 \end{bmatrix}. \tag{58}$$

We denote $\mathbf{s} = \mathrm{diag}(\mathbf{S})$. The compound set of channels may be parameterized by

$$\mathcal{S}(R_{\mathrm{BT}}) = \left\{ \mathbf{S} : \sum_{i=1}^{K} \frac{1}{2} \log\left(1 + s_i^2\right) = R_{\mathrm{BT}} \right\}. \tag{59}$$

We note that we may associate with each diagonal element a "rate" corresponding to an individual source

$$R_i = \frac{1}{2} \log\left(1 + s_i^2\right). \tag{60}$$

Thus, the compound class of sources may equivalently be represented by the set of rates

$$\mathcal{R}(R_{\mathrm{BT}}) = \left\{ (R_1, R_2, \ldots, R_K) \in \mathbb{R}^K : \sum_{i=1}^{K} R_i = R_{\mathrm{BT}} \right\}. \tag{61}$$

We define a "quantized" rate-tuple set as follows. The interval $[0, R_{\mathrm{BT}}]$ is divided into N sub-intervals, each of length $\Delta = R_{\mathrm{BT}}/N$. Thus, the resolution is determined by the parameter Δ. The quantized rate-tuples belong to the grid

$$\begin{aligned}\mathcal{R}^{\Delta}(R_{\mathrm{BT}}) = \Big\{ \left(R_1^{\Delta}, R_2^{\Delta}, \ldots, R_K^{\Delta}\right) \in \Delta \cdot R_{\mathrm{BT}} \cdot {\mathbb{Z}^+}^K : \\ \sum_{i=1}^{K} R_i^{\Delta} = R_{\mathrm{BT}} \Big\}. \end{aligned} \tag{62}$$

We may similarly define the (non-uniformly) quantized set $\mathcal{S}^{\Delta}(R_{\mathrm{BT}})$ of diagonal matrices such that the diagonal entries satisfy $s_{i,\Delta}^2 = 2^{2R_i^{\Delta}} - 1$, $i = 1, \ldots, K$, where $\mathbf{R}^{\Delta} \in \mathcal{R}^{\Delta}(R_{\mathrm{BT}})$.

Theorem 4. *For any Gaussian vector of independent sources with covariance matrix* $\mathbf{S}$ *such that* $\mathbf{S} \in \mathcal{S}(R_{\mathrm{BT}})$*, the rate of IF source coding with a given precoding matrix* $\mathbf{P}$ *is upper bound by*

$$R_{\mathrm{IF}}(\mathbf{S}, \mathbf{P}) \leq \max_{\mathbf{S}^{\Delta} \in \mathcal{S}^{\Delta}(R_{\mathrm{BT}})} R_{\mathrm{IF}}\left(\mathbf{S}^{\Delta}, \mathbf{P}\right) + K \log \eta \tag{63}$$

where

$$\eta = \frac{2^{2\left(\frac{R_{\mathrm{BT}}}{K}\right)} - 1}{2^{2\left(\frac{R_{\mathrm{BT}}}{K} - (K-1)\Delta R_{\mathrm{BT}}\right)} - 1}. \tag{64}$$

Proof. Assume we have a covariance matrix $\mathbf{S}$ in the compound class. Hence, its associated rate-tuple satisfies $(R_1, R_2, \ldots, R_K) \in \mathcal{R}(R_{\mathrm{BT}})$. Assume without loss of generality that

$$R_1 \leq R_2 \leq \ldots \leq R_K. \tag{65}$$

By (23), the rate of IF source coding associated with a specific integer linear combination vector $\mathbf{a}$ is

$$R_{\mathrm{IF}}(\mathbf{S}, \mathbf{P}; \mathbf{a}) = \frac{1}{2} \log\left(\mathbf{a}^T \mathbf{P} \left(\mathbf{I} + \mathbf{S}\right) \mathbf{P}^T \mathbf{a}\right). \tag{66}$$

Denote

$$\mathbf{v} = \mathbf{a}^T \mathbf{P}. \tag{67}$$

We will need the following two lemmas, whose proofs appear in Appendices E and F, respectively.

Lemma 3. *For any diagonal covariance matrix* $\mathbf{S}$*, associated with a rate-tuple* $(R_1, R_2, \ldots, R_K) \in \mathcal{R}(R_{\mathrm{BT}})$*, there exists a* *diagonal covariance matrix* $\mathbf{S}^{\Delta}$*, associated with a rate-tuple* $(R_1^{\Delta}, R_2^{\Delta}, \ldots, R_K^{\Delta}) \in \mathcal{R}^{\Delta}(R_{\mathrm{BT}})$*, such that*

$$s_i \leq \eta^2 s_{i,\Delta}^2 \tag{68}$$

for $1 \leq i \leq K$*, where* η *is defined in (64).*

Lemma 4. *Consider a Gaussian vector with a diagonal covariance matrix* $\mathbf{S}$ *and let* $\mathbf{a}$ *be an integer vector. Then for any* $\beta \geq 1$*, we have*

$$R_{\mathrm{IF}}(\beta^2\mathbf{S}, \mathbf{P}; \mathbf{a}) \leq \log(\beta) + R_{\mathrm{IF}}(\mathbf{S}, \mathbf{P}; \mathbf{a}). \tag{69}$$

Using Lemma 3, and denoting by $\mathbf{S}^{\Delta}$ the covariance matrix associated with $\mathbf{S}$, and whose existence is guaranteed by the lemma, it follows from (66) that

$$\begin{aligned} R_{\mathrm{IF}}(\eta^2\mathbf{S}^{\Delta}, \mathbf{P}; \mathbf{a}) &= \frac{1}{2}\log\left(\sum_{i=1}^{K} v_i^2(1+\eta^2 s_{i,\Delta}^2)\right) \\ &\geq \frac{1}{2}\log\left(\sum_{i=1}^{K} v_i^2(1+s_i^2)\right) \\ &= R_{\mathrm{IF}}(\mathbf{S}, \mathbf{P}; \mathbf{a}). \end{aligned} \tag{70}$$

Using Lemma 4, we further have that

$$R_{\mathrm{IF}}(\eta^2\mathbf{S}^{\Delta}, \mathbf{P}; \mathbf{a}) \leq \frac{1}{2}\log(\eta) + R_{\mathrm{IF}}(\mathbf{S}^{\Delta}, \mathbf{P}; \mathbf{a}). \tag{71}$$

Combining (70) and (71), we obtain

$$\begin{aligned} R_{\mathrm{IF}}(\mathbf{S}, \mathbf{P}; \mathbf{a}) &\leq R_{\mathrm{IF}}(\eta^2\mathbf{S}^{\Delta}, \mathbf{P}; \mathbf{a}) \\ &\leq \log(\eta) + R_{\mathrm{IF}}(\mathbf{S}^{\Delta}, \mathbf{P}; \mathbf{a}). \end{aligned} \tag{72}$$

Recalling (20), we have

$$R_{\mathrm{IF}}(\mathbf{S}^{\Delta}, \mathbf{P}) = \frac{K}{2}\log\left(\min_{\substack{\mathbf{A}\in\mathbb{Z}^{K\times K}\\ \det\mathbf{A}\neq 0}} \max_{k=1,\ldots,K} R_{\mathrm{IF}}(\mathbf{S}^{\Delta}, \mathbf{P}; \mathbf{a}_k)\right). \tag{73}$$

