

# The Approximate Sum Capacity of the Symmetric Gaussian $K$ -User Interference Channel

Or Ordentlich, Uri Erez, *Member, IEEE*, and Bobak Nazer, *Member, IEEE*

**Abstract**—Interference alignment has emerged as a powerful tool in the analysis of multiuser networks. Despite considerable recent progress, the capacity region of the Gaussian  $K$ -user interference channel is still unknown in general, in part due to the challenges associated with alignment on the signal scale using lattice codes. This paper develops a new framework for lattice interference alignment, based on the compute-and-forward approach. Within this framework, each receiver decodes by first recovering two or more linear combinations of the transmitted codewords with integer-valued coefficients and then solving these linear combinations for its desired codeword. For the special case of symmetric channel gains, this framework is used to derive the approximate sum capacity of the Gaussian interference channel, up to an explicitly defined outage set of the channel gains. The key contributions are the capacity lower bounds for the weak through strong interference regimes, where each receiver should jointly decode its own codeword along with part of the interfering codewords. As part of the analysis, it is shown that decoding  $K$  linear combinations of the codewords can approach the sum capacity of the  $K$ -user Gaussian multiple-access channel up to a gap of no more than  $K/2 \log K$  bits.

**Index Terms**—Interference channels, multiple access, lattice codes, interference alignment.

## I. INTRODUCTION

**H**ANDLING interference efficiently is a major challenge in multi-user wireless communication. Recently, it has become clear that this challenge can sometimes be overcome via *interference alignment* [1], [2]. For instance, consider the  $K$ -user Gaussian interference channel, where  $K$  transmitter-receiver pairs wish to communicate simultaneously. Through the use of clever encoding strategies, it is possible to align the transmitted signals so that each receiver only observes its desired signal along with a single effective interferer. As a result, each user can achieve roughly half the rate that would be available were there no interference whatsoever, i.e.,  $K/2$  degrees-of-freedom (DoF) are available.

Manuscript received June 1, 2012; revised August 30, 2013; accepted November 30, 2013. Date of publication April 11, 2014; date of current version May 15, 2014. O. Ordentlich was supported in part by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities, in part by a Fellowship through the Yitzhak and Chaya Weinstein Research Institute for Signal Processing, Tel Aviv University, and in part by the Feder Family Award. U. Erez was supported by the Israel Science Foundation under Grant 1557/10. B. Nazer was supported by the National Science Foundation under Grant CCF-1253918.

O. Ordentlich and U. Erez are with Tel Aviv University, Tel Aviv 69978, Israel (e-mail: ordent@eng.tau.ac.il; uri@eng.tau.ac.il).

B. Nazer is with the Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215 USA (e-mail: bobak@bu.edu).

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.2014.2316136

However, many schemes, such as the Cadambe-Jafar framework [2] and ergodic interference alignment [3], require a large number of independent channel realizations to achieve near-perfect alignment. In certain settings, this level of channel diversity may not be attainable; ideally, we would like to achieve alignment over a single channel realization.

The capacity region of the (static) Gaussian  $K$ -user interference channel [4] is unknown in general, although significant progress has been made recently, in part due to the discovery of interference alignment and the shift from exact capacity results to capacity approximations [5]–[7]. It has been shown by Motahari *et al.* that  $K/2$  DoF are achievable for almost all channel realizations [8] but it is an open question as to whether this result translates to real gains outside of the very high signal-to-noise ratio (SNR) regime. One promising direction is the use of lattice codes [9]–[11], as they can enable alignment on the signal scale. By taking advantage of the fact that the sum of lattice codewords is itself a lattice codeword, a receiver can treat several users as one effective user, thereby reducing the number of effective interferers. A compelling example of this approach is the derivation of the approximate capacity of the many-to-one interference channel by Bresler, Parekh, and Tse [7]. For fully connected channels, much less is known, owing to the difficulty of choosing lattices that simultaneously align at several receivers.

In some cases, focusing on the special case of symmetric channel gains has yielded important insights. For instance, in the two-user case, Etkin, Tse, and Wang [5] used the symmetric interference channel to develop the notion of generalized degrees-of-freedom. This in turn revealed five operating regimes, based on relative interference strength:

- *Noisy*: Each receiver treats interference as noise, which is optimal for sufficiently weak interference [12]–[14].
  - *Weak and Moderately Weak*: Each transmitter sends a public and a private codeword following the scheme of Han and Kobayashi [15]. Each receiver jointly decodes both public codewords and its desired private codeword while treating the interfering private codeword as noise.
  - *Strong*: Each receiver jointly decodes both users' codewords. This regime and its capacity was discovered by Sato [16] as well as Han and Kobayashi [15].
  - *Very Strong*: Each receiver decodes and subtracts the interference before recovering its desired codeword. This regime and its capacity was discovered by Carleial [17].
- Using these regimes as a guideline, they were able to approximate the capacity region to within half a bit per real channel dimension.

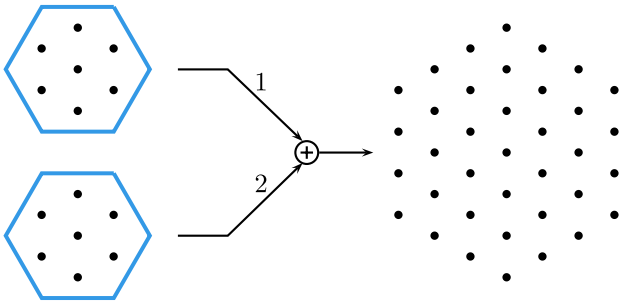


Fig. 1. Two transmitters employ the same 7-symbol lattice code over the channel  $\mathbf{x}_1 + 2\mathbf{x}_2$ . The effective constellation seen by the receiver contains only 37 points, which means that the receiver cannot always uniquely identify which pair of symbols was transmitted.

In this paper, we focus on the special case of the symmetric (real) Gaussian  $K$ -user interference channel. Each receiver observes

$$\mathbf{y}_k = \mathbf{x}_k + g \sum_{\ell \neq k} \mathbf{x}_\ell + \mathbf{z}_k \quad (1)$$

where  $\mathbf{x}_k$  is the codeword sent by the  $k$ th transmitter,  $g$  is the cross-channel gain, and  $\mathbf{z}_k$  is additive white Gaussian noise. Building on the compute-and-forward strategy [18], we propose a framework for lattice-based interference alignment whose performance can be evaluated both numerically and analytically at any SNR. Within our framework, each receiver first decodes integer linear combinations of the codewords and only afterwards solves these for its desired codeword. As we will argue, this choice of receiver architecture allows us to circumvent some of the difficulties encountered in the analysis of a direct decoding strategy. Below, we summarize the main technical contributions of the paper in the context of prior work.

### A. Paper Overview

One of the appealing properties of the symmetric Gaussian interference channel is that, if each transmitter draws its codeword  $\mathbf{x}_k$  from the same lattice codebook, the sum of the  $K - 1$  interfering codewords at each receiver  $\sum_{\ell \neq k} \mathbf{x}_\ell$  will align into a single effective codeword. This is due to the fact that lattices are closed under addition, i.e., the sum of any lattice codewords is itself a lattice codeword. The difficulty is that, depending on the value of the cross-channel gain  $g$ , the desired codeword may also align with the interference, since it is drawn from the same lattice codebook. The achievable rate is thus closely linked to the behavior of signal scale alignment, which makes this channel an ideal setting to gain a deeper understanding of this phenomenon at finite SNR. In Section II-A, we provide a formal problem statement.

When  $|g|$  is sufficiently large, it is easy for the receiver to distinguish its desired codeword from the aligned interfering codewords. Specifically, in the very strong regime ( $|g| > \sqrt{\text{SNR}}$ ), the sum of the interfering codewords acts as the cloud center from a classical superposition codebook [19] and the desired codebook acts as the cloud. Thus, as proposed by Sridharan *et al.* [20], the receiver can employ a successive cancellation strategy: first decode the sum of the interference  $\sum_{\ell \neq k} \mathbf{x}_\ell$ , then subtract it from its channel observation  $\mathbf{y}_k$ , and

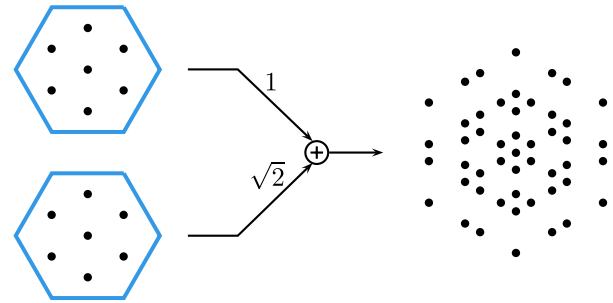


Fig. 2. Two transmitters employ the same 7-symbol lattice code over the channel  $\mathbf{x}_1 + \sqrt{2}\mathbf{x}_2$ . The effective constellation seen by the receiver consists of 49 points, which enables the receiver to determine which pair of symbols was transmitted.

finally decode  $\mathbf{x}_k$  from the resulting interference-free effective channel. We review this approach within the context of our framework in Section VI-A.

As the magnitude of  $g$  decreases below  $\sqrt{\text{SNR}}$ , the codebooks corresponding to the desired codeword and the aligned interference will start to overlap from the receiver's perspective. For certain values of  $g$ ,  $\mathbf{x}_k$  and  $\sum_{\ell \neq k} \mathbf{x}_\ell$  will align, which in turn significantly reduces the achievable rates. For example, in Figure 1, we illustrate the effective codebook corresponding to the linear combination  $\mathbf{x}_1 + 2\mathbf{x}_2$  where  $\mathbf{x}_1$  and  $\mathbf{x}_2$  are drawn from the same lattice codebook. There are only 37 points in this effective codebook, meaning that it is not always possible to uniquely determine which of the 49 possible pairs of codewords was transmitted, regardless of the SNR. However, for the linear combination  $\mathbf{x}_1 + \sqrt{2}\mathbf{x}_2$  shown in Figure 2, there are 49 points in the effective codebook, each corresponding to a unique codeword pair, even though the interference strength has decreased.

Thus, while employing the same lattice codebook at each transmitter aligns the interference at every receiver, it sometimes has the unintended effect of aligning the desired signal as well. When this occurs, the rate must be reduced until the desired codewords can be uniquely identified. We now summarize several recent papers that have aimed to quantify this effect. Etkin and Ordentlich [21] showed that, for the Gaussian  $K$ -user interference channel, the DoF is strictly less than  $K/2$  if all channel gains are rational. They also demonstrated, using a scalar lattice codebook, that if the diagonal elements are irrational algebraic numbers and the off-diagonals are rational,  $K/2$  DoF is achievable. Subsequently, Motahari *et al.* [8] proposed the “real interference alignment” framework. In particular, they argued that scalar lattice codewords can be uniquely identified from a linear combination (in the high SNR limit) provided that the coefficients are rationally independent.<sup>1</sup> Using this framework, they demonstrated that, for the Gaussian  $K$ -user interference channel,  $K/2$  DoF is achievable for almost all channel matrices by embedding the asymptotic alignment framework of [2] into a single dimension. This result was generalized by Wu *et al.* using Rényi's information dimension [22].

<sup>1</sup>The coefficients  $h_1, \dots, h_K \in \mathbb{R}$  are said to be *rationally independent* if there is no non-trivial choice of integers  $q_1, \dots, q_K$  such that  $q_1 h_1 + \dots + q_K h_K = 0$ .

For finite SNRs, [23] derived lower bounds on the achievable symmetric rate for a two-user multiple-access channel  $\mathbf{x}_1 + g\mathbf{x}_2 + \mathbf{z}$  where each user employs the same linear code over  $\mathbb{Z}_p$  for some prime  $p$ . The sensitivity of the bounds to the rationality of  $g$  at different SNRs was investigated, and the bounds were used to obtain achievable rate regions for Gaussian  $K$ -user interference channels with integer-valued off-diagonal channel gains. For the two-user Gaussian X channel,<sup>2</sup> Niesen and Maddah-Ali [24] approximated the sum capacity via an “outage set” characterization. Their coding scheme is guided by a variation on the deterministic model [6] and consists of a scalar lattice constellation combined with a random i.i.d. outer code. From one perspective, for any  $c > 0$ , their scheme approximates the sum capacity to within a constant gap of  $c + 66$  bits up to an outage set of channel matrices of measure roughly  $2^{-c/2}$ .

1) *Novel Coding Strategies*: The prior work described above attempts to directly bound the minimum distance in the effective codebook that results from the linear combination of the transmitters’ lattice codebooks. This is a challenging task, even for scalar lattices, and limits the analytical and numerical results to relatively high SNRs. In this paper, we take an alternative approach: we lower bound the achievable rate by the rate required to decode enough integer linear combinations to reveal the desired messages. For instance, in the strong regime  $1 \leq g \leq \sqrt{\text{SNR}}$ , each receiver *first decodes two linear combinations* of the form

$$a_{11}\mathbf{x}_k + a_{12} \sum_{\ell \neq k} \mathbf{x}_\ell \quad a_{21}\mathbf{x}_k + a_{22} \sum_{\ell \neq k} \mathbf{x}_\ell$$

where  $a_{11}, a_{12}, a_{21}$ , and  $a_{22}$  are integer-valued coefficients. If the vectors  $\mathbf{a}_1 = [a_{11} \ a_{12}]^T$  and  $\mathbf{a}_2 = [a_{21} \ a_{22}]^T$  are linearly independent, then each receiver can solve for its desired codeword  $\mathbf{x}_k$ . The rates at which these linear combinations can be decoded can be determined directly via the compute-and-forward framework [18], which we review in Section III. Since this framework employs high-dimensional nested lattice codes that can approach the point-to-point AWGN capacity, we can obtain analytical and numerical results for any finite SNR.

In Figure 3, we have plotted the symmetric rate of this scheme (Theorem 9) at SNRs 15 and 25dB with respect to the cross-gain  $g$  for the symmetric Gaussian 3-user interference channel. Alongside, we have plotted the symmetric rate for the two-user upper bound described in Section V, i.e., the rate that would be achievable if each receiver only encountered a single interferer. At 15dB, it is clear that the desired codeword aligns with the interference only at integer-valued cross-gains. At 25dB, alignment also occurs at  $g = 3/2, 5/2$ , and  $7/2$ , i.e., rationals with denominator 2. In other words, the number of channel gains where the rate saturates depends on the SNR.

