On the Robustness of Lattice Interference Alignment

Or Ordentlich and Uri Erez, *Member, IEEE*

***Abstract*—A static (constant channel gains) real K-user interference channel is considered, where all interference (cross) channel gains are integers. For such channels, previous results demonstrate that the number of degrees of freedom is very sensitive to slight variations in the direct channel gains. In this paper, we derive an achievable rate region for such channels that is valid for finite SNR. At moderate values of SNR, the derived rate region is robust to slight variations in the direct channel gains. At asymptotic high SNR conditions, known results on the degrees of freedom are recovered. The new rate region is based on lattice interference alignment. The result is established via a new coding theorem for the two-user Gaussian multiple-access channel where both users use a single linear code.**

***Index Terms*—Interference alignment, interference channel, linear codes, multiple-access channel (MAC).**

I. Introduction

An important open problem in network information theory is determining the capacity region of the interference channel. The interference channel is a communication model where multiple pairs of transmitters and receivers utilize the same communication medium. As a result, each user receives the output of a multiple-access channel (MAC), i.e., it suffers from interference from transmissions intended for other users.

An important special case of this channel model is the Gaussian interference channel, where each receiver sees a linear combination of its intended signal and the signals transmitted by the interfering users plus an additive white Gaussian noise (AWGN). For the case where only two users are sharing the same medium, i.e., the interference at each receiver is generated by only one user, the capacity region was characterized up to half a bit only recently [1]. The achievability part utilizes the Han–Kobayashi [2] scheme which is shown to be nearly optimal in the two-user case. The results of [1] are rather disappointing in the sense that they imply that for a wide range of channel parameters, either treating the interference as noise, or alternating access to the medium (i.e., time sharing) between the two transmitter–receiver pairs, is a reasonable approach. In particular, time sharing yields the maximal degrees of freedom (DoFs) afforded by the channel (i.e., one), where the number of DoFs is defined as the ratio between the maximal possible sum rate and $\frac{1}{2}\log(\mathsf{SNR})$ in the limit where the SNR goes to infinity.

An interesting aspect of the interference channel is that the two-user case does not capture the quintessential features of the general (K-user) interference channel, as has recently been demonstrated in the framework of Gaussian interference channels. In particular, while one may have suspected that the channel would be interference limited, i.e., that time sharing would be optimal at high SNR, it has been demonstrated that this is not the case. Rather, it has been shown [3]–[6] that the correct "extension" of the two-user results is that in general, $K/2$ DoF are afforded by the K-user Gaussian interference channel.

A. Related Work

The works of [3]–[6] have revealed that the Han–Kobayashi approach is inadequate for $K > 2$, and a new approach, namely, interference alignment, was needed to achieve the DoF afforded by the (general) K-user interference channel. See [7] for a comprehensive survey.

The first applications of interference alignment for Gaussian interference channels included the time-varying single-input single-output (SISO) K-user interference channel [3], [4], the multiple-input multiple-output (MIMO) X channel [8], [9], and K-user MIMO interference channels with constant channel gains [3], [10], [11] and time-varying channel gains [10]. In these works, the alignment schemes rely on the diversity in the channel gains. The focus of this paper is the real *static* (constant channel gains) SISO K-user Gaussian interference channel, for which another form of interference alignment has proven to play a key role. In this case, it was shown in [5] and [6] that by taking the transmitted signal to belong to the (1-D) integer lattice, it is possible to align the interference so that it remains confined to this lattice. As a result, the minimum distance of the received constellation at each receiver does not decrease with K, and when the SNR approaches infinity, each receiver can decode its intended signal with rate $\approx \frac{1}{4}\log \mathsf{SNR}$, yielding a total of $K/2$ DoF. Thus, linear constellations, i.e., a PAM constellation in the 1-D case play a key role in interference alignment for static channels.

Specifically, it was shown in [5] that if at each receiver, the channel gains corresponding to the interferers are rational, whereas the direct channel gains corresponding to the intended signal are irrational algebraic, $K/2$ DoFs are achievable. Even more interestingly, the authors of [5] have shown that if the direct channel gains are rational as well, the DoFs of the

Manuscript received June 18, 2011; revised July 22, 2012; accepted November 21, 2012. Date of publication January 30, 2013; date of current version April 17, 2013. This work was supported in part by the Israel Science Foundation under Grant 1557/12 and in part by the Binational Science Foundation under Grant 2008455. O. Ordentlich was supported in part by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities, in part by a fellowship from The Yitzhak and Chaya Weinstein Research Institute for Signal Processing at Tel Aviv University, and in part by the Feder Family Award. This paper was presented in part at the 2011 IEEE Information Theory Workshop.

The authors are with the Department of Electrical Engineering-Systems, Tel Aviv University, Ramat Aviv 69978, Israel. (e-mail: ordent@eng.tau.ac.il; uri@eng.tau.ac.il).

Communicated by S. A. Jafar, Associate Editor for Communications.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2012.2235523

channel are strictly smaller than $K/2$. Later, the authors of [6] proved that the DoF of the static interference channel are $K/2$ for almost all sets of channel gains. Recently, a general single-letter formula for the number of DoF the interference channel offers was derived by Wu *et al.* [12].

The results of [5] and [6] imply that in asymptotic high SNR conditions. the capacity characterization of the interference channel is extremely sensitive to slight variations in the channel gains. Such a behavior is highly undesirable, and puts into question the feasibility of lattice interference alignment for static channels. The main motivation for this paper is to explore the robustness of lattice interference alignment at a nonasymptotic setting, and to find out whether the aforementioned phenomena are a by-product of the asymptotic definition of the DoF or an inherent feature of lattice interference alignment.

As noted above, linear/lattice codes play an important role in coding for static Gaussian interference channels. This feature of the K-user Gaussian interference channel is shared with a growing number of problems in network information theory where lattice strategies have been shown to be an important ingredient. There are several examples where lattice strategies achieve better performance than the best known random coding strategies. In particular, Philosof *et al.* introduced lattice interference alignment in the context of the doubly-dirty MAC [13], i.e., to a Gaussia MAC with multiple interference signals, each known to a different transmitter. Other network scenarios where lattices play a key role are the two-way (or multiple-way) relay problem [14] and the compute-and-forward approach to relay networks [15].

Lattice interference alignment for the interference channel was first proposed by Bresler *et al.* in [16], [17], where an approximate characterization of the capacity region for the many-to-one and one-to-many interference channels was derived. Lattice interference alignment was later utilized by Sridharan *et al.* in [18] where a coding scheme where all users transmit points from the *same* lattice was introduced. If at each receiver all the gains corresponding to the interferers are integers, the sum of the interferences is a point in the same lattice, and thus, the interference from $K-1$ users is confined to one lattice. Under very strong interference conditions, which are defined in [18] and play the same role as the well known very high interference condition [19] for the two-user case, the decoder can first decode the sum of interferers while treating the desired lattice point as noise, then subtract the decoded interference, and finally decode the intended codeword. Later, in [20], this scheme was combined with a layered coding scheme in order to show that lattice interference alignment can yield substantial gains and, in particular, achieve more that one DoF in some cases for a broader (but still quite limited) class of channels.

The works of [18] and [20] allowed for important progress toward the understanding of lattice interference alignment at finite SNR. Nonetheless, these results are limited since they essentially rely on using superposition coding with a judicious choice of power allocation such that a very strong interference condition holds, in conjunction with successive decoding. In the decoding procedure, a single layer is decoded at every step, while the other layers are treated as noise. At each step of the successive decoding procedure, the decoder sees an equivalent point-to-point AWGN channel where lattice codes are used. The performance of lattice codes over point-to-point AWGN channels is well understood, and therefore, the scheme of [18] and [20] can be analyzed with relative ease. For special classes of channel gains, it is possible to design a layered codebook that is simultaneously good for all receivers. For a wide range of channel parameters, however, such a layered scheme is not beneficial, as also noted in [20].

B. Summary of Results

The main contribution of this work is in providing a general framework for lattice interference alignment that is not confined to successive decoding. A coding theorem is established for a two-user MAC where both users use the same linear codebook.

Specifically, if the interference is aligned to a lattice, but the very strong interference condition is not satisfied, the decoder can still perform *joint decoding* of the interference codeword and the desired codeword. A major obstacle however arises when one attempts to jointly decode both codewords: the alignment of all interferers into one lattice point, which occurs simultaneously at all receivers, is only possible due to the fact that all users transmit lattice points from the *same* lattice. Thus, if joint decoding is applied, each decoder sees a two-user Gaussia MAC where both users use the same linear code. The capacity region of the MAC without this restriction is derived based on joint typicality arguments, which assume that either the pair of transmitted codewords is statistically independent of any pair of competing codewords or one of the codewords in the transmitted pair is the same as in the competing pair, and the other codeword in the competing pair is statistically independent of that in the transmitted pair. This assumption, which is valid when both users use different random codes, is no longer valid when both users use the same linear code.

The fact that the number of DoF of many families of interference channels is $K/2$ implies that for certain channel conditions, it is sufficient to transform the channel seen by each receiver into a two-user MAC in order to approach the optimal performance, i.e., half of the resources are dedicated to the interference and the other half to the intended signal. In light of this observation, along with the proven advantages of lattice codes in creating alignment between different users, it is important to study the achievable performance of the two-user Gaussia MAC where the same lattice code is used by all transmitters.

In this paper, we first address the question of finding an achievable symmetric rate for the Gaussian (modulo-additive) MAC with two users that are "forced" to use the same linear code. We then employ this new ingredient in order to analyze an interference alignment scheme, suitable for a class of interference channels, which we refer to as the integer-interference channel, where all cross gains are integers (or rationals). In contrast to the results obtained in previous work, here, the analysis is not asymptotic in the SNR.

While the proposed coding scheme does not require asymptotic conditions, we show that it is asymptotically optimal in a DoF sense, i.e., it achieves $K/2$ DoFs for the integer-interference channel for almost all direct channel gains. The rate region achieved by the scheme enables to shed light on the effect of

the direct channel gains being rational or irrational, which has to date only been understood for asymptotic high SNR conditions. In the proposed scheme, rational direct channel gains of the form r/q limit the achievable symmetric rate to be smaller than $\log q$, which is not a serious limitation if q is large and the SNR is moderate, but does indeed pose a limitation in the limit of very high SNR.

Moreover, previous results [5], [6] state that the DoF of an interference channel with integer interference gains are discontinuous at rational direct channel gains. Such a result is quite displeasing and calls into question the applicability of interference alignment for static channels at nonasymptotic conditions, i.e., raises questions with respect to (w.r.t.) the robustness of lattice interference alignment. The results of this work demonstrate the behavior of the rate when the direct channel gains approach a given set of rational numbers. The (derived achievable) rate is everywhere continuous, as is to be expected, in the direct channel gains for any SNR, but the variation, i.e., sensitivity to the direct channel gain, increases with the SNR.

While the presented scheme is only valid for channels where all the (nondirect) interference gains are integers (or rationals), we believe that the results are an important step toward the understanding of the robustness of lattice interference alignment for general static interference channels at finite SNR.

Since the first appearance of this paper, there have been significant progress in the understanding of lattice interference alignment. Specifically, in [21], our main result regarding the achievable rate over a two-user MAC where both users transmit from the same lattice code has been strengthened and extended to the K-user MAC where all users share the same lattice code. This allowed to characterize the sum capacity of the symmetric K-user Gaussian interference channel to within a constant gap for all channel gains outside some outage set [21]. Another subsequent work with flavor similar to ours is [22] which finds the approximate capacity of the SISO two-user X-channel for all channel gains outside some outage set.

The rest of this paper is organized as follows. In Section II, some notations used throughout the paper are defined. In Section III, an achievable symmetric rate is derived for a two-user Gaussian (modulo-additive) MAC where both users use the same linear code. Section IV presents an interference alignment scheme for finite SNR. Section V discusses the effect of the direct channel gains being rational versus irrational on the performance of interference alignment. In Section VI, possible approaches for interference alignment when the interference gains are not restricted to be integers (or rational) are discussed. The paper concludes with Section VII.

II. Notational Conventions

Throughout the paper, we use the following notational conventions. Random variables are denoted by uppercase letters and their realizations by lowercase letters. For example, X is a random variable, whereas x is a specific value it can take. We use boldface letters to denote vectors, e.g., $\mathbf{x}$ denotes a vector with entries x_i.

A number of distinct modulo operations are used extensively throughout the paper. The notation $x_{\mathrm{mod}[a,b)}$ denotes reducing $x \in \mathbb{R}$ modulo the interval $[a,b)$. That is, $x_{\mathrm{mod}[a,b)}$ is equal to

$$x - m \cdot (b-a)$$

where $m \in \mathbb{Z}$ is the (unique) integer such that

$$x - m \cdot (b-a) \in [a,b).$$

Similarly, for $x \in \mathbb{Z}$, $x_{\mathrm{mod}\, p}$ is defined to equal

$$x - m \cdot p$$

where $m \in \mathbb{Z}$ is the unique integer such that $x - m \cdot p \in \mathbb{Z}_p$.

If $\mathbf{x}$ is a vector, the notation $\mathbf{x}_{\mathrm{mod}[a,b)}$ is understood to mean reducing each component of $\mathbf{x}$ modulo the interval $[a,b)$. We define the basic interval $\mathcal{I}$ as $[-L/2, L/2)$ where $L = \sqrt{12}$. Reducing x modulo the interval $\mathcal{I}$ is denoted by x^*, i.e.,

$$x^* \triangleq [x]_{\mathrm{mod}[-L/2,L/2)}.$$

The Euclidean norm of a vector $\mathbf{x}$ is denoted by $\|\mathbf{x}\|$. The notation $\lfloor x \rceil$ denotes rounding x to the nearest integer. We denote the set of all prime numbers by $\mathcal{P}$. All logarithms in the paper are to the base 2, and therefore, all rates are expressed in bits per (real) channel use. All signals considered in this paper are real valued.

III. Achievable Symmetric Rate for the Two-User Gaussia MAC With a Single Linear Codebook

A. Problem Statement

We consider the modulo-additive MAC

$$Y = [X_1 + \gamma X_2 + Z]^* \tag{1}$$

where Z is an i.i.d. Gaussian noise with zero mean and variance $\mathbb{E}\left[Z^2\right] = 1/\mathsf{SNR}$. We are interested in characterizing the achievable rate region for this channel where both users are forced to use the *same linear* code, and where both users are subject to the power constraint

$$\frac{1}{n}\mathbb{E}\left[\|\mathbf{X}\|^2\right] \le 1.$$

Note that a random variable uniformly distributed over $\mathcal{I}$ has unit power.

An (n,R) code for this model is defined by *one* encoding function (for both users)

$$f : \left\{1, \ldots, 2^{nR}\right\} \to \mathcal{I}^n$$

and a decoding function

$$g : \mathcal{I}^n \to \left\{1, \ldots, 2^{nR}\right\} \times \left\{1, \ldots, 2^{nR}\right\}.$$

The linearity constraint on the encoding function f is expressed by the condition that for any $w_1, w_2 \in \left\{1, \ldots, 2^{nR}\right\}$, there exists a $w_3 \in \left\{1, \ldots, 2^{nR}\right\}$ such that

$$[f(w_1) + f(w_2)]^* = f(w_3). \tag{2}$$

User 1 chooses a message $w_1 \in \{1,\dots,2^{nR}\}$ and transmits $\mathbf{x}_1 = f(w_1)$, and user 2 chooses a message $w_2 \in \{1,\dots,2^{nR}\}$ and transmits $\mathbf{x}_2 = f(w_2)$. The decoder upon receiving

$$\mathbf{y} = [\mathbf{x}_1 + \gamma \mathbf{x}_2 + \mathbf{z}]^*$$

generates estimates for the transmitted messages

$$\{\hat{w}_1, \hat{w}_2\} = g(\mathbf{y}).$$

The error probability for decoding the transmitted messages is denoted by

$$P_e \stackrel{\triangle}{=} \mathbb{E}\left[\Pr\left(\{\hat{w}_1, \hat{w}_2\} \neq \{w_1, w_2\}\right)\right]$$

where the expectation is taken with respect to the uniform distribution on all pairs of messages w_1, w_2. We say that a symmetric rate R is achievable if for any $\epsilon > 0$ and n large enough (depending on ϵ), there exists an (n, R) linear code such that $P_e < \epsilon$.

B. Connection to the Standard Gaussia MAC and Previous Results

The channel (1) can be viewed as a degraded version of the Gaussia MAC

$$\tilde{Y} = X_1 + \gamma X_2 + Z \tag{3}$$

as Y can be obtained from $\tilde{Y}$ by the transformation

$$Y = \tilde{Y}^*.$$

It follows that the achievable symmetric rate for our channel model with the constraint that both users use the same linear code is upper bounded by that of the channel (3) where each user can use any codebook with rate R.

The capacity region of the Gaussia MAC (3) was characterized by Ahlswede [23] through the following equations:

$$\begin{aligned} R_1 &< \frac{1}{2}\log\left(1+\mathsf{SNR}\right) \\ R_2 &< \frac{1}{2}\log\left(1+\gamma^2\mathsf{SNR}\right) \\ R_1+R_2 &< \frac{1}{2}\log\left(1+(1+\gamma^2)\mathsf{SNR}\right). \end{aligned}$$

It follows that the symmetric capacity (i.e., the maximum achievable $R = R_1 = R_2$) is given by

$$\begin{aligned} C = \min\Big\{ &\frac{1}{2}\log\left(1+\mathsf{SNR}\right), \frac{1}{2}\log\left(1+\gamma^2\mathsf{SNR}\right) \\ &\frac{1}{4}\log\left(1+(1+\gamma^2)\mathsf{SNR}\right)\Big\}. \end{aligned} \tag{4}$$

The achievability part of the capacity theorem is proved using two *different random* codebooks, whereas we restrict both users to use the *same linear* (over the group $\mathcal{I}$ with the addition operation) codebook.

C. Main Result and Discussion

The main result of this section is the following theorem.

Theorem 1 (MAC With One Linear Code): For the setting described in Section III-A (a two-user Gaussia MAC where both users use the same linear code), the following symmetric rate is achievable

$$\begin{aligned} R < \max_{p\in\mathcal{P}(\gamma,\mathsf{SNR})} \min \\ \Bigg\{ &-\frac{1}{2}\log\left(\frac{1}{p^2}+\sqrt{\frac{2\pi/3}{\mathsf{SNR}}}+\frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\delta^2(p,\gamma)}+2e^{-\frac{3\mathsf{SNR}}{8}}\right), \\ &-\log\left(\frac{1}{p}+\sqrt{\frac{2\pi/3}{\delta^2(p,\gamma)\mathsf{SNR}}}+2e^{-\frac{3\mathsf{SNR}}{8}}\right)\Bigg\} \end{aligned} \tag{5}$$

where

$$\delta(p,\gamma) = \min_{l\in\mathbb{Z}_p\setminus\{0\}} l\cdot\left|\gamma - \frac{\lfloor l\gamma\rceil}{l}\right| \tag{6}$$

and[1]

$$\mathcal{P}(\gamma,\mathsf{SNR}) = \left[p\in\mathcal{P}\middle| e^{-\frac{3\mathsf{SNR}}{2p^2}\left(\gamma_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}\right)^2} < 1-2p\cdot e^{-\frac{3\mathsf{SNR}}{8}}\right]. \tag{7}$$

Discussion: Inspecting the equations describing the achievable rate region of Theorem 1, the role of the optimization parameter p and the factor $\delta(p,\gamma)$ may seem at first strange. In the scheme that achieves this rate region, the parameter p is the size of the PAM constellation over which the n-dimensional linear code is constructed (in a Construction A [24] manner). The factor $\delta(p,\gamma)$ is a measure of how accurately γ can be approximated by a rational number with a denominator smaller than p. See Fig. 1 for an illustration of the function's behavior for different values of p.