Denoting $\mathbf{A}^{\Delta}$ as the optimal integer matrix for the quantized channel $\mathbf{S}^{\Delta}$ (which is not necessarily the optimal matrix for $\mathbf{S}$), it follows that

$$R_{\mathrm{IF}}(\mathbf{S}, \mathbf{P}) \leq R_{\mathrm{IF}}(\mathbf{S}, \mathbf{P}; \mathbf{A}^{\Delta}). \tag{74}$$

Now assuming $\mathbf{a}_k^T$ are the rows of $\mathbf{A}^{\Delta}$, by (72) we have

$$\begin{aligned} R_{\mathrm{IF}}(\mathbf{S}, \mathbf{P}; \mathbf{A}^{\Delta}) &= K \max_{k=1,\ldots,K} R_{\mathrm{IF}}(\mathbf{S}, \mathbf{P}; \mathbf{a}_k) \\ &\leq K \max_{k=1,\ldots,K} \left(\log(\eta^2) + R_{\mathrm{IF}}(\mathbf{S}^{\Delta}, \mathbf{P}; \mathbf{a}_k)\right) \\ &= K\log(\eta) + R_{\mathrm{IF}}(\mathbf{S}^{\Delta}, \mathbf{P}; \mathbf{A}^{\Delta}). \end{aligned} \tag{75}$$

Thus, we conclude that

$$R_{\mathrm{IF}}(\mathbf{S}, \mathbf{P}) \leq R_{\mathrm{IF}}(\mathbf{S}^{\Delta}, \mathbf{P}) + K\log(\eta) \tag{76}$$

and therefore, for any $\mathbf{S} \in \mathcal{S}(R_{\mathrm{BT}})$, we have

$$R_{\mathrm{IF}}(\mathbf{S}, \mathbf{P}) \leq \max_{\mathbf{S}^{\Delta}\in\mathbb{S}^{\Delta}} R_{\mathrm{IF}}(\mathbf{S}^{\Delta}, \mathbf{P}) + K\log(\eta). \tag{77}$$

This concludes the proof of the theorem.

□

As an example for the achievable performance, show Fig. 5 gives the empirical worst-case performance for two and three (real) sources that is achieved when using IF source coding and a fixed precoding matrix over the grid $\mathcal{R}^{\Delta}(R_{\mathrm{BT}})$ for $\Delta = 0.01$. The precoding matrix was taken from [12]. The explicit precoding matrix used for two sources is $\mathbf{P} = \mathrm{cyclo}_2$ where

$$\mathrm{cyclo}_2 = \begin{bmatrix} -0.5257311121 & -0.8506508083 \\ -0.8506508083 & 0.5257311121 \end{bmatrix}, \tag{78}$$

and for three sources $\mathbf{P} = \text{cyclo}_3$ where

$$\text{cyclo}_3 = \begin{bmatrix} -0.3279852776 & -0.5910090485 & -0.7369762291 \\ -0.7369762291 & -0.3279852776 & 0.5910090485 \\ -0.5910090485 & 0.7369762291 & -0.3279852776 \end{bmatrix}. \tag{79}$$

Rather than plotting the gap from the Berger-Tung benchmark, we plot the efficiency $\eta = \frac{R_{IF}}{R_{\text{BT}}}$, i.e., the ratio of the (worst-case) rate of IF source coding and R_{BT}. We also plot the upper bound on $R_{\text{IF}}(\mathbf{S}, \mathbf{P})$ given by Theorem 4.

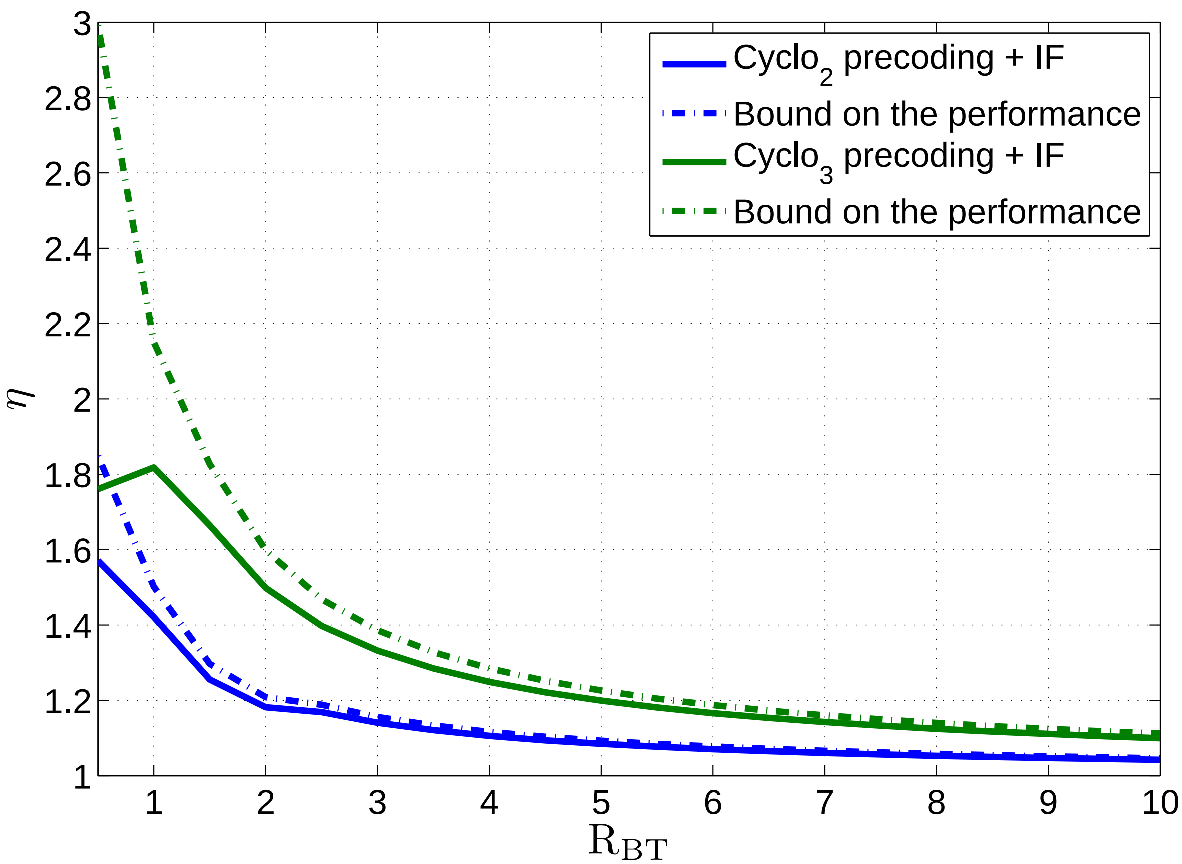

Fig. 5. Empirical and guaranteed (upper bound) worst-case efficiency of IF source coding for two and three uncorrelated Gaussian sources, when using the precoding matrices cyclo_2 and cyclo_3 given in (78)-(79), respectively, and taking $\Delta = 0.01$ for the calculation of Theorem 4.

Appendix A
Proof of Theorem 1

Following the footsteps of Lemma 2 in [3] and adopting the same geometric interpretation described there, we may interpret $\Pr\left(\|\mathbf{D}^{-1/2}\mathbf{U}^T\mathbf{a}\| < \sqrt{\beta}\right)$ as the ratio of the surface area of an ellipsoid that is inside a ball with radius $\sqrt{\beta}$ and the surface area of the entire ellipsoid. The axes of this ellipsoid are defined by $x_i = \frac{\|\mathbf{a}\|}{\sqrt{d_i}}$.