We also propose a lattice version of the Han-Kobayashi scheme [15] for the weak and moderately weak regimes: each transmitter splits its information into a public lattice codeword  $\mathbf{x}_{k1}$  and a private lattice codeword  $\mathbf{x}_{k2}$ . Each receiver recovers

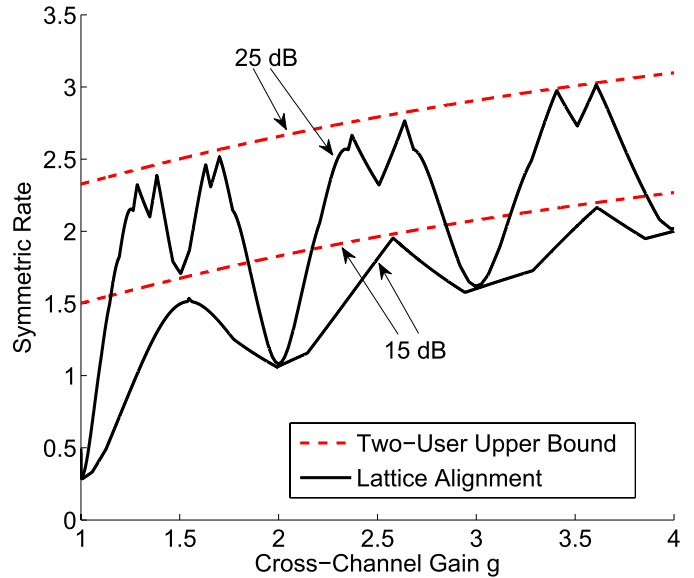


Fig. 3. Achievable symmetric rate for the symmetric Gaussian 3-user interference channel from Theorem 9.

its desired information by first decoding three linear combinations of the form

$$a_{m1}\mathbf{x}_{k1} + a_{m2}\mathbf{x}_{k2} + a_{m3} \sum_{\ell \neq k} \mathbf{x}_{\ell 1} \quad m = 1, 2, 3$$

for integer-valued coefficients  $a_{m1}, a_{m2}, a_{m3}$  that suffice to solve for the desired public codeword  $\mathbf{x}_{k1}$ , the desired private codeword  $\mathbf{x}_{k2}$ , and the sum of the public interfering codewords  $\sum_{\ell \neq k} \mathbf{x}_{\ell 1}$ . (The private interfering codewords are treated as noise.)

Within the standard compute-and-forward framework, the rate of each codeword should be set according to the lowest computation rate across all desired linear combinations. In Section IV, we propose an *algebraic successive cancellation* decoding strategy that can achieve higher rates. Consider a single receiver that decodes  $K$  linearly independent combinations of  $K$  lattice codewords in a given order. Each linear combination is associated with a certain computation rate, which we set as the rate of one of the codewords. After decoding each linear combination, the receiver can cancel out the effect of one codeword from its channel observation to reduce the effective rate. As we show in Theorem 11, for the lattice Han-Kobayashi scheme, this allows each user to attain the sum of the second and third highest computation rates (as opposed to twice the third highest).

Overall, these two lattice strategies, when combined with successive cancellation for the very strong regime and treating interference as noise for the noisy regime, yield an achievable rate region for the symmetric Gaussian  $K$ -user interference channel. To evaluate this rate region, we only need to optimize over the integer coefficients of the linear combinations. See Section III-D for a discussion on how the space of integer-coefficients can be explored numerically. In Figure 7, we have plotted the resulting lower bound on the symmetric capacity along with the two-user upper bound from Section V.

<sup>2</sup>In the X channel scenario, each transmitter has an independent message for each receiver.

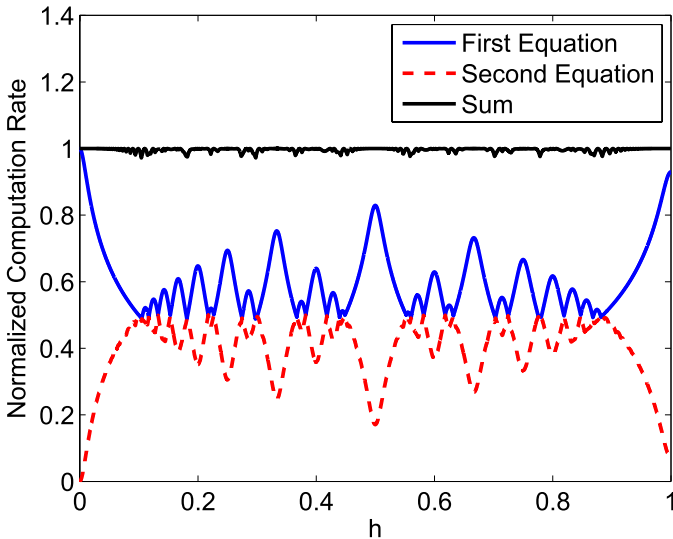


Fig. 4. Computation rates for the best two linearly independent integer linear combinations vs.  $h$  for the channel  $\mathbf{y} = \mathbf{x}_1 + h\mathbf{x}_2 + \mathbf{z}$  at SNR=40dB. The sum of these computation rates is nearly equal to the multiple-access sum capacity. All rates are normalized by this sum capacity  $1/2 \log(1 + (1 + h^2)\text{SNR})$ .

2) *Analytical Bounds:* We also develop new tools for deriving closed-form lower bounds for the rate achievable via lattice alignment. These tools and specifically the compute-and-forward transform, derived in Section IV-A, may be of independent interest. Consider again  $K$  transmitted codewords and a receiver that decodes  $K$  linear combinations according to the  $K$  highest computation rates with linearly independent coefficient vectors. While the computation rate for each of these  $K$  combinations is very sensitive to the exact values of the channel gains, the *sum of the computation rates is equal to the multiple-access sum capacity up to a constant gap* that is independent of the channel gains and the SNR as we show in Theorem 3. See Figure 4 for a plot of this behavior for  $K = 2$ . That is, lattice-based multiple-access can operate near the boundary of the capacity region. We also argue in Section VII that the degrees-of-freedom associated to each of these  $K$  linear combinations is  $1/K$  for almost all channel gains.

Interestingly, this sum capacity lower bound is very helpful in deriving closed-form lower bounds. For instance, in the strong regime, each user attains the rate associated with the second best linear combination. Thus, to bound the achievable rate, we should obtain an upper bound on the second best integer approximation of the real-valued channel gain  $g$ . Instead of attacking this problem directly, we instead develop a lower bound on the best integer approximation of  $g$  and combine this with the sum capacity lower bound to obtain our upper bound. More details are given in Section VI-B and similar bounds are developed for the moderately weak and weak regime in Sections VI-C and VI-D, respectively.

When compared with the two-user upper bound, these lower bounds yield an approximation of the sum capacity in all regimes that we summarize in Theorem 1. As in [24], our approximation is stated in terms of outage sets, i.e., for a given constant gap, we exclude a certain measure of channel gains. This outage set can be understood in terms of the quality of

the best integer approximation of  $g$ , and is characterized as part of the analysis in Section VI.

### B. Related Work

Interference alignment has generated a great deal of excitement, due to the promise of higher throughputs in wireless networks [1], [2] as well as other applications, including coding for distributed storage [25]. See the recent monograph by Jafar for a comprehensive survey [26]. Of particular note is a series of recent papers that delineate the degrees-of-freedom limits of linear beamforming strategies for alignment over a finite number of channel realizations [27], [28]. Beamforming strategies can only approach perfect alignment asymptotically, whereas lattice-based schemes can achieve  $K/2$  degrees-of-freedom over a single channel realization [8]. However, lattice-based alignment at finite SNR has to date been limited to special cases, such as symmetric [20], [23], [29], integer [30], and many-to-one interference channels [7], [31]. Capacity approximations are also available for one-to-many [7] and cyclic interference channels [32], although these coding schemes do not employ alignment. Bandemer and El Gamal have recently proposed a class of three-user deterministic channels where the interfering signals are passed through a function on their way to the receiver, which, in a certain sense, models interference alignment [33]. They develop a new rate region based on interference decoding for this model.

Nested lattice codes have been thoroughly studied as a framework for efficient source and channel coding with side information [9], [10], [34]. Recently, it has become clear that the inherent linear structure of lattices can enable many interesting new schemes, including distributed dirty paper coding [35], distributed source coding of linear functions [36]–[38], distributed antenna systems [39], [40], and physical-layer network coding [18], [41]–[44], to name a few. See [45] for a comprehensive survey. The origins of these schemes can be traced to the work of Körner and Marton [46], who showed that linear binning is optimal for the distributed compression of the parity of a doubly symmetric binary source.

## II. SYMMETRIC GAUSSIAN $K$ -USER INTERFERENCE CHANNEL

### A. Problem Statement

We begin with some notational conventions. We will denote vectors with boldface lowercase letters and matrices with boldface uppercase letters. For instance,  $\mathbf{a} \in \mathbb{Z}^K$  and  $\mathbf{A} \in \mathbb{Z}^{K \times K}$ . Let  $\|\mathbf{a}\| = \sqrt{\sum_{k=1}^K a_k^2}$  denote the  $\ell_2$ -norm of the vector  $\mathbf{a}$ . Also, let  $\mathbf{0}$  denote the zero vector and  $\mathbf{I}_{K \times K}$  denote the identity matrix of size  $K$ . We use  $\lfloor \cdot \rfloor$  to denote rounding to the nearest integer,  $\lfloor \cdot \rfloor$  to denote the floor operation and  $\lceil \cdot \rceil$  for the ceiling operation. In general, the letters  $a$  and  $b$  are used in this paper whenever the variables they describe are integer valued. All logarithms are to base 2. We also occasionally use the notation  $\log^+(x) \triangleq \max(0, \log(x))$ . All measures in this paper are Lebesgue measures.

There are  $K$  transmitter-receiver pairs that wish to simultaneously communicate across a shared channel over  $n$

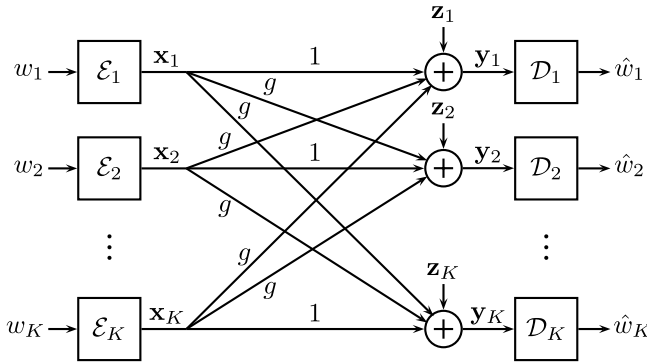


Fig. 5. Block diagram of a symmetric Gaussian  $K$ -user interference channel.

time slots, where the channel gains are constant over all  $n$  channel uses. We assume a real-valued channel model throughout.

*Definition 1 (Messages):* Each transmitter has a message  $w_k$  drawn independently and uniformly over  $\{1, 2, \dots, 2^{nR_{\text{SYM}}}\}$ .

*Definition 2 (Encoders):* Each transmitter is equipped with an encoder,  $\mathcal{E}_k: \{1, 2, \dots, 2^{nR_{\text{SYM}}}\} \rightarrow \mathbb{R}^n$ , that maps its message into a length- $n$  channel input  $\mathbf{x}_k = \mathcal{E}_k(w_k)$  that satisfies the power constraint,

$$\|\mathbf{x}_k\|^2 \leq n\text{SNR}$$

where  $\text{SNR} > 0$  is the signal-to-noise ratio.

*Definition 3 (Channel Model):* The channel output at each receiver is a noisy linear combination of its desired signal and the sum of the interfering terms, of the form

$$\mathbf{y}_k = \mathbf{x}_k + g \sum_{\ell \neq k} \mathbf{x}_\ell + \mathbf{z}_k, \quad (2)$$

where  $g > 0$  parametrizes the interference strength and  $\mathbf{z}_k$  is an i.i.d. Gaussian vector with mean 0 and variance 1. See Figure 5 for an illustration. We define the *interference-to-noise ratio* to be

$$\text{INR} \triangleq g^2 \text{SNR}$$

and the *interference level* to be

$$\alpha \triangleq \frac{\log(\text{INR})}{\log(\text{SNR})}.$$

*Remark 1:* Note that our definition of INR ignores the fact that there are  $K - 1$  interferers observed at each receiver. This is for two reasons. First, this definition parallels that of the two-user case [5], which will make it easier to compare the two rate regions. Second, the receivers will often be able to treat the interference as stemming from a single effective transmitter, via interference alignment. Of course, this is not the case when the receiver treats the interference as noise, as discussed in Section VI-E.

*Definition 4 (Decoders):* Each receiver is equipped with a decoder,  $\mathcal{D}_k: \mathbb{R}^n \rightarrow \{1, 2, \dots, 2^{nR_{\text{SYM}}}\}$ , that produces an estimate  $\hat{w}_k = \mathcal{D}_k(\mathbf{y}_k)$  of its desired message  $w_k$ .

*Definition 5 (Symmetric Capacity):* A symmetric rate  $R_{\text{SYM}}$  is *achievable* if, for any  $\epsilon > 0$  and  $n$  large enough, there exist encoders and decoders that can attain probability

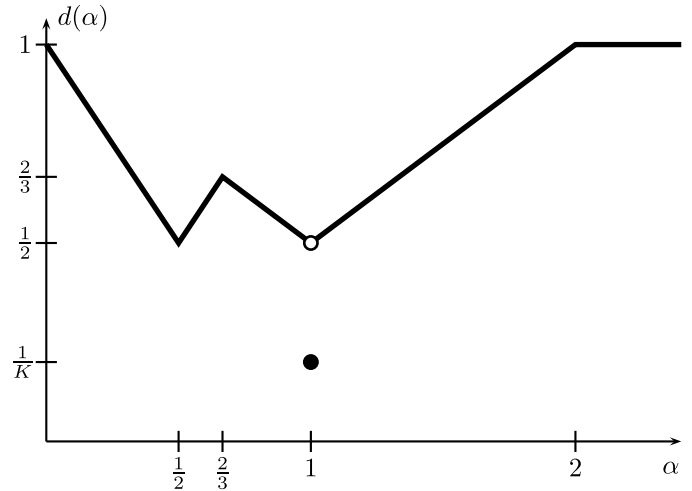


Fig. 6. Generalized degrees-of-freedom for the symmetric Gaussian  $K$ -user interference channel.

of error at most  $\epsilon$ ,

$$\Pr(\{\hat{w}_1 \neq w_1\} \cup \dots \cup \{\hat{w}_K \neq w_K\}) < \epsilon.$$

The *symmetric capacity*  $C_{\text{SYM}}$  is the supremum of all achievable symmetric rates.