Note that above some moderate value of SNR, the expressions dominating (5) are

$$\frac{1}{p^2}+\sqrt{\frac{2\pi/3}{\mathsf{SNR}}} \tag{8}$$

and

$$\frac{1}{p}+\sqrt{\frac{2\pi/3}{\delta^2(p,\gamma)\mathsf{SNR}}}. \tag{9}$$

Since the parameter p dictates the support of the minimization space in (6), $\delta(p,\gamma)$ is monotonically nonincreasing in p. On the one hand, increasing p decreases the terms $1/p$ and $1/p^2$, but on the other hand, this reduces the "effective SNR" $\delta^2(p,\gamma)\mathsf{SNR}$ in (9). Thus, the choice of p should balance the two effects. For example, if γ is a rational number that can be written in the form

[1]Replacing the constraint $p\in\mathcal{P}(\gamma,\mathsf{SNR})$ with the constraint $p\in\mathcal{P}$ results in a negligible change in the rate region described by (5) and (6) for values of γ that are not very "close" (w.r.t. SNR) to $\pm 1/2$.

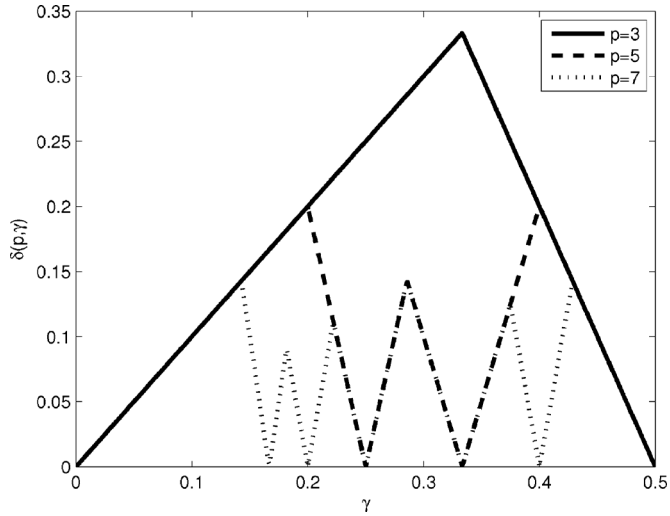

Fig. 1. Function $\delta(p,\gamma)$ for $p=3$, $p=5$, and $p=7$.

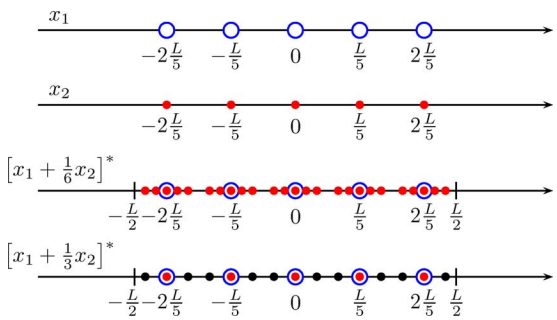

Fig. 2. Illustration of the ambiguity problem for rational channel coefficients with small denominator. In this example, a 5-PAM constellation is used by both transmitters. If $\gamma=1/6$, each pair of inputs x_1, x_2 results in a different channel output. However, for $\gamma=1/3$, different pairs of x_1, x_2 may cause the same channel output. Channel outputs that are ambiguous, i.e., that can be the result of different inputs, are marked by full black circles in the figure.

$\gamma=r/q$, then for any $p>q$, we have $\delta(p,\gamma)=0$. If this is the case, only values of $p\le q$ yield nontrivial rates in (5), which in turn implies that for any value of SNR, the rate of (5) is smaller than $\log(q)$.

This saturation phenomenon of the achievable rates for rational values of γ is most intuitively understood through the case of uncoded transmission. For instance, assume both users transmit from the PAM constellation

$$\mathcal{C}=\frac{L}{p}\left\{-\frac{p-1}{2},\ldots,0,\ldots,\frac{p-1}{2}\right\}.$$

Note that according to the definitions of Section III-A, the constellation $\mathcal{C}$ forms a 1-D linear code of rate $\log(p)$. Even in the noiseless case, if $\gamma=r/q$ with $q<p$, the individual symbols transmitted by the two users cannot always be distinguished from the channel's output. This point is illustrated by Fig. 2 for $p=5$ and $\gamma=1/3$.

It is tempting to think that the ambiguity problem is caused by the cardinality of the constellation being too small. In other words, one may hope that the problem can be overcome by constructing a high-dimensional linear code over a larger alphabet with cardinality $p\gg q$. However, as the following example implies, this is not necessarily the case.

Example: In order to informally show why the rate saturation phenomenon (why $R<\log(q)$ for any SNR) does not disappear when n-dimensional linear codes over large constellations are used, we consider the case of $\gamma=1/3$. We assume that for any $a\in\mathbb{Z}\setminus\{0\}$, the linear codebook $\mathcal{C}$ that is used satisfies the property

$$\mathcal{C}=[a\mathcal{C}]^*. \tag{10}$$

That is, for any pair of distinct codewords $\mathbf{x}_1,\mathbf{x}_2\in\mathcal{C}$ and nonzero integer a, we have $[a\mathbf{x}_1]^*\neq[a\mathbf{x}_2]^*$. The ensemble of codebooks we consider in the proof of Theorem 1 satisfies this property.

Assume that the first user transmitted the codewords $\mathbf{x}_1$ and the second user transmitted the codeword $\mathbf{x}_2$.

From the linearity of the code (2) combined with the property (10), for any choice of $\mathbf{x}_2$, there exists a codeword $\mathbf{x}_3$ such that

$$\mathbf{x}_2=[3\mathbf{x}_3]^*.$$

Furthermore, for any codeword $\mathbf{x}_4$, there exists a codeword $\mathbf{x}_5$ such that

$$\mathbf{x}_5=[3\mathbf{x}_4]^*.$$

Now consider the pair of competing codewords

$$\mathbf{x}_{\mathcal{E}1}=[\mathbf{x}_1+\mathbf{x}_3+\mathbf{x}_4]^*$$

and

$$\mathbf{x}_{\mathcal{E}2}=[-\mathbf{x}_5]^*$$

which exists by the linearity of the codebook. Note that while $\mathbf{x}_3$ is determined by $\mathbf{x}_2$, the codeword $\mathbf{x}_4$ can be any of the 2^{nR} codewords in $\mathcal{C}$, regardless of the choice of $\mathbf{x}_1$ and $\mathbf{x}_2$. Thus, there are 2^{nR} different pairs of competing codewords from this form.

After passing through the channel (1), the "distance" between the transmitted pair of codewords and the "competing" pair of codewords is

$$\begin{aligned}\mathbf{u}&=\left[\mathbf{x}_1+\frac{1}{3}\mathbf{x}_2-\mathbf{x}_{\mathcal{E}1}-\frac{1}{3}\mathbf{x}_{\mathcal{E}2}\right]^*\\&=\left[\mathbf{x}_1+\frac{1}{3}\mathbf{x}_2-\mathbf{x}_1-\mathbf{x}_3-\mathbf{x}_4-\frac{1}{3}[-\mathbf{x}_5]^*\right]^*\\&=\left[\left(\frac{1}{3}[3\mathbf{x}_3]^*-\mathbf{x}_3\right)-\left(\mathbf{x}_4+\frac{1}{3}[-3\mathbf{x}_4]^*\right)\right]^*.\end{aligned} \tag{11}$$

The term

$$\frac{1}{3}[3\mathbf{x}_3]^*-\mathbf{x}_3$$

which is dictated by the codeword $\mathbf{x}_2$ transmitted by the second user can only take values in $\{-L/3,0,L/3\}^n$. For each of the 2^{nR} different choices of $\mathbf{x}_4$, the term

$$\mathbf{x}_4+\frac{1}{3}[-3\mathbf{x}_4]^* \tag{12}$$

also results in a sequence in $\{-L/3,0,L/3\}^n$. If these sequences were all different, then the fact that there are only 3^n different sequences in $\{-L/3,0,L/3\}^n$ would imply that if $R>\log(3)$, there must exist a $\mathbf{x}_4\in\mathcal{C}$ for which $\mathbf{u}=\mathbf{0}$ and an error always occurs. In general, there is no guarantee that

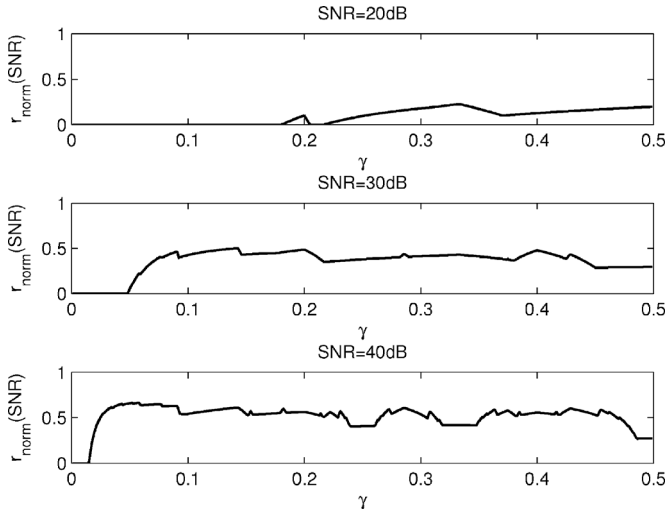

Fig. 3. $r_{\mathrm{norm}}(\mathsf{SNR})$ plotted as a function of γ for $\mathsf{SNR} = 20,\ 30, \mathrm{and} 40$ dB.

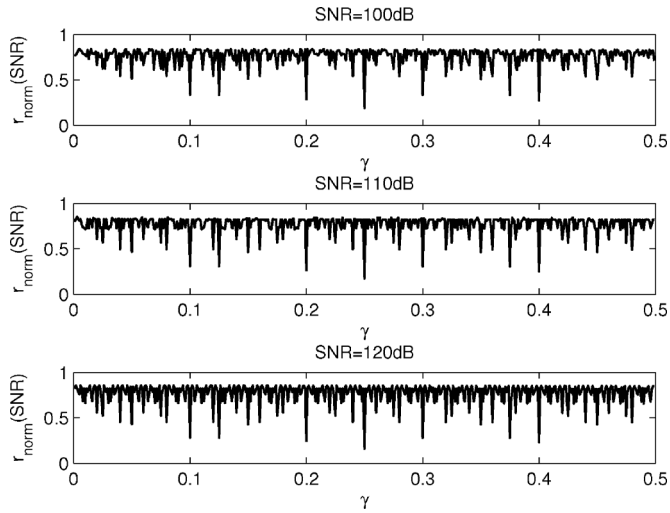

Fig. 4. $r_{\mathrm{norm}}(\mathsf{SNR})$ plotted as a function of γ for $\mathsf{SNR} = 100,\ 110, \mathrm{and} 120$ dB.

the 2^{nR} sequences of (12) are distinct. However, for a random ensemble of linear codes where $\mathbf{x}_4$ is uniformly distributed over a dense grid in $\mathcal{I}$, the probability that one of the choices of $\mathbf{x}_4$ will null the distance $\mathbf{u}$ is high. Thus, it is difficult to construct a linear codebook for which the described form of competing codewords does not incur an error.

We emphasize that this example does not prove that it is impossible to achieve rates substantially greater than $\log(q)$ when both users share the same linear code. Rather, it gives intuition to why this might be the case.

As already mentioned, Theorem 1 was improved upon in [21] using a pair of "good" nested lattice codes with dithering. The dithering allows to achieve rates somewhat higher than $\log(q)$ for high SNR, but the results of [21] still exhibit a saturation phenomenon for rational channel gains. Note that in the problem formulation from Section III-A, dithering is not allowed. However, for purposes of interference alignment, which is the main motivation for this work, letting all users transmit from the same linear code with different dithers for each user does not pose a problem.

Comparison With Random Codebooks: In order to better understand the performance of our coding scheme, we compare the maximum symmetric rate it achieves, which we refer to as $R_{\mathrm{lin}}(\mathsf{SNR})$, with that achieved by a coding scheme that utilizes two different random codebooks. We refer to the latter symmetric rate as $R_{\mathrm{rand}}(\mathsf{SNR})$, which is given by (4). Define the normalized rate

$$r_{\mathrm{norm}}(\mathsf{SNR}) \stackrel{\triangle}{=} \frac{R_{\mathrm{lin}}(\mathsf{SNR})}{R_{\mathrm{rand}}(\mathsf{SNR})}. \tag{13}$$

Fig. 3 depicts $r_{\mathrm{norm}}(\mathsf{SNR})$ as a function of $\gamma \in [0, 0.5)$ for a range of moderate to high values of SNR, specifically $\mathsf{SNR} = 20, 30, \text{ and } 40$ dB. Fig. 4 depicts $r_{\mathrm{norm}}(\mathsf{SNR})$ as a function of $\gamma \in [0, 0.5)$ for extremely high values of SNR, namely $\mathsf{SNR} = 100, 110, \text{ and } 120$ dB.

Figs. 3 and 4 demonstrate the sensitivity of the rate to the channel gains. For a range of "reasonable" values of SNR, the rate changes rather smoothly with γ. For extremely high SNR, however, a slight change in the value of γ may dramatically change the achievable rate.

The figures also suggest that for almost every value of γ, the normalized rate $r_{\mathrm{norm}}(\mathsf{SNR})$ approaches one as the SNR tends to infinity. Thus, the symmetric rate achieved when both users are using the same linear code scales with the SNR as $R_{\mathrm{rand}}(\mathsf{SNR})$ for asymptotic high SNR conditions.

Note that Theorem 1 does not take into account shaping issues, since it uses a 1-D lattice as the coarse lattice. We have chosen not to pursue shaping in this paper in order to simplify the analysis. Moreover, the main contribution of this paper is in characterizing the losses resulting from using the same linear code for both users, which outweigh the shaping loss, that can be upper bounded by a constant fraction of a bit. In subsequent work [21], nested lattice codes with a "good" coarse lattice (as opposed to the 1-D coarse lattice used here) were used to improve the rate region obtained here.

D. Proof of Theorem 1

We first describe the process of the code generation, the encoding and the decoding procedures, and then turn to analyze the error probability in decoding the transmitted messages.

Construction of Linear Codebook Ensemble: We begin by describing the generation process of the ensemble of linear codebooks considered, which is a variant of the well-known Construction A (see, e.g., [24]). A generating matrix G of dimensions $k \times n$ is used, where all elements of G are independently and randomly drawn according to the uniform distribution over the prime field $\mathbb{Z}_p$. We set

$$k = \frac{R}{\log p} n.$$

The set $\tilde{\mathcal{C}}$ is generated by multiplying all k-tuple vectors with elements from the field $\mathbb{Z}_p$ by the matrix G (where all operations are over $\mathbb{Z}_p$)

$$\tilde{\mathcal{C}} = \left\{ \tilde{\mathbf{c}} = \mathbf{w}^T G \mid \mathbf{w} \in \mathbb{Z}_p^{k \times 1} \right\}.$$

We refer to the vectors $\mathbf{w}$ as message vectors, and note that there are 2^{nR} such vectors, each corresponding to one of the possible messages.

Finally, the codebook $\mathcal{C}$ is generated from the set $\tilde{\mathcal{C}}$ by properly scaling and shifting it such that it meets the power constraint

$$\mathcal{C} = \left[\frac{L}{p} \cdot \tilde{\mathcal{C}}\right]^*.$$

The ensemble of codebooks created by the aforementioned procedure satisfies the following properties.

1) For any set of linearly independent message vectors, $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_l$, the corresponding codewords $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_l$ are statistically independent.
2) Each codeword $\mathbf{X}$ (except for the zero codeword) has a memoryless distribution
$$P(\mathbf{X}) = \prod_{t=1}^{n} P(X_t).$$
3) For any $\mathbf{w} \neq \mathbf{0}$, the corresponding codeword is uniformly distributed over the constellation
$$\Lambda = \frac{L}{p}\left[\mathbb{Z}_p^n\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}.$$
4) Each codeword $\mathbf{X}$ in the ensemble satisfies the power constraint
$$\frac{1}{n}\mathbb{E}\left[\|\mathbf{X}\|^2\right] \leq 1.$$
5) Each codebook in the ensemble satisfies the linearity constraint as defined in (2).

Encoding: Suppose a codebook from the aforementioned ensemble, which is completely characterized by the matrix G, has been chosen. User i uniformly draws a message vector $\mathbf{w}_i$, and transmits

$$\mathbf{x}_i = \left[\frac{L}{p}\mathbf{w}_i^T G\right]^*. \tag{14}$$

The channel output is thus

$$\mathbf{y} = \left[\mathbf{x}_1 + \gamma\mathbf{x}_2 + \mathbf{z}\right]^*.$$

Decoding: Given the encoding matrix G and the channel output $\mathbf{y}$, the decoder searches the pair of codewords $\{\mathbf{x}_i, \mathbf{x}_j\}$ for which

$$\boldsymbol{\psi}(i,j) = \left[\mathbf{x}_i + \gamma\mathbf{x}_j\right]^* \tag{15}$$

is closest to $\mathbf{y}$ in the following sense

$$\left(\hat{i}, \hat{j}\right) = \arg\min_{i,j}\left(\sum_{t=1}^{n}\left(\left[y_t - \psi_t(i,j)\right]^*\right)^2\right). \tag{16}$$

If there is more than one pair of indices satisfying (16), an error is declared.

The decoder only searches over the pairs of codewords corresponding to message vectors $\{\mathbf{w}_i, \mathbf{w}_j\}$ that are linearly independent (over $\mathbb{Z}_p$). This constraint on the decoder facilitates the analysis of the error probability since it means that when G is assumed to be random, the decoder only searches over the pairs of codewords $\{\mathbf{X}_i, \mathbf{X}_j\}$ which are statistically independent. The aforementioned constraint implies that if the users had chosen message vectors $\{\mathbf{w}_i, \mathbf{w}_j\}$ which are linearly dependent, an error event occurs. For the rest of the analysis, we assume that indeed the chosen message vectors are linearly independent, and as a consequence, $\{\mathbf{X}_i, \mathbf{X}_j\}$ are statistically independent when G is assumed to be random. We account for the probability of the error event that occurs when this is not the case, in the final step of the proof.

We note that the decision rule (16) is an approximation of the maximum-likelihood decoder which searches for a pair of codewords $\{\mathbf{x}_i, \mathbf{x}_j\}$ that satisfies

$$\Pr\left(\mathbf{y}|\mathbf{x}_i, \mathbf{x}_j\right) > \Pr\left(\mathbf{y}|\mathbf{x}_m, \mathbf{x}_l\right) \quad \forall (m,l) \neq (i,j).$$

We choose the suboptimal decoder (16) rather than the optimal maximum-likelihood decision rule in order to simplify the analysis.