Denote the vector $\mathbf{o}_{\|\mathbf{a}\|}$ as a vector drawn from the CRE with norm $\|\mathbf{a}\|$. Using Lemma 1 in [3] and since we assume that $\mathbf{U}^T$ is drawn from the CRE, we have that the right hand side of (43) is equal to

$$\begin{aligned} &\sum_{\mathbf{a}\in\mathbb{A}(\beta,1/d_{\max},K)} \Pr\left(\|\mathbf{D}^{-1/2}\mathbf{o}_{\|\mathbf{a}\|} < \sqrt{\beta}\right) \\ &= \sum_{\mathbf{a}\in\mathbb{A}(\beta,1/d_{\max},K)} \frac{\text{CAP}_{\text{ell}}(\text{x}_1,\cdots,\text{x}_\text{K})}{L(x_1,\cdots,x_K)} \end{aligned} \tag{80}$$

where

$$\text{CAP}_{\text{ell}}(\text{x}_1,\cdots,\text{x}_\text{K}) < K\frac{\pi^{K/2}}{\Gamma(1+K/2)}\sqrt{\beta}^{K-1} \triangleq \overline{\text{CAP}_{\text{ell}}}, \tag{81}$$

and

$$L(x_1,\cdots,x_K) > \frac{\pi^{K/2}}{\Gamma(1+K/2)}\frac{\|\mathbf{a}\|^K}{\prod_{i=1}^K \sqrt{d_i}}\sum_{i=1}^K \frac{\sqrt{d_i}}{\|\mathbf{a}\|} \tag{82}$$

$$> \frac{\pi^{K/2}}{\Gamma(1+K/2)}\frac{\|\mathbf{a}\|^{K-1}\sqrt{d_{\max}}}{2^{R_{\mathrm{BT}}}} \triangleq \underline{L}. \tag{83}$$

Substituting (81) and (83) in (80), we obtain

$$\begin{aligned}&\sum_{\mathbf{a}\in\mathbb{A}(\beta,1/d_{\max},K)} \frac{\mathrm{CAP_{ell}}(\mathrm{x_1},\cdots,\mathrm{x_K})}{L(x_1,\cdots,x_K)}\\ &< \sum_{\mathbf{a}\in\mathbb{A}(\beta,1/d_{\max},K)} \frac{\overline{\mathrm{CAP_{ell}}}}{\underline{L}}\\ &= \sum_{\mathbf{a}\in\mathbb{A}(\beta,1/d_{\max},K)} \frac{K\sqrt{\beta}^{K-1}}{\frac{\|\mathbf{a}\|^{K-1}}{2^{R_{\mathrm{BT}}}}\sqrt{d_{\max}}}. \end{aligned} \tag{84}$$

Recalling (see (37)) that $\beta = \alpha(K)2^{-\frac{2}{K}(R_{\mathrm{BT}}+\Delta R)}$, we obtain

$$\begin{aligned}&\Pr\left(R_{\mathrm{IF-SUC}}(\mathbf{D},\mathbf{U}) > C+\Delta R\right)\\ &< \sum_{\mathbf{a}\in\mathbb{A}(\beta,1/d_{\max},K)} \frac{K\alpha(K)^{\frac{K-1}{2}}2^{-\frac{K-1}{K}(R_{\mathrm{BT}}+\Delta R)}}{\frac{\|\mathbf{a}\|^{K-1}}{2^{R_{\mathrm{BT}}}}\sqrt{d_{\max}}}. \end{aligned} \tag{85}$$

Appendix B
Proof of Theorem 1

To establish Theorem 1, we follow the footsteps of the proof of Theorem 1 in [3] to obtain

$$\begin{aligned}&\sum_{\mathbf{a}\in\mathbb{A}(\beta,1/d_{\max},K)} \Pr\left(\|\mathbf{D}^{-1/2}\mathbf{U}^T\mathbf{a}\| < \sqrt{\beta}\right)\\ &\le \sum_{\mathbf{a}\in\mathbb{A}(\beta,1/d_{\max},K)} \frac{K\sqrt{\beta}^{K-1}}{\frac{\|\mathbf{a}\|^{K-1}}{2^{R_{\mathrm{BT}}}}\sqrt{d_{\max}}} \end{aligned} \tag{86}$$

where $\mathbb{A}(\beta,1/d_{\max},K)$ and β are defined in (28) and (37), respectively. Noting that

$$\mathbb{A}(\beta,1/d_{\max},K) \subseteq \left\{\mathbf{a} : \|\mathbf{a}\| \le \left\lfloor\sqrt{\beta d_{\max}}\right\rfloor + 1\right\},$$

the summation in (86) can be bounded as

$$\le \sum_{k=0}^{\lfloor\sqrt{\beta d_{\max}}\rfloor} \sum_{k<\|\mathbf{a}\|\le k+1} \frac{K\sqrt{\beta}^{K-1}}{\frac{k^{K-1}}{2^{R_{\mathrm{BT}}}}\sqrt{d_{\max}}} \tag{87}$$

We apply Lemma 1 in [13] (a bound for the number of integer vectors contained in a ball of a given radius). Using this bound while noting that when $\|\mathbf{a}\| = 1$ there are exactly K integer vectors, the right hand side of (87) may be further bounded as

$$\begin{aligned}&\le \frac{K\sqrt{\beta}^{K-1}2^{R_{\mathrm{BT}}}}{\sqrt{d_{\max}}}\mathrm{Vol}(\mathcal{B}_K(1)) \times \left[K + \sum_{k=1}^{\lfloor\sqrt{\beta d_{\max}}\rfloor}\right.\\ &\left.\frac{\left(k+1+\frac{\sqrt{K}}{2}\right)^K - \left(\max\left(k-\frac{\sqrt{K}}{2},0\right)\right)^K}{k^{K-1}}\right] \end{aligned} \tag{88}$$

where we note that (88) trivially holds when $\left\lfloor\sqrt{\beta d_{\max}}\right\rfloor = 0$ since $\mathbb{A}(\beta,1/d_{\max},K)$ is the empty set in this case and hence the left hand side of (80) evaluates to zero.

The right hand side of (88) can further be rewritten as

$$\frac{K\sqrt{\beta}^{K-1}2^{R_{\mathrm{BT}}}}{\sqrt{d_{\max}}}\mathrm{Vol}(\mathcal{B}_K(1)) \times \left[\underbrace{K}_{\mathrm{I}} + \underbrace{\sum_{k=1}^{\left\lfloor \frac{\sqrt{K}}{2} \right\rfloor} \frac{\left(k+1+\frac{\sqrt{K}}{2}\right)^K}{k^{K-1}}}_{\mathrm{II}} + \right.$$

$$\left. \underbrace{\sum_{k=\left\lfloor \frac{\sqrt{K}}{2} \right\rfloor+1}^{\left\lfloor \sqrt{\beta d_{\max}} \right\rfloor} \frac{\left[\left(k+1+\frac{\sqrt{K}}{2}\right)^K - \left(k-\frac{\sqrt{K}}{2}\right)^K\right]}{k^{K-1}}}_{\mathrm{III}} \right]. \tag{89}$$