*Remark 2:* Due to the symmetry of the channel, the symmetric capacity is equal to the sum capacity, normalized by the number of users. To see this, assume that the users employ different rates and that a rate tuple  $(R_1, R_2, \dots, R_K)$  is achievable. Since each transmitter-receiver pair sees the same effective channel, we can simply exchange the encoders and decoders to achieve the rate tuple  $(R_{\pi(1)}, R_{\pi(2)}, \dots, R_{\pi(K)})$  for any permutation  $\pi$ . By time-sharing across all permutations, we find that each user can achieve  $\frac{1}{K} \sum_{k=1}^K R_k$ , corresponding to a symmetric rate. Thus, the sum of any achievable rate tuple is upper bounded by  $K C_{\text{SYM}}$ .

*Definition 6 (Generalized Degrees-of-Freedom):* The *generalized degrees-of-freedom* (GDoF) specifies the fraction of the point-to-point Gaussian capacity that can be attained per user for a given interference level  $\alpha \geq 0$  as SNR tends to infinity,

$$d(\alpha) = \lim_{\text{SNR} \rightarrow \infty} \frac{C_{\text{SYM}}}{\frac{1}{2} \log(1 + \text{SNR})}.$$

### B. Approximate Sum Capacity

As shown by Jafar and Vishwanath [47, Th. 3.1], the GDoF of the symmetric  $K$ -user interference channel is identical to that of the two-user channel, except for a singularity at  $\alpha = 1$ ,

$$d(\alpha) = \begin{cases} 1 - \alpha & 0 \leq \alpha < \frac{1}{2} \text{ (noisy)} \\ \alpha & \frac{1}{2} \leq \alpha < \frac{2}{3} \text{ (weak)} \\ 1 - \frac{\alpha}{2} & \frac{2}{3} \leq \alpha < 1 \text{ (moderately weak)} \\ \frac{1}{K} & \alpha = 1 \\ \frac{\alpha}{2} & 1 < \alpha < 2 \text{ (strong)} \\ 1 & \alpha \geq 2 \text{ (very strong)}. \end{cases} \quad (3)$$

See Figure 6 for a plot. Notice that since SNR is taken to infinity, the GDoF characterization treats all channel gains

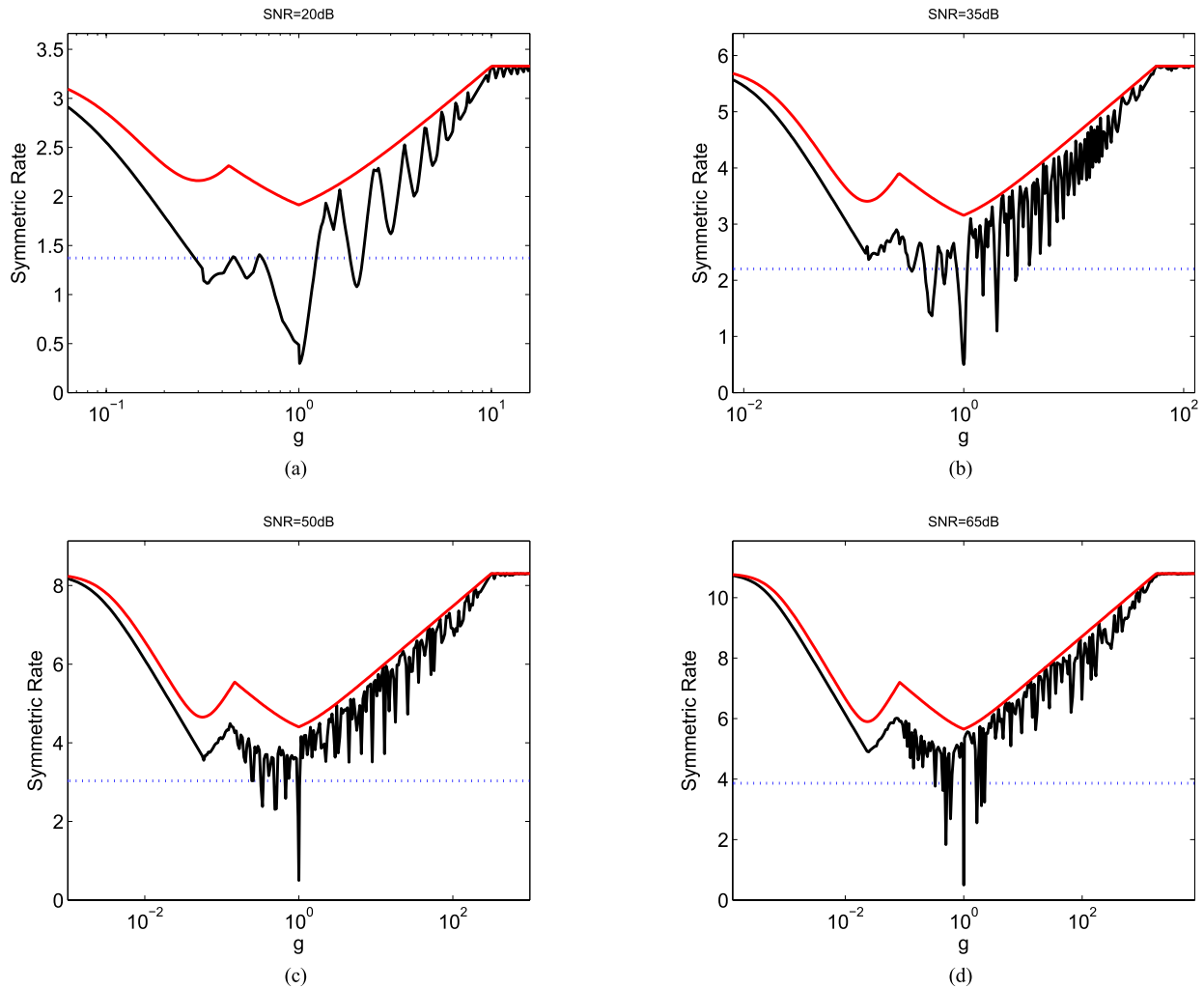


Fig. 7. Upper and lower bounds on the symmetric capacity of a 3-user symmetric Gaussian interference channel with respect to the cross-gain  $g$ . The upper bound (red line) is given by (46) and the lower bound (black line) is the maximum of the achievable rates from Theorem 9 and Corollary 3, which were computed numerically, and Theorem 10. The lower bounds from Theorem 1 are not plotted in this figure. For reference, we have also plotted the symmetric rate achievable via time-division (dotted blue line). (a) SNR = 20dB. (b) SNR = 35dB. (c) SNR = 50dB. (d) SNR = 65dB.

$g$  that do not scale with SNR as a single point at  $\alpha = 1$ . A finer view of this regime is possible at high SNR by simply setting  $g$  to be some fixed value and then taking SNR to infinity, corresponding to the standard notion of degrees-of-freedom. Surprisingly, this degrees-of-freedom characterization is discontinuous at rational values of  $g$  [21]. This presents an obstacle towards a clean capacity approximation at finite SNR.

To overcome this difficulty, our approximations allow for the possibility of an *outage set*, which is explicitly characterized. Specifically, in the regime around  $\alpha = 1$ , our capacity results take the following shape: for any constant  $c > 0$ , the capacity is approximated within at most  $c + 9 + \log K$  bits over the entire range of SNR, and all channel gains  $g$ , except for a set of measure  $\mu(c)$  which vanishes rapidly with  $c$ . This type of capacity approximation has also been used by Niesen and Maddah-Ali for the two-user Gaussian X channel [24] and seems to arise from the capacity region itself, not just the lower bound. That is, it appears that the capacity may in fact simultaneously vary rapidly with the fine scale of the channel gains (e.g., the distance to an appropriately

scaled integer) and slowly on the coarse scale (e.g., relative interference strength). In the high SNR limit, this behavior shows up as a discontinuity on the rationals but, at reasonable SNRs, our achievable scheme shows that this variation is in fact fairly smooth. The theorem below captures our capacity approximations in a simple form. All upper bounds in the theorem are based on [5] and [47]. The lower bound for the noisy interference regime is straightforward and the lower bound for the very strong interference regime is (a slight variation of) that of [20]. Our contribution is in the lower bounds for the weak and strong interference regimes.

*Theorem 1:* The symmetric capacity of the symmetric Gaussian  $K$ -user interference channel can be lower and upper bounded as follows:

- *Noisy Interference Regime*,  $0 \leq \alpha < \frac{1}{2}$ ,
 
$$\frac{1}{2} \log \left( 1 + \frac{\text{SNR}}{1 + \text{INR}} \right) - \frac{1}{2} \log(K - 1) \leq C_{\text{SYM}} < \frac{1}{2} \log \left( 1 + \frac{\text{SNR}}{1 + \text{INR}} \right) + 1$$

- *Weak Interference Regime*,  $\frac{1}{2} \leq \alpha < \frac{2}{3}$ ,  

$$\frac{1}{2} \log^+(\text{INR}) - \frac{7}{2} - \log(K) \leq C_{\text{SYM}} \leq \frac{1}{2} \log^+(\text{INR}) + 1$$
for all channel gains.

- *Moderately Weak Interference Regime*,  $\frac{2}{3} \leq \alpha < 1$ ,

$$\frac{1}{2} \log^+ \left( \frac{\text{SNR}}{\sqrt{\text{INR}}} \right) - c - 8 - \log(K) \leq C_{\text{SYM}} \leq \frac{1}{2} \log^+ \left( \frac{\text{SNR}}{\sqrt{\text{INR}}} \right) + 1$$

for all channel gains except for an outage set of measure  $\mu < 2^{-c}$  for any  $c > 0$ .

- *Strong Interference Regime*,  $1 \leq \alpha < 2$ ,

$$\frac{1}{4} \log^+(\text{INR}) - \frac{c}{2} - 3 \leq C_{\text{SYM}} \leq \frac{1}{4} \log^+(\text{INR}) + 1$$

for all channel gains except for an outage set whose measure is a fraction of  $2^{-c}$  of the interval  $1 < |g| < \sqrt{\text{SNR}}$ , for any  $c > 0$ .

- *Very Strong Interference Regime*,  $\alpha \geq 2$ ,

$$\frac{1}{2} \log(1 + \text{SNR}) - 1 \leq C_{\text{SYM}} \leq \frac{1}{2} \log(1 + \text{SNR})$$

*Remark 3:* Our characterization of the outage set in the strong and moderately weak interference regimes is in fact somewhat stronger than the characterization given in Theorem 1. Specifically, for the strong interference regime we show that, for any integer  $b$  in the range  $[1, \sqrt{\text{SNR}})$  and constant gap  $c > 0$ , the measure of the set of channel coefficients in the interval  $g \in [b, b+1)$  for which our inner bound does not hold is smaller than  $2^{-c}$ . Similarly, for the moderately weak interference regime we show that, for any integer  $b$  in the range  $[1, 1/6 \log(\text{SNR}))$  and constant gap  $c > 0$ , the measure of the set of channel coefficients the interval  $g \in [2^{-b}, 2^{-b+1})$  for which our inner bound does not hold is smaller than  $2^{-(c+b)}$ . Using this refined characterization, our results can be interpreted in the following way: For all values of  $\alpha$  except for an outage set with Lebesgue measure smaller than  $2^{-c}$ , the symmetric capacity of the symmetric Gaussian  $K$ -user interference channel is

$$C_{\text{SYM}} = \frac{d(\alpha)}{2} \log(\text{SNR}) \pm \delta(K, c),$$

where  $0 \leq \delta(K, c) < c + \log(K) + 10$  and  $d(\alpha)$  is given in (3).

### III. PRELIMINARIES

In this section, we give some basic definitions and results that will be extensively used in the sequel.

#### A. $K$ -User Gaussian MAC

Consider the  $K$ -user Gaussian multiple-access channel (MAC)

$$\mathbf{y} = \sum_{k=1}^K h_k \mathbf{x}_k + \mathbf{z}, \quad (4)$$

where the vector  $\mathbf{h} = [h_1 \cdots h_K]^T \in \mathbb{R}^K$  represents the channel gains,  $\mathbf{x}_k \in \mathbb{R}^n$ ,  $k = 1, \dots, K$ , are the channel inputs,

$\mathbf{z} \in \mathbb{R}^n$  is additive white Gaussian noise (AWGN) with zero mean and unit variance and  $\mathbf{y} \in \mathbb{R}^n$  is the channel output. Without loss of generality, we assume all  $K$  users are subject to the same power constraint<sup>3</sup>

$$\|\mathbf{x}_k\|^2 \leq n\text{SNR}, \quad k = 1, \dots, K. \quad (5)$$

The capacity region of the channel (4) is known (see e.g., [48, Th. 15.3.6]) to be the set of all rate tuples  $(R_1, \dots, R_K)$  satisfying

$$\sum_{k \in \mathcal{S}} R_k < \frac{1}{2} \log \left( 1 + \text{SNR} \sum_{k \in \mathcal{S}} |h_k|^2 \right) \quad (6)$$

for all subsets  $\mathcal{S} \subseteq \{1, \dots, K\}$ . The achievability part of the capacity theorem is established using i.i.d. Gaussian codebooks for all users. Motivated by lattice interference alignment, we are interested in establishing the achievability of certain rate tuples under the constraint that the codebooks employed by the  $K$  users form a chain of nested lattice codes.

*Remark 4:* Recall that the corner points of the capacity region are achievable via successive interference cancellation, either using i.i.d. Gaussian codebooks [48, Sec. 15.3.6] or nested lattice codebooks [18, Sec. VII.A]. Time-sharing between these corner points suffices to reach any point in the capacity region. However, this time-sharing approach does not suffice for an interference channel, as each receiver will require a different time allocation between users.

#### B. Nested Lattice Codes

We employ the nested lattice framework originally proposed in [10]. A lattice  $\Lambda$  is a discrete subgroup of  $\mathbb{R}^n$  which is closed under reflection and real addition. Formally, for any  $\mathbf{t}_1, \mathbf{t}_2 \in \Lambda$ , we have that  $-\mathbf{t}_1, -\mathbf{t}_2 \in \Lambda$  and  $\mathbf{t}_1 + \mathbf{t}_2 \in \Lambda$ . Note that by definition the zero vector  $\mathbf{0}$  is always a member of the lattice. Any lattice  $\Lambda$  in  $\mathbb{R}^n$  is spanned by some  $n \times n$  matrix  $\mathbf{G}$  such that

$$\Lambda = \{\mathbf{t} = \mathbf{G}\mathbf{q} : \mathbf{q} \in \mathbb{Z}^n\}.$$

We say that a lattice is full-rank if its spanning matrix  $\mathbf{G}$  is full-rank.