Analysis of Error Probability: We analyze the average error probability over the ensemble of codebooks, i.e., we assume the generating matrix G is random, and average over all possible realizations of G.

Assume that the message vectors $\{\mathbf{w}_1, \mathbf{w}_2\}$ were chosen by users 1 and 2, respectively, such that codeword $\mathbf{X}_1$ was transmitted by user 1, and $\mathbf{X}_2$ by user 2. We first analyze the pairwise error probability, i.e., the probability of the decoder preferring a (different) specific pair of message vectors $\{\mathbf{w}_{\mathcal{E}1}, \mathbf{w}_{\mathcal{E}2}\}$, corresponding to the pair of codewords $\{\mathbf{X}_{\mathcal{E}1}, \mathbf{X}_{\mathcal{E}2}\}$, over the transmitted pair.

As we recall, due to the linear structure of the codebook, linear dependences within the set of chosen and "competing" message vectors $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_{\mathcal{E}1}, \mathbf{w}_{\mathcal{E}2}\}$ result in statistical dependences within the set of transmitted and "competing" codewords $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_{\mathcal{E}1}, \mathbf{X}_{\mathcal{E}2}\}$. We are interested in the average pairwise error probability associated with each pair of "competing" message vectors $\{\mathbf{w}_{\mathcal{E}1}, \mathbf{w}_{\mathcal{E}2}\}$. Thus, the average pairwise error probability has to be analyzed w.r.t. each one of the possible statistical dependences. We develop upper bounds on the average pairwise error probability associated with each type of statistical dependence, and then invoke the union bound in order to establish an upper bound on $\mathbb{E}[P_e]$, the average probability of the decoder not deciding on the correct pair of transmitted codewords. Using this bound, an achievable rate region is obtained.

Denote the pairwise error probability from the pair of message vectors $\{\mathbf{w}_1, \mathbf{w}_2\}$ to the pair $\{\mathbf{w}_{\mathcal{E}1}, \mathbf{w}_{\mathcal{E}2}\}$, for a given codebook in the ensemble, by $P_{e,\mathrm{pair}}$, and the average pairwise error probability over the ensemble by $\mathbb{E}\left[P_{e,\mathrm{pair}}\right]$.

We begin by deriving a general expression that upper bounds the average pairwise error probability $\mathbb{E}\left[P_{e,\mathrm{pair}}\right]$ and then evaluate it for each of the possible statistical dependences within the set $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_{\mathcal{E}1}, \mathbf{X}_{\mathcal{E}2}\}$.

The decoder makes an error to the pair $\{\mathbf{X}_{\mathcal{E}1}, \mathbf{X}_{\mathcal{E}2}\}$ only if

$$\sum_{t=1}^{n}\left(\left[Y_t - \Psi_t(1,2)\right]^*\right)^2 \geq \sum_{t=1}^{n}\left(\left[Y_t - \Psi_t(\mathcal{E}1,\mathcal{E}2)\right]^*\right)^2 \tag{17}$$

where $\Psi(i,j)$ is defined in (15). The condition in (17) is equivalent to

$$\sum_{t=1}^{n}(Z_t^*)^2 \geq \sum_{t=1}^{n}\left([Z_t+\Psi_t(1,2)-\Psi_t(\mathcal{E}1,\mathcal{E}2)]^*\right)^2. \quad (18)$$

Define the *pairwise difference* random variable

$$\begin{aligned} U_t &= [\Psi_t(1,2)-\Psi_t(\mathcal{E}1,\mathcal{E}2)]^* \\ &= [X_{1,t}+\gamma X_{2,t}-X_{\mathcal{E}1,t}-\gamma X_{\mathcal{E}2,t}]^* \end{aligned} \quad (19)$$

and vector

$$\mathbf{U}=[U_1\ U_2\ \cdots\ U_n]. \quad (20)$$

Note that the distribution of the pairwise difference vector $\mathbf{U}$ encapsulates the statistical dependences in the set of codewords $\{\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_{\mathcal{E}1},\mathbf{X}_{\mathcal{E}2}\}$. We first express our upper bounds on the average pairwise error probability as a function of the random vector $\mathbf{U}$, and only then account for the fact that the statistics of $\mathbf{U}$ vary with the different types of statistical dependences between the transmitted and the "competing" pairs of codewords.

Substituting (19) into (18), we have that an error occurs only if

$$\sum_{t=1}^{n}(Z_t^*)^2 \geq \sum_{t=1}^{n}\left([Z_t+U_t]^*\right)^2. \quad (21)$$

Given a specific codebook from the ensemble was chosen, $\mathbf{U}$ is deterministic, and (21) implies

$$\begin{aligned} P_{e,\text{pair}} &= \Pr\left(\|\mathbf{Z}^*\|^2 \geq \left\|[\mathbf{Z}+\mathbf{u}]^*\right\|^2\right) \\ &= \Pr\left(\|\mathbf{Z}^*\|^2 \geq \min_{\mathbf{v}\in L\mathbb{Z}^n}\|\mathbf{Z}^*+\mathbf{u}+\mathbf{v}\|^2\right). \end{aligned} \quad (22)$$

Let $\mathbb{T}^n=\{-1,0,1\}^n$. Since every coordinate of the vectors $\mathbf{u}$ and $\mathbf{Z}^*$ has an absolute value smaller than $L/2$, the value of $\mathbf{v}\in L\mathbb{Z}^n$ that minimizes the expression $\|\mathbf{Z}^*+\mathbf{u}+\mathbf{v}\|^2$ cannot have an absolute value greater than L in any component, and it suffices to limit the search for it to $L\mathbb{T}^n$. Hence, (22) simplifies to

$$P_{e,\text{pair}} = \Pr\left(\|\mathbf{Z}^*\|^2 \geq \min_{\mathbf{v}\in L\mathbb{T}^n}\|\mathbf{Z}^*+\mathbf{u}+\mathbf{v}\|^2\right). \quad (23)$$

We now state a simple lemma that enables us to replace the folded Gaussian noise $\mathbf{Z}^*$ in (23) with a simple Gaussian noise $\mathbf{Z}$.

Lemma 1: For $\mathbf{z}\in\mathbb{R}^n$, $\mathbf{u}\in\mathcal{I}^n$, and the events

$$E_1=\left\{\|\mathbf{z}^*\|^2 \geq \min_{\mathbf{v}\in L\mathbb{T}^n}\|\mathbf{z}^*+\mathbf{u}+\mathbf{v}\|^2\right\} \quad (24)$$

and

$$E_2=\left\{\|\mathbf{z}\|^2 \geq \min_{\tilde{\mathbf{v}}\in L\mathbb{T}^n}\|\mathbf{z}+\mathbf{u}+\tilde{\mathbf{v}}\|^2\right\} \quad (25)$$

the following relation holds:

$$E_1\subseteq E_2.$$

Proof: See Appendix I. ∎

The next lemma provides an upper bound on $\mathbb{E}[P_{e,\text{pair}}]$, the average pairwise error probability over the ensemble, that depends on the statistics of U only through $\mathbb{E}\left[\exp\left\{-\frac{\mathsf{SNR}}{8}U^2\right\}\right]$.

Lemma 2: The average pairwise error probability over the ensemble is upper bounded by

$$\mathbb{E}[P_{e,\text{pair}}]\leq\Omega^n$$

where

$$\Omega=\mathbb{E}\left[\exp\left\{-\frac{\mathsf{SNR}}{8}U^2\right\}\right]+2\exp\left\{-\frac{3\mathsf{SNR}}{8}\right\}. \quad (26)$$

Proof: Using Lemma 1, we have

$$\begin{aligned} P_{e,\text{pair}} &= \Pr\left(\|\mathbf{Z}^*\|^2 \geq \min_{\mathbf{v}\in L\mathbb{T}^n}\|\mathbf{Z}^*+\mathbf{u}+\mathbf{v}\|^2\right) \\ &\leq \Pr\left(\|\mathbf{Z}\|^2 \geq \min_{\tilde{\mathbf{v}}\in L\mathbb{T}^n}\|\mathbf{Z}+\mathbf{u}+\tilde{\mathbf{v}}\|^2\right) \\ &= \Pr\left(\bigcup_{\tilde{\mathbf{v}}\in L\mathbb{T}^n}\left(\|\mathbf{Z}\|^2 \geq \|\mathbf{Z}+\mathbf{u}+\tilde{\mathbf{v}}\|^2\right)\right). \end{aligned} \quad (27)$$

Using the union bound, (27) can be further bounded by

$$\begin{aligned} P_{e,\text{pair}} &\leq \sum_{\tilde{\mathbf{v}}\in L\mathbb{T}^n}\Pr\left(\|\mathbf{Z}\|^2 \geq \|\mathbf{Z}+\mathbf{u}+\tilde{\mathbf{v}}\|^2\right) \\ &= \sum_{\tilde{\mathbf{v}}\in L\mathbb{T}^n}\Pr\left(-(\mathbf{u}+\tilde{\mathbf{v}})^T\mathbf{Z} \geq \frac{1}{2}\|\mathbf{u}+\tilde{\mathbf{v}}\|^2\right). \end{aligned} \quad (28)$$

Since $\mathbf{Z}$ is a vector of i.i.d. Gaussian components with zero mean and variance $1/\mathsf{SNR}$, the random variable $-(\mathbf{u}+\tilde{\mathbf{v}})^T\mathbf{Z}$ is Gaussian with zero mean and variance $\|\mathbf{u}+\tilde{\mathbf{v}}\|^2/\mathsf{SNR}$. Using the notation

$$Q(\tau)=\int_{\tau}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\tau^2\right\}d\tau$$

and recalling that

$$Q(\tau)\leq\exp\left\{-\frac{1}{2}\tau^2\right\}$$

(28) becomes

$$\begin{aligned} P_{e,\text{pair}} &\leq \sum_{\tilde{\mathbf{v}}\in L\mathbb{T}^n}Q\left(\frac{\sqrt{\mathsf{SNR}}}{2}\|\mathbf{u}+\tilde{\mathbf{v}}\|\right) \\ &\leq \sum_{\tilde{\mathbf{v}}\in L\mathbb{T}^n}\exp\left\{-\frac{\mathsf{SNR}}{8}\|\mathbf{u}+\tilde{\mathbf{v}}\|^2\right\}. \end{aligned}$$

In order to find the average (over the ensemble) pairwise error probability, we need to average $P_{e,\text{pair}}$ according to the distribution of $\mathbf{U}$

$$\mathbb{E}[P_{e,\text{pair}}]\leq\sum_{\tilde{\mathbf{v}}\in L\mathbb{T}^n}\mathbb{E}\left[\exp\left\{-\frac{\mathsf{SNR}}{8}\|\mathbf{U}+\tilde{\mathbf{v}}\|^2\right\}\right]. \quad (29)$$

Since the code generation is memoryless, $\mathbf{U}$ is also memoryless, and (29) can be rewritten as

$$\mathbb{E}[P_{e,\text{pair}}] \leq \sum_{\tilde{\mathbf{v}} \in L\mathbb{T}^n} \prod_{t=1}^{n} \mathbb{E}\left[\exp\left\{-\frac{\mathsf{SNR}}{8}(U_t + \tilde{v}_t)^2\right\}\right]. \quad (30)$$

Equation (30) can be further simplified by

$$\begin{aligned}
&\mathbb{E}[P_{e,\text{pair}}] \\
&\quad \leq \prod_{t=1}^{n} \sum_{\tilde{v} \in L\mathbb{T}} \mathbb{E}\left[\exp\left\{-\frac{\mathsf{SNR}}{8}(U_t + \tilde{v})^2\right\}\right] \\
&\quad = \left(\sum_{\tilde{v} \in L\mathbb{T}} \mathbb{E}\left[\exp\left\{-\frac{\mathsf{SNR}}{8}(U + \tilde{v})^2\right\}\right]\right)^n \quad (31) \\
&\quad = \left(\mathbb{E}\left[\exp\left\{-\frac{\mathsf{SNR}}{8}(U - L)^2\right\}\right]\right. \\
&\quad + \mathbb{E}\left[\exp\left\{-\frac{\mathsf{SNR}}{8}U^2\right\}\right] \\
&\quad \left. + \mathbb{E}\left[\exp\left\{-\frac{\mathsf{SNR}}{8}(U + L)^2\right\}\right]\right)^n \\
&\quad \leq \left(\mathbb{E}\left[\exp\left\{-\frac{\mathsf{SNR}}{8}U^2\right\}\right] + 2\exp\left\{-\frac{\mathsf{SNR}}{8}\frac{L^2}{4}\right\}\right)^n \quad (32) \\
&\quad = \left(\mathbb{E}\left[\exp\left\{-\frac{\mathsf{SNR}}{8}U^2\right\}\right] + 2\exp\left\{-\frac{3\mathsf{SNR}}{8}\right\}\right)^n \\
&\quad = \Omega^n
\end{aligned}$$

where (31) follows from the fact that the random variables $\{U_t\}_{t=1}^n$ are identically distributed, and (32) is true since $|U| \leq L/2$, and thus, $(U+L)^2 \geq L^2/4$ as well as $(U-L)^2 \geq L^2/4$. ■

In order to obtain an explicit upper bound on the average (over the ensemble) pairwise error probability, we are left with the task of calculating Ω, or equivalently calculating $\mathbb{E}\left[\exp\left\{-\frac{\mathsf{SNR}}{8}U^2\right\}\right]$.

We recall that $\mathbf{U}$ is a deterministic function of the pair of transmitted codewords $\{\mathbf{X}_1, \mathbf{X}_2\}$ and the pair of "competing" codewords $\{\mathbf{X}_{\mathcal{E}1}, \mathbf{X}_{\mathcal{E}2}\}$, where each one of the codewords is generated as specified in (14). The statistical dependences within the set of codewords $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_{\mathcal{E}1}, \mathbf{X}_{\mathcal{E}2}\}$ correspond to the linear dependences within the set of message vectors $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_{\mathcal{E}1}, \mathbf{w}_{\mathcal{E}2}\}$. Since we assumed the message vectors $\{\mathbf{w}_1, \mathbf{w}_2\}$ are linearly independent,[2] there are only four possible cases of linear dependences within the set $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_{\mathcal{E}1}, \mathbf{w}_{\mathcal{E}2}\}$.

Case A: The four vectors $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_{\mathcal{E}1}, \mathbf{w}_{\mathcal{E}2}\}$ are linearly independent.

Case B: The vectors $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_{\mathcal{E}1}\}$ are linearly independent and $\mathbf{w}_{\mathcal{E}2}$ is a linear combination of them.

Case C: The vectors $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_{\mathcal{E}2}\}$ are linearly independent and $\mathbf{w}_{\mathcal{E}1}$ is a linear combination of them.

Case D: The vectors $\{\mathbf{w}_1, \mathbf{w}_2\}$ are linearly independent and both $\mathbf{w}_{\mathcal{E}1}$ and $\mathbf{w}_{\mathcal{E}2}$ are linear combination of them.

Each case of statistical dependences induces a different distribution on U. Thus, for the calculation of $\mathbb{E}\left[\exp\left\{-\frac{\mathsf{SNR}}{8}U^2\right\}\right]$, each case should be considered separately. To that end, we now give upper bounds on Ω for the four different possible cases. The derivations of these bounds are given in Appendix III, and rely on some auxiliary lemmas brought in Appendix II

Case A: The codewords $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_{\mathcal{E}1}, \mathbf{X}_{\mathcal{E}2}\}$ are all statistically independent. Given $\{\mathbf{w}_1, \mathbf{w}_2\}$, there are less than 2^{2nR} pairs of competing message vectors $\{\mathbf{w}_{\mathcal{E}1}, \mathbf{w}_{\mathcal{E}2}\}$ that incur this kind of statistical dependence. Denote by Ω_A the value of Ω associated with case A. We have

$$\Omega_A < \frac{1}{p^2} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\delta^2(p,\gamma)} + 2e^{-\frac{3\mathsf{SNR}}{8}} \quad (33)$$

where

$$\delta(p,\gamma) = \min_{l \in \mathbb{Z}_p \setminus \{0\}} l \cdot \left|\gamma - \frac{\lfloor l\gamma \rceil}{l}\right|. \quad (34)$$

Case B: The codewords $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_{\mathcal{E}1}\}$ are statistically independent and

$$\mathbf{X}_{\mathcal{E}2} = [a\mathbf{X}_1 + b\mathbf{X}_2 + c\mathbf{X}_{\mathcal{E}1}]^*$$

where a, b, and c can take any value in $\mathbb{Z}_p$. Given $\{\mathbf{w}_1, \mathbf{w}_2\}$, there are no more than $p^3 2^{nR}$ pairs of competing message vectors $\{\mathbf{w}_{\mathcal{E}1}, \mathbf{w}_{\mathcal{E}2}\}$ that incur this kind of statistical dependence. Denote by Ω_B the value of Ω associated with case B. We have

$$\Omega_B < \frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\gamma)\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}} \quad (35)$$

where $\delta(p,\gamma)$ is as in (34).

Case C: The codewords $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_{\mathcal{E}2}\}$ are statistically independent and

$$\mathbf{X}_{\mathcal{E}1} = [a\mathbf{X}_1 + b\mathbf{X}_2 + c\mathbf{X}_{\mathcal{E}2}]^*$$

where a, b, and c can take any value in $\mathbb{Z}_p$. Given $\{\mathbf{w}_1, \mathbf{w}_2\}$, there are no more than $p^3 2^{nR}$ pairs of competing message vectors $\{\mathbf{w}_{\mathcal{E}1}, \mathbf{w}_{\mathcal{E}2}\}$ that incur this kind of statistical dependence. Denote by Ω_C the value of Ω associated with case C. We have

$$\Omega_C < \frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\gamma)\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}} \quad (36)$$

where $\delta(p,\gamma)$ is as in (34). Note that although the bounds (35) and (36) are identical, cases B and C are not identical (i.e., there is no symmetry) since the two codewords $\mathbf{X}_{\mathcal{E}1}$ and $\mathbf{X}_{\mathcal{E}2}$ play a different role in the pairwise difference vector $\mathbf{U}$, as $\mathbf{X}_{\mathcal{E}2}$ is multiplied by γ while $\mathbf{X}_{\mathcal{E}1}$ is not.

Case D: The codewords $\{\mathbf{X}_1, \mathbf{X}_2\}$ are statistically independent, whereas

$$\mathbf{X}_{\mathcal{E}1} = [a\mathbf{X}_1 + b\mathbf{X}_2]^*$$

and

$$\mathbf{X}_{\mathcal{E}2} = [c\mathbf{X}_1 + d\mathbf{X}_2]^*$$

where a, b, c, and d can take any value in $\mathbb{Z}_p$, except for $a = 1, b = 0, c = 0, d = 1$ (in which case $\mathbf{X}_{\mathcal{E}1} = \mathbf{X}_1$ and $\mathbf{X}_{\mathcal{E}2} = \mathbf{X}_2$). Given $\{\mathbf{w}_1, \mathbf{w}_2\}$, there are no more than p^4 pairs

[2]The message vectors $\{\mathbf{w}_{\mathcal{E}1}, \mathbf{w}_{\mathcal{E}2}\}$ are also linearly independent, as the decoder only searches over the pairs of linearly independent message vectors.

of competing message vectors $\{\mathbf{w}_{\mathcal{E}1}, \mathbf{w}_{\mathcal{E}2}\}$ that incur this kind of statistical dependence. Denote by Ω_D the value of Ω associated with case D. We have

$$\Omega_D < \max\left\{\frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\gamma)\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}}, \frac{p-1}{p} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\left(\gamma_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}\right)^2} + 2e^{-\frac{3\mathsf{SNR}}{8}}\right\} \tag{37}$$

where $\delta(p,\gamma)$ is given in (34).