We search for c_1 and c_2 (independent of k) such that

$$\left(k+1+\frac{\sqrt{K}}{2}\right)^K \leq c_1 k^{K-1} \tag{90}$$

for $1 \leq k \leq \left\lfloor \frac{\sqrt{K}}{2} \right\rfloor$, and

$$\left[\left(k+1+\frac{\sqrt{K}}{2}\right)^K - \left(k-\frac{\sqrt{K}}{2}\right)^K\right] \leq c_2 k^{K-1} \tag{91}$$

for $k \geq 1$, since it will then follow that

$$\begin{aligned} II + III &\leq \sum_{k=1}^{\left\lfloor \sqrt{\beta d_{\max}} \right\rfloor} \max(c_1, c_2) \\ &= \left\lfloor \sqrt{\beta d_{\max}} \right\rfloor \max(c_1, c_2). \end{aligned} \tag{92}$$

We note that since (again assuming $\left\lfloor \sqrt{\beta d_{\max}} \right\rfloor \geq 1$)

$$K \leq \sum_{k=1}^{\left\lfloor \sqrt{\beta d_{\max}} \right\rfloor} K, \tag{93}$$

it will thus follow that

$$\begin{aligned} I + II + III &\leq \sum_{k=1}^{\left\lfloor \sqrt{\beta d_{\max}} \right\rfloor} \left[K + \max(c_1, c_2)\right] \\ &= \left\lfloor \sqrt{\beta d_{\max}} \right\rfloor \left[K + \max(c_1, c_2)\right]. \end{aligned} \tag{94}$$

An explicit derivation for c_1 and c_2 appears in Appendix B of [3] (where $2N_t$ should be replaced with K), from which we obtain

$$\begin{aligned} c_1 &= \left(1+\sqrt{K}\right)^K \\ c_2 &= \left[\left(2+\frac{\sqrt{K}}{2}\right)^K - \left(1-\frac{\sqrt{K}}{2}\right)^K\right]. \end{aligned} \tag{95}$$

In is also shown in Appendix B of [3] that for $K \geq 4$, $c_2 \leq c_1$ holds. For $1 \leq K < 4$ we observe that $c_1 \leq c_2$. This is so since $K < 4$ implies that $1 - \frac{\sqrt{K}}{2} > 0$, and hence indeed for $K < 4$ we have

$$\begin{aligned} c_2 &= \left[\left(2+\frac{\sqrt{K}}{2}\right)^K - \left(1-\frac{\sqrt{K}}{2}\right)^K\right] \\ &= \left(1+\sqrt{K}+1-\frac{\sqrt{K}}{2}\right)^K - \left(1-\frac{\sqrt{K}}{2}\right)^K \end{aligned}$$

$$
\begin{aligned}
&\geq (1+\sqrt{K})^K \\
&= c_1. \qquad (96)
\end{aligned}
$$

Recalling now (94) and denoting

$$
c_{\max} = \begin{cases} c_2 & 1 \leq K < 4 \\ c_1 & K \geq 4 \end{cases} , \qquad (97)
$$

it follows that

$$
I + II + III \leq \left\lfloor \sqrt{\beta d_{\max}} \right\rfloor (K + c_{\max}). \qquad (98)
$$

Using (98) and substituting the volume of a unit ball $\mathrm{Vol}(\mathcal{B}_K(1)) = \frac{\pi^{K/2}}{\Gamma(K/2+1)}$, it follows that right hand side of (88) is upper bounded by

$$
\begin{aligned}
&\frac{K\sqrt{\beta}^{K-1}2^{R_{\mathrm{BT}}}}{\sqrt{d_{\max}}}\frac{\pi^{K/2}}{\Gamma(K/2+1)}\sqrt{\beta d_{\max}}(K+c_{\max}) \qquad (99)\\
&= K\sqrt{\beta}^K 2^{R_{\mathrm{BT}}}(K+c_{\max})\frac{\pi^{K/2}}{\Gamma(K/2+1)}. \qquad (100)
\end{aligned}
$$

Finally, we substitute β, as defined in (37), into (100) to obtain

$$
\begin{aligned}
&\sum_{\mathbf{a}\in\mathbb{A}(\beta,1/d_{\max},K)} \Pr\left(\|\mathbf{D}^{-1/2}\mathbf{U}^T\mathbf{a}\| < \sqrt{\beta}\right) \\
&\leq K\left(\alpha(K)2^{-\frac{2}{K}(R_{\mathrm{BT}}+\Delta R)}\right)^{K/2} 2^{R_{\mathrm{BT}}}(K+c_{\max}) \\
&\leq K\left(\alpha(K)\right)^{K/2} 2^{-(R_{\mathrm{BT}}+\Delta R)}2^{R_{\mathrm{BT}}}(K+c_{\max}) \\
&\leq K\left(\alpha(K)\right)^{K/2}(K+c_{\max})2^{-\Delta R} \\
&= c(K)2^{-\Delta R}, \qquad (101)
\end{aligned}
$$

where $c(K)$ is as defined in (47).

Appendix C
Proof of Theorem 2

We first recall a theorem of Minkowski [14, Theorem 1.5] that upper bounds the product of the successive minima.

Theorem 5 (Minkowski)**.** *For any lattice $\Lambda(\mathbf{F}^T)$ that is spanned by a full rank $K \times K$ matrix $\mathbf{F}^T$*

$$
\prod_{m=1}^{K} \lambda_m^2(\mathbf{F}^T) \leq K^K (\det \mathbf{F}^T)^2. \qquad (102)
$$

To prove Theorem 2, we further need the following two lemmas.

Lemma 5. *For a Gaussian source vector with covariance matrix $\mathbf{K_{xx}} \in \mathcal{K}(R_{\mathrm{BT}})$, and for any full-rank integer matrix $\mathbf{A}$, the sum-rate of IF-SUC satisfies*

$$
\sum_{m=1}^{K} R_{\mathrm{IF-SUC},m}(\mathbf{K_{xx}};\mathbf{A}) = R_{\mathrm{BT}} + \log|\det(A)|, \qquad (103)
$$

where $R_{\mathrm{IF-SUC},m}(\mathbf{K_{xx}};\mathbf{A})$ is defined in (18)*.*

Proof.

$$
\begin{aligned}
\sum_{m=1}^{K} R_{\mathrm{IF-SUC},m}(\mathbf{K_{xx}};\mathbf{A}) &= \frac{1}{2}\sum_{m=1}^{K}\log(\ell_{m,m}^2) \\
&= \frac{1}{2}\log\left(\prod_{m=1}^{K}\ell_{m,m}^2\right) \\
&= \frac{1}{2}\log\det(\mathbf{L}\mathbf{L}^T)
\end{aligned}
$$

$$= \frac{1}{2} \log\det\left(\mathbf{A}\left(\mathbf{I} + \frac{1}{d}\mathbf{K_{xx}}\right)\mathbf{A}^T\right)$$
$$= R_{\mathrm{BT}} + \log|\det(A)| \tag{104}$$

□

Theorem 3 in [15] shows that for successive IF (used for channel coding) there is no loss (in terms of achievable rate) in restricting $\mathbf{A}$ to the class of unimodular matrices. We note that same claim holds also in our framework (that of successive integer-source source coding) by replacing $\mathbf{G}$, the matrix spanning the lattice which was defined in Theorem 3 as

$$\left(\mathbf{I} + \mathsf{SNR}\mathbf{H}\mathbf{H}^T\right)^{-1} = \mathbf{G}\mathbf{G}^T \tag{105}$$

with $\mathbf{F}$, as defined in (12), and noting that the optimal $\mathbf{A}$ can be expressed (in both cases) as

$$\mathbf{A}^{\mathrm{opt}}_{\mathrm{SIC/SUC}} = \min_{\substack{\mathbf{A}\in\mathbb{Z}^{K\times K}\\ \det\mathbf{A}\neq 0}} \max_{k=1,\ldots,K} \ell^2_{k,k}, \tag{106}$$

where $\ell_{k,k}$ are the diagonal elements of the corresponding $\mathbf{L}$ matrix (derived from the Cholesky decomposition) in each case.