We denote the nearest neighbor quantizer associated with the lattice  $\Lambda$  by

$$Q_\Lambda(\mathbf{x}) = \arg \min_{\mathbf{t} \in \Lambda} \|\mathbf{x} - \mathbf{t}\|. \quad (7)$$

The Voronoi region of  $\Lambda$ , denoted by  $\mathcal{V}$ , is the set of all points in  $\mathbb{R}^n$  which are quantized to the zero vector, where ties in (7) are broken in a systematic manner. The modulo operation returns the quantization error w.r.t. the lattice,

$$[\mathbf{x}] \bmod \Lambda = \mathbf{x} - Q_\Lambda(\mathbf{x}),$$

and satisfies the distributive law,

$$[a[\mathbf{x}] \bmod \Lambda + b[\mathbf{y}] \bmod \Lambda] \bmod \Lambda = [a\mathbf{x} + b\mathbf{y}] \bmod \Lambda,$$

for all  $a, b \in \mathbb{Z}$  and  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ .

A lattice  $\Lambda$  is said to be nested in  $\Lambda_1$  if  $\Lambda \subseteq \Lambda_1$ . The coding schemes presented in this paper utilize a chain of  $K+1$  nested

<sup>3</sup>As otherwise the different powers can be absorbed into the channel gains.

lattices satisfying

$$\Lambda \subseteq \Lambda_K \subseteq \dots \subseteq \Lambda_1. \quad (8)$$

From these lattices, we construct  $K$  codebooks, one for each user. Specifically, user  $k$  is allocated the codebook  $\mathcal{L}_k = \Lambda_{\theta(k)} \cap \mathcal{V}$ , where  $\mathcal{V}$  is the Voronoi region of  $\Lambda$  and the function  $\theta(k) : \{1, \dots, K\} \rightarrow \{1, \dots, K\}$  maps between users and lattices. The rate of each codebook  $\mathcal{L}_k$  is

$$R_k = \frac{1}{n} \log |\Lambda_{\theta(k)} \cap \mathcal{V}|.$$

User  $k$  encodes its message into a lattice point from its codebook,  $\mathbf{t}_k \in \mathcal{L}_k$ . Each user also has a random<sup>4</sup> dither vector  $\mathbf{d}_k$  which is generated independently and uniformly over  $\mathcal{V}$ . These dithers are made available to the decoder. The signal transmitted by user  $k$  is

$$\mathbf{x}_k = [\mathbf{t}_k - \mathbf{d}_k] \bmod \Lambda.$$

*Remark 5:* The nested lattice construction from [10] employs Construction A. To create each fine lattice, this procedure first embeds codewords drawn from a linear code into the unit cube, and then applies the generator matrix for the coarse lattice  $\Lambda$ . As shown in [10], this ensemble of nested lattice codes can approach the capacity of a point-to-point Gaussian channel. If the integers  $\mathbb{Z}^n$  are selected as the coarse lattice, the resulting nested lattice code is equivalent to a linear code coupled with a pulse amplitude modulation (PAM) constellation. Furthermore, the  $\bmod \Lambda$  operation simplifies to the quantization error from rounding to the integers. It can be shown that the cost of this simplification is only the shaping gain, which corresponds to at most  $1/2 \log(2\pi e/12) \simeq 0.255$  bits per channel use [49].

### C. Compute-and-Forward

Our objective is to communicate over the MAC using the compute-and-forward scheme from [18]. See Figure 8 for an illustration. To this end, the receiver first decodes a linearly independent set of  $K$  integer linear combinations of the lattice codewords. Afterwards, it solves this set of linear combinations for the lattice codewords. Assume that the receiver is interested in decoding the integer linear combination

$$\mathbf{v} = \left[ \sum_{k=1}^K a_k \mathbf{t}_k \right] \bmod \Lambda$$

with coefficient vector  $\mathbf{a} = [a_1 \dots a_K]^T \in \mathbb{Z}^K$ . Following the scheme of [18], the receiver scales the observation  $\mathbf{y}$  by a factor  $\beta$ , removes the dithers, and reduces modulo  $\Lambda$  to get

$$\begin{aligned} \mathbf{s} &= \left[ \beta \mathbf{y} + \sum_{k=1}^K a_k \mathbf{d}_k \right] \bmod \Lambda \\ &= \left[ \sum_{k=1}^K a_k \mathbf{x}_k + \sum_{k=1}^K a_k \mathbf{d}_k + \sum_{k=1}^K (\beta h_k - a_k) \mathbf{x}_k + \beta \mathbf{z} \right] \bmod \Lambda \\ &= [\mathbf{v} + \mathbf{z}_{\text{eff}}(\mathbf{h}, \mathbf{a}, \beta)] \bmod \Lambda, \end{aligned} \quad (9)$$

<sup>4</sup>It can be shown that these random dithers can be replaced with deterministic ones, meaning that no common randomness is required.

where

$$\mathbf{z}_{\text{eff}}(\mathbf{h}, \mathbf{a}, \beta) = \sum_{k=1}^K (\beta h_k - a_k) \mathbf{x}_k + \beta \mathbf{z} \quad (10)$$

is effective noise. From [18], we have that  $\mathbf{z}_{\text{eff}}(\mathbf{h}, \mathbf{a}, \beta)$  is statistically independent of  $\mathbf{v}$  and its effective variance, defined as

$$\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, \beta) \triangleq \frac{1}{n} \mathbb{E} \|\mathbf{z}_{\text{eff}}(\mathbf{h}, \mathbf{a}, \beta)\|^2 \quad (11)$$

is

$$\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, \beta) = \|\beta \mathbf{h} - \mathbf{a}\|^2 \cdot \text{SNR} + \beta^2. \quad (12)$$

Let  $k^* = \min_{k: a_k \neq 0} \theta(k)$  be the index of the densest lattice participating in the integer linear combination  $\mathbf{v}$ . The receiver produces an estimate for  $\mathbf{v}$  by applying to  $\mathbf{s}$  the lattice quantizer associated with  $\Lambda_{k^*}$ ,

$$\hat{\mathbf{v}} = [Q_{\Lambda_{k^*}}(\mathbf{s})] \bmod \Lambda. \quad (13)$$

Let  $\mathcal{V}_{k^*}$  be the Voronoi region of  $\Lambda_{k^*}$ , and note that the probability of decoding error is upper bounded by the probability that the effective noise lies outside the Voronoi region of  $\Lambda_{k^*}$ ,

$$\Pr(\hat{\mathbf{v}} \neq \mathbf{v}) \leq \Pr(\mathbf{z}_{\text{eff}}(\mathbf{h}, \mathbf{a}, \beta) \notin \mathcal{V}_{k^*}). \quad (14)$$

The next theorem summarizes and reformulates relevant results from Sections IV.C, IV.D, and V.A of [18].

*Theorem 2:* For any  $\epsilon > 0$  and  $n$  large enough there exists a chain of  $n$ -dimensional nested lattices  $\Lambda \subseteq \Lambda_K \subseteq \dots \subseteq \Lambda_1$  forming the set of codebooks  $\mathcal{L}_1, \dots, \mathcal{L}_K$  having rates  $R_1, \dots, R_K$  and satisfying the power constraint (5) such that:

- For all channel vectors  $\mathbf{h} \in \mathbb{R}^K$  and coefficient vectors  $\mathbf{a} \in \mathbb{Z}^K$ , the average error probability in decoding the integer linear combination  $\mathbf{v} = [\sum_{k=1}^K a_k \mathbf{t}_k] \bmod \Lambda$  of transmitted lattice points  $\mathbf{t}_k \in \mathcal{L}_k$  can be made smaller than  $\epsilon$  so long as the message rates do not exceed the computation rate,

$$R_k < R_{\text{comp}}(\mathbf{h}, \mathbf{a}, \beta) \triangleq \frac{1}{2} \log \left( \frac{\text{SNR}}{\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, \beta)} \right), \quad (15)$$

- for all  $k$  such that  $a_k \neq 0$  and some  $\beta \in \mathbb{R}$ .
- The codebooks  $\mathcal{L}_1, \dots, \mathcal{L}_K$  are isomorphic to some set of linear codebooks  $\mathcal{C}_1, \dots, \mathcal{C}_K$  over the finite field  $\mathbb{Z}_p$ , where  $p$  is a sufficiently large prime number.
- For the same  $p$ , the equation  $[p \cdot \mathbf{t}] \bmod \Lambda = \mathbf{0}$  holds  $\forall \mathbf{t} \in \Lambda_k, k = 1, \dots, K$ .

*Corollary 1:* Given  $K$  integer linear combinations  $\mathbf{V} = [\mathbf{v}_1 \dots \mathbf{v}_K]$  with coefficient vectors  $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_K]^T$ , the lattice points  $\mathbf{t}_1, \dots, \mathbf{t}_K$  can be recovered if  $[\mathbf{A}] \bmod p$  is full rank over  $\mathbb{Z}_p$ .

*Remark 6:* By taking the blocklength  $n$  and field size  $p$  to be large enough, it can be shown that, for a fixed channel vector  $\mathbf{h}$  and finite SNR, it suffices to check whether  $\mathbf{A}$  is full rank over the reals. See [18, Sec. VI] for an in-depth discussion.

*Remark 7:* Note that it is also possible to map both the messages and the integer linear combinations into an appropriately chosen finite field. That is, the messages can be written as vec-



tors with elements that take values in a prime-sized finite field, and the receiver ultimately recovers linear combinations of the messages over the same finite field. See [18] for more details.

It follows from Theorem 2(a) that in order to maximize the computation rate  $R_{\text{comp}}(\mathbf{h}, \mathbf{a}, \beta)$  for a given coefficient vector, one has to minimize  $\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, \beta)$  over  $\beta$ . It is seen from (12) that the expression for  $\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, \beta)$  is equal to the mean squared error (MSE) for linear estimation of  $\tilde{X} = \sum_{k=1}^K a_k X_k$  from  $Y = \sum_{k=1}^K h_k X_k + Z$  where  $\{X_k\}_{k=1}^K$  are i.i.d. random variables with zero mean and variance SNR and  $Z$  is statistically independent of  $\{X_k\}_{k=1}^K$  with zero mean and unit variance. Hence the minimizing value of  $\beta$  is the linear minimum mean squared error (MMSE) estimation coefficient of  $\tilde{X}$  from  $Y$ . This value of  $\beta$  was found in [18, Th. 2] and the resulting MSE is given by

$$\begin{aligned} \sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}) &\triangleq \min_{\beta \in \mathbb{R}} \sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, \beta) \\ &= \text{SNR} \left( \|\mathbf{a}\|^2 - \frac{\text{SNR}(\mathbf{h}^T \mathbf{a})^2}{1 + \text{SNR}\|\mathbf{h}\|^2} \right) \\ &= \text{SNR} \mathbf{a}^T \left( \mathbf{I}_{K \times K} - \frac{\text{SNR} \mathbf{h} \mathbf{h}^T}{1 + \text{SNR}\|\mathbf{h}\|^2} \right) \mathbf{a} \\ &= \mathbf{a}^T (\text{SNR}^{-1} \mathbf{I}_{K \times K} + \mathbf{h} \mathbf{h}^T)^{-1} \mathbf{a} \quad (16) \\ &= \left\| (\text{SNR}^{-1} \mathbf{I}_{K \times K} + \mathbf{h} \mathbf{h}^T)^{-1/2} \mathbf{a} \right\|^2, \quad (17) \end{aligned}$$

where (16) can be verified using Woodbury's matrix identity (i.e., the Matrix Inversion Lemma) [50, Th. 18.2.8]. Accordingly, we define

$$\begin{aligned} R_{\text{comp}}(\mathbf{h}, \mathbf{a}) &\triangleq \max_{\beta \in \mathbb{R}} R_{\text{comp}}(\mathbf{h}, \mathbf{a}, \beta) \\ &= \frac{1}{2} \log \left( \frac{\text{SNR}}{\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a})} \right). \quad (18) \end{aligned}$$

In the sequel, we will require that the receiver decodes  $K$  linearly independent integer linear combinations. However, the specific values of the coefficient vectors for these linear combinations are not important as long as they form a full-rank set. Therefore, we are free to choose these coefficients such as to maximize the corresponding computation rate.

Define the matrix

$$\mathbf{F} \triangleq (\text{SNR}^{-1} \mathbf{I}_{K \times K} + \mathbf{h} \mathbf{h}^T)^{-1/2}, \quad (19)$$

and the lattice  $\Lambda(\mathbf{F}) = \{\boldsymbol{\nu} = \mathbf{F} \mathbf{a} : \mathbf{a} \in \mathbb{Z}^K\}$ . Notice that this  $K$ -dimensional lattice is induced by the channel matrix, not the  $n$ -dimensional coding scheme. The effective variance for the coefficient vector  $\mathbf{a}$  is

$$\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}) = \|\mathbf{F} \mathbf{a}\|^2, \quad (20)$$

and hence  $\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a})$  is the length of the lattice vector corresponding to the integer-valued vector  $\mathbf{a}$ . It follows that the problem of finding the  $K$  linearly independent integer-valued vectors that result in the highest computation rates is equivalent to finding a set of shortest independent vectors  $\{\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_K\}$  in the lattice  $\Lambda(\mathbf{F})$ , and then taking the integer coefficient vectors as  $\mathbf{a}_m = \mathbf{F}^{-1} \cdot \boldsymbol{\nu}_m$ . The lengths of the shortest linearly independent vectors in a lattice are called *successive minima*, as defined next.

*Definition 7 (Successive Minima):* Let  $\Lambda(\mathbf{F})$  be a full-rank lattice in  $\mathbb{R}^K$  spanned by the matrix  $\mathbf{F} \in \mathbb{R}^{K \times K}$ . For  $m = 1, \dots, K$ , we define the  $m$ th successive minimum as

$$\lambda_m(\mathbf{F}) \triangleq \inf \left\{ r : \dim \left( \text{span} \left( \Lambda(\mathbf{F}) \cap \mathcal{B}(\mathbf{0}, r) \right) \right) \geq m \right\}$$

where  $\mathcal{B}(\mathbf{0}, r) = \{\mathbf{x} \in \mathbb{R}^K : \|\mathbf{x}\| \leq r\}$  is the closed ball of radius  $r$  around  $\mathbf{0}$ . In words, the  $m$ th successive minimum of a lattice is the minimal radius of a ball centered around  $\mathbf{0}$  that contains  $m$  linearly independent lattice points.

The following definition identifies the  $K$  linearly independent coefficient vectors which yield the highest computation rates.