We can now establish the theorem. Denote by $\mathbb{E}\left[P_{e,\mathrm{pair},i}\right]$, $i = A, B, C, D$, the average error probability associated with each case of statistical dependences. We recall that the decoder in our scheme only searches over the pairs of codewords corresponding to message vectors $\{\mathbf{w}_i, \mathbf{w}_j\}$ that are linearly independent. Thus, an error event occurs if the message vectors $\{\mathbf{w}_1, \mathbf{w}_2\}$ chosen by the users are linearly dependant. Denote by $P_{e,E}$ the probability of this event (which is independent of the codebook). By basic combinatorics

$$P_{e,E} = (p+1)\cdot 2^{-nR} - p\cdot 2^{-2nR} < 2p\cdot 2^{-nR}.$$

Using the union bound, the average error probability over the ensemble can be upper bounded by

$$\begin{aligned}\mathbb{E}\left[P_e\right] &\le 2^{2nR}\cdot\mathbb{E}[P_{e,\mathrm{pair},A}] + p^3\cdot 2^{nR}\cdot\mathbb{E}[P_{e,\mathrm{pair},B}] \\ &\quad + p^3\cdot 2^{nR}\cdot\mathbb{E}[P_{e,\mathrm{pair},C}] + p^4\cdot\mathbb{E}[P_{e,\mathrm{pair},D}] + P_{e,E} \\ &\le 2^{2n\left(R+\frac{1}{2}\log\Omega_A\right)} \\ &\quad + 2^{n\left(R+3\frac{\log p}{n}+\log\Omega_B\right)} \\ &\quad + 2^{n\left(R+3\frac{\log p}{n}+\log\Omega_C\right)} \\ &\quad + 2^{n\left(\frac{4\log p}{n}+\log\Omega_D\right)} \\ &\quad + 2^{n\left(-R+\frac{\log 2p}{n}\right)}.\end{aligned}$$

Holding p constant and taking n to infinity, we see that the average error probability goes to zero if

$$R < -\frac{1}{2}\log\Omega_A \tag{38}$$

$$R < -\log\Omega_B \tag{39}$$

$$R < -\log\Omega_C \tag{40}$$

$$0 > \log\Omega_D. \tag{41}$$

The conditions (38)–(40) imply that the rate should be taken to satisfy

$$R < \min\left\{-\frac{1}{2}\log\left(\frac{1}{p^2} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\delta^2(p,\gamma)} + 2e^{-\frac{3\mathsf{SNR}}{8}}\right) - \log\left(\frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\gamma)\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}}\right)\right\} \tag{42}$$

whereas condition (41) implies

$$\frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\gamma)\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}} < 1 \tag{43}$$

and

$$\frac{p-1}{p} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\left(\gamma_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}\right)^2} + 2e^{-\frac{3\mathsf{SNR}}{8}} < 1. \tag{44}$$

Condition (43) is satisfied for any positive rate, since it is contained in (42). Condition (43) is equivalent to

$$e^{-\frac{3\mathsf{SNR}}{2p^2}\left(\gamma_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}\right)^2} < 1 - 2pe^{-\frac{3\mathsf{SNR}}{8}}. \tag{45}$$

Since any prime value of p that satisfies (45) is valid, we can maximize (42) over all prime values of p satisfying (45), i.e., over all values in $\mathcal{P}(\gamma,\mathsf{SNR})$ as defined in (7), which yields (5). Finally, since there must be at least one codebook in the ensemble with a smaller (or equal) error probability than the average over the ensemble, the theorem is proved.

IV. Application to Interference Alignment

In the previous section, we found an achievable symmetric rate for the Gaussian modulo-additive MAC where both users use the same linear codebook. The motivation for developing such a coding scheme is to enable a finite SNR analysis of lattice interference alignment.

Assume a receiver observes a linear combination of codewords transmitted by several users (corrupted by noise) and is interested in decoding only one of the codewords, namely the received signal is

$$\mathbf{y} = h_1\mathbf{x}_1 + \sum_{k=2}^{K} h_k\mathbf{x}_k + \mathbf{z}$$

where $\{h_k\}_{k=1}^K$ are the channel gains, $\mathbf{x}_1$ is the desired codeword, $\{\mathbf{x}_k\}_{k=2}^K$ are the interfering codewords, and $\mathbf{z}$ is a vector of i.i.d. Gaussian noise.

One approach is to treat all the interfering codewords as noise. This approach would not be effective when the total power of the interference is on the order of that of the desired codeword (or stronger). Another possible approach would be trying to decode all the codewords $\{\mathbf{x}_k\}_{k=1}^K$, thus treating the channel as a MAC with K users. It is well known (see, for example, [25]) that at high SNR, time sharing is essentially optimal for the Gaussia MAC. In particular, for such a channel, if all users are working at the same rate, the symmetric rate scales like

$$\frac{1}{2K}\log(\mathsf{SNR}).$$

Since the decoder is only interested in one of the codewords, it seems wasteful to decode all of the interferers as well. For this reason, it is desirable to *align* all interferers to one codeword, as was first noticed in [17]. After alignment is performed, the receiver only has to decode two codewords: the desired codeword $\mathbf{x}_1$, and the aligned interference codeword.

A linear code, as defined in (2), facilitates the task of aligning the $K-1$ interfering codewords into one codeword. Specifically, if all interfering codewords $\{\mathbf{x}_k\}_{k=2}^K$ are taken from the same linear code $\mathcal{C}$, and the channel gains $\{h_k\}_{k=2}^K$ associated with the interfering codewords are all integers, we have

$$\left[\sum_{k=2}^K h_k \mathbf{x}_k\right]^* = \mathbf{x}_{\mathrm{IF}} \in \mathcal{C}$$

and therefore the received vector can be reduced modulo the interval $\mathcal{I}$ to yield

$$\begin{aligned}\mathbf{y}^* &= \left[h_1\mathbf{x}_1 + \left[\sum_{k=2}^K h_k\mathbf{x}_k\right]^* + \mathbf{z}\right]^* \\ &= \left[h_1\mathbf{x}_1 + \mathbf{x}_{\mathrm{IF}} + \mathbf{z}\right]^*. \end{aligned} \tag{46}$$

Since $\mathbf{x}_1$ and $\mathbf{x}_{\mathrm{IF}}$ are both members of the same linear codebook, the equivalent channel in (46) satisfies the conditions of Theorem 1, and we can find an achievable symmetric rate for it.

At this point, it is worth noting the advantage of joint decoding over successive decoding. A successive decoding procedure, as used in [18] and [20], can decode both codewords only if a very strong interference condition is satisfied, i.e., one of the codewords can be treated as noise while decoding the other codeword. For a wide range of values of h_1, successive decoding does not allow for positive transmission rates. The result of the previous section provides a nontrivial achievable rate region for a much wider range of values of h_1.

We next give a formal definition for the Gaussian interference channel, and then use the results of the previous section in order to derive achievable rates for certain classes of interference channels.

A. K-User Gaussian Interference Channel

The K-user Gaussian interference channel consists of K pairs of transmitters and receivers, where each transmitter k tries to convey one message w_k out of a set of 2^{nR_k} possible messages to its corresponding receiver. Specifically, the signal observed by receiver j is

$$Y_j = h_{jj}X_j + \sum_{k=1,k\neq j}^K h_{jk}X_k + Z_j$$

where h_{jk} is the channel gain from transmitter k to receiver j, and Z_j is the Gaussian noise present at receiver j. All transmitters and receivers have perfect knowledge of all channel gains. We assume that the Gaussian noise at each receiver is i.i.d. with zero mean and variance $1/\mathsf{SNR}$ and that the noises at different receivers are statistically independent. We assume all transmitters are subject to the same power constraint

$$\frac{1}{n}\mathbb{E}\left[\|\mathbf{x}_k\|^2\right] \leq 1.$$

Each transmitter k has an encoding function

$$f_k : \left\{1,\ldots,2^{nR_k}\right\} \to \mathbb{R}^n$$

such that the signal transmitted by user k during n channel uses is

$$\mathbf{x}_k = f_k(w_k).$$

Receiver j recovers the message using a decoding function

$$g_j : \mathbb{R}^n \to \left\{1,\ldots,2^{nR_j}\right\}.$$

Let

$$\hat{w}_j = g_j(\mathbf{y}_j)$$

be the estimate receiver j produces for the message transmitted by transmitter j. We define the error probability as the probability that at least one of the receivers did not decode its intended message correctly

$$P_{e,\mathrm{IF}} = \mathbb{E}\left[\Pr\left(\{\hat{w}_1,\ldots,\hat{w}_K\} \neq \{w_1,\ldots,w_K\}\right)\right]$$

where the expectation here is over a uniform distribution on the messages.

We say that a rate tuple $\{R_1,\ldots,R_K\}$ is achievable if there exists a set of encoding and decoding functions such that $P_{e,\mathrm{IF}}$ vanishes as n goes to infinity, and that a symmetric rate R_{sym} is achievable if the rate tuple $\{R_{\mathrm{sym}},\ldots,R_{\mathrm{sym}}\}$ is achievable. In the sequel, we focus on the achievable symmetric rate.

B. Integer-Interference Channel

We restrict attention to a special family of K-user Gaussian interference channels which we refer to as the integer-interference channel. In this family, all the channel gains corresponding to interferers are integers, i.e., for all $j \neq k$

$$h_{jk} = a_{jk} \in \mathbb{Z}.$$

Note that the symmetric Gaussian K-user interference channel falls in this family of channels. The following theorem establishes an achievable symmetric rate for the integer-interference channel.

Theorem 2: For the K-user integer-interference channel, the following symmetric rate is achievable:

$$\begin{aligned} R_{\mathrm{sym}} < \max_{p\in\bigcap_{i=1}^K \mathcal{P}(h_{ii},\mathsf{SNR})} \min_{j\in\{1,\ldots,K\}} \min \\ \left\{ -\frac{1}{2}\log\left(\frac{1}{p^2} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\delta^2(p,h_{jj})} + 2e^{-\frac{3\mathsf{SNR}}{8}}\right), \right. \\ \left. -\log\left(\frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,h_{jj})\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}}\right)\right\} \end{aligned} \tag{47}$$

where $\delta(\cdot,\cdot)$ is defined in (6), and $\mathcal{P}(\gamma,\mathsf{SNR})$ is defined in (7).

Proof: We begin by recalling the encoding and decoding procedures.

Encoding: Consider the linear ensemble of codebooks of rate R_{sym} over $\mathbb{Z}_p$ described in Section III, where p is taken as the maximizing value in (47). Choose the codebook $\mathcal{C}$ which

achieves the smallest average error probability, $P_{e,\mathrm{IF}}$. Each user encodes its message using this codebook.

Decoding: Each receiver first reduces its observation modulo the interval $\mathcal{I}$. The equivalent channel receiver j sees is therefore

$$\begin{aligned}\mathbf{y}_j^* &= \left[h_{jj}\mathbf{x}_j + \sum_{k=1,k\neq j}^{K} a_{jk}\mathbf{x}_k + \mathbf{z}_j\right]^* \\ &= [h_{jj}\mathbf{x}_j + \mathbf{x}_{\mathrm{IF},j} + \mathbf{z}_j]^*\end{aligned}$$

where

$$\mathbf{x}_{\mathrm{IF},j} = \left[\sum_{k=1,k\neq j}^{K} a_{jk}\mathbf{x}_k\right]^*.$$

The linearity of the codebook implies that $\mathbf{x}_{\mathrm{IF},j} \in \mathcal{C}$.

From Theorem 1, we know that $\mathbf{x}_j$ and $\mathbf{x}_{\mathrm{IF},j}$ can be decoded reliably (by receiver j) as long as the symmetric rate R_{sym} satisfies

1)

$$R_{\mathrm{sym}} < -\frac{1}{2}\log\left(\frac{1}{p^2} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\delta^2(h_{jj},p)} + 2e^{-\frac{3\mathsf{SNR}}{8}}\right) \tag{48}$$

2)

$$R_{\mathrm{sym}} < -\log\left(\frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,h_{jj})\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}}\right) \tag{49}$$

and

$$p \in \mathcal{P}(\gamma, \mathsf{SNR}). \tag{50}$$

The codebook $\mathcal{C}$ satisfies conditions (48)–(50) for every[3] $1 \leq j \leq K$, and thus, the theorem is proved. ■

As we shall see in the next section, the achievable sum rate from Theorem 2 scales like $^{K}/_{4}\log(\mathsf{SNR})$ for almost all direct channel gains. Recall that the symmetric rate from Theorem 1 for the two-user MAC does not increase with the SNR when the channel gain γ is rational. For the integer-interference channel, this translates to a bounded symmetric rate if one of the direct channel gains is rational. Thus, while for finite values of SNR the symmetric rate is not necessarily sensitive to the direct channel gains being rational or irrational, the asymptotic behavior is poor for rational channel gains. Integer-interference channels with rational direct channel gains were considered in [26] and in [27], and finite-SNR achievable rate regions were developed for such channels. The coding schemes used in these works combined lattice codes with an additional algebraic coding procedure. In certain cases, these schemes achieve better DoF than time sharing. Thus, in the case of rational direct channel gains with small denominator, for which our scheme is less effective, the schemes of [26] and [27] may be used.

[3]The existence of a codebook $\mathcal{C}$ that is simultaneously good for all K equivalent MACs is guaranteed for any finite number of users K, as this is a compound channel model.

C. Integer-Interference Channel: DoFs

Theorem 2 provides an achievable symmetric rate for the integer-interference channel that is valid for any SNR. We now show that in the limit where the SNR goes to infinity, for almost all direct channel gains, the coding scheme achieves $K/2$ DoFs, which is the upper bound established in [28]. This sanity check shows that for the integer-interference channel, the proposed scheme is optimal in a DoF sense for almost all channel gains, and (partially) recovers the asymptotic results of [5].

Before giving a formal definition to the number of DoF, we need a few preliminary definitions. We define an interference channel code $\mathcal{C}'$ as a set of encoders $\{f_k\}_{k=1}^K$ and decoders $\{g_k\}_{k=1}^K$. We define an interference channel coding scheme as a family of interference channel codes $\{\mathcal{C}'(\mathsf{SNR})\}$, and define $\mathcal{R}'(\mathsf{SNR})$ as the set of all rate-tuples that are achievable for the interference channel code $\mathcal{C}'(\mathsf{SNR})$.

Definition 1: An interference channel coding scheme $\{\mathcal{C}'(\mathsf{SNR})\}$ is said to achieve d degrees of freedom if

$$\lim_{\mathsf{SNR}\to\infty} \max_{R_1,\ldots,R_K \in \mathcal{R}'(\mathsf{SNR})} \frac{\sum_{k=1}^K R_k}{\frac{1}{2}\log(1+\mathsf{SNR})} = d.$$

Theorem 3: The lattice interference alignment scheme from Section IV-B achieves $K/2$ DoF for almost every integer-interference channel.

For the proof, we will need the following theorem from the field of Diophantine approximations, which is due to Khinchin.

Theorem 4 (Khinchin): For almost every $\gamma \in \mathbb{R}$, the number of solutions to the inequality

$$\left|\gamma - \frac{a}{l}\right| \leq \Phi(l)$$

for $a \in \mathbb{Z}$ and $l \in \mathbb{N}$ is finite if the series

$$\sum_{l=1}^{\infty} l\Phi(l)$$

converges, and infinite if it diverges.

Proof: See, e.g., [29]. ■

Using Khinchin's Theorem, we now prove Theorem 3.

Proof of Theorem 3: Setting $\Phi(l) = l^{-2-\epsilon_1}$ in Theorem 4, it follows that for any $\epsilon_1 > 0$, and almost every $\gamma \in \mathbb{R}$, there exist an integer $l^*(\gamma, \epsilon_1)$ for which there are no solutions to the inequality

$$l \cdot \left|\gamma - \frac{\lfloor l\gamma \rceil}{l}\right| \leq l^{-1-\epsilon_1} \tag{51}$$

in the range $l > l^*(\gamma, \epsilon_1)$. Moreover, for such γ, there exist a constant $c_1 > 0$ for which

$$l \cdot \left|\gamma - \frac{\lfloor l\gamma \rceil}{l}\right| \geq c_1 \tag{52}$$

for every $l \leq l^*(\gamma, \epsilon_1)$. Combining (51) with (52), it follows that for almost every $\gamma \in \mathbb{R}$, there exist a positive integer $\tilde{l}^*(\gamma, \epsilon_1)$, such that if $p > \tilde{l}^*(\gamma, \epsilon_1)$, there are no solutions to the inequality

$$l \cdot \left|\gamma - \frac{\lfloor l\gamma \rceil}{l}\right| \leq p^{-1-\epsilon_1}$$

in the range $l \in \mathbb{Z}_p \backslash \{0\}$. Thus, for any $\epsilon_1 > 0$ and p large enough, we have

$$\delta(p,\gamma) \geq p^{-1-\epsilon_1} \tag{53}$$

for almost every $\gamma \in \mathbb{R}$.

We would now like to show that

$$\lim_{\mathsf{SNR}\to\infty} \frac{R_{\mathrm{sym}}}{\frac{1}{2}\log(1+\mathsf{SNR})} = \frac{1}{2} \tag{54}$$

where R_{sym} is the symmetric rate from (47).

Set $p = \mathsf{SNR}^{1/4-\epsilon_2}$ (where $\epsilon_2 > 0$ is chosen such that p is a prime number). From (53), we see that for this choice, for high enough SNR, and almost every $\gamma \in \mathbb{R}$, we have

$$\delta(p,\gamma) > \mathsf{SNR}^{-1/4+\epsilon_1\epsilon_2+\epsilon_2-\epsilon_1/4}. \tag{55}$$

We now use (55) in order to find lower bounds on the maximal achievable symmetric rate of Theorem 2 for asymptotic SNR conditions.