Having established that the optimal $\mathbf{A}$ is unimodular, it now follows that $\sum_{m=1}^{K} R_{\mathrm{IF-SUC},m}(\mathbf{K_{xx}}; \mathbf{A}^{\mathrm{opt}}_{\mathrm{SUC}}) = R_{\mathrm{BT}}$.

We are now ready to prove the following lemma that is analogous to Theorem 3 in [16].

Lemma 6. *For a Gaussian source vector with covariance matrix $\mathbf{K_{xx}} \in \mathcal{K}(R_{\mathrm{BT}})$, and for the optimal integer matrix $\mathbf{A}$, the sum-rate of IF is upper bounded as*

$$\sum_{m=1}^{K} R_{\mathrm{IF},m}(\mathbf{K_{xx}}) \le R_{\mathrm{BT}} + \frac{K}{2}\log(K). \tag{107}$$

where $R_{m,\mathrm{IF}}(\mathbf{K_{xx}})$ is the rate of the mth equation (corresponding to the mth row of $\mathbf{A}$) as defined in (11).

Proof.

$$\sum_{m=1}^{K} R_{\mathrm{IF},m}(\mathbf{K_{xx}}) = \sum_{m=1}^{K} \frac{1}{2}\log\left(\lambda_m^2(\mathbf{F}^T)\right)$$
$$= \frac{1}{2}\log\left(\prod_{m=1}^{K}\lambda_m^2(\mathbf{F}^T)\right)$$
$$\le \frac{1}{2}\log\left(K^K|\det(F)|^2\right)$$
$$= R_{\mathrm{BT}} + \frac{K}{2}\log(K). \tag{108}$$

where the inequality is due to Theorem 5 (Minkowski's Theorem). □

Now, for the case of two sources we have by Lemma 5

$$R_{\mathrm{IF-SUC},1}(\mathbf{K_{xx}}) + R_{\mathrm{IF-SUC},2}(\mathbf{K_{xx}}) = R_{\mathrm{BT}}, \tag{109}$$

or equivalently

$$R_{\mathrm{IF-SUC},2}(\mathbf{K_{xx}}) = R_{\mathrm{BT}} - R_{\mathrm{IF-SUC},1}(\mathbf{K_{xx}}). \tag{110}$$

We further have by lemma 6 that

$$R_{\mathrm{IF},1}(\mathbf{K_{xx}}) + R_{\mathrm{IF},2}(\mathbf{K_{xx}}) \le R_{\mathrm{BT}} + 1. \tag{111}$$

We note that the (optimal) integer matrix $\mathbf{A}$ used for IF in (111) is in general different than the (optimal) matrix $\mathbf{A}$ used for IF-SUC in (109)-(110). Nonetheless, when applying IF-SUC, one decodes first the equation with the *lowest* rate. Since for this equation SUC has no effect, it follows that the first row of $\mathbf{A}$ is the same in both cases and hence

$$R_{\mathrm{IF-SUC},1}(\mathbf{K_{xx}}) = R_{\mathrm{IF},1}(\mathbf{K_{xx}}). \tag{112}$$

Since source 1 is decoded first, it follows that $R_{\mathrm{IF},1} > R_{\mathrm{IF},2}$ and hence

$$R_{\mathrm{IF},1}(\mathbf{K_{xx}}) \le \frac{R_{\mathrm{BT}} + 1}{2}. \tag{113}$$

Therefore,

$$\begin{aligned}R_{\mathrm{IF-SUC}}(\mathbf{K_{xx}}) &= 2\max\left(R_{\mathrm{IF-SUC},1}(\mathbf{K_{xx}}), R_{\mathrm{IF-SUC},2}(\mathbf{K_{xx}})\right)\\ &\leq 2\max\left(\frac{R_{\mathrm{BT}}+1}{2}, R_{\mathrm{BT}} - R_{\mathrm{IF},1}(\mathbf{K_{xx}})\right)\\ &= \max(R_{\mathrm{BT}}+1, 2R_{\mathrm{BT}} - 2R_{\mathrm{IF},1}(\mathbf{K_{xx}})). \end{aligned} \tag{114}$$

Henceforth, we analyze the outage for $R_{\mathrm{BT}} > 1$ and target rates that are no smaller than $R_{\mathrm{BT}}+1$, so that the inequality $2R_{\mathrm{BT}} - 2R_{\mathrm{IF},1}(\mathbf{K_{xx}}) > R_{\mathrm{BT}}+1$ is satisfied. Thus, we consider excess rate values satisfying $\Delta R > 1$. Our goal is to bound

$$\begin{aligned}&\Pr\left(R_{\mathrm{IF-SUC}}(\mathbf{K_{xx}}) \geq R_{\mathrm{BT}} + \Delta R\right)\\ &= \Pr\left(2R_{\mathrm{BT}} - 2R_{\mathrm{IF},1}(\mathbf{K_{xx}}) \geq R_{\mathrm{BT}} + \Delta R\right)\\ &= \Pr\left(R_{\mathrm{IF},1}(\mathbf{K_{xx}}) \leq \frac{R_{\mathrm{BT}} - \Delta R}{2}\right)\\ &= \Pr\left(\frac{1}{2}\log(\lambda_1^2(\mathbf{F}^T)) \leq \frac{R_{\mathrm{BT}} - \Delta R}{2}\right)\\ &= \Pr\left(\lambda_1^2(\mathbf{F}^T) \leq 2^{R_{\mathrm{BT}} - \Delta R}\right)\end{aligned} \tag{115}$$

Let $\beta = 2^{R_{\mathrm{BT}} - \Delta R}$. We wish to bound (115), or equivalently

$$\Pr\left(\lambda_1^2(\mathbf{F}^T) \leq \beta\right) = \Pr\left(\lambda_1(\mathbf{F}^T) \leq \sqrt{\beta}\right) \tag{116}$$

for a given matrix $\mathbf{D}$ corresponding to $\mathbf{K_{xx}}$ (via the relation (21)). Note that the event $\lambda_1(\Lambda) < \sqrt{\beta}$ is equivalent to the event

$$\bigcup_{\mathbf{a}\in\mathbb{Z}^{\mathbf{K}}\setminus\{\mathbf{0}\}} ||\mathbf{D}^{1/2}\mathbf{U}^T\mathbf{a}|| < \sqrt{\beta}. \tag{117}$$

Applying the union bound yields

$$\Pr\left(\lambda_1(\Lambda) < \sqrt{\beta}\right) \leq \sum_{\mathbf{a}\in\mathbb{Z}^K\setminus\{\mathbf{0}\}} \Pr\left(||\mathbf{D}^{1/2}\mathbf{U}^T\mathbf{a}|| < \sqrt{\beta}\right). \tag{118}$$

Using the same derivation as in Lemma 2 in [3], we get

$$\begin{aligned}&\Pr\left(R_{\mathrm{IF-SUC}}(\mathbf{D},\mathbf{U}) \geq R_{\mathrm{BT}} + \Delta R\right)\\ &\leq \sum_{\mathbf{a}\in\mathbb{A}(\beta,d_{\min},K)} \frac{K\sqrt{\beta}^{K-1}}{\|\mathbf{a}\|^{K-1}2^{R_{\mathrm{BT}}}\frac{1}{\sqrt{d_{\min}}}}.\end{aligned} \tag{119}$$