*Definition 8:* Let  $\mathbf{F}$  be the matrix defined in (19). We say that an ordered set of integer coefficient vectors  $\{\mathbf{a}_1, \dots, \mathbf{a}_K\}$  with corresponding computation rates  $R_{\text{comp}, m} \triangleq R_{\text{comp}}(\mathbf{h}, \mathbf{a}_m)$  is *optimal* if the  $K$  vectors are linearly independent and  $\|\mathbf{F} \mathbf{a}_m\| = \lambda_m(\mathbf{F})$  for any  $m = 1, \dots, K$ . Note, that such a set always exists by definition of successive minima, and that it is not unique. For example, if  $\{\mathbf{a}_1, \dots, \mathbf{a}_K\}$  is an optimal set of coefficient vectors, so is the set  $\{-\mathbf{a}_1, \dots, -\mathbf{a}_K\}$ . Note also that the optimal computation rates satisfy  $R_{\text{comp}, 1} \geq \dots \geq R_{\text{comp}, K}$ .

*Remark 8:* Several recent papers have proposed families of constellations and codes that are well-suited for low-complexity implementations of compute-and-forward [40], [44], [51]–[55]. These codes could serve as building blocks for a practical implementation of our alignment scheme.

## D. Numerical Evaluations

The optimal coefficient vectors and computation rates from Definition 8 play an important role in the achievable rate regions derived in this paper. The problem of determining the optimal coefficient vectors is that of finding the set of  $K$  linearly independent *integer-valued* vectors that minimizes the effective noise (17). As discussed above, this problem is equivalent to finding the shortest  $K$  linearly independent lattice vectors in the lattice  $\Lambda(\mathbf{F})$  spanned by the matrix  $\mathbf{F}$  defined in (19).

It is shown in [18, Lemma 1] that only integer vectors  $\mathbf{a} \in \mathbb{Z}^K$  that satisfy the condition

$$\|\mathbf{a}\|^2 < 1 + \|\mathbf{h}\|^2 \text{SNR} \quad (21)$$

yield positive rates. Therefore, in our considerations it suffices to enumerate all integer vectors (other than the zero vector) that satisfy (21), and then exhaustively search over these vectors in order to find the optimal set. At moderate values of SNR this task is computationally reasonable. Nevertheless, it is sometimes simpler to find a set of short linearly independent lattice vectors in  $\Lambda(\mathbf{F})$ , which is not necessarily optimal, in order to obtain lower bounds on the set of optimal computation rates. A simple low-complexity algorithm for computing a short lattice basis (which forms a set of  $K$  linearly independent lattice vectors) is the LLL algorithm [56].<sup>5</sup> In producing the

<sup>5</sup>Pseudocode for the LLL algorithm can be found, e.g., in [57].

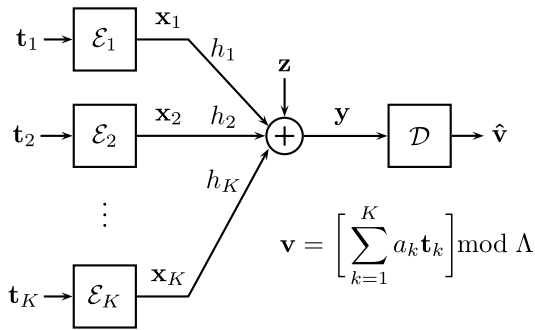


Fig. 8. Compute-and-forward on a Gaussian multiple-access channel. The transmitters send lattice points  $\mathbf{t}_k$  and the receiver decodes an integer combination of them, modulo the coarse lattice  $\Lambda$ . The rate is determined by how closely the equation coefficients  $a_k$  match the channel coefficients  $h_k$ .

figures for this paper we have employed the LLL algorithm, meaning that the plotted achievable rates in Figure 7 are in fact lower bounds on the rates given by Theorems 9 and 11.

We note that a similar procedure for finding the optimal coefficient vectors was also described in [44], where the optimal coefficient vectors are termed dominated solutions.

#### IV. MULTIPLE-ACCESS VIA COMPUTE-AND-FORWARD

This section introduces a new coding technique for reliable communication over the  $K$ -user Gaussian multiple-access channel. The basic idea is to first decode a linearly independent set of  $K$  integer linear combinations of the transmitted codewords, and then solve these for the transmitted messages. As we will argue, under certain technical conditions, it is possible to map the users' rates to the computation rates in a one-to-one fashion. We begin this section with a high-level overview of the scheme, which is illustrated in Figures 9 and 10.

Each user  $k$  maps its message to a lattice point  $\mathbf{t}_k$  in its codebook  $\mathcal{L}_k$  and transmits a dithered version of it. The  $K$  lattice codebooks utilized by the different users form a chain of nested lattices as in (8). Assume for now that the users are ordered with descending rates  $R_1 \geq R_2 \geq \dots \geq R_K$ , i.e.,  $\theta(k) = k$  for  $k = 1, \dots, K$ . The receiver, which sees a noisy real-valued linear combination of the transmitted codewords, begins by decoding the integer linear combination  $\mathbf{v}_1 = [\sum a_{1k} \mathbf{t}_k] \bmod \Lambda$  that yields the highest computation rate  $R_{\text{comp},1}$ . Using the compute-and-forward framework, this is possible if  $R_1 < R_{\text{comp},1}$ . Then, it proceeds to decode the integer linear combination  $\mathbf{v}_2 = [\sum a_{2k} \mathbf{t}_k] \bmod \Lambda$  that yields the second highest computation rate  $R_{\text{comp},2}$ . In general,  $\mathbf{t}_1$  participates in this linear combination and the condition for correct decoding of  $\mathbf{v}_2$  is therefore  $R_1 < R_{\text{comp},2}$ . Nevertheless, this condition can be relaxed using the linear combination  $\mathbf{v}_1$  that has already been decoded. Specifically, after scaling of the channel output and removing the dithers, the receiver has a noisy observation

$$\mathbf{s}_2 = [\mathbf{v}_2 + \mathbf{z}_{\text{eff}}(\mathbf{h}, \mathbf{a}_2)] \bmod \Lambda$$

of the desired linear combination  $\mathbf{v}_2$ . If  $\mathbf{t}_1$  participates in  $\mathbf{v}_1$ , it is possible to cancel out  $\mathbf{t}_1$  from the second linear combination by adding a scaled version of  $\mathbf{v}_1$  to  $\mathbf{s}_2$ . Namely,

the receiver adds  $r_{21} \mathbf{v}_1$  to  $\mathbf{s}_2$ , where  $r_{21}$  is an integer chosen such that  $[(a_{21} + r_{21} a_{11})] \bmod p = 0$ , which assures that  $[(a_{21} + r_{21} a_{11}) \mathbf{t}_1] \bmod \Lambda = \mathbf{0}$  for any  $\mathbf{t}_1 \in \mathcal{L}_1$ . After reducing mod  $\Lambda$ , this yields

$$\begin{aligned} \mathbf{s}_2^{\text{SI}} &= [\mathbf{v}_2 + r_{21} \mathbf{v}_1 + \mathbf{z}_{\text{eff}}(\mathbf{h}, \mathbf{a}_2)] \bmod \Lambda \\ &= [\tilde{\mathbf{v}}_2 + \mathbf{z}_{\text{eff}}(\mathbf{h}, \mathbf{a}_2)] \bmod \Lambda, \\ \tilde{\mathbf{v}}_2 &= \left[ \sum_{k=2}^K (a_{2k} + r_{21} a_{1k}) \mathbf{t}_k \right] \bmod \Lambda. \end{aligned}$$

Note that  $\mathbf{t}_1$  does not participate in  $\tilde{\mathbf{v}}_2$ . Since the effective noise  $\mathbf{z}_{\text{eff}}(\mathbf{h}, \mathbf{a}_2)$  is unchanged by this process, the receiver can decode  $\tilde{\mathbf{v}}_2$  as long as  $R_2 < R_{\text{comp},2}$ . Now, the receiver can obtain  $\mathbf{v}_2$  by subtracting  $r_{21} \mathbf{v}_1$  from  $\tilde{\mathbf{v}}_2$  and reducing mod  $\Lambda$ .<sup>6</sup> The receiver decodes the remaining linear combinations in a similar manner, i.e., before decoding the  $m$ th linear combination  $\mathbf{v}_m$  with computation rate  $R_{\text{comp},m}$  the receiver adds to

$$\mathbf{s}_m = [\mathbf{v}_m + \mathbf{z}_{\text{eff}}(\mathbf{h}, \mathbf{a}_m)] \bmod \Lambda$$

an integer linear combination  $[\sum_{\ell=1}^{m-1} r_{m\ell} \mathbf{v}_\ell] \bmod \Lambda$  of its previously decoded linear combinations. The coefficients  $r_{m1}, \dots, r_{m,m-1} \in \mathbb{Z}$  are chosen such that the effect of  $\mathbf{t}_1, \dots, \mathbf{t}_{m-1}$  is canceled out from  $\mathbf{v}_m$ . Assuming that such coefficients exist, the receiver can decode  $\tilde{\mathbf{v}}_m = [\mathbf{v}_m + \sum_{\ell=1}^{m-1} r_{m\ell} \mathbf{v}_\ell] \bmod \Lambda$  as long as  $R_m < R_{\text{comp},m}$ .

Lemma 2, stated in Appendix A, establishes that for any set of  $K$  linearly independent coefficient vectors  $\{\mathbf{a}_1, \dots, \mathbf{a}_K\}$  there indeed always exist integer-valued coefficients  $\{r_{m\ell}\}$  such that in the  $m$ th decoding step the receiver can cancel out  $m - 1$  lattice points from the desired linear combination  $\mathbf{v}_m$ , using the previously decoded linear combinations  $\{\mathbf{v}_1, \dots, \mathbf{v}_{m-1}\}$ . The procedure for finding these coefficients is reminiscent of the Gaussian elimination procedure of a full-rank matrix. One of the basic operations in Gaussian elimination is row switching. In our considerations, this would correspond to using a linear combination that has not been decoded yet for eliminating lattice points from another linear combination, which is clearly not possible. Therefore, a major difference between our procedure for finding a good set of coefficients  $\{r_{ij}\}$  and Gaussian elimination is that row switching is not permitted. This will sometimes constrain the order in which we can cancel out users from linear combinations. Nevertheless, there always exists at least one valid successive cancellation order. In other words, we can always cancel out the effect of  $m - 1$  users from  $\mathbf{v}_m$  using the decoded linear combination  $\{\mathbf{v}_1, \dots, \mathbf{v}_{m-1}\}$ , but we cannot always control which of the  $K$  users to cancel. As a result, there always exists at least one permutation vector  $\pi$  such that all  $K$  linear combinations can be decoded as long as

$$R_{\pi(m)} < R_{\text{comp},m}, \quad m = 1, \dots, K. \quad (22)$$

It follows that a sum-rate of  $\sum_{m=1}^K R_{\text{comp},m}$  is achievable

<sup>6</sup>The operation of extracting  $\mathbf{v}_2$  from  $\tilde{\mathbf{v}}_2$  is in fact not necessary as the receiver is only interested in decoding any linearly independent set of  $K$  integer linear combinations. We describe this step only to simplify the exposition of the scheme.

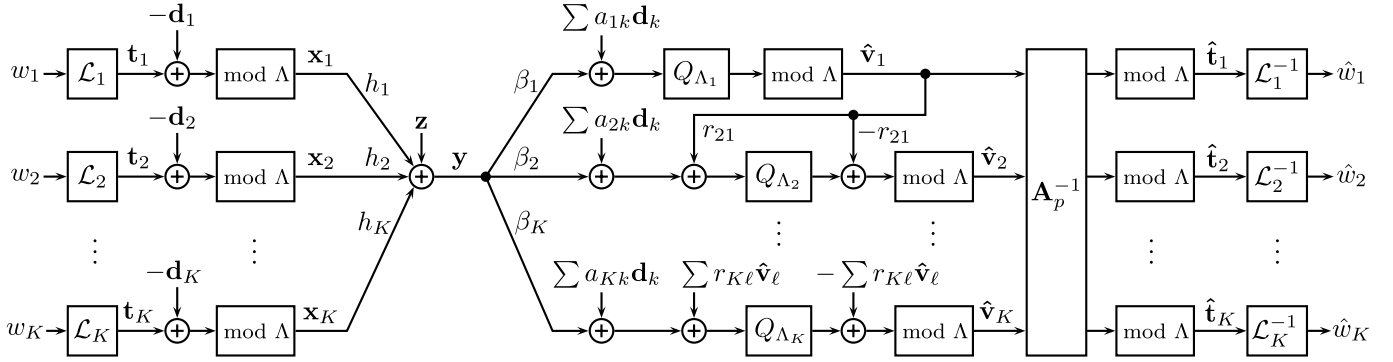


Fig. 9. System diagram of the nested lattice encoding and decoding operations employed as part of the compute-and-forward transform. Each message  $w_k$  is mapped to a lattice codeword  $\mathbf{t}_k$  according to codebook  $\mathcal{L}_k$ , dithered, and transmitted as  $\mathbf{x}_k$ . The multiple-access channel scales codeword  $k$  by  $h_k$  and outputs the sum plus Gaussian noise  $\mathbf{z}$ . The decoder attempts to recover a linearly independent set of  $K$  integer linear combinations with coefficients  $\mathbf{A} = \{\mathbf{a}_{mk}\}$ . For the figure, we have assumed that  $R_1 \geq R_2 \geq \dots \geq R_K$  and that  $R_m < R_{\text{comp}}(\mathbf{h}, \mathbf{a}_m, \beta_m)$ . To decode the first linear combination  $\mathbf{v}_1 = [\sum a_{1k} \mathbf{t}_k] \bmod \Lambda$ , the receiver scales  $\mathbf{y}$  by  $\beta_1$ , removes the dithers, quantizes using  $Q_{\Lambda_1}$ , and takes  $\bmod \Lambda$ . For the second linear combination  $\mathbf{v}_2 = [\sum a_{2k} \mathbf{t}_k] \bmod \Lambda$ , the decoder scales by  $\beta_2$ , removes the dithers, and then eliminates the lattice point  $\mathbf{t}_1$  using its estimate of the first linear combination  $\hat{\mathbf{v}}_1$  so that the rate of the remaining lattice points is at most  $R_2$ . It then quantizes using  $Q_{\Lambda_2}$ , adds back in  $\hat{\mathbf{v}}_1$ , and takes  $\bmod \Lambda$ . Decoding proceeds in this fashion, using a form of successive interference cancellation to keep the rates of the lattice points below the computation rates. Afterwards, the receiver solves for the original lattice points by multiplying by  $\mathbf{A}_p^{-1}$ , which is the inverse of  $\mathbf{A}$  over  $\mathbb{Z}_p$ , and taking  $\bmod \Lambda$ . Finally, it maps these estimates  $\hat{\mathbf{t}}_k$  of the transmitted lattice points back to the corresponding messages.