For almost every $h_{jj} \in \mathbb{R}$, the argument of the logarithm in (48) can be upper bounded (after some straightforward algebra) by

$$\begin{aligned}&\frac{1}{p^2} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + \frac{1}{p} e^{-\frac{3\mathsf{SNR}}{2p^2}\delta^2(p,h_{jj})} + 2e^{-\frac{3\mathsf{SNR}}{8}} \\ &\quad < \mathsf{SNR}^{-1/2+2\epsilon_2} + \sqrt{\frac{2\pi}{3}}\mathsf{SNR}^{-1/2} \\ &\quad + \mathsf{SNR}^{-1/4+\epsilon_2} e^{-\frac{3}{2}\mathsf{SNR}^{4\epsilon_2+2\epsilon_1\epsilon_2-\epsilon_1/2}} + 2e^{-\frac{3\mathsf{SNR}}{8}}\end{aligned} \tag{56}$$

and the argument of the logarithm in (49) by

$$\begin{aligned}&\frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,h_{jj})\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}} \\ &\quad < \mathsf{SNR}^{-1/4+\epsilon_2} + \sqrt{\frac{2\pi}{3}}\mathsf{SNR}^{-1/4-\epsilon_2-\epsilon_1\epsilon_2+\epsilon_1/4} + 2e^{-\frac{3\mathsf{SNR}}{8}}.\end{aligned} \tag{57}$$

Taking $\epsilon_1 = \epsilon$, and $\epsilon/2 < \epsilon_2(\mathsf{SNR}) < \epsilon$ chosen such that p is prime ensures that for any $\epsilon > 0$, for SNR high enough, (56) and (57) can be bounded by

$$\begin{aligned}&\frac{1}{p^2} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + \frac{1}{p} e^{-\frac{3\mathsf{SNR}}{2p^2}\delta^2(p,h_{jj})} + 2e^{-\frac{3\mathsf{SNR}}{8}} \\ &\quad < 2\mathsf{SNR}^{-1/2+2\epsilon}\end{aligned} \tag{58}$$

and

$$\frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,h_{jj})\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}} < 2\mathsf{SNR}^{-1/4+\epsilon}. \tag{59}$$

Since (58) and (59) continue to hold simultaneously for almost all sets of direct channel gains in $\mathbb{R}^K$, for any $\epsilon > 0$, for SNR high enough, the symmetric rate

$$R_{\mathrm{sym}} = \left(\frac{1}{4} - \epsilon\right)\log(\mathsf{SNR}) - 1 \tag{60}$$

is achievable for almost every $h_{jj} \in \mathbb{R}$. Taking ϵ to zero gives (54), from which the theorem follows. ■

D. Example

Consider the five-user integer-interference channel where the channel gains are the entries of the matrix

$$H = \begin{pmatrix} h & 1 & 2 & 3 & 4 \\ 5 & h & 3 & 6 & 7 \\ 2 & 11 & h & 1 & 3 \\ 3 & 7 & 6 & h & 9 \\ 11 & 2 & 6 & 4 & h \end{pmatrix}. \tag{61}$$

The integers in the off-diagonal elements of H are arbitrary. For this channel, we plot the achievable sum rate of our scheme (which is five times the symmetric rate R_{sym}), and for reference, we also plot the sum rate a time-sharing scheme would have achieved. One more curve we plot for reference is the curve

$$\frac{K}{2}\frac{1}{2}\log(1+(1+h^2)\mathsf{SNR}) \tag{62}$$

which at high SNR corresponds to the sum rate that could have been achieved if the symmetric rate for a two-user Gaussia MAC with one linear code was the same as that of the same channel with two random codes, in other words, if r_{norm} given in (13) were 1. In the absence of explicit upper bounds for the K-user interference channel with finite SNR, (62) serves as a reasonable benchmark to the best performance one can expect to achieve, which is based on the known fact that the number of DoF the channel offers is $K/2$.

We consider two different values of h: $h = 0.707$, and $h = \sqrt{2}/2$.[4] The results are shown in Fig. 5. In Fig. 6, we plot the same curves for $h = 0.24$ and $h = \sqrt{7}/11$.

These examples show the advantages of lattice interference alignment over time sharing for high enough SNR. For a larger number of users, interference alignment is preferable over time sharing for lower values of SNR. We note that in this paper we have not considered further optimizations for the achievable rate, such as combining it with time sharing of powers, or superposition (layered coding schemes), which may result in higher gains. Time sharing of powers, i.e., having all users transmit with more power for some of the time and remain silent for the rest of the time is important at low SNR values, since our coding scheme only achieves positive rates above some SNR threshold.

An important insight from Figs. 5 and 6 is the sensitivity of interference alignment to the channel gain h. Even though in each figure we have used values of h that are very close, the

[4]Our coding scheme would have the same performance for $h+m$, $m \in \mathbb{Z}$ as well; however, the reference curves do change when adding integers to h.

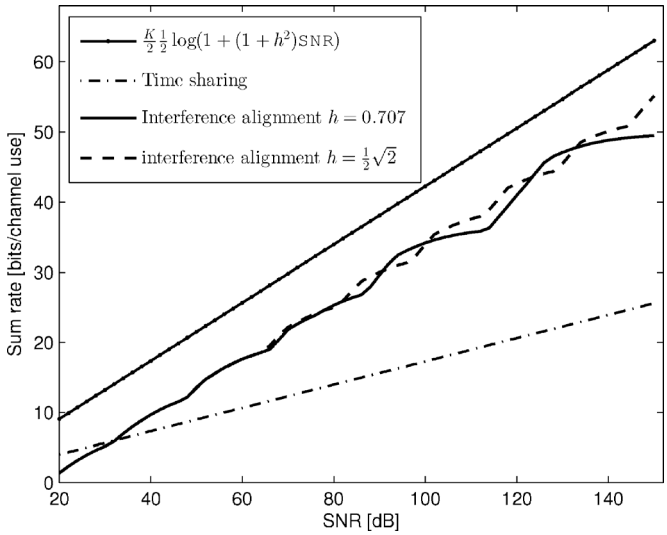

Fig. 5. Achievable sum rate for the integer-interference channel (61) for $h = 0.707$ and $h = \sqrt{2}/2$.

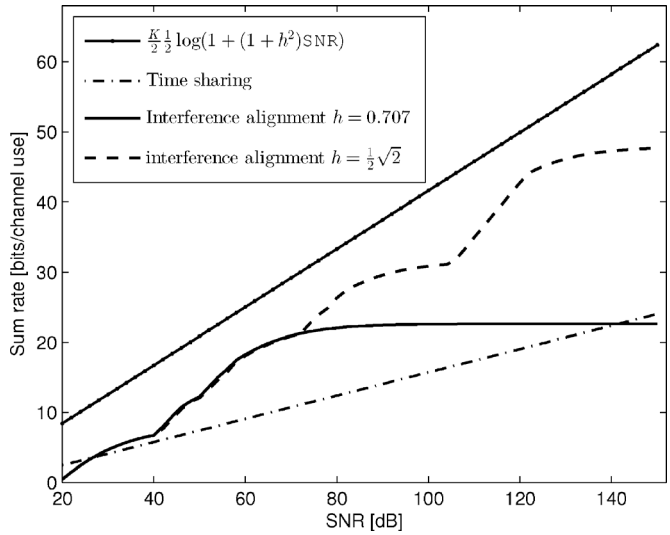

Fig. 6. Achievable sum rate for the integer-interference channel (61) for $h = 0.24$ and $h = \sqrt{7}/11$.

performance of the scheme differs significantly when the SNR is very high. This sensitivity is the subject of the next section.

V. Quantifying Rationality

Previous results regarding interference alignment for the static K-user Gaussian interference channel were mainly focused on the DoF of the channel. In [5], it is proved that the number of DoF of a K-user integer-interference channel are $K/2$ when the direct channel gains are irrational algebraic[5] and is strictly smaller than $K/2$ when the channel gains are rational. This result is also supported by the results of [6] where it is shown that the DoF of a K-user (not necessarily integer) interference channel is $K/2$, unless there exist some rational connections between the channel coefficients.

Clearly, from an engineering perspective, the dependence of the DoF on whether the direct channel gains are rational or irrational is very displeasing. It is therefore important to understand the effect of the channel gains being rational at finite SNR.

Theorem 2 sheds light on this matter. Specifically, it quantifies the loss associated with a rational direct channel gain h_{jj} in the symmetric rate as a function of how large the denominator of h_{jj} is w.r.t. the SNR. Specifically, if h_{jj} is a rational number of the form $h_{jj} = r/q$, the symmetric rate achieved by the proposed coding scheme can never exceed $\log(q)$. This is evident from the presence of the factor $\delta(p, h_{jj})$ in (47). For any $p > q$, we have $\delta(p, r/q) = 0$. Thus, in order to get positive rates, we must choose $p \leq q$ and the symmetric rate would be smaller than $\log(q)$ for any SNR. For this reason, a small denominator q limits the symmetric rate even for low values of SNR. However, a large value of q limits performance only at high SNR. This phenomenon can be seen in Fig. 5 where at a certain SNR point, the symmetric rate corresponding to $h = 0.707$ (which is a rational number) saturates. In Fig. 6, this effect is even more pronounced for $h = 0.24$ as the denominator of h in this case is $q = 25$, rather than $q = 1000$ which is the case for $h = 0.707$. It is also seen from Figs. 5 and 6 that for low values of SNR, the symmetric rates corresponding to the irrational values of h and their quantized rational versions are nearly indistinguishable.

Another question that arises from the results of [5] and [6] is how the rate behaves when the direct gains approach a rational number. Theorem 2 provides an answer to this question as well. If $h_{jj} - r/q$ is small, then $\delta(h_{jj}, p)$ would also be small for $p > q$, which would result in an effectively lower SNR. However, the function $\delta(p, h_{jj})$ is continuous in the second variable, and thus, letting h_{jj} approach r/q (where $q < p$) results in a *continuous* decrease of the effective SNR.

VI. Noninteger Interference Channels

We have seen that for the integer-interference channel, the result of Theorem 1 was very useful for finding a new achievable rate region. The requirement that at each receiver all the channel gains corresponding to interferers are integers is necessary because the codebook we use is only closed under addition of integer-valued multiplications of codewords, which allows us to align all interfering signals into one codeword.

Unfortunately, the integer-interference channel model does not capture the essence of the physical wireless medium. Under realistic statistical models for the interference channel, the probability of getting an integer-interference channel is clearly zero.

It is thus desirable to transform a general interference channel into an integer-interference channel by applying certain operations at the transmitters and the receivers.[6] Specifically, it is necessary that at each receiver, the ratios between all interference gains be rational, and then, an appropriate scaling at each receiver can transform the channel into an integer-interference channel.

For the general K-user interference channel with arbitrary channel gains, assume that each receiver had scaled its observation such that one of the interference gains equals 1. We would

[5]In fact, it was proved in [30] (and also in [12]) that the number of DoF is $K/2$ for all irrational direct channel gains.

[6]We do not discuss the possibility of adding more antennas at the receivers or the transmitters which could also assist in the problem.

like to "shape" the other interfering gains seen at each receiver to be integers as well, using operations at the transmitters. It turns out that by using power back-off at each transmitter, i.e., each transmitter k scales its codeword by a factor of $\alpha_k \leq 1$ prior to transmission, it is possible to transform only $K-1$ (in addition to the channel gains that were equalized to 1 by the receivers) of the total of $K(K-1)$ interfering channel gains into integers. It follows that perfect alignment, i.e., alignment of all interferers at all receivers simultaneously, is not possible (by these methods) even for $K=3$ users.

One solution to this problem is performing partial alignment, as described in [6], which is suitable for almost every set of channel gains. This method roughly transforms the channel seen by each receiver into a MAC with a large number of inputs, where about half of the inputs correspond to the information transmitted by the desired transmitter, and the other half to interferences. At asymptotic (high) SNR conditions, it was shown in [6] that this approach achieves $K/2$ DoF, i.e., each receiver is capable of decoding all the inputs corresponding to its intended messages in the presence of the interfering inputs.

An extension of the partial interference alignment approach proposed in [6] to the nonasymptotic SNR regime, in the same manner we extended the results of [5] for the integer-interference channel to nonasymptotic SNR, requires an extension of Theorem 1 to more than two users. After the appearance of this work, such an extension was derived in [21]. This extension enables us to translate the DoF results of [6] into achievable rate regions. Nevertheless, the large number of layers used by the partial alignment scheme of [6] makes this rate region inferior to that achieved by time sharing for most channel parameters of practical interest.

A different approach that is yet to be exhausted for the static interference channel is using the time dimension in conjunction with power back-off in order to achieve alignment, and allow for joint decoding of the intended message and some function of the interferers at each receiver.

Such an approach can be thought of as the interference channel's dual of space-time codes, and we refer to it as "power-time" codes. An example of such a power-time code is given in Section VI-A. The power-time code in the example is suitable for the three-user interference channel (with arbitrary channel coefficients), and allows us to achieve $9/8$ DoF for almost all channel gains. While it is already known from [6] that the number of DoF offered by this channel is $3/2$, our "power-time" approach gives an explicit expression for the symmetric rate for finite SNR.

A. Example of a Three-User Interference Channel Power-Time Code

In this section, we introduce a coding scheme that utilizes both power back-off at the transmitters, and the time domain, in order to allow for perfect interference alignment (with some loss in the number of DoF). We illustrate the scheme by an example which is useful for the general three-user interference channel and achieves 9/8 DoFs out of the 3/2 DoF afforded by the channel for almost all channel gains (see [6]).

We consider the channel

$$H = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}.$$

We use the channel $4n$ times, in order to transmit three codewords of length n by each user. We refer to n consecutive channel uses as a frame. The actions taken by the transmitters and the receivers vary from frame to frame, as will be described in detail. All transmitters and receivers use the same linear codebook $\mathcal{C}$ of rate R_{sym} and length n during all frames. The codeword transmitted by user k at frame t is denoted by $\mathbf{x}_{k,t}$.

We assume that $h_{13} \geq h_{12}, h_{22} \geq h_{23}$, and $h_{32} \geq h_{31}$. There is no loss of generality in this assumption, as the scheme we now describe can be easily modified for different ratios between the channel gains. For all frames, receiver 1 scales its observation by $1/h_{12}$, receiver 2 scales its observation by $1/h_{23}$, and receiver 3 scales its observation by $1/h_{31}$ such that the equivalent channel is

$$\tilde{H} = \begin{pmatrix} \tilde{h}_{11} & 1 & \tilde{h}_{13} \\ \tilde{h}_{21} & \tilde{h}_{22} & 1 \\ 1 & \tilde{h}_{32} & \tilde{h}_{33} \end{pmatrix}$$

where $\tilde{h}_{1j} = h_{1j}/h_{12}$, $\tilde{h}_{2j} = h_{2j}/h_{23}$, and $\tilde{h}_{3j} = h_{3j}/h_{31}$.

We describe the operations taken by the transmitters and the receivers at each frame.

1) *Frame* 1: User 1 transmits the codeword $\mathbf{x}_{1,1}$, user 2 transmits the codeword $\mathbf{x}_{2,1}$, and user 3 transmits the codeword $\mathbf{x}_{3,1}$.
   User 3 scales its codeword by the factor $\alpha_3 = 1/\tilde{h}_{13}$, and all other transmitters do not scale their codewords. The equivalent channel is thus

$$\tilde{H}_1 = \begin{pmatrix} \tilde{h}_{11} & 1 & 1 \\ \tilde{h}_{21} & \tilde{h}_{22} & 1/\tilde{h}_{13} \\ 1 & \tilde{h}_{32} & \tilde{h}_{33}/\tilde{h}_{13} \end{pmatrix}.$$

   Due to the perfect alignment at receiver 1, it can decode $\mathbf{x}_{1,1}$, the codeword transmitted by user 1. The other receivers cannot decode their codewords at this stage.
2) *Frame* 2: User 1 transmits the codeword $\mathbf{x}_{1,2}$, user 2 transmits the codeword $\mathbf{x}_{2,2}$, and user 3 transmits the codeword $\mathbf{x}_{3,2}$.
   User 1 scales its codeword by the factor $\alpha_1 = 1/\tilde{h}_{21}$, and all other transmitters do not scale their codewords. The equivalent channel is thus

$$\tilde{H}_2 = \begin{pmatrix} \tilde{h}_{11}/\tilde{h}_{21} & 1 & \tilde{h}_{13} \\ 1 & \tilde{h}_{22} & 1 \\ 1/\tilde{h}_{21} & \tilde{h}_{32} & \tilde{h}_{33} \end{pmatrix}.$$

   Due to the perfect alignment at receiver 2, it can decode $\mathbf{x}_{2,2}$, the codeword transmitted by user 2. The other receivers cannot decode their codewords at this stage.
3) *Frame* 3: User 1 transmits the codeword $\mathbf{x}_{1,3}$, user 2 transmits the codeword $\mathbf{x}_{2,3}$, and user 3 transmits the codeword $\mathbf{x}_{3,3}$.

User 2 scales its codeword by the factor $\alpha_2 = 1/\tilde{h}_{32}$, and all other transmitters do not scale their codewords. The equivalent channel is thus

$$\tilde{H}_3 = \begin{pmatrix} \tilde{h}_{11} & 1/\tilde{h}_{32} & \tilde{h}_{13} \\ \tilde{h}_{21} & \tilde{h}_{22}/\tilde{h}_{32} & 1 \\ 1 & 1 & \tilde{h}_{33} \end{pmatrix}.$$

Due to the perfect alignment at receiver 3, it can decode $\mathbf{x}_{3,3}$, the codeword transmitted by user 3. The other receivers cannot decode their codewords at this stage.

4) *Frame* 4: In this frame, each user repeats a codeword it has already transmitted in one of the previous frames. Specifically, user 1 transmits the codeword $\mathbf{x}_{1,1}$, user 2 transmits the codeword $\mathbf{x}_{2,2}$, and user 3 transmits the codeword $\mathbf{x}_{3,3}$. None of the users scale their codewords, such that the equivalent channel is $\tilde{H}$. Now, receiver k observes the signal

$$\mathbf{y}_{k,4} = \sum_{j=1}^{3} \tilde{h}_{jk}\mathbf{x}_{j,j} + \mathbf{z}_k$$

where $\mathbf{z}_k$ is the Gaussian noise present at receiver k.
Since receiver k had already decoded the codeword $\mathbf{x}_{k,k}$ in the kth frame, it can subtract $\tilde{h}_{kk}\mathbf{x}_{k,k}$ from $\mathbf{y}_{k,4}$ which results in the equivalent two-user MAC

$$\bar{\mathbf{y}}_{k,4} = \mathbf{y}_{k,4} - \tilde{h}_{kk}\mathbf{x}_{k,k} = \sum_{j=1, j\neq k}^{3} \tilde{h}_{jk}\mathbf{x}_{j,j} + \mathbf{z}_k.$$

User k can now decode the two codewords transmitted by the other users during the fourth frame. For instance, in this step, user 1 decodes the codewords $\mathbf{x}_{2,2}$ and $\mathbf{x}_{3,3}$.
Now that user 1 has the side information $\mathbf{x}_{2,2}$ and $\mathbf{x}_{3,3}$, it can return to its observations from the second and the third frames. It can subtract from $\mathbf{y}_{1,2}$ the term $\mathbf{x}_{2,2}$, leaving it with a two-user MAC which allows it to decode $\mathbf{x}_{1,2}$. In the same manner, it can subtract the term $\tilde{h}_{13}\mathbf{x}_{3,3}$ from $\mathbf{y}_{1,3}$, leaving it with a two-user MAC which allows it to decode $\mathbf{x}_{1,3}$.
The same procedure is done by each one of the other decoders.