Since we are analyzing the case of two sources, we have

$$\begin{aligned}&\Pr\left(R_{\mathrm{IF-SUC}}(\mathbf{D},\mathbf{U}) \geq R_{\mathrm{BT}} + \Delta R\right)\\ &\leq \sum_{\mathbf{a}\in\mathbb{A}(\beta,d_{\min},K)} \frac{2\sqrt{\beta}}{\|\mathbf{a}\|2^{R_{\mathrm{BT}}}\frac{1}{\sqrt{d_{\min}}}}.\end{aligned} \tag{120}$$

Applying a similar argument as appears in Appendix B (as part of the proof of Theorem 1), and noting that $K < 4$ implies that $c_{\max} = c_2$, we get

$$\begin{aligned}&\Pr\left(R_{\mathrm{IF-SUC}}(\mathbf{D},\mathbf{U}) \geq R_{\mathrm{BT}} + \Delta R)\right)\\ &\qquad\leq \frac{2\beta}{2^{R_{\mathrm{BT}}}}\left(2+3+3\sqrt{2}\right)\pi.\end{aligned} \tag{121}$$

Finally, substituting β as defined in (37), we obtain

$$\begin{aligned}&\Pr\left(R_{\mathrm{IF-SUC}}(\mathbf{D},\mathbf{U}) \geq R_{\mathrm{BT}} + \Delta R\right)\\ &\leq \frac{2\pi\cdot 2^{R_{\mathrm{BT}}-\Delta R}}{2^{R_{\mathrm{BT}}}}\left(2+3+3\sqrt{2}\right)\\ &= 2\pi\left(5+3\sqrt{2}\right)2^{-\Delta R}.\end{aligned} \tag{122}$$

APPENDIX D
ADDITIVE (WORST-CASE) BOUND FOR NVD SPACE-TIME PRECODED SOURCES

Combining space-time precoding and integer forcing in the context of channel coding was suggested in [17], as we next briefly recall. We then present the necessary modifications for the case of source coding.

We derive below an additive bound using a unitary precoding matrix satisfying a non-vanishing determinant (NVD) property. As the theory of NVD space-time codes has been developed over the complex field, it will prove convenient for us to employ complex precoding matrices. To this end, we may assume that we stack samples from two time slots where the samples stacked at the first time slot represent the real part of a complex number and the samples stacked at the second time slot represent the imaginary part of a complex number. Hence, we have

$$\mathbf{K}_{\hat{x}\hat{x}} = \mathbf{I}_{2\times 2} \otimes \mathbf{K}_{xx}, \tag{123}$$

where $\otimes$ is the Kronecker product. We note that the Berger-Tung benchmark of this stacked source vector is

$$R_{\mathrm{BT}}(\mathbf{K}_{\hat{x}\hat{x}}) = 2R_{\mathrm{BT}}(\mathbf{K}_{xx}). \tag{124}$$

Next, in order to allow space-time precoding, we stack T times the K "complex" outputs of the K sources and let $\bar{\boldsymbol{x}}_c \in \mathbb{R}^{2KT\times 1}$ denote the effective source vector. Its corrlation matrix, which we refer to as the effective covariance matrix, takes the form

$$\mathcal{K} = \mathbf{I}_{T\times T} \otimes \mathbf{K}_{\hat{x}\hat{x}}. \tag{125}$$

We assume a precoding matrix, that in principle can be either deterministic or random, is applied to the effective source vector. We analyze performance for the case where the precoding matrix $\mathbf{P}_{st,c}$ is deterministic, specifically a precoding matrix induced by a perefect space-time block code, operating on the stacked source $\bar{\boldsymbol{x}}_c$ having covariance matrix as given in (125). An explanation on how to extract the precoding matrix from a space-time code can be found in Section IV in [18].

We denote the corresponding precoding matrix over the reals as $\mathbf{P}_{st} \in \mathbb{R}^{2KT\times 2KT}$. We denote

$$\mathbf{I} + \mathbf{P}_{st}\mathcal{K}\mathbf{P}_{st}^H = \mathcal{F}\mathcal{F}^T. \tag{126}$$

As we assume that the precoding matrix is unitary, the Berger-Tung benchmark (normalized by the total number of time extensions used) remains unchanged, i.e.,

$$\begin{aligned}\frac{1}{2T}\frac{1}{2}\log\det\left(\mathbf{I} + \mathbf{P}_{st}\mathcal{K}\mathbf{P}_{st}^H\right) &= \frac{1}{2T}\frac{1}{2}\log\left(\det\mathcal{F}^T\right)^2 \\ &= R_{\mathrm{BT}}.\end{aligned} \tag{127}$$

As noted above, we assume that the generating matrix of a perfect code [19], [20] is employed as a precoding matrix.

A space-time code is called perfect if:

- It is full rate;
- It satisfies the non-vanishing determinant (NVD) condition;
- The code's generating matrix is unitary.

Let $\delta_{\min}$ denote the minimal non-vanishing determinant of this code. Such codes further use the minimal number of time extensions possible, i.e., $T = K$. Thus, we have a total of K^2 stacked complex samples. Subsisting $2K^2$ as the dimension (number of real samples jointly processsed) in (15), the rate of IF source coding for the time-stacked samples is given by

$$R_{\mathrm{IF}}(\mathcal{K}, \mathbf{P}_{st}) = (2K^2)\frac{1}{2}\log\left(\lambda^2_{2K^2}(\mathcal{F}^T)\right). \tag{128}$$

To bound $\lambda^2_{2K^2}(\mathcal{F}^T)$, we note that for every $2K^2$-dimensional lattice, we have

$$\lambda_1^{2(2K^2-1)}(\mathcal{F}^T)\lambda^2_{2K^2}(\mathcal{F}^T) \le \prod_{m=1}^{2K^2}\lambda_m^2(\mathcal{F}^T). \tag{129}$$

Using Minkowski's theorem (Theorem 5 in Appendix C), it follows that

$$\lambda^2_{2K^2}(\mathcal{F}^T) \le (2K^2)^{2K^2}(\det\mathcal{F}^T)^2\frac{1}{\lambda_1^{2(2K^2-1)}(\mathcal{F}^T)}. \tag{130}$$

Hence, the rate of IF source coding (normalized by the number of time extensions) can be bounded as:

$$\begin{aligned}&\frac{1}{2K}R_{\mathrm{IF}}(\mathcal{K}, \mathbf{P}_{st}) \\ &= \frac{1}{2K}(2K^2)\frac{1}{2}\log\left(\lambda^2_{2K^2}(\mathcal{F}^T)\right) \\ &\le \frac{K}{2}\log\left((2K^2)^{2K^2}(\det\mathcal{F}^T)^2\frac{1}{\lambda_1^{2(2K^2-1)}(\mathcal{F}^T)}\right) \\ &= K^3\log(2K^2) + \frac{K}{2}\log(\det(\mathcal{F}^T)^2) - \frac{K(2K^2-1)}{2}\log(\lambda_1^2(\mathcal{F}^T))\end{aligned}$$

$$= K^3 \log(2K^2) + 2K^2 R_{\mathrm{BT}} - \frac{K(2K^2-1)}{2} \log(\lambda_1^2(\mathcal{F}^T)). \tag{131}$$

We next use the results derived in [18] for channel coding (using NVD precoding). We note that since the covariance matrix is positive semi-definite, it may be written as

$$\mathbf{K}_{xx} = \mathbf{H}\mathbf{H}^T. \tag{132}$$

The covariance matrix of the stacked source vector may be written as

$$\mathbf{K}_{\hat{x}\hat{x}} = \hat{\mathbf{H}}\hat{\mathbf{H}}^T, \tag{133}$$

where we may take $\hat{\mathbf{H}} = \mathbf{I}_{2\times 2} \otimes \mathbf{H}$.