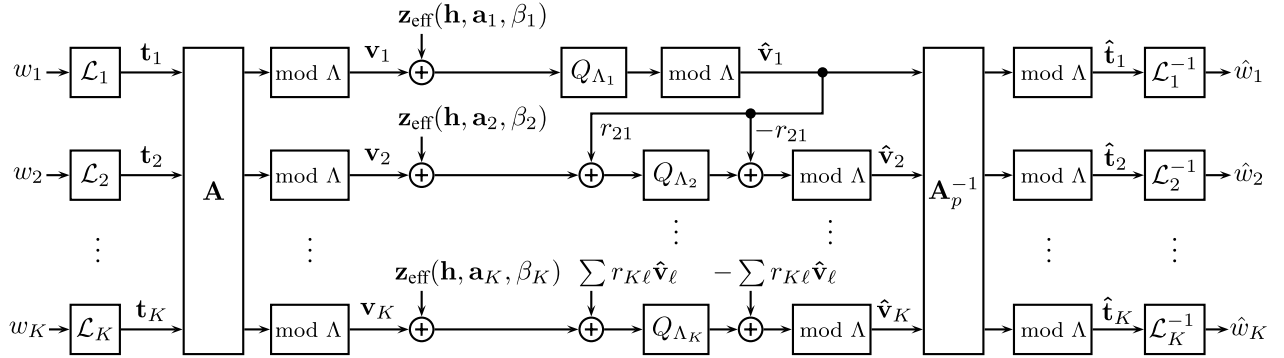


Fig. 10. Effective MIMO channel induced by the compute-and-forward transform of a Gaussian multiple-access channel. The channel output  $\mathbf{y} = \sum h_k \mathbf{x}_k + \mathbf{z}$  is converted into a linearly independent set of  $K$  integer linear combinations  $\mathbf{v}_m = [\sum a_{mk} \mathbf{t}_k] \bmod \Lambda$  plus effective noise  $\mathbf{z}_{\text{eff}}(\mathbf{h}, \mathbf{a}_m, \beta_m) = \beta_m \mathbf{z} + \sum (\beta_m h_k - a_{mk}) \mathbf{x}_k$ . As in Figure 9, these linear combinations can be decoded using a version of successive cancellation.

over the  $K$ -user MAC with our scheme, in which all users employ nested lattice codebooks. As we shall see, this sum rate is within a constant gap, smaller than  $K/2 \log(K)$  bits, from the sum capacity of the MAC, for all channel gains and SNR.

#### A. The Compute-and-Forward Transform

We first introduce a transformation of a MAC to a multiple-input multiple-output (MIMO) mod- $\Lambda$  channel, where the  $K \times K$  channel matrix is integer-valued. This transformation, dubbed the *compute-and-forward transform*, will play an important role in our decoding scheme for the interference channel.

*Definition 9:* Let  $\{\mathbf{a}_1, \dots, \mathbf{a}_K\}$  be a set of optimal integer coefficient vectors (see Definition 8),  $\beta_1, \dots, \beta_K$  the corresponding optimal scaling factors, and  $R_{\text{comp},1} \geq \dots \geq R_{\text{comp},K}$  the corresponding optimal computation rates. We define the *compute-and-forward transform* of the MAC with

nested lattice codes as

$$\begin{aligned} \mathbf{S} &= \begin{pmatrix} \mathbf{s}_1 \\ \vdots \\ \mathbf{s}_K \end{pmatrix} = \begin{pmatrix} [\beta_1 \mathbf{y} + \sum_{k=1}^K a_{1k} \mathbf{d}_k] \bmod \Lambda \\ \vdots \\ [\beta_K \mathbf{y} + \sum_{k=1}^K a_{Kk} \mathbf{d}_k] \bmod \Lambda \end{pmatrix} \\ &= \left[ \mathbf{A} \begin{pmatrix} \mathbf{t}_1 \\ \vdots \\ \mathbf{t}_K \end{pmatrix} + \mathbf{Z}_{\text{eff}} \right] \bmod \Lambda, \end{aligned} \quad (23)$$

where we have written the channel output  $\mathbf{y}$ , dithers  $\mathbf{d}_k$ , and lattice codewords  $\mathbf{t}_k$  as length- $n$  row vectors. We also denote  $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_K]^T$  and  $\mathbf{Z}_{\text{eff}} = [\mathbf{z}_{\text{eff},1}^T \dots \mathbf{z}_{\text{eff},K}^T]^T$ .

*Remark 9:* The transform is not unique as the set of optimal integer coefficient vectors is not unique. Nevertheless, the set of optimal computation rates is unique. As we shall see, the set of optimal computation rates dictates the rates attained over the transformed channel. Therefore, we use the term *the compute-and-forward transform* of the channel, with the understanding that although there

may be multiple options for the transform, they are all equivalent.

The  $m$ th output  $\mathbf{s}_m$  of the transformed channel corresponds to an integer linear combination plus effective noise. Due to Theorem 2, each such linear combination can be reliably decoded as long as all lattice points participating in it belong to codes of rates smaller than  $R_{\text{comp},m}$ . We now lower bound the sum of  $K$  optimal computation rates, and in the sequel we show that this sum can be translated to a valid MAC sum rate.

*Theorem 3:* The sum of optimal computation rates is lower bounded by

$$\sum_{m=1}^K R_{\text{comp},m} \geq \frac{1}{2} \log(1 + \|\mathbf{h}\|^2 \text{SNR}) - \frac{K}{2} \log(K). \quad (24)$$

The proof makes use of the following well-known theorem due to Minkowski [58, Th. 1.5], that upper bounds the product of successive minima.

*Theorem 4 (Minkowski):* For any lattice  $\Lambda(\mathbf{F})$  which is spanned by a full-rank  $K \times K$  matrix  $\mathbf{F}$

$$\prod_{m=1}^K \lambda_m^2(\mathbf{F}) \leq K^K |\det(\mathbf{F})|^2. \quad (25)$$

We are now ready to prove Theorem 3.

*Proof of Theorem 3:* Let  $\Lambda(\mathbf{F})$  be a lattice spanned by the matrix  $\mathbf{F}$  from (19), and let  $\lambda_1(\mathbf{F}), \dots, \lambda_K(\mathbf{F})$  be its  $K$  successive minima. Let  $\mathbf{a}_1, \dots, \mathbf{a}_K \in \mathbb{Z}^K$  denote the optimal coefficient vectors. By Definition 8 and (20) we have  $\|\mathbf{F} \mathbf{a}_m\| = \lambda_m(\mathbf{F})$  for  $m = 1, \dots, K$ . The sum of optimal computation rates is

$$\begin{aligned} \sum_{m=1}^K R_{\text{comp},m} &= \sum_{m=1}^K R_{\text{comp}}(\mathbf{h}, \mathbf{a}_m) \\ &= \sum_{m=1}^K \frac{1}{2} \log\left(\frac{\text{SNR}}{\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}_m)}\right) \\ &= \frac{K}{2} \log(\text{SNR}) - \frac{1}{2} \log\left(\prod_{m=1}^K \|\mathbf{F} \mathbf{a}_m\|^2\right) \\ &= \frac{K}{2} \log(\text{SNR}) - \frac{1}{2} \log\left(\prod_{m=1}^K \lambda_m^2(\mathbf{F})\right). \end{aligned}$$

Applying Theorem 4 to the product  $\prod_{m=1}^K \lambda_m^2(\mathbf{F})$  yields

$$\sum_{m=1}^K R_{\text{comp},m} \geq \frac{K}{2} \log(\text{SNR}) - \frac{1}{2} \log(K^K |\det(\mathbf{F})|^2). \quad (26)$$

Using Sylvester's determinant identity (see e.g., [50])

$$\det(\mathbf{I}_{K \times K} + \text{SNR} \mathbf{h} \mathbf{h}^T) = \det(1 + \|\mathbf{h}\|^2 \text{SNR}),$$

we have that

$$|\det(\mathbf{F})|^2 = \frac{\text{SNR}^K}{1 + \|\mathbf{h}\|^2 \text{SNR}}. \quad (27)$$

Substituting (27) into (26) proves the theorem.  $\blacksquare$

*Remark 10:* It is possible to avoid the loss of the constant factor  $K/2 \log K$  in (24) using successive compute-and-

forward, as described in [59] and [60]. However, in this case the operational interpretation of the sum of computation rates becomes more involved than that described in the sequel. See [60] for more details.

Next, we give an operational meaning to the  $K$  optimal computation rates.

### B. Multiple-Access Sum Capacity to Within a Constant Gap

We now show that the compute-and-forward transform can be used for achieving several rate tuples within a constant gap from the boundary of the capacity region of the  $K$ -user MAC. To establish this result, we introduce a decoding technique that we will refer to as *algebraic successive cancellation*. Namely, each decoded linear combination will be used to cancel out the effect of one user from the linear combinations that have yet to be decoded. We first illustrate the coding scheme by an example, and then formalize our result in Theorem 5.

*Example 1:* Consider the two-user MAC

$$\mathbf{y} = \sqrt{5} \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z},$$

at SNR = 15dB. It can be shown using (17) and (18) that the compute-and-forward transform of this channel is

$$\begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} = \left[ \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{z}_{\text{eff},1} \\ \mathbf{z}_{\text{eff},2} \end{pmatrix} \right] \bmod \Lambda$$

with  $R_{\text{comp},1} \simeq 2.409$  bits and  $R_{\text{comp},2} \simeq 1.372$  bits. Note that  $(R_{\text{comp},1} + R_{\text{comp},2}) / (1/2 \log(1 + \|\mathbf{h}\|^2 \text{SNR})) \simeq 0.998$ . We use a chain of three nested lattices  $\Lambda \subseteq \Lambda_2 \subseteq \Lambda_1$  that satisfy the conditions of Theorem 2 in order to construct the codebooks  $\mathcal{L}_1 = \Lambda_1 \cap \mathcal{V}$  with rate  $R_1$  arbitrarily close to  $R_{\text{comp},1}$  for user 1 and  $\mathcal{L}_2 = \Lambda_2 \cap \mathcal{V}$  with rate  $R_2$  arbitrarily close to  $R_{\text{comp},2}$  for user 2.

From Theorem 2(a), we know that  $\mathbf{v}_1 = [2\mathbf{t}_1 + \mathbf{t}_2] \bmod \Lambda$  can be decoded from  $\mathbf{s}_1$  since  $R_1$  and  $R_2$  are smaller than  $R_{\text{comp},1}$ . However, Theorem 2 does not guarantee that  $\mathbf{v}_2 = [3\mathbf{t}_1 + \mathbf{t}_2] \bmod \Lambda$  can be decoded directly from  $\mathbf{s}_2$  since the first user employs a codebook with a rate  $R_1 \approx R_{\text{comp},1}$  which is higher than the second computation rate  $R_{\text{comp},2}$ . To circumvent this issue, we use the estimate  $\hat{\mathbf{v}}_1$  of the linear combination  $\mathbf{v}_1$  as side information in order to cancel out the lattice point  $\mathbf{t}_1 \in \Lambda_1 \cap \mathcal{V}$  from  $\mathbf{s}_2$ . Note that Theorem 2(c) guarantees that  $[p \cdot \mathbf{t}_k] \bmod \Lambda = \mathbf{0}$ ,  $k = 1, 2$  for some sufficiently large prime number  $p$ . Let  $2^{-1} \in \mathbb{Z}$  be an integer that satisfies  $[2^{-1} \cdot 2] \bmod p = 1$ . The receiver computes

$$\begin{aligned} \mathbf{s}_2^{\text{SI}} &= [\mathbf{s}_2 - 3 \cdot 2^{-1} \hat{\mathbf{v}}_1] \bmod \Lambda \\ &\stackrel{(a)}{=} [(3 - 3 \cdot 2^{-1} \cdot 2) \mathbf{t}_1 + (1 - 3 \cdot 2^{-1}) \mathbf{t}_2 + \mathbf{z}_{\text{eff},2}] \bmod \Lambda \\ &\stackrel{(b)}{=} [[1 - 3 \cdot 2^{-1}] \bmod p \cdot \mathbf{t}_2 + \mathbf{z}_{\text{eff},2}] \bmod \Lambda \\ &= [\tilde{a}_{12} \cdot \mathbf{t}_2 + \mathbf{z}_{\text{eff},2}] \bmod \Lambda, \end{aligned} \quad (28)$$

where  $\tilde{a}_{12} = [1 - 3 \cdot 2^{-1}] \bmod p$ . Step (a) in (28) follows from the distributive law. Step (b) follows since  $3 - 3 \cdot 2^{-1} \cdot 2 = M \cdot p$  for some  $M \in \mathbb{Z}$ . Thus,

$$\begin{aligned} [(3 - 3 \cdot 2^{-1} \cdot 2)\mathbf{t}_1] \bmod \Lambda &= [M \cdot p \cdot \mathbf{t}_1] \bmod \Lambda \\ &= [M \cdot [p \cdot \mathbf{t}_1] \bmod \Lambda] \bmod \Lambda \\ &= \mathbf{0}, \end{aligned}$$

where the last equality is justified by Theorem 2(c).

Now only  $\mathbf{t}_2$  participates in the linear combination  $\tilde{\mathbf{v}}_2 = [\tilde{a}_{12}\mathbf{t}_2] \bmod \Lambda$  and, since  $R_2$  is smaller than  $R_{\text{comp},2}$ , Theorem 2 guarantees that it can be decoded from  $\mathbf{s}_2^{\text{SI}}$ . This is accomplished by quantizing onto  $\Lambda_2$  and reducing modulo  $\Lambda$ ,

$$\hat{\tilde{\mathbf{v}}}_2 = [Q_{\Lambda_2}(\mathbf{s}_2^{\text{SI}})] \bmod \Lambda.$$

After decoding both linear combinations  $\mathbf{v}_1$  and  $\tilde{\mathbf{v}}_2$  the receiver can solve for the transmitted lattice points  $\mathbf{t}_1$  and  $\mathbf{t}_2$ , as the two linear combinations are full-rank over  $\mathbb{Z}_p$ . We have therefore shown that the rate region  $R_1 < R_{\text{comp},1}$  and  $R_2 < R_{\text{comp},2}$  is achievable. In a similar manner, we can show that the rate region  $R_1 < R_{\text{comp},2}$  and  $R_2 < R_{\text{comp},1}$  is achievable with this scheme.

In order to formally characterize the achievable rate region, we will need the following definition which identifies the orders for which algebraic successive cancellation can be performed.