The described power-time code results in three different codewords, each with a rate that scales like $\frac{1}{4}\log \mathsf{SNR}$, that were decoded by each decoder. Taking into account the (symbol) rate of the power-time code, which is 3/4 (since the fourth channel use is "wasted"), we get a sum rate that scales like $\frac{9}{8}\frac{1}{2}\log \mathsf{SNR}$, which means that the number of DoF is 9/8 . More importantly, using Theorem 1, we can find an achievable symmetric rate for this power-time code for any SNR. From our DoF analysis of Section IV-C, we know that for almost any channel realization, there exist a certain value of SNR from which the described power-time code outperforms time sharing. While this value of SNR may be very high, it is finite, and can be explicitly computed for a given channel realization. To the best of our knowledge, this is the first nontrivial example for a scheme that outperforms time sharing at finite values of SNR over a general fully connected *real* K-user interference channel. Note that for general complex fully connected three-user interference channels, Cadambe *et al.* [31] proposed an asymmetric complex signaling scheme that achieves 6/5 DoF and gives a finite SNR achievable rate region.

VII. Concluding Remarks

In this work, we have made a first attempt toward characterizing the performance of lattice-based interference alignment schemes at finite SNR. To do that, we recognized that a fundamental feature of such schemes is that each receiver sees an induced MAC where all inputs are from the same lattice codebook. A new coding theorem was derived for the two-user MAC where both users use the same lattice codebook. While the rate expressions given by this theorem are not very tight, they seem to capture the essence of the problem.

The new coding theorem was utilized for establishing a new achievable rate region for the integer-interference channel, which is based on lattice interference alignment, and is valid for any value of SNR. Previous DoF results for this family of channels suggested that at high SNR conditions, its capacity becomes very sensitive to the exact values of the channel gains.

While at high SNR the rate region derived here agrees with those results, and is indeed very sensitive to the channel gains, at moderate values of SNR, it is rather robust. This shows that lattice interference alignment is a legit candidate for coding over fully connected static interference channels at SNR conditions of practical interest. Indeed, subsequent work [21] used lattice interference alignment in order to approximate the capacity of the symmetric Gaussian K-user interference channel for all channel gains outside some outage set.

The channel model used here, as well as in [21], is such that all interfering lattice codewords are received aligned in all receivers. For general interference channels, this is not the case. Thus, efficient precoding schemes for creating alignment need to be developed in order to make lattice interference alignment attractive. For very high SNR, this problem was solved in [6], but the problem remains open for moderate values of SNR.

Appendix I
Proof of Lemma 1

Proof: Let

$$\mathbf{v}' = \arg \min_{\mathbf{v}\in L\mathbb{Z}^n} \|\mathbf{z} + \mathbf{u} + \mathbf{v}\|^2$$

and denote by $\mathcal{S}$ the set of indices for which the absolute value of $\mathbf{v}'$ is not greater than L, and by $\bar{\mathcal{S}}$ the set of indices for which it is greater than L, namely

$$\mathcal{S} = \{s : |v'(s)| \leq L\}$$
$$\bar{\mathcal{S}} = \{s : |v'(s)| > L\}.$$

Let $\mathbf{u}_{\mathcal{S}}, \mathbf{v}_{\mathcal{S}}, \tilde{\mathbf{v}}_{\mathcal{S}}, \mathbf{z}_{\mathcal{S}}, \mathbf{z}^*_{\mathcal{S}}$ be subvectors of $\mathbf{u}, \mathbf{v}, \tilde{\mathbf{v}}, \mathbf{z}, \mathbf{z}^*$ in the indices $\mathcal{S}$, and $\mathbf{u}_{\bar{\mathcal{S}}}, \mathbf{v}_{\bar{\mathcal{S}}}, \tilde{\mathbf{v}}_{\bar{\mathcal{S}}}, \mathbf{z}_{\bar{\mathcal{S}}}, \mathbf{z}^*_{\bar{\mathcal{S}}}$ be their subvectors in the indices $\bar{\mathcal{S}}$. It suffices to show the next two inequalities

$$\|\mathbf{z}_{\mathcal{S}}\|^2 - \min_{\tilde{\mathbf{v}}_{\mathcal{S}}\in L\mathbb{T}^{|\mathcal{S}|}} \|\mathbf{z}_{\mathcal{S}} + \mathbf{u}_{\mathcal{S}} + \tilde{\mathbf{v}}_{\mathcal{S}}\|^2 \geq$$
$$\|\mathbf{z}^*_{\mathcal{S}}\|^2 - \min_{\mathbf{v}_{\mathcal{S}}\in L\mathbb{T}^{|\mathcal{S}|}} \|\mathbf{z}^*_{\mathcal{S}} + \mathbf{u}_{\mathcal{S}} + \mathbf{v}_{\mathcal{S}}\|^2 \qquad (63)$$

and

$$\|\mathbf{z}_{\bar{\mathcal{S}}}\|^2 - \min_{\tilde{\mathbf{v}}_{\bar{\mathcal{S}}} \in L\mathbb{T}^{|\bar{\mathcal{S}}|}} \|\mathbf{z}_{\bar{\mathcal{S}}} + \mathbf{u}_{\bar{\mathcal{S}}} + \tilde{\mathbf{v}}_{\bar{\mathcal{S}}}\|^2 \geq \|\mathbf{z}^*_{\bar{\mathcal{S}}}\|^2 - \min_{\mathbf{v}_{\bar{\mathcal{S}}} \in L\mathbb{T}^{|\bar{\mathcal{S}}|}} \|\mathbf{z}^*_{\bar{\mathcal{S}}} + \mathbf{u}_{\bar{\mathcal{S}}} + \mathbf{v}_{\bar{\mathcal{S}}}\|^2. \quad (64)$$

The first inequality (63) is true since for every index $s \in \mathcal{S}$, we have

$$\min_{\tilde{v}_s \in L\mathbb{T}} (z_s + u_s + \tilde{v}_s)^2 = \min_{v_s \in L\mathbb{T}} (z^*_s + u_s + v_s)^2$$

and since $z_s^2 \geq (z_s^*)^2$.

In order to see that (64) is true, we note that for any index $\bar{s}$ (and in particular $\bar{s} \in \bar{\mathcal{S}}$), we have

$$(z^*_{\bar{s}})^2 - \min_{v_{\bar{s}} \in L\mathbb{T}} (z^*_{\bar{s}} + u_{\bar{s}} + v_{\bar{s}})^2 \leq \left(\frac{L}{2}\right)^2.$$

On the other hand, for any $\bar{s} \in \bar{\mathcal{S}}$, the inequality $|z_{\bar{s}}| > L$ must hold, since otherwise the value of $v_{\bar{s}} \in L\mathbb{Z}$ minimizing $(z_{\bar{s}} + u_{\bar{s}} + v_{\bar{s}})^2$ cannot be greater than L. Moreover, since $|u_{\bar{s}}| \leq L/2$, it follows that

$$\min_{\tilde{v}_{\bar{s}} \in L\mathbb{T}} (z_{\bar{s}} + u_{\bar{s}} + \tilde{v}_{\bar{s}})^2 \leq \left(|z_{\bar{s}}| - \frac{L}{2}\right)^2.$$

Finally, for every $\bar{s} \in \bar{\mathcal{S}}$, we can write

$$\begin{aligned}(z_{\bar{s}})^2 - \min_{\tilde{v}_{\bar{s}} \in L\mathbb{T}} (z_{\bar{s}} + u_{\bar{s}} + \tilde{v}_{\bar{s}})^2 &\geq (z_{\bar{s}})^2 - \left(|z_{\bar{s}}| - \frac{L}{2}\right)^2 \\ &= \frac{L}{2}\left(2|z_{\bar{s}}| - \frac{L}{2}\right) > \left(\frac{L}{2}\right)^2 \\ &\geq (z^*_{\bar{s}})^2 - \min_{v_{\bar{s}} \in L\mathbb{T}} (z^*_{\bar{s}} + u_{\bar{s}} + v_{\bar{s}})^2\end{aligned}$$

which establishes (64). ■

Appendix II
Preliminary Lemmas

In this appendix, we state four lemmas that are repeatedly used in Appendix III for the derivations of error probabilities for the different cases of statistical dependences. The proofs are rather cumbersome and are given in Appendix IV.

Lemma 3: Let X be a random variable uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$, and let Θ be some random variable statistically independent of X. Define the random variable U by

$$U = [X + \Theta]^*.$$

The following inequality holds:

$$\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}\right] \leq \frac{1}{p} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}}.$$

Lemma 4: Let X be a random variable uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$, and Θ be some random variable statistically independent of X. Define

$$U = [(f + \varepsilon)X + \Theta]^*$$

where f is a constant in $\mathbb{Z}_p$, and $\varepsilon \in \mathbb{R}$. For any $f \in \mathbb{Z}_p$

$$\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}\right] \leq \frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\varepsilon)\mathsf{SNR}}}$$

where

$$\delta(p,\varepsilon) = \min_{l \in \mathbb{Z}_p \backslash \{0\}} l \cdot \left|\varepsilon - \frac{\lfloor l\varepsilon \rceil}{l}\right|. \quad (65)$$

Lemma 5: Let X_1, X_2, and X_3 be three i.i.d. random variables uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$, and $\gamma \in \mathbb{R}$ some arbitrary constant. Define

$$U = [X_1 + \gamma(X_2 - X_3)]^*. \quad (66)$$

Then

$$\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}\right] \leq \frac{1}{p^2} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\delta^2(p,\gamma)}$$

where $\delta(\cdot,\cdot)$ is defined in (65).

Lemma 6: Let X be a random variable uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$, and $\varepsilon \in \mathbb{R}$ some arbitrary constant. Define

$$U = [fX + \varepsilon X - \varepsilon[rX]^*]^* \quad (67)$$

where $f \in \mathbb{Z}_p \backslash \{0\}$ and $r \in \mathbb{Z}_p \backslash \{0,1\}$. Then

$$\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}\right] \leq \frac{p-1}{p} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\left(\varepsilon_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}\right)^2}.$$

We also state two properties that are extensively used (sometimes implicitly) throughout the calculations in Appendix III.

Property 1: For any $X \in \mathbb{R}$ and $a \in \mathbb{Z}$

$$\left[a[X]^*\right]^* = [aX]^*.$$

Property 2: For any $\alpha > 0$ and $\beta > 0$

$$[\alpha X]_{\mathrm{mod}[-\frac{\beta}{2},\frac{\beta}{2})} = \alpha[X]_{\mathrm{mod}[-\frac{\beta}{2\alpha},\frac{\beta}{2\alpha})}.$$

Appendix III
Derivation of the Error Probabilities for the Different Cases

In this appendix, we analyze the different cases of statistical dependences, and give upper bounds for the error probability in each one.

1) Case A: In this case, we have

$$U_A = [X_1 + \gamma X_2 - X_{\mathcal{E}1} - \gamma X_{\mathcal{E}2}]^*$$

where X_1, X_2, $X_{\mathcal{E}1}$, and $X_{\mathcal{E}2}$ are four i.i.d. random variables uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$. Let

$$\bar{X}_1 = [X_1 - X_{\mathcal{E}1}]^*$$

and note that $\bar{X}_1$ is uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$, and is statistically independent of X_2 and $X_{\mathcal{E}2}$. We have

$$U_A = \left[\bar{X}_1 + \gamma(X_2 - X_{\mathcal{E}2})\right]^*.$$

Applying Lemma 5 gives

$$\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}\right] \le \frac{1}{p^2} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\delta^2(p,\gamma)}. \tag{68}$$

Denote by Ω_A the value of Ω associated with case A. Substituting (68) into (26), we have

$$\Omega_A < \frac{1}{p^2} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\delta^2(p,\gamma)} + 2e^{-\frac{3\mathsf{SNR}}{8}}. \tag{69}$$

2) Case B: In this case

$$\begin{aligned}
U_B &= \Big[X_1 + \gamma X_2 - X_{\mathcal{E}1} - \gamma\left[aX_1 + bX_2 + cX_{\mathcal{E}1}\right]^*\Big]^* \\
&= \Big[X_1 + \lfloor\gamma\rceil X_2 + (\gamma - \lfloor\gamma\rceil)X_2 - X_{\mathcal{E}1} \\
&\quad - \lfloor\gamma\rceil\left[aX_1 + bX_2 + cX_{\mathcal{E}1}\right]^* \\
&\quad - (\gamma - \lfloor\gamma\rceil)\left[aX_1 + bX_2 + cX_{\mathcal{E}1}\right]^*\Big]^* \\
&= \Big[(1 - \lfloor\gamma\rceil a)X_1 + (\lfloor\gamma\rceil - \lfloor\gamma\rceil b)X_2 + (-1 - \lfloor\gamma\rceil c)X_{\mathcal{E}1} \\
&\quad + (\gamma - \lfloor\gamma\rceil)\left(X_2 - \left[aX_1 + bX_2 + cX_{\mathcal{E}1}\right]^*\right)\Big]^*
\end{aligned}$$

where X_1, X_2, and $X_{\mathcal{E}1}$ are three i.i.d. random variables uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$. Define

$$\begin{aligned}
d_1 &= [1 - \lfloor\gamma\rceil a]_{\mathrm{mod}\,p} \\
d_2 &= [\lfloor\gamma\rceil - \lfloor\gamma\rceil b]_{\mathrm{mod}\,p} \\
d_3 &= [-1 - \lfloor\gamma\rceil c]_{\mathrm{mod}\,p}
\end{aligned}$$

and

$$\varepsilon = \gamma - \lfloor\gamma\rceil \in [-\frac{1}{2}, \frac{1}{2}).$$

Using these notations, we have

$$\begin{aligned}
U_B = \Big[&[d_1X_1 + d_2X_2 + d_3X_{\mathcal{E}1}]^* \\
&+ \varepsilon\left(X_2 - [aX_1 + bX_2 + cX_{\mathcal{E}1}]^*\right)\Big]^*.
\end{aligned} \tag{70}$$

We now show that if the vectors $[d_1\ d_3]$ and $[a\ c]$ are linearly independent (over $\mathbb{Z}_p$), the first and the second elements in the sum are statistically independent. To that end, we prove the following lemma.

Lemma 7: Let A be a full-rank deterministic matrix with dimensions $M \times M$ over $\mathbb{Z}_p$ where p is a prime number, and $\mathbf{X}$ a vector with dimensions $M \times 1$ containing elements which are i.i.d. random variables uniformly distributed over $\mathbb{Z}_p$.

Let $\bar{\mathbf{X}} = A\mathbf{X}$ (with operations over $\mathbb{Z}_p$). Then, the elements of $\bar{\mathbf{X}}$ are also i.i.d. random variables uniformly distributed over $\mathbb{Z}_p$.

Proof: Since A is full rank, for any vector $\bar{\mathbf{x}}$, there exists one and only one vector $\mathbf{x}$ that satisfies $\bar{\mathbf{x}} = A\mathbf{x}$. Since the elements of $\mathbf{X}$ are i.i.d. and uniformly distributed, the vector $\mathbf{X}$ is uniformly distributed over $\mathbb{Z}_p^M$, which implies that

$$\Pr\left(\bar{\mathbf{X}} = \bar{\mathbf{x}}\right) = \Pr\left(\mathbf{X} = \mathbf{x}\right) = \left(\frac{1}{p}\right)^M$$

for every vector $\bar{\mathbf{x}} \in \mathbb{Z}_p^M$. This in turn implies that the elements of $\bar{\mathbf{X}}$ are i.i.d. with uniform distribution over $\mathbb{Z}_p$. ■

Let

$$\begin{pmatrix}\bar{X}_1\\ \bar{X}_2\\ \bar{X}_3\end{pmatrix} = \begin{pmatrix}d1 & d2 & d3\\ 0 & 1 & 0\\ a & b & c\end{pmatrix}\begin{pmatrix}X_1\\ X_2\\ X_{\mathcal{E}1}\end{pmatrix} \mod \mathcal{I}. \tag{71}$$

With this notation, (70) can be rewritten as

$$U_B = \left[\bar{X}_1 + \varepsilon(\bar{X}_2 - \bar{X}_3)\right]^*.$$

We now have to distinguish between the case where $[d_1\ d_3]$ and $[a\ c]$ are linearly independent which will be referred to as *Case B*1, and the case where they are linearly dependent. The case where $[d_1\ d_3]$ and $[a\ c]$ are linearly dependent and $[a\ c] \neq [0\ 0]$ will be called *Case B*2, and the case where $[a\ c] = [0\ 0]$ will be called *Case B*3.

Case B1: If $[d_1\ d_3]$ and $[a\ c]$ are linearly independent, the matrix in (71) is full rank, and it follows from Lemma 7 that $\left\{\bar{X}_1, \bar{X}_2, \bar{X}_3\right\}$ are each uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$ and are statistically independent. In this case

$$U_{B1} = \left[\bar{X}_1 + \Theta_{B1}\right]^*$$

where

$$\Theta_{B1} = \varepsilon(\bar{X}_2 - \bar{X}_3)$$

is statistically independent of $\bar{X}_1$.

Applying Lemma 3 gives

$$\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U_{B1}^2}\right] \le \frac{1}{p} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}}. \tag{72}$$

Denote by Ω_{B1} the value of Ω associated with case $B1$. Substituting (72) into (26), we have

$$\Omega_{B1} < \frac{1}{p} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}}. \tag{73}$$

Case B2: We now consider the case where $[d_1\ d_3]$ and $[a\ c]$ are linearly dependent and $[a\ c] \neq [0\ 0]$. In this case, we have

$$[d_1X_1 + d_3X_{\mathcal{E}1}]^* = [r\,(aX_1 + cX_{\mathcal{E}1})]^*$$

for some $r \in \mathbb{Z}_p$. Let

$$X_3 = [aX_1 + cX_{\mathcal{E}1}]^*$$

and note that X_3 is uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$ and is statistically independent of X_2. We can rewrite U_{B2} as

$$U_{B2} = \left[rX_3 + d_2X_2 + \varepsilon\left(X_2 - [X_3 + bX_2]^*\right)\right]^*.$$

Now let

$$X_4 = [X_3 + bX_2]^*$$

and note that X_4 is statistically independent of X_2. Using this notation, we have

$$\begin{aligned} U_{B2} &= [rX_4 + (d_2 - rb)\,X_2 + \varepsilon\,(X_2 - X_4)]^* \\ &= [(f+\varepsilon)\,X_2 + (r-\varepsilon)\,X_4]^* \end{aligned} \tag{74}$$

where

$$f = [d_2 - rb]_{\mathrm{mod}\,p}.$$

Let

$$\Theta_{B2} = (r-\varepsilon)\,X_4$$

and note that Θ_{B2} is statistically independent of X_2. Using Lemma 4, we have

$$\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U_{B2}^2}\right] \le \frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\varepsilon)\mathsf{SNR}}} \tag{75}$$

which means that

$$\Omega_{B2} < \frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\varepsilon)\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}}. \tag{76}$$

We further note that

$$\begin{aligned} \delta(p,\gamma) &= \min_{l\in\mathbb{Z}_p\setminus\{0\}} l\cdot\left|\gamma - \frac{\lfloor l\gamma\rceil}{l}\right| \\ &= \min_{l\in\mathbb{Z}_p\setminus\{0\}} l\cdot\left|\varepsilon - \frac{\lfloor l\varepsilon\rceil}{l}\right| = \delta(p,\varepsilon). \end{aligned}$$

Thus, (76) is equivalent to

$$\Omega_{B2} < \frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\gamma)\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}}. \tag{77}$$

Case B3: Since $a = c = 0$, it follows that $d_1 = 1$ and $d_3 = [-1]_{\mathrm{mod}\,p} = p-1$. Thus, (70) can be written as

$$U_{B3} = \left[X_1 + d_2X_2 + (p-1)X_{\mathcal{E}1} + \varepsilon\left(X_2 - [bX_2]^*\right)\right]^*.$$

Letting

$$\Theta_{B3} = \left[d_2X_2 + (p-1)X_{\mathcal{E}1} + \varepsilon\left(X_2 - [bX_2]^*\right)\right]^*$$

and using the fact that X_1 is uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$ and is statistically independent of Θ_{B3}, Lemma 3 can be applied, which yields

$$\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U_{B3}^2}\right] \le \frac{1}{p} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}}. \tag{78}$$

Denote by Ω_{B3} the value of Ω associated with case $B3$. Substituting (78) into (26) gives

$$\Omega_{B3} < \frac{1}{p} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}}. \tag{79}$$

Since $\delta^2(p,\gamma) < 1$, combining (73) with (77) and (79) yields

$$\Omega_B < \frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\varepsilon)\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}} \tag{80}$$

for all possible values of a, b, and c.