There are many such choices of $\mathbf{H}$ and any such choice corresponds to a channel matrix $\hat{\mathbf{H}}$ that can be viewed as the real representation of a complex channel matrix (which in the present case is real, i.e., has no imaginary part) in the context of [18]. The effective covariance matrix can similarly be rewritten as

$$\mathcal{K} = \mathcal{H}\mathcal{H}^T, \tag{134}$$

where $\mathcal{H} = \mathbf{I}_{T\times T} \otimes \hat{\mathbf{H}}$.

Using the channel coding terminology of [18], we further define the minimum distance at the receiver $d_{\min}(\mathbf{H}, L)$ as

$$d_{\min}(\mathbf{H}, L) \triangleq \min_{\mathbf{a}\in \mathrm{QAM}^{\mathbf{K}}(\mathbf{L})\setminus \mathbf{0}} \|\mathbf{H}\mathbf{a}\|, \tag{135}$$

where

$$\begin{aligned} \mathrm{QAM}(L) \triangleq &\{-L, -L+1, \ldots, L-1, L\} \\ &+ i\{-L, -L+1, \ldots, L-1, L\}. \end{aligned} \tag{136}$$

Setting $\mathsf{SNR} = 1$ and for $\hat{\mathbf{H}} \in \mathbb{R}^{2K\times 2K}$ (which is the real representation of $\mathbf{H}_c \in \mathbb{C}^{K\times K}$), Lemma 2 in [18] states that

$$\begin{aligned} &\frac{1}{4K^4} \min_{\mathbf{a}\in\mathbb{Z}^{2K}\setminus 0} \mathbf{a}^T \left(\mathbf{I} + \mathbf{P}_{st}^H \mathcal{H}^H \mathcal{H} \mathbf{P}_{st}\right) \mathbf{a} \\ &\geq \frac{1}{4K^4} \min_{L=1,2,\cdots} \left(L^2 + \mathsf{SNR} d_{\min}(\mathcal{H}\mathbf{P}_{st}, L)\right). \end{aligned} \tag{137}$$

Using Corollary 1 in [18], we get

$$\begin{aligned} &\frac{1}{4K^4} \min_{\mathbf{a}\in\mathbb{Z}^{2K}\setminus 0} \mathbf{a}^T \left(\mathbf{I} + \mathbf{P}_{st}^H \mathcal{H}^H \mathcal{H} \mathbf{P}_{st}\right) \mathbf{a} \\ &\geq \frac{1}{4K^4} \min_{L=1,2,\cdots} \left(L^2 + \left[\delta_{\min}^{\frac{1}{K}} 2^{\frac{C_{\mathrm{WI}}}{K}} - 2K^2 L^2\right]^+\right) \\ &\geq \frac{1}{4K^4} \min_{L=1,2,\cdots} \left(L^2 + \left[\frac{\delta_{\min}^{\frac{1}{K}} 2^{\frac{C_{\mathrm{WI}}}{K}}}{2K^2} - L^2\right]^+\right) \\ &\geq \frac{1}{8K^6} \delta_{\min}^{\frac{1}{K}} 2^{\frac{C_{\mathrm{WI}}}{K}}, \end{aligned} \tag{138}$$

where $C_{\mathrm{WI}} = \frac{1}{2}\log\det\left(\mathbf{I} + \mathbf{H}^H\mathbf{H}\right)$ is the mutual information of $\mathbf{H}$. Since C_{WI} is the rate of a $2K\times 2K$ real matrix (resulting from a $K\times K$ complex matrix), it equals $R_{\mathrm{BT}}(\mathbf{K}_{\hat{x}\hat{x}})$ defined in (124). Hence, we obtain

$$\min_{\mathbf{a}\in\mathbb{Z}^{2K}\setminus 0} \mathbf{a}^T \left(\mathbf{I} + \mathbf{P}_{st}^H \mathcal{H}^H \mathcal{H} \mathbf{P}_{st}^H\right) \mathbf{a} \geq \frac{1}{2K^2} \delta_{\min}^{\frac{1}{K}} 2^{\frac{2R_{\mathrm{BT}}}{K}} \tag{139}$$

which in turn yields

$$\begin{aligned} \lambda_1^2(\mathcal{F}^T) &= \min_{\mathbf{a}\in\mathbb{Z}^{2K}\setminus 0} \mathbf{a}^T \left(\mathbf{I} + \mathbf{P}_{st} \mathcal{H}\mathcal{H}^T \mathbf{P}_{st}^T\right) \mathbf{a} \\ &\geq \frac{1}{2K^2} \delta_{\min}^{\frac{1}{K}} 2^{\frac{2R_{\mathrm{BT}}}{K}}. \end{aligned} \tag{140}$$

Finally, plugging the bound (140) into (131), we arrive at

$$\begin{aligned} &\frac{1}{2K} R_{\mathrm{IF}}(\mathcal{K}, \mathbf{P}_{st}) \\ &\leq 2K^2 R_{\mathrm{BT}} + K^3 \log(2K^2) + \frac{K(2K^2-1)}{2} \log(2K^2) \end{aligned}$$

$$+\frac{K(2K^2-1)}{2K}\log(\frac{1}{\delta_{\min}})-\frac{K(2K^2-1)}{2K}2R_{\mathrm{BT}}$$
$$\leq R_{\mathrm{BT}}+2K^3\log(2K^2)+K^2\log\frac{1}{\delta_{\min}}. \tag{141}$$

Appendix E
Proof of Lemma 3

A Gaussian source component with a specific rate R_i can be transformed to a different Gaussian source with rate R_i^Δ by appropriate scaling. Specifically, scaling each source component i by

$$\alpha_i=\sqrt{\frac{2^{2R_i^\Delta}-1}{2^{2R_i}-1}} \tag{142}$$

results in parallel (uncorrelated) sources

$$\hat{x_i}=\alpha_i x_i \tag{143}$$

with variances

$$\hat{s}_i^2=\alpha_i^2 s_i^2. \tag{144}$$

By (60), we therefore have

$$\begin{aligned}\hat{R}_i&=\frac{1}{2}\log\left(1+\hat{s}_i^2\right)\\&=\frac{1}{2}\log\left(1+\alpha_i^2 s_i^2\right)\\&=\frac{1}{2}\log\left(1+\frac{2^{2R_i^\Delta}-1}{2^{2R_i}-1}s_i^2\right)\\&=\frac{1}{2}\log\left(1+\frac{2^{2R_i^\Delta}-1}{2^{2R_i}-1}(2^{2R_i}-1)\right)\\&=\frac{1}{2}\log\left(1+2^{2R_i^\Delta}-1\right)\\&=R_i^\Delta.\end{aligned} \tag{145}$$