*Definition 10:* For a full-rank  $K \times K$  matrix  $\mathbf{A}$  with integer-valued entries we define the *pseudo-triangularization* process, which transforms the matrix  $\mathbf{A}$  to a matrix  $\tilde{\mathbf{A}}$  which is upper triangular up to column permutation  $\pi = [\pi(1) \pi(2) \cdots \pi(K)]$ . This is accomplished by left-multiplying  $\mathbf{A}$  by a lower triangular matrix  $\mathbf{L}$  with unit diagonal, such that  $\tilde{\mathbf{A}} = \mathbf{L}\mathbf{A}$  is upper triangular up to column permutation  $\pi$ . Although the matrix  $\mathbf{A}$  is integer valued, the matrices  $\mathbf{L}$  and  $\tilde{\mathbf{A}}$  need not necessarily be integer valued. Note that the pseudo-triangularization process is reminiscent of Gaussian elimination except that row switching and row multiplication are prohibited. It is also closely connected to the LU decomposition where only column pivoting is permitted.

*Example 2:* The  $2 \times 2$  matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$$

from Example 1 can be pseudo-triangularized with two different permutation vectors

$$\begin{aligned} \tilde{\mathbf{A}} &= \begin{pmatrix} 1 & 0 \\ -\frac{3}{2} & 1 \end{pmatrix} \cdot \mathbf{A} \\ &= \begin{pmatrix} 2 & 1 \\ 0 & -\frac{1}{2} \end{pmatrix}, \quad \pi = [1 \ 2], \end{aligned}$$

or

$$\begin{aligned} \tilde{\mathbf{A}} &= \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \cdot \mathbf{A} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}, \quad \pi = [2 \ 1]. \end{aligned}$$

*Remark 11:* Any full-rank matrix can be triangularized using the Gaussian elimination process, and therefore any full-rank matrix can be pseudo-triangularized with at least one permutation vector  $\pi$ . In particular, since for any MAC

the integer-valued matrix  $\mathbf{A}$  from the compute-and-forward transform is full-rank, it can always be pseudo-triangularized with at least one permutation vector  $\pi$ . There are full-rank matrices that can be pseudo-triangularized with several different permutation vectors, such as  $\mathbf{A}$  from Example 2. However, there are also full-rank matrices  $\mathbf{A}$  that can be pseudo-triangularized with only one permutation vector  $\pi$ . An example of such a matrix is the identity matrix  $\mathbf{I}_{K \times K}$ .

The next theorem gives an achievable rate region for the MAC under the compute-and-forward transform. The proof is given in Appendix A and follows along the same lines as Example 1.

*Theorem 5:* Consider the MAC (4). For any  $\epsilon > 0$  and  $n$  large enough, there exists a chain of  $n$ -dimensional nested lattices  $\Lambda \subseteq \Lambda_K \subseteq \cdots \subseteq \Lambda_1$  forming the set of codebooks  $\mathcal{L}_1, \dots, \mathcal{L}_K$  with rates  $R_1, \dots, R_K$  such that for all  $\mathbf{h} \in \mathbb{R}^K$ , if:

- 1) each user  $k$  encodes its message using the codebook  $\mathcal{L}_k$ ,
- 2) the integer-valued matrix from the compute-and-forward transform of the MAC (4) can be pseudo-triangularized with the permutation vector  $\pi$ , and the optimal computation rates are  $R_{\text{comp},1} \geq \cdots \geq R_{\text{comp},K}$ ,
- 3) all rates  $R_1, \dots, R_K$  satisfy

$$R_k < R_{\text{comp},\pi^{-1}(k)}, \quad \text{for } k = 1, \dots, K \quad (29)$$

where  $\pi^{-1}$  is the inverse permutation vector of  $\pi$ ,

then all messages can be decoded with error probability smaller than  $\epsilon$ .

Combining Theorems 3 and 5 gives the following theorem.

*Theorem 6:* The sum rate achieved by the compute-and-forward transform has a gap of no more than  $K/2 \log K$  bits from the sum capacity of the MAC.

*Proof:* Let  $R_{\text{comp},1} \geq \cdots \geq R_{\text{comp},K}$  be the optimal computation rates in the compute-and-forward transform of the MAC (4). The integer-valued matrix from the compute-and-forward transform can be pseudo-triangularized with at least one permutation vector  $\pi$ . By Theorem 5, the rate tuple

$$R_k = R_{\text{comp},\pi^{-1}(k)} - \delta, \quad \text{for } k = 1, \dots, K \quad (30)$$

is achievable for any  $\delta > 0$ . For this rate tuple we have

$$\begin{aligned} \sum_{k=1}^K R_k &= \sum_{k=1}^K (R_{\text{comp},\pi^{-1}(k)} - \delta) \\ &= \sum_{k=1}^K R_{\text{comp},k} - K\delta \\ &\geq \frac{1}{2} \log(1 + \|\mathbf{h}\|^2 \text{SNR}) - \frac{K}{2} \log(K) - K\delta, \end{aligned} \quad (31)$$

where (31) follows from Theorem 3. Since this is true for any  $\delta > 0$ , the result follows. ■

### C. Effective Multiple-Access Channel

A channel that often arises in the context of lattice interference alignment is a  $K$ -user Gaussian multiple-access channel (MAC) with integer-valued ratios between some of the channel

coefficients. Specifically, the output of such a channel can be written as

$$\mathbf{y} = \sum_{\ell=1}^L g_{\ell} \left( \sum_{i \in \mathcal{K}_{\ell}} b_i \mathbf{x}_i \right) + \mathbf{z}, \quad (32)$$

where  $\mathcal{K}_1, \dots, \mathcal{K}_L$  are disjoint subsets of  $\{1, \dots, K\}$ . We assume that the  $b_i \in \mathbb{Z}$  are non-zero integers, which opens up the possibility of lattice alignment.

The channel (32) may describe the signal seen by a receiver in an interference network, perhaps after appropriate precoding at the transmitters. In such networks, each receiver is only interested in the messages from some of the users while the others act as interferers. Hence, it is beneficial to align several interfering users into one effective interferer, by taking advantage of the fact that the sum of lattice codewords is itself a lattice codeword.

*Definition 11 (Effective Users):* For the MAC specified by (32), we define  $L$  effective users

$$\mathbf{x}_{\text{eff},\ell} \triangleq \sum_{i \in \mathcal{K}_{\ell}} b_i \mathbf{x}_i, \quad \ell = 1, \dots, L.$$

*Definition 12 (Effective MAC):* The  $K$ -user MAC (32) induces the effective  $L$ -user MAC

$$\mathbf{y} = \sum_{\ell=1}^L g_{\ell} \mathbf{x}_{\text{eff},\ell} + \mathbf{z}, \quad (33)$$

with the vector of effective channel coefficients  $\mathbf{g} = [g_1 \dots g_L]^T \in \mathbb{R}^L$ . The effective channel is further characterized by the effective users' weights

$$b_{\text{eff},\ell}^2 \triangleq \sum_{i \in \mathcal{K}_{\ell}} b_i^2$$

for  $\ell = 1, \dots, L$ , and the effective (diagonal) weight matrix

$$\mathbf{B} \triangleq \text{diag}(b_{\text{eff},1}^2, \dots, b_{\text{eff},L}^2). \quad (34)$$

*Definition 13 (Effective Lattice Points):* Let  $\mathbf{t}_i$  be the lattice point transmitted by user  $i$ . We define the effective lattice point corresponding to effective user  $\ell$  as

$$\mathbf{t}_{\text{eff},\ell} = \left[ \sum_{i \in \mathcal{K}_{\ell}} b_i \mathbf{t}_i \right] \bmod \Lambda.$$

Let  $\theta_{\text{eff}}(\ell) = \min_{i \in \mathcal{K}_{\ell}} \theta(i)$  (where  $\theta(\cdot)$  is the mapping between users and fine lattices defined in Section III) be the index of the densest lattice contributing to  $\mathbf{t}_{\text{eff},\ell}$ . Since all lattices are nested, it follows that  $\mathbf{t}_{\text{eff},\ell} \in \Lambda_{\theta_{\text{eff}}(\ell)}$ .

*Example 3: (Symmetric  $K$ -User Interference Channel):* Consider the symmetric  $K$ -user interference channel (2). The channel seen by the  $k$ th receiver is of the form of (32) with  $g_1 = 1$ ,  $g_2 = g$ ,  $\mathcal{K}_1 = \{k\}$ ,  $\mathcal{K}_2 = \{1, \dots, K\} \setminus k$ , and  $b_i = 1$  for  $i = 1, \dots, K$ . If each of the  $K$  users transmits a single codeword drawn from a common nested lattice code, the channel becomes an effective two-user MAC,

$$\mathbf{y}_k = \mathbf{x}_{\text{eff},k,1} + g \mathbf{x}_{\text{eff},k,2} + \mathbf{z}_k,$$

where the effective users are  $\mathbf{x}_{\text{eff},k,1} = \mathbf{x}_k$  and  $\mathbf{x}_{\text{eff},k,2} = \sum_{i \neq k} \mathbf{x}_i$ , and the effective users' weights are  $b_{\text{eff},1}^2 = 1$  and

$b_{\text{eff},2}^2 = K - 1$ . The effective lattice points are  $\mathbf{t}_{\text{eff},k,1} = \mathbf{t}_k$  and  $\mathbf{t}_{\text{eff},k,2} = [\sum_{i \neq k} \mathbf{t}_i] \bmod \Lambda$ .

Our achievable schemes for the symmetric  $K$ -user interference channel, developed in Section VI, are based on transforming the  $K$ -user MAC seen by each receiver into an effective MAC with less effective users. We will develop two schemes: One transforms the channel into an effective two-user MAC as in the example above. The other, which mimics the Han-Kobayashi approach, transforms the channel into an effective three-user MAC.

When lattice interference alignment schemes are designed properly, the message intended for the receiver is mapped into a separate effective user, while multiple interfering users are folded into a smaller number of effective users. In this case, it suffices for the receiver to decode only the  $L$  effective lattice points corresponding to the effective users, rather than the  $K$  lattice points transmitted by all users. In our considerations, the effective lattice points are recovered by first decoding  $L$  integer linear combinations of the form

$$\begin{aligned} \mathbf{v} &= \left[ \sum_{\ell=1}^L a_{\ell} \sum_{i \in \mathcal{K}_{\ell}} b_i \mathbf{t}_i \right] \bmod \Lambda \\ &= \left[ \sum_{\ell=1}^L a_{\ell} \mathbf{t}_{\text{eff},\ell} \right] \bmod \Lambda \end{aligned} \quad (35)$$

with linearly independent coefficient vectors, and then solving for  $\mathbf{t}_{\text{eff},1}, \dots, \mathbf{t}_{\text{eff},L}$ .

As in Section III, in order to decode an integer linear combination  $\mathbf{v}$ , the receiver first scales its observation by a factor  $\beta$ , removes the dithers, and reduces modulo  $\Lambda$ , which yields

$$\begin{aligned} \mathbf{s} &= \left[ \beta \mathbf{y} + \sum_{\ell=1}^L a_{\ell} \sum_{i \in \mathcal{K}_{\ell}} b_i \mathbf{d}_i \right] \bmod \Lambda \\ &= \left[ \sum_{\ell=1}^L a_{\ell} \mathbf{x}_{\text{eff},\ell} + \sum_{\ell=1}^L a_{\ell} \sum_{i \in \mathcal{K}_{\ell}} b_i \mathbf{d}_i \right. \\ &\quad \left. + \sum_{\ell=1}^L (\beta g_{\ell} - a_{\ell}) \mathbf{x}_{\text{eff},\ell} + \beta \mathbf{z} \right] \bmod \Lambda \\ &= [\mathbf{v} + \mathbf{z}_{\text{eff}}(\mathbf{g}, \mathbf{a}, \beta, \{b_i\})] \bmod \Lambda, \end{aligned} \quad (36)$$

where

$$\begin{aligned} \mathbf{z}_{\text{eff}}(\mathbf{g}, \mathbf{a}, \beta, \{b_i\}) &= \beta \mathbf{y} - \sum_{\ell=1}^L a_{\ell} \mathbf{x}_{\text{eff},\ell} \\ &= \sum_{\ell=1}^L (\beta g_{\ell} - a_{\ell}) \sum_{i \in \mathcal{K}_{\ell}} b_i \mathbf{x}_i + \beta \mathbf{z} \end{aligned} \quad (37)$$

is effective noise which is statistically independent of  $\mathbf{v}$ . Its effective variance is

$$\sigma_{\text{eff}}^2(\mathbf{g}, \mathbf{a}, \beta, \mathbf{B}) = \text{SNR} \sum_{\ell=1}^L (\beta g_{\ell} - a_{\ell})^2 b_{\text{eff},\ell}^2 + \beta^2, \quad (38)$$

where  $\mathbf{B}$  is defined in (34). Let  $\ell^* = \min_{\ell: a_{\ell} \neq 0} \theta_{\text{eff}}(\ell)$  be the index of the densest lattice participating in the linear combination  $\mathbf{v}$ . Since all lattices are nested, then  $\mathbf{v} \in \Lambda_{\ell^*}$ .

The receiver produces an estimate for  $\mathbf{v}$  by applying to  $\mathbf{s}$  the lattice quantizer associated with  $\Lambda_{\ell^*}$ ,

$$\hat{\mathbf{v}} = [Q_{\Lambda_{\ell^*}}(\mathbf{s})] \bmod \Lambda. \quad (39)$$

It follows from Theorem 2 that there exists a chain of  $K + 1$  nested lattices which allows to decode  $\mathbf{v}$  with a vanishing error probability so long as

$$R_i < R_{\text{comp}}(\mathbf{g}, \mathbf{a}, \beta, \mathbf{B}) = \frac{1}{2} \log \left( \frac{\text{SNR}}{\sigma_{\text{eff}}^2(\mathbf{g}, \mathbf{a}, \beta, \mathbf{B})} \right), \quad (40)$$

for all  $i \in \bigcup_{\ell: a_\ell \neq 0} \mathcal{K}_\ell$ .