3) Case C: In this case

$$\begin{aligned} U &= \left[X_1 + \gamma X_2 - [aX_1 + bX_2 + cX_{\mathcal{E}2}]^* - \gamma X_{\mathcal{E}2}\right]^* \\ &= [(1-a)X_1 + (\gamma-b)X_2 + (-\gamma-c)X_{\mathcal{E}2}]^* \end{aligned} \tag{81}$$

where X_1, X_2, and $X_{\mathcal{E}2}$ are three i.i.d. random variables uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$.

We distinguish between the case where $a \neq 1$, which we refer to as *Case C*1, and the case where $a = 1$, which we refer to as *Case C*2.

Case C1: Since $a \neq 1$, the random variable $[(1-a)X_1]^*$ is uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$. We further define the random variable

$$\Theta_{C1} = (\gamma-b)X_2 + (-\gamma-c)X_{\mathcal{E}2}$$

which is statistically independent of $[(1-a)X_1]^*$. Applying Lemma 3 yields

$$\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U_{C1}^2}\right] \le \frac{1}{p} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}}. \tag{82}$$

Denote by Ω_{C1} the value of Ω associated with case $C1$. Substituting (82) into (26), we have

$$\Omega_{C1} < \frac{1}{p} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}}. \tag{83}$$

Case C2: Since $a = 1$, (81) can be rewritten as

$$\begin{aligned} U_{C2} &= [(\gamma-b)X_2 + (-\gamma-c)X_{\mathcal{E}2}]^* \\ &= [(\varepsilon + \lfloor\gamma\rceil - b)X_2 + (-\varepsilon - \lfloor\gamma\rceil - c)X_{\mathcal{E}2}]^* \\ &= [(f+\varepsilon)X_2 + (r-\varepsilon)X_{\mathcal{E}2}]^* \end{aligned} \tag{84}$$

where

$$\begin{aligned} \varepsilon &= \gamma - \lfloor\gamma\rceil \\ f &= [\lfloor\gamma\rceil - b]_{\mathrm{mod}\,p} \end{aligned}$$

and

$$r = [-\lfloor\gamma\rfloor - c]_{\mathrm{mod}p}\,.$$

Letting

$$\Theta_{C2} = (r-\varepsilon)X_{\mathcal{E}2}$$

and applying Lemma 4 gives

$$\Omega_{C2} < \frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\gamma)\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}}. \tag{85}$$

Combining (83) with (85), and using the fact that $\delta(p,\gamma) < 1$, yields

$$\Omega_{C} < \frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\gamma)\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}} \tag{86}$$

for all possible values of a, b, and c.

4) Case D: In this case, we have

$$\begin{aligned} U_D &= \left[X_1 + \gamma X_2 - aX_1 - bX_2 - \gamma\left[cX_1 + dX_2\right]^*\right]^* \\ &= \Big[(1-a-\lfloor\gamma\rfloor c)X_1 + (\lfloor\gamma\rfloor - b - \lfloor\gamma\rfloor d)X_2 \\ &\quad + \varepsilon\left(X_2 - [cX_1 + dX_2]^*\right)\Big]^* \end{aligned} \tag{87}$$

where X_1 and X_2 are two i.i.d. random variables uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$, and $\varepsilon = \gamma - \lfloor\gamma\rfloor$ as earlier. Further, letting

$$d_1 = [1 - a - \lfloor\gamma\rfloor c]_{\mathrm{mod}p}$$

and

$$d_2 = [\lfloor\gamma\rfloor - b - \lfloor\gamma\rfloor d]_{\mathrm{mod}p}$$

we have

$$U_D = \left[d_1X_1 + d_2X_2 + \varepsilon\left(X_2 - [cX_1 + dX_2]^*\right)\right]^*.$$

This is the most complicated case in terms of the number of different combinations of a, b, c, and d we have to consider. *Case D*1 corresponds to $c \neq 0$, *Case D*2 to $\{c = 0, a \neq 1\}$, *Case D*3 to $\{c = 0, a = 1, d = 1\}$, and, finally, the case where $\{c = 0, a = 1, d \geq 2\}$ will be referred to as *Case D*4. The case where $\{c = 0, a = 1, d = 0\}$ does not have to be considered as in this case $\mathbf{w}_{\mathcal{E}2} = 0$, and hence, the message vectors $\mathbf{w}_{\mathcal{E}1}$ and $\mathbf{w}_{\mathcal{E}2}$ are linearly dependent. As we recall, the decoder in our scheme does not consider such pairs of message vectors.

We denote by $\Omega_{Di},\ i = 1,\ldots,4$, the value of Ω associated with Case Di.

Case D1: Define

$$X_3 = [cX_1 + dX_2]^*$$

which is statistically independent of X_2 as $c \neq 0$. Now

$$\begin{aligned} U_{D1} &= [d_1X_1 + d_2X_2 + \varepsilon(X_2 - X_3)]^* \\ &= \left[(d_1c^{-1})cX_1 + d_2X_2 + \varepsilon(X_2 - X_3)\right]^* \\ &= \Big[(d_1c^{-1})(cX_1 + dX_2) \\ &\quad + (d_2 - d_1c^{-1}d)X_2 + \varepsilon(X_2 - X_3)\Big]^* \end{aligned} \tag{88}$$

where $c^{-1} \in \mathbb{Z}_p$ is the inverse element of c in the field $\mathbb{Z}_p$. Let

$$r = \left[d_1c^{-1}\right]_{\mathrm{mod}p}$$

and

$$f = \left[d_2 - d_1c^{-1}d\right]_{\mathrm{mod}p}.$$

With these notations, (88) can be written as

$$\begin{aligned} U_{D1} &= [rX_3 + fX_2 + \varepsilon(X_2 - X_3)]^* \\ &= [(f+\varepsilon)X_2 + (r-\varepsilon)X_3]^*. \end{aligned} \tag{89}$$

Letting

$$\Theta_{D1} = (r-\varepsilon)X_3$$

and applying Lemma 4 gives

$$\Omega_{D1} < \frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\gamma)\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}}. \tag{90}$$

Case D2: Since $c = 0$, (87) becomes

$$U_{D2} = \left[(1-a)X_1 + d_2X_2 + \varepsilon(X_2 - [dX_2]^*)\right]^*.$$

We define

$$\Theta_{D2} = d_2X_2 + \varepsilon(X_2 - [dX_2]^*)$$

which is statistically independent of $(1-a)X_1$. Since $a \neq 1$, the random variable $[(a-1)X_1]^*$ is uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$. We can therefore apply Lemma 3 and get

$$\Omega_{D2} < \frac{1}{p} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}}. \tag{91}$$

Case D3: Substituting $c = 0$, $a = 1$, and $d = 1$ into (87) gives

$$U_{D3} = [-bX_2]^*.$$

Applying Lemma 3 with $\Theta = 0$ gives

$$\Omega_{D3} < \frac{1}{p} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}} \tag{92}$$

for any $b \neq 0$. The case $b = 0$ is not interesting because it implies $\mathbf{w}_{\mathcal{E}1} = \mathbf{w}_1$ and $\mathbf{w}_{\mathcal{E}2} = \mathbf{w}_2$, and an error does not occur.

Case D4: We are left only with the case $a = 1, c = 0, d \geq 2$ for which

$$U_{D4} = \left[d_2X_2 + \varepsilon X_2 - \varepsilon\left[dX_2\right]^*\right]^*. \tag{93}$$

Since X_2 is uniformly distributed over $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$, $d_2 \in \mathbb{Z}_p\backslash\{0\}$, and $d \in \mathbb{Z}_p\backslash\{0,1\}$, we can apply Lemma 6 which gives

$$\Omega_{D4} < \frac{p-1}{p} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\left(\varepsilon_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}\right)^2} + 2e^{-\frac{3\mathsf{SNR}}{8}}. \tag{94}$$

Combining (90), (91), (92), (94), and the fact that $\delta(p,\gamma) < 1$, we conclude that for all values of a, b, c, and d that are consid-

ered by the decoder (except for $a = 1, b = 0, c = 0, d = 1$ which does not incur an error event), we have[7]

$$\Omega_D < \max\left\{\frac{1}{p} + \sqrt{\frac{2\pi/3}{\delta^2(p,\gamma)\mathsf{SNR}}} + 2e^{-\frac{3\mathsf{SNR}}{8}}, \right.$$
$$\left. \frac{p-1}{p} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\left(\gamma_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}\right)^2} + 2e^{-\frac{3\mathsf{SNR}}{8}}\right\}. \quad (95)$$

Appendix IV
Proofs of Preliminary Lemmas

The aim of this section is to prove Lemmas 3–6. We begin by deriving some auxiliary lemmas that will be used.

Lemma 8:

a) For any $\theta \in \left[-\frac{1}{2}, \frac{1}{2}\right)$ and $\rho > 0$

$$\sum_{k=-\infty}^{\infty} e^{-\rho(k+\theta)^2} < e^{-\rho\theta^2} + \int_{-\infty}^{\infty} e^{-\rho x^2} dx. \quad (96)$$

b) For any $\theta \in \mathbb{R}$ and $\rho > 0$

$$\sum_{k=-\infty}^{\infty} e^{-\rho(k+\theta)^2} < e^{-\rho(\theta-\lfloor\theta\rceil)^2} + \sqrt{\frac{\pi}{\rho}} \le 1 + \sqrt{\frac{\pi}{\rho}}. \quad (97)$$

Proof: In order to prove part (a) of the lemma, we write

$$\sum_{k=-\infty}^{\infty} e^{-\rho(k+\theta)^2} = e^{-\rho\theta^2} + \sum_{k=-\infty}^{-1} e^{-\rho(k+\theta)^2} + \sum_{k=1}^{\infty} e^{-\rho(k+\theta)^2}$$
$$= e^{-\rho\theta^2} + \sum_{k=1}^{\infty} e^{-\rho(k-\theta)^2} + \sum_{k=1}^{\infty} e^{-\rho(k+\theta)^2}. \quad (98)$$

We have

$$\int_{\theta}^{\infty} e^{-\rho x^2} dx > \sum_{k=1}^{\infty} \min_{x\in[k-1+\theta,k+\theta]} e^{-\rho x^2}$$
$$= \sum_{k=1}^{\infty} e^{-\rho(k+\theta)^2} \quad (99)$$

and

$$\int_{-\infty}^{\theta} e^{-\rho x^2} dx = \int_{-\theta}^{\infty} e^{-\rho x^2} dx > \sum_{k=1}^{\infty} \min_{x\in[k-1-\theta,k-\theta]} e^{-\rho x^2}$$
$$= \sum_{k=1}^{\infty} e^{-\rho(k-\theta)^2}. \quad (100)$$

[7]We also used the fact that $\gamma_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})} = \varepsilon_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}$.

Substituting (99) and (100) into (98) yields

$$\sum_{k=-\infty}^{\infty} e^{-\rho(k+\theta)^2} < e^{-\rho\theta^2} + \int_{-\infty}^{\theta} e^{-\rho x^2} dx + \int_{\theta}^{\infty} e^{-\rho x^2} dx$$
$$= e^{-\rho\theta^2} + \int_{-\infty}^{\infty} e^{-\rho x^2} dx \quad (101)$$

which establishes the first part of the lemma.

In order to prove part (b), we have

$$\sum_{k=-\infty}^{\infty} e^{-\rho(k+\theta)^2} = \sum_{k=-\infty}^{\infty} e^{-\rho(k+\lfloor\theta\rceil+(\theta-\lfloor\theta\rceil))^2}. \quad (102)$$

Letting $\tilde{\theta} = \theta - \lfloor\theta\rceil \in \left[-\frac{1}{2}, \frac{1}{2}\right)$ and $\tilde{k} = k + \lceil\theta\rfloor$, we have

$$\sum_{k=-\infty}^{\infty} e^{-\rho(k+\theta)^2} = \sum_{\tilde{k}=-\infty}^{\infty} e^{-\rho(\tilde{k}+\tilde{\theta})^2}$$
$$< e^{-\rho\tilde{\theta}^2} + \int_{-\infty}^{\infty} e^{-\rho x^2} dx$$
$$= e^{-\rho\tilde{\theta}^2} + \sqrt{\frac{\pi}{\rho}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\frac{1}{2\rho}}} e^{-\frac{1}{2}\frac{x^2}{1/2\rho}} dx$$
$$= e^{-\rho\tilde{\theta}^2} + \sqrt{\frac{\pi}{\rho}} \quad (103)$$

where we have used part (a) of the lemma for the first inequality. ■

Lemma 9: For any $\theta \in \mathbb{R}$ and a prime number p

$$[\mathbb{Z}_p + \theta]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})} \equiv \mathbb{Z}_p - \frac{p-1}{2} + \tilde{\theta} \quad (104)$$

where $\tilde{\theta} = \theta - \lfloor\theta\rceil \in \left[-\frac{1}{2}, \frac{1}{2}\right)$, and the notation $\equiv$ stands for equality between sets of points (constellations).

Proof:

$$[\mathbb{Z}_p + \theta]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})} \equiv [\mathbb{Z}_p + \lfloor\theta\rceil + \theta - \lfloor\theta\rceil]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}$$
$$\equiv \left[[\mathbb{Z}_p + \lfloor\theta\rceil]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})} + \tilde{\theta}\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}$$
$$\equiv \left[\mathbb{Z}_p - \frac{p-1}{2} + \tilde{\theta}\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}$$
$$\equiv \mathbb{Z}_p - \frac{p-1}{2} + \tilde{\theta}. \quad (105)$$

■

We are now ready to prove Lemma 3.

Proof of Lemma 3:

$$\begin{aligned}
\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}|\Theta=\theta\right]\right] \\
&\le \max_{\theta\in\mathbb{R}} \mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}|\Theta=\theta\right] \\
&= \max_{\theta\in\mathbb{R}} \frac{1}{p}\sum_{k=0}^{p-1} e^{-\frac{\mathsf{SNR}}{8}\left(\left[\frac{L}{p}k+\theta\right]_{\mathrm{mod}[-\frac{L}{2},\frac{L}{2})}\right)^2} \\
&= \max_{\theta\in\mathbb{R}} \frac{1}{p}\sum_{k=0}^{p-1} e^{-\frac{\mathsf{SNR}L^2}{8p^2}\left(\left[k+\frac{p}{L}\theta\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}\right)^2} \\
&\le \max_{\theta\in\mathbb{R}} \frac{1}{p}\sum_{\tilde{k}=-\frac{p-1}{2}}^{\frac{p-1}{2}} e^{-\frac{\mathsf{SNR}L^2}{8p^2}\left(\tilde{k}+\frac{p}{L}\theta-\lfloor\frac{p}{L}\theta\rceil\right)^2} \quad (106) \\
&\le \max_{\tilde{\theta}\in[-0.5,0.5)} \frac{1}{p}\sum_{\tilde{k}=-\infty}^{\infty} e^{-\frac{\mathsf{SNR}L^2}{8p^2}\left(\tilde{k}+\tilde{\theta}\right)^2} \\
&\le \frac{1}{p}\left(1+\sqrt{\frac{\pi\cdot 8p^2}{12\mathsf{SNR}}}\right) \quad (107) \\
&= \frac{1}{p}+\sqrt{\frac{2\pi/3}{\mathsf{SNR}}}
\end{aligned}$$

where (106) follows from Lemma 9, and (107) follows from part (b) of Lemma 8, and the fact that $L^2 = 12$. ■

Before proceeding to the proof of Lemma 4, we need to derive two more simple results.

Lemma 10: Let $\mathcal{D}$ be some constellation of finite cardinality $|\mathcal{D}|$, with minimum distance

$$d_{\min} = \min_{x_1,x_2\in\mathcal{D},x_1\neq x_2} |x_1 - x_2|.$$

Then

$$\sum_{x\in\mathcal{D}} e^{-\rho x^2} < \sum_{k=-\infty}^{\infty} e^{-\rho(k\cdot d_{\min}+\theta)^2} \quad (108)$$

for some $\theta \in \mathbb{R}$.

Proof: Let

$$d_0 = \arg\min_{x\in\mathcal{D}} |x|.$$

Each point $x \in \mathcal{D}$ in the constellation can be written in terms of its distance from d_0, i.e., as

$$x = d_0 + \Delta(x)$$

where $\Delta(x) = x - d_0$.

Let us sort the points of $\mathcal{D}$ in ascending order by

$$d_0-\Delta_{-M_{\mathrm{neg}}} \le \cdots \le d_0-\Delta_{-1} \le d_0 \le d_0+\Delta_1 \cdots \le d_0+\Delta_{M_{\mathrm{pos}}}$$

where the different Δ's correspond to the distance of each point from d_0. Note that there are M_{pos} points in $\mathcal{D}$ with magnitude greater than d_0 and M_{neg} points in $\mathcal{D}$ with magnitude smaller than d_0. Moreover, $\Delta_k \ge k d_{\min}$ for $k = 1,\ldots,M_{\mathrm{pos}}$ and $\Delta_{-k} \ge k d_{\min}$ for $k = 1,\ldots,M_{\mathrm{neg}}$.

We have

$$\begin{aligned}
\sum_{x\in\mathcal{D}} e^{-\rho x^2} &= \sum_{k=1}^{M_{\mathrm{neg}}} e^{-\rho(d_0-\Delta_{-k})^2} + e^{-\rho d_0^2} + \sum_{k=1}^{M_{\mathrm{pos}}} e^{-\rho(d_0+\Delta_k)^2} \\
&\le \sum_{k=1}^{M_{\mathrm{neg}}} e^{-\rho(d_0-k\cdot d_{\min})^2} + \sum_{k=0}^{M_{\mathrm{pos}}} e^{-\rho(d_0+k\cdot d_{\min})^2} \\
&\le \sum_{k=-\infty}^{\infty} e^{-\rho(d_0+k\cdot d_{\min})^2}.
\end{aligned}$$

Setting $\theta = d_0$, the lemma is proved. ■

Remark 1: We note that the bound (108) is rather loose, and is one of the weakest links in the chain of bounds we use for obtaining an upper bound on the average pairwise error probability $\mathbb{E}(P_{e,\mathrm{pair}})$.