We associate with any rate tuple $(R_1,R_2,\ldots,R_M)\in\mathcal{R}(R_{\mathrm{BT}})$ a rate tuple $(R_1^\Delta,R_2^\Delta,\ldots,R_M^\Delta)\in\mathcal{R}^\Delta(R_{\mathrm{BT}})$, according to the following transformation

$$\begin{aligned}R_i^\Delta&=\left\lceil\frac{R_i}{\Delta R_{\mathrm{BT}}}\right\rceil\cdot\Delta R_{\mathrm{BT}},\quad i=1,2,\ldots,K-1,\\R_K^\Delta&=R_K-\sum_{i=1}^{K-1}(R_i^\Delta-R_i).\end{aligned} \tag{146}$$

For $i=K$ we have

$$R_K^\Delta\geq R_K-(K-1)\Delta R_{\mathrm{BT}}. \tag{147}$$

It follows that the scaling factor needed to achieve R_K^Δ is bounded by

$$\begin{aligned}\alpha_K^2&=\frac{2^{2R_K^\Delta}-1}{2^{2R_K}-1}\\&\geq\frac{2^{2(R_K-(K-1)\Delta R_{\mathrm{BT}})}-1}{2^{2R_K}-1}\\&\geq\frac{2^{2\left(\frac{R_{\mathrm{BT}}}{K}-(K-1)\Delta R_{\mathrm{BT}}\right)}-1}{2^{2\left(\frac{R_{\mathrm{BT}}}{K}\right)}-1}.\end{aligned} \tag{148}$$

where (148) follows since it is readily verified that the function $\frac{ax-1}{x-1}$ is monotonically increasing in x, for $x\geq 1$ and $0\leq a\leq 1$. Denoting $\eta=\frac{1}{\alpha_k}$, it follows from (148) that

$$\eta=\frac{2^{2\left(\frac{R_{\mathrm{BT}}}{K}\right)}-1}{2^{2\left(\frac{R_{\mathrm{BT}}}{K}-(K-1)\Delta R_{\mathrm{BT}}\right)}-1}. \tag{149}$$

Now for $1 \le i \le K-1$ we have by (146) that $R_i^{\Delta} \ge R_i$. Hence, for such i it trivially holds that (since $\eta \ge 1$)

$$\begin{aligned} s_i^2 &\le s_{i,\Delta}^2 && (150)\\ &\le \eta^2 s_{i,\Delta}^2. && (151) \end{aligned}$$

Thus, the lemma follows by observing that from (148) it follows that $s_K^2 \le \eta^2 s_{K,\Delta}^2$ as well.

Appendix F
Proof of Lemma 4

Recalling (67), we note that (66) can be written as

$$R_{\mathrm{IF}}(\mathbf{S},\mathbf{P};\mathbf{a}) = \frac{1}{2}\log\left(\sum_{i=1}^{K} v_i^2(1+s_i^2)\right). \tag{152}$$

Therefore, when scaling the Gaussian input vector by a factor of $\beta \ge 1$, we have

$$\begin{aligned} R_{\mathrm{IF}}(\beta^2\mathbf{S},\mathbf{P};\mathbf{a}) &= \frac{1}{2}\log\left(\sum_{i=1}^{K} v_i^2(1+\beta^2 s_i^2)\right)\\ &\le \frac{1}{2}\log\left(\beta^2\sum_{i=1}^{K} v_i^2(1+s_i^2)\right)\\ &= \frac{1}{2}\log\left(\beta^2\right) + R_{\mathrm{IF}}(\mathbf{S},\mathbf{P};\mathbf{a})\\ &= \log\left(\beta\right) + R_{\mathrm{IF}}(\mathbf{S},\mathbf{P};\mathbf{a}). \end{aligned} \tag{153}$$

References

[1] O. Ordentlich and U. Erez, "Integer-forcing source coding," *IEEE Transactions on Information Theory*, vol. 63, no. 2, pp. 1253–1269, 2017.

[2] L. G. Ordóñez, I. E. Aguerri, and M. Guillaud, "Integer forcing analog-to-digital conversion for massive MIMO systems," in *Signals, Systems and Computers, 2016 50th Asilomar Conference on*. IEEE, 2016, pp. 11–15.

[3] E. Domanovitz and U. Erez, "Outage behavior of integer forcing with random unitary pre-processing," *IEEE Transactions on Information Theory*, vol. PP, no. 99, pp. 1–1, 2017.

[4] A. Edelman and N. R. Rao, *Random matrix theory*. Cambridge University Press, 2005, vol. 14.

[5] S. Tung, "Multiterminal source coding (Ph.D. Thesis)," *Information Theory, IEEE Transactions on*, vol. 24, no. 6, pp. 787–787, November 1978.

[6] O. Fischler, "Universal precoding for parallel Gaussian channels," Master's thesis, Tel Aviv University, Tel Aviv, 2014. [Online]. Available: https://www.eng.tau.ac.il/~uri/theses/fischler_msc.pdf

[7] W. He and B. Nazer, "Integer-forcing source coding: Successive cancellation and source-channel duality," in *Information Theory (ISIT), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 155–159.

[8] M. Mehta, *Random matrices and the statistical theory of energy level*. Academic Press, 1967.

[9] J. Lagarias, J. Lenstra, H.W., and C. Schnorr, "Korkin-Zolotarev bases and successive minima of a lattice and its reciprocal lattice," *Combinatorica*, vol. 10, no. 4, pp. 333–348, 1990.

[10] H. Blichfeldt, "The minimum value of quadratic forms, and the closest packing of spheres," *Mathematische Annalen*, vol. 101, no. 1, pp. 605–608, 1929.

[11] O. Ordentlich, Private communication, 2017.

[12] F. Oggier and E. Viterbo, "Algebraic number theory and code design for rayleigh fading channels," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 3, pp. 333–415, 2004.

[13] O. Ordentlich and U. Erez, "A simple proof for the existence of good pairs of nested lattices," *IEEE Transactions on Information Theory*, vol. 62, no. 8, pp. 4439–4453, 2016.

[14] D. Micciancio and S. Goldwasser, *Complexity of lattice problems: a cryptographic perspective*. Springer Science & Business Media, 2012, vol. 671.

[15] O. Ordentlich, U. Erez, and B. Nazer, "Successive Integer-Forcing and its Sum-Rate Optimality," *CoRR*, vol. abs/1307.2105, 2013. [Online]. Available: http://arxiv.org/abs/1307.2105

[16] ——, "The approximate sum capacity of the symmetric gaussian k-user interference channel," *IEEE Transactions on Information Theory*, vol. 60, no. 6, pp. 3450–3482, 2014.

[17] E. Domanovitz and U. Erez, "Combining space-time block modulation with integer forcing receivers," in *Electrical Electronics Engineers in Israel (IEEEI), 2012 IEEE 27th Convention of*, Nov 2012, pp. 1–4.

[18] O. Ordentlich and U. Erez, "Precoded Integer-Forcing Universally Achieves the MIMO Capacity to Within a Constant Gap," *Information Theory, IEEE Transactions on*, vol. 61, no. 1, pp. 323–340, Jan 2015.

[19] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, "Perfect space–time block codes," *IEEE Transactions on information theory*, vol. 52, no. 9, pp. 3885–3902, 2006.

[20] P. Elia, K. Kumar, S. Pawar, P. Kumar, and H. feng Lu, "Explicit Space Time Codes Achieving the Diversity Multiplexing Gain Tradeoff," *Information Theory, IEEE Transactions on*, vol. 52, no. 9, pp. 3869–3884, Sept 2006.