Lemma 11: Let

$$X \equiv \left[\frac{L}{p}\mathbb{Z}_p\right]^* \equiv \frac{L}{p}\left[\mathbb{Z}_p\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}$$

and let $\mathcal{D} \equiv [(f+\varepsilon)X+\theta]^*$ where $f \in \mathbb{Z}_p$, and $\varepsilon, \theta \in \mathbb{R}$ are arbitrary constants. The minimum distance in the constellation $\mathcal{D}$ is lower bounded by

$$d_{\min} \ge \frac{L}{p}\min_{l\in\mathbb{Z}_p\setminus\{0\}} l\cdot\left|\varepsilon - \frac{\lfloor l\varepsilon\rceil}{l}\right|.$$

Proof: The distance between any pair of distinct constellation points can be written as

$$\begin{aligned}
&\left|\,[(f+\varepsilon)x_1+\theta]^* - [(f+\varepsilon)x_2+\theta]^*\,\right| \\
&\ge \left|\,\left[[(f+\varepsilon)x_1+\theta]^* - [(f+\varepsilon)x_2+\theta]^*\right]^*\,\right| \\
&= \left|\,[f(x_1-x_2)+\varepsilon(x_1-x_2)]^*\,\right| \quad (109)
\end{aligned}$$

where x_1 and x_2 are two distinct points in $\left[\frac{L}{p}\mathbb{Z}_p\right]^*$. Letting

$$\tilde{x}_1 = \frac{p}{L}x_1 \in [\mathbb{Z}_p]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}$$

and

$$\tilde{x}_2 = \frac{p}{L}x_2 \in [\mathbb{Z}_p]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}$$

we can further bound (109) as

$$\begin{aligned}
&\left| [f(x_1-x_2)+\varepsilon(x_1-x_2)]^* \right| \\
&= \frac{L}{p}\left| [f(\tilde{x}_1-\tilde{x}_2)+\varepsilon(\tilde{x}_1-\tilde{x}_2)]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})} \right| \\
&\geq \frac{L}{p}\left| [f(\tilde{x}_1-\tilde{x}_2)+\varepsilon(\tilde{x}_1-\tilde{x}_2)]_{\mathrm{mod}[-\frac{1}{2},\frac{1}{2})} \right| \\
&= \frac{L}{p}\left| [\varepsilon(\tilde{x}_1-\tilde{x}_2)]_{\mathrm{mod}[-\frac{1}{2},\frac{1}{2})} \right| \\
&\geq \frac{L}{p}\min_{l\in\mathbb{Z}_p\backslash\{0\}} \left| l\varepsilon - \lfloor l\varepsilon \rceil \right| \qquad (110)\\
&\geq \frac{L}{p}\min_{l\in\mathbb{Z}_p\backslash\{0\}} l\cdot\left|\varepsilon - \frac{\lfloor l\varepsilon\rceil}{l}\right|
\end{aligned}$$

where inequality (110) is true since $0 < |\tilde{x}_1-\tilde{x}_2| \leq p-1$. Recall that

$$\delta(p,\varepsilon) = \min_{l\in\mathbb{Z}_p\backslash\{0\}} l\cdot\left|\varepsilon - \frac{\lfloor l\varepsilon\rceil}{l}\right|$$

and hence

$$d_{\min} \geq \frac{L}{p}\,\delta(p,\varepsilon).$$

■

Aided by Lemmas 10 and 11, we can now prove Lemma 4.

Proof of Lemma 4:

$$\begin{aligned}
\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}|\Theta=\theta\right]\right] \\
&\leq \max_{\theta\in\mathbb{R}} \mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}|\Theta=\theta\right] \\
&= \max_{\theta\in\mathbb{R}} \frac{1}{p}\sum_{k=0}^{p-1} e^{-\frac{\mathsf{SNR}}{8}\left(\left[(f+\varepsilon)\frac{L}{p}k+\theta\right]_{\mathrm{mod}[-\frac{L}{2},\frac{L}{2})}\right)^2} \\
&< \max_{\tilde{\theta}\in\mathbb{R}} \frac{1}{p}\sum_{k=-\infty}^{\infty} e^{-\frac{\mathsf{SNR}}{8}\left(k\cdot d_{\min}+\tilde{\theta}\right)^2} \qquad (111)\\
&= \max_{\tilde{\theta}\in\mathbb{R}} \frac{1}{p}\sum_{k=-\infty}^{\infty} e^{-\frac{\mathsf{SNR}\cdot d_{\min}^2}{8}\left(k+\frac{\tilde{\theta}}{d_{\min}}\right)^2} \\
&< \frac{1}{p}\left(1+\sqrt{\frac{8\pi}{\mathsf{SNR}\cdot d_{\min}^2}}\right) \qquad (112)\\
&\leq \frac{1}{p}\left(1+\sqrt{\frac{8\pi p^2}{\mathsf{SNR}\cdot\delta^2(p,\varepsilon)L^2}}\right) \qquad (113)\\
&= \frac{1}{p}+\sqrt{\frac{2\pi/3}{\delta^2(p,\varepsilon)\mathsf{SNR}}}
\end{aligned}$$

where (111) follows from Lemma 10, (112) follows from part (b) of Lemma 8, and (113) from Lemma 11. ■

We use a similar technique for the proof of Lemma 5.

Proof of Lemma 5: Let

$$\tilde{X}_i = \frac{p}{L}X_i, \text{ for } i=1,2,3$$

such that $\tilde{X}_1$, $\tilde{X}_2$ and $\tilde{X}_3$ are three i.i.d. random variables uniformly distributed over $[\mathbb{Z}_p]_{\mathrm{mod}[-p/2,p/2)}$. With this notation, (66) can be written as

$$U = \frac{L}{p}\left[\tilde{X}_1+\gamma\left(\tilde{X}_2-\tilde{X}_3\right)\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}.$$

Further, let

$$\Theta = \tilde{X}_2-\tilde{X}_3$$

and note that for $\theta\in[0,\pm1,\ldots,\pm(p-1)]$, we have

$$\Pr(\Theta=\theta) = \frac{p-|\theta|}{p^2} \leq \frac{1}{p}. \qquad (114)$$

Now

$$\begin{aligned}
\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}|\Theta=\theta\right]\right] \\
&= \sum_{\theta=-(p-1)}^{p-1}\Pr(\Theta=\theta)\,\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}|\Theta=\theta\right]. \qquad (115)
\end{aligned}$$

For any value of θ, we have

$$\begin{aligned}
&\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}|\Theta=\theta\right] \\
&= \mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}\frac{L^2}{p^2}\left(\left[\tilde{X}_1+\gamma\theta\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}\right)^2}\right] \\
&= \frac{1}{p}\sum_{k=-\frac{p-1}{2}}^{\frac{p-1}{2}} e^{-\frac{3\mathsf{SNR}}{2p^2}(k+(\gamma\theta-\lfloor\gamma\theta\rceil))^2} \qquad (116)\\
&\leq \frac{1}{p}\left(e^{-\frac{3\mathsf{SNR}}{2p^2}(\gamma\theta-\lfloor\gamma\theta\rceil)^2}+\sqrt{\frac{p^2 2\pi/3}{\mathsf{SNR}}}\right) \qquad (117)\\
&= \sqrt{\frac{2\pi/3}{\mathsf{SNR}}}+\frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}(\gamma\theta-\lfloor\gamma\theta\rceil)^2} \qquad (118)
\end{aligned}$$

where (116) follows from Lemma 9 and (117) follows from part (b) of Lemma 8.

We note that for $\theta=0$, (118) becomes

$$\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}|\Theta=0\right] \leq \sqrt{\frac{2\pi/3}{\mathsf{SNR}}}+\frac{1}{p} \qquad (119)$$

and for any $\theta\neq0$

$$\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}|\Theta\neq0\right] \leq \sqrt{\frac{2\pi/3}{\mathsf{SNR}}}+\frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\delta^2(p,\gamma)} \qquad (120)$$

which follows directly from the definition of $\delta(p,\gamma)$ and the fact that $|\theta|\in\mathbb{Z}_p\backslash\{0\}$.

Substituting (114), (119), and (120) into (115) yields

$$\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}\right] \leq \frac{1}{p^2} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\delta^2(p,\gamma)} + \sqrt{\frac{2\pi/3}{\mathsf{SNR}}}.$$

■

Proof of Lemma 6: The first step in proving the lemma is showing that for any $r \in \mathbb{Z}_p\backslash\{0,1\}$, there exists at least one value of $x \in \left[\frac{L}{p}\mathbb{Z}_p\right]^*$ for which $|u| > \frac{L}{p}|\varepsilon_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}|$.

For any value of $x \in \left[\frac{L}{p}\mathbb{Z}_p\right]^*$, define $\tilde{x} = \frac{p}{L}x \in [\mathbb{Z}_p]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}$. We have

$$\begin{aligned}|u| &= \left|[fx + \varepsilon x - \varepsilon[rx]^*]^*\right| \\ &= \frac{L}{p}\left|\left[f\tilde{x} + \varepsilon\tilde{x} - \varepsilon\left[r\tilde{x}\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}\right| \\ &\geq \frac{L}{p}\left|\left[f\tilde{x} + \varepsilon\tilde{x} - \varepsilon\left[r\tilde{x}\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}\right]_{\mathrm{mod}[-\frac{1}{2},\frac{1}{2})}\right| \\ &= \frac{L}{p}\left|\left[\varepsilon\left(\tilde{x} - \left[r\tilde{x}\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}\right)\right]_{\mathrm{mod}[-\frac{1}{2},\frac{1}{2})}\right|. \quad (121)\end{aligned}$$

We focus on the expression

$$\tilde{x} - \left[r\tilde{x}\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}. \quad (122)$$

We show that for any value of $2 \leq r < p-1$, there exists a value of $\tilde{x} \in [\mathbb{Z}_p]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}$ for which (122) equals 1. In order to see that, we observe that the following equation

$$\begin{aligned}&\left[\tilde{x} - \left[r\tilde{x}\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})}\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})} \\ &\quad = \left[(1-r)\tilde{x}\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})} = 1 \quad (123)\end{aligned}$$

has a (single) solution for any $r \in \mathbb{Z}_p\backslash\{1\}$. Further, since

$$\tilde{x} - \left[r\tilde{x}\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})} \in [-(p-1),(p-1)] \quad (124)$$

it can be deduced that for any $r \in \mathbb{Z}_p\backslash\{1\}$, there exists a (single) value of $\tilde{x}$ for which (122) equals either 1 or $-(p-1)$. Since (122) equals $-(p-1)$ only if

$$\tilde{x} = -\left[r\tilde{x}\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})} = -\frac{p-1}{2}$$

which is possible only for $r = p-1$, we conclude that for any $2 \leq r < p-1$, there is a value of x, which we denote x^1, for which (122) equals 1. Substituting x^1 into (121) yields

$$|u| \geq \frac{L}{p}\left|\varepsilon_{\mathrm{mod}[-\frac{1}{2},\frac{1}{2})}\right| \geq \frac{L}{p}\left|\varepsilon_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}\right|. \quad (125)$$

We are left with the case $r = p-1$, for which

$$\tilde{x} - \left[r\tilde{x}\right]_{\mathrm{mod}[-\frac{p}{2},\frac{p}{2})} = \tilde{x} - [-\tilde{x}] = 2\tilde{x}. \quad (126)$$

Substituting (126) into (121) gives

$$|u| \geq \frac{L}{p}\left|[2\tilde{x}\varepsilon]_{\mathrm{mod}[-\frac{1}{2},\frac{1}{2})}\right|.$$

It follows that for $r = p-1$ and $\tilde{x} = 1$

$$|u| \geq \frac{L}{p}\left|[2\varepsilon]_{\mathrm{mod}[-\frac{1}{2},\frac{1}{2})}\right| \geq \frac{L}{p}\left|\varepsilon_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}\right|. \quad (127)$$

Combining (125) and (127), we conclude that for any $r \in \mathbb{Z}_p\backslash\{0,1\}$, there exists at least one value of x for which

$$|u| \geq \frac{L}{p}\left|\varepsilon_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}\right|. \quad (128)$$

Now, since at least one of the equiprobable p possible values of U is bigger (in absolute value) than $\frac{L}{p}\left|\varepsilon_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}\right|$, $\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}\right]$ can be upper bounded by

$$\begin{aligned}\mathbb{E}\left[e^{-\frac{\mathsf{SNR}}{8}U^2}\right] &= \frac{1}{p}\sum_u e^{-\frac{\mathsf{SNR}}{8}U^2} \\ &\leq \frac{p-1}{p} + \frac{1}{p}e^{-\frac{L^2\mathsf{SNR}}{8p^2}\left(\varepsilon_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}\right)^2} \\ &= \frac{p-1}{p} + \frac{1}{p}e^{-\frac{3\mathsf{SNR}}{2p^2}\left(\varepsilon_{\mathrm{mod}[-\frac{1}{4},\frac{1}{4})}\right)^2}. \quad (129)\end{aligned}$$

■

ACKNOWLEDGMENT

The authors would like to express their deep gratitude to Ayal Hitron and Ronen Dar for their helpful technical comments.

REFERENCES

[1] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 5534–5562, Dec. 2008.

[2] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 1, pp. 49–60, Jan. 1981.

[3] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[4] B. Nazer, M. Gastpar, S. Jafar, and S. Vishwanath, "Ergodic interference alignment," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6355–6371, Oct. 2012.

[5] R. Etkin and E. Ordentlich, "The degrees-of-freedom of the K-user Gaussian interference channel is discontinuous at rational channel coefficients," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4932–4946, Nov. 2009, submitted for publication.

[6] A. S. Motahari, S. O. Gharan, M. Maddah-Ali, and A. K. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," IEEE Trans. Inf. Theory 2009, submitted for publication [Online]. Available: http://arxiv.org/abs/0908.2282

[7] S. A. Jafar, "Interference alignment—A new look at signal dimensions in a communication network," in *Foundations and Trends in Communications and Information Theory*. Norwell, MA: Now Publishers, 2011, vol. 7, pp. 1–134.

[8] M. Maddah-Ali, A. Motahari, and A. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.

[9] S. Jafar and S. Shamai, "Degrees of freedom region of the MIMO X channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.

[10] T. Gou and S. Jafar, "Degrees of freedom of the K user $M \times N$ MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6040–6057, Dec. 2010.

[11] C. Yetis, T. Gou, S. Jafar, and A. Kayran, "On feasibility of interference alignment in MIMO interference networks," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4771–4782, Sep. 2010.
[12] Y. Wu, S. Shamai (Shitz), and S. Verdú, "Degrees of freedom of the interference channel: A general formula," in *Proc. IEEE Int. Symp. Inf. Theory*, St. Petersburg, Russia, Aug. , pp. 1362–1366, submitted for publication.
[13] T. Philosof, A. Khisti, U. Erez, and R. Zamir, "Lattice strategies for the dirty multiple access channel," in *Proc. Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 386–390.
[14] M. P. Wilson, K. Narayanan, H. Pfister, and A. Sprintson, "Joint physical layer coding and network coding for bidirectional relaying," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5641–5645, Nov. 2010.
[15] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.
[16] G. Bresler, A. Parekh, and D. N. C. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," in *Proc. 42nd Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Sep. 2007, pp. 791–801.
[17] G. Bresler, A. Parekh, and D. N. C. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4566–4592, Sep. 2010.
[18] S. Sridharan, A. Jafarian, S. Vishwanath, and S. A. Jafar, "Capacity of symmetric K-user Gaussian very strong interference channels," in *Proc. IEEE Global Telecommun. Conf.*, New Orleans, LA, Dec. 2008, pp. 1–5.
[19] A. B. Carleial, "A case where interference does not reduce capacity," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 5, pp. 569–570, Sep. 1975.
[20] S. Sridharan, A. Jafarian, S. Vishwanath, S. A. Jafar, and S. Shamai (Shitz), "A layered lattice coding scheme for a class of three user Gaussian interference channels," in *Proc. 46nd Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Sep. 2008, pp. 531–538.
[21] O. Ordentlich, U. Erez, and B. Nazer, "The approximate sum capacity of the symmetric Gaussian K-user interference channel," IEEE Trans. Inf. Theory May 2012, submitted for publication [Online]. Available: http://arxiv.org/abs/1206.0197
[22] U. Niesen and M. A. Maddah-Ali, "Interference alignment: From degrees-of-freedom to constant-gap capacity approximations," IEEE Trans. Inf. Theory Dec. 2011, submitted for publication [Online]. Available: http://arxiv.org/abs/1112.4879
[23] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd Int. Symp. Inf. Theory*, Tsaghkadzor, Armenia, 1971, pp. 23–52.
[24] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. New York: Springer-Verlag, 1988.
[25] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[26] A. Jafarian, J. Jose, and S. Vishwanath, "Algebraic lattice alignment for k-user interference channels," in *Proc. 47th Annu. Allerton Conf. Commun., Control, Comput.*, Oct. 2009, pp. 88–93.
[27] A. Jafarian and S. Vishwanath, "Achievable rates for K-user Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4367–4380, Jul. 2012.
[28] A. Host-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," in *Proc. Int. Symp. Inf. Theory*, Adelaide, SA, Australia, Sep. 2005, pp. 2065–2069.
[29] W. M. Schmidt, *Diophantine Approximation*. New York: Springer-Verlag, 1980.
[30] A. Motahari, A. Khandani, and S. Gharan, "On the degrees of freedom of the 3-user Gaussian interference channel: The symmetric case," in *Proc. IEEE Int. Symp. Inf. Theory*, 2009, pp. 1914–1918.
[31] V. Cadambe, S. Jafar, and C. Wang, "Interference alignment with asymmetric complex signaling—Settling the Høst-Madsen–Nosratinia conjecture," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4552–4565, Sep. 2010.

Or Ordentlich received the B.Sc. degree (cum laude) and the M.Sc. degree (summa cum laude) in 2010 and 2011, respectively, in electrical engineering from Tel Aviv University, Israel. He is currently working toward the Ph.D. degree at Tel Aviv University.

Or is the recipient the Adams Fellowship awarded by the Israel Academy of Sciences and Humanities, the Advanced Communication Center (ACC) Feder Family award for outstanding research work in the field of communication technologies (2011), and the Weinstein Prize for research in signal processing (2011).

Uri Erez (M'09) was born in Tel-Aviv, Israel, on October 27, 1971. He received the B.Sc. degree in mathematics and physics and the M.Sc. and Ph.D. degrees in electrical engineering from Tel-Aviv University in 1996, 1999, and 2003, respectively. During 2003–2004, he was a Postdoctoral Associate at the Signals, Information and Algorithms Laboratory at the Massachusetts Institute of Technology (MIT), Cambridge. Since 2005, he has been with the Department of Electrical Engineering-Systems at Tel-Aviv University. His research interests are in the general areas of information theory and digital communication. He served in the years 2009–2011 as Associate Editor for Coding Techniques for the IEEE TRANSACTIONS ON INFORMATION THEORY.