The expression for  $\sigma_{\text{eff}}^2(\mathbf{g}, \mathbf{a}, \beta, \mathbf{B})$  is equal to the MSE for linear estimation of  $X_{\text{eff}} = \sum_{\ell=1}^L a_\ell X_{\text{eff},\ell}$  from  $Y = \sum_{\ell=1}^L g_\ell X_{\text{eff},\ell} + Z$  where  $\{X_{\text{eff},\ell}\}_{\ell=1}^L$  are statistically independent random variables with zero mean and variances  $b_{\text{eff},\ell}^2 \text{SNR}$  respectively and  $Z$  is statistically independent of  $\{X_{\text{eff},\ell}\}_{\ell=1}^L$  with zero mean and unit variance. Hence, the minimizing value of  $\beta$  is the linear MMSE estimation coefficient of  $\tilde{X}$  from  $Y$ . A straightforward calculation shows that the minimizing value of  $\beta$  is

$$\beta = \frac{\mathbb{E}(\tilde{X}_{\text{eff}} Y)}{\text{Var}(Y)} = \frac{\text{SNR} \mathbf{g}^T \mathbf{B} \mathbf{a}}{1 + \text{SNR} \mathbf{g}^T \mathbf{B} \mathbf{g}}$$

and the MSE it achieves is

$$\begin{aligned} \sigma_{\text{eff}}^2(\mathbf{g}, \mathbf{a}, \mathbf{B}) &\triangleq \min_{\beta \in \mathbb{R}} \sigma_{\text{eff}}^2(\mathbf{g}, \mathbf{a}, \beta, \mathbf{B}) \\ &= \text{SNR} \mathbf{a}^T \left( \mathbf{B} - \frac{\text{SNR} \mathbf{B} \mathbf{g} \mathbf{g}^T \mathbf{B}}{1 + \text{SNR} \cdot \mathbf{g}^T \mathbf{B} \mathbf{g}} \right) \mathbf{a} \quad (41) \\ &= \mathbf{a}^T (\text{SNR}^{-1} \mathbf{B}^{-1} + \mathbf{g} \mathbf{g}^T)^{-1} \mathbf{a} \quad (42) \\ &= \left\| (\text{SNR}^{-1} \mathbf{B}^{-1} + \mathbf{g} \mathbf{g}^T)^{-1/2} \mathbf{a} \right\|^2, \end{aligned}$$

where again (42) can be verified using Woodbury's matrix identity [50, Th. 18.2.8]. Accordingly, we define

$$R_{\text{comp}}(\mathbf{g}, \mathbf{a}, \mathbf{B}) \triangleq \frac{1}{2} \log \left( \frac{\text{SNR}}{\sigma_{\text{eff}}^2(\mathbf{g}, \mathbf{a}, \mathbf{B})} \right). \quad (43)$$

As in Section III, we define the set of optimal  $L$  coefficient vectors for the equivalent channel (33) as the  $L$  linearly independent vectors  $\{\mathbf{a}_1, \dots, \mathbf{a}_L\}$  that yield the highest computation rates  $R_{\text{comp},1} = R_{\text{comp}}(\mathbf{g}, \mathbf{a}_1, \mathbf{B}) \geq \dots \geq R_{\text{comp},L} = R_{\text{comp}}(\mathbf{g}, \mathbf{a}_L, \mathbf{B})$  (see Definition 8). The compute-and-forward transform of the effective  $L$ -user MAC is

$$\begin{aligned} \mathbf{S} &= \begin{pmatrix} \left[ \beta_1 \mathbf{y} + \sum_{\ell=1}^L a_{1\ell} \sum_{i \in \mathcal{K}_\ell} b_i \mathbf{d}_i \right] \bmod \Lambda \\ \vdots \\ \left[ \beta_L \mathbf{y} + \sum_{\ell=1}^L a_{L\ell} \sum_{i \in \mathcal{K}_\ell} b_i \mathbf{d}_i \right] \bmod \Lambda \end{pmatrix} \\ &= \left[ \mathbf{A} \begin{pmatrix} \mathbf{t}_{\text{eff},1} \\ \vdots \\ \mathbf{t}_{\text{eff},L} \end{pmatrix} + \mathbf{Z}_{\text{eff}} \right] \bmod \Lambda, \quad (44) \end{aligned}$$

where  $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_L]^T$  and  $\mathbf{Z}_{\text{eff}} = [\mathbf{z}_{\text{eff},1}^T \dots \mathbf{z}_{\text{eff},L}^T]^T$ .

The next two theorems are simple extensions of Theorems 3 and 5. Their proofs are given in Appendix B.

*Theorem 7:* The sum of optimal computation rates for the effective  $L$ -user MAC (33) is lower bounded by

$$\sum_{\ell=1}^L R_{\text{comp},\ell} \geq \frac{1}{2} \log \left( \frac{1 + \text{SNR} \sum_{\ell=1}^L g_\ell^2 b_{\text{eff},\ell}^2}{\det(\mathbf{B})} \right) - \frac{L}{2} \log(L).$$

*Theorem 8:* Consider the effective  $L$ -user MAC (33), induced from the  $K$ -user MAC (32), characterized by the effective channel vector  $\mathbf{g}$  and the effective weight matrix  $\mathbf{B}$ . For any  $\epsilon > 0$  and  $n$  large enough there exists a chain of  $n$ -dimensional nested lattices  $\Lambda \subseteq \Lambda_L \subseteq \dots \subseteq \Lambda_1$  forming the set of codebooks  $\mathcal{L}_1, \dots, \mathcal{L}_L$  with rates  $R_1, \dots, R_L$  such that for all  $\mathbf{g} \in \mathbb{R}^L$  and  $\mathbf{B}$ , if:

- 1) each user  $i \in \mathcal{K}_\ell$  encodes its message using the codebook  $\mathcal{L}_\ell$  or a codebook nested in  $\mathcal{L}_\ell$ ,
- 2) the integer-valued matrix from the compute-and-forward transform of the effective MAC (33) can be pseudo-triangularized with the permutation vector  $\pi$ , and the optimal computation rates are  $R_{\text{comp},1} \geq \dots \geq R_{\text{comp},L}$ ,
- 3) all rates  $R_1, \dots, R_L$  satisfy

$$R_\ell < R_{\text{comp},\pi^{-1}(\ell)}, \quad \text{for } \ell = 1, \dots, L \quad (45)$$

where  $\pi^{-1}$  is the inverse permutation vector of  $\pi$ ,

then all effective lattice points  $\mathbf{t}_{\text{eff},\ell}$  can be decoded with error probability smaller than  $\epsilon$ .

*Corollary 2 (Achievable Symmetric Rate):* Consider the effective  $L$ -user MAC (33), induced from the  $K$ -user MAC (32), characterized by channel coefficients  $\mathbf{g}$  and the effective weight matrix  $\mathbf{B}$ . There exists a pair of  $n$ -dimensional nested lattices  $\Lambda \subseteq \Lambda_1$  forming the codebook  $\mathcal{L}$  of rate  $R$  such that for all  $\mathbf{g} \in \mathbb{R}^L$  and  $\mathbf{B}$ , if

- 1) all users encode their messages using  $\mathcal{L}$  (or codebooks nested in  $\mathcal{L}$ ),
- 2) The  $L$ th optimal computation rate in the compute-and-forward transform of (33) is  $R_{\text{comp},L}$ ,
- 3)  $R < R_{\text{comp},L}$ ,

then, for  $n$  large enough, all effective lattice points  $\mathbf{t}_{\text{eff},\ell}$  can be decoded with an arbitrarily small error probability.

*Remark 12:* Corollary 2 is easily obtained from Theorem 8. However, it can also be established without incorporating the compute-and-forward transform machinery. Indeed, if all users transmit from the same lattice codebook with rate smaller than  $R_{\text{comp},L}$ , by Theorem 2, each of the  $L$  linear combinations with optimal coefficient vectors can be decoded (without using algebraic successive decoding as in the compute-and-forward transform approach). Then, the decoded linear combinations can be solved for the effective lattice points.

In Section VI, we introduce two achievable schemes for the  $K$ -user Gaussian interference channel. One of them is a simple transmission scheme where all users transmit from the same nested lattice code. The result of Corollary 2 suffices to establish the rates achieved by this scheme. In the second achievable scheme, which mimics the Han-Kobayashi scheme for the two-user interference channel, each user transmits a superposition of codewords taken from two nested lattice

codebooks. In this case, Corollary 2 does not suffice and Theorem 8, which uses the compute-and-forward transform machinery, is needed.

In Section VI, we leverage these achievability results to lower bound the capacity of the symmetric Gaussian  $K$ -user interference channel.

## V. SYMMETRIC CAPACITY UPPER BOUNDS

In this section, we state an upper bound on the symmetric capacity of the symmetric  $K$ -user Gaussian interference channel. We follow the same arguments given in [47] for showing that the symmetric capacity of the symmetric  $K$ -user interference channel is upper bounded by that of the symmetric two-user interference channel. Namely, eliminating all but two users, say users 1 and 2, the symmetric capacity is upper bounded by the results of [5]. This is simply because removing interferers cannot decrease the symmetric rates for users 1 and 2. Thus, the upper bounds from [5] hold for the symmetric rates of user 1 and 2 in the  $K$ -user symmetric interference channel. Repeating the same argument for each pair of users we see that the upper bounds on  $C_{\text{SYM}}$  developed in [5] for  $K = 2$  continue to hold for all  $K > 2$  as well. Therefore, the symmetric capacity of the symmetric  $K$ -user Gaussian interference channel is upper bounded as [5]

$$C_{\text{SYM}} \leq \begin{cases} \frac{1}{2} \log \left( 1 + \text{INR} + \frac{\text{SNR}}{1 + \text{INR}} \right) & 0 \leq \alpha < \frac{2}{3} \\ \frac{1}{4} \log (1 + \text{SNR}) \\ \quad + \frac{1}{4} \log \left( 1 + \frac{\text{SNR}}{1 + \text{INR}} \right) & \frac{2}{3} \leq \alpha < 1 \\ \frac{1}{4} \log (1 + \text{SNR} + \text{INR}) & 1 \leq \alpha < 2 \\ \frac{1}{2} \log (1 + \text{SNR}) & 2 \leq \alpha. \end{cases} \quad (46)$$

Since we are only after an approximate capacity characterization, we further upper bound  $C_{\text{SYM}}$  as

$$C_{\text{SYM}} \leq \begin{cases} \frac{1}{2} \log \left( 1 + \frac{\text{SNR}}{1 + \text{INR}} \right) + 1 & 0 \leq \alpha < \frac{1}{2} \\ \frac{1}{2} \log^+ (\text{INR}) + 1 & \frac{1}{2} \leq \alpha < \frac{2}{3} \\ \frac{1}{2} \log^+ \left( \frac{\text{SNR}}{\sqrt{\text{INR}}} \right) + 1 & \frac{2}{3} \leq \alpha < 1 \\ \frac{1}{4} \log^+ (\text{INR}) + 1 & 1 \leq \alpha < 2 \\ \frac{1}{2} \log (1 + \text{SNR}) & 2 \leq \alpha. \end{cases} \quad (47)$$

for all values of SNR.

## VI. ACHIEVABLE SCHEMES

This section introduces two simple achievable schemes for reliable communication over the symmetric  $K$ -user interference channel that are based on nested lattice codes. These schemes are then shown to approximately achieve  $C_{\text{SYM}}$ , the symmetric capacity of the channel, for all channel gains  $g$ , except for an outage set of bounded measure. This outage set is explicitly characterized.

We begin by describing the two schemes and deriving their achievable symmetric rates. These rates are given in terms of the optimal computation rates corresponding to a certain effective multiple access channel, i.e., the rates are given as a solution to an optimization problem. This optimization problem, which amounts to finding the optimal coefficient

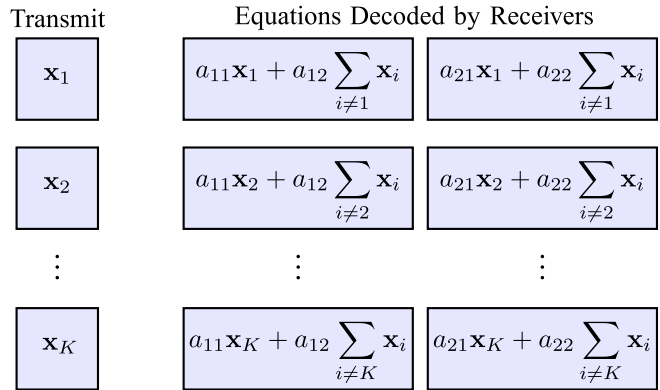


Fig. 11. Illustration of the single-layer lattice scheme. Each transmitter sends a codeword drawn from a common lattice. Each receiver decodes two equations of the codewords, which it can then solve for its desired message.

vectors, can be efficiently solved numerically, as described in Section III-D. Figure 7 shows our achievable rates for the three-user symmetric interference channel as a function of the interference level  $g$ , for several values of SNR. It is evident that the obtained rates significantly improve over time-sharing even for moderate values of SNR.

In order to establish the approximate optimality of these schemes, we derive explicit lower bounds on the rates they achieve which depend only on the SNR and INR. As in the two-user case, the symmetric capacity exhibits a different behavior for different regimes of interference strength, characterized by the parameter  $\alpha$ .

We now present the two achievable schemes. The first achieves the approximate symmetric capacity in the noisy, strong, and very strong interference regimes, while the second achieves the approximate symmetric capacity in the weak and moderately weak interference regimes.

*First scheme - A single-layer lattice code:* A pair of nested lattices  $\Lambda \subseteq \Lambda_1$  is utilized to construct the codebook  $\mathcal{L} = \Lambda_1 \cap \Lambda$  of rate  $R_{\text{SYM}}$ . All users encode their messages using this codebook. Since all interferers arrive at the  $k$ th receiver with the same gain, they will be aligned into one effective lattice point. Thus, the  $K$ -user MAC seen by the  $k$ th receiver becomes an effective two-user MAC of the form defined in Section IV-C (see Example 3)

$$\mathbf{y}_k = \mathbf{x}_{\text{eff},k1} + g\mathbf{x}_{\text{eff},k2} + \mathbf{z}_k, \quad (48)$$

where  $\mathbf{x}_{\text{eff},k1} = \mathbf{x}_k$ ,  $\mathbf{x}_{\text{eff},k2} = \sum_{i \neq k} \mathbf{x}_i$  are the effective users,  $b_{\text{eff},1}^2 = 1$ ,  $b_{\text{eff},2}^2 = K - 1$  are the effective users' weights and  $\mathbf{g} = [1 \ g]^T$  is the vector of channel gains.

The next theorem gives an achievable rate region for the  $K$ -user interference channel when each receiver jointly decodes both the effective user  $\mathbf{x}_{\text{eff},k1}$  which carries the desired information, and the effective user  $\mathbf{x}_{\text{eff},k2}$  which carries the sum of interfering codewords. The theorem relies on decoding two independent linear combinations of the effective lattice points. See Figure 11 for an illustration. This is in contrast to the successive decoding technique used in [20], where first the interference is decoded and removed, and only then the desired lattice point is decoded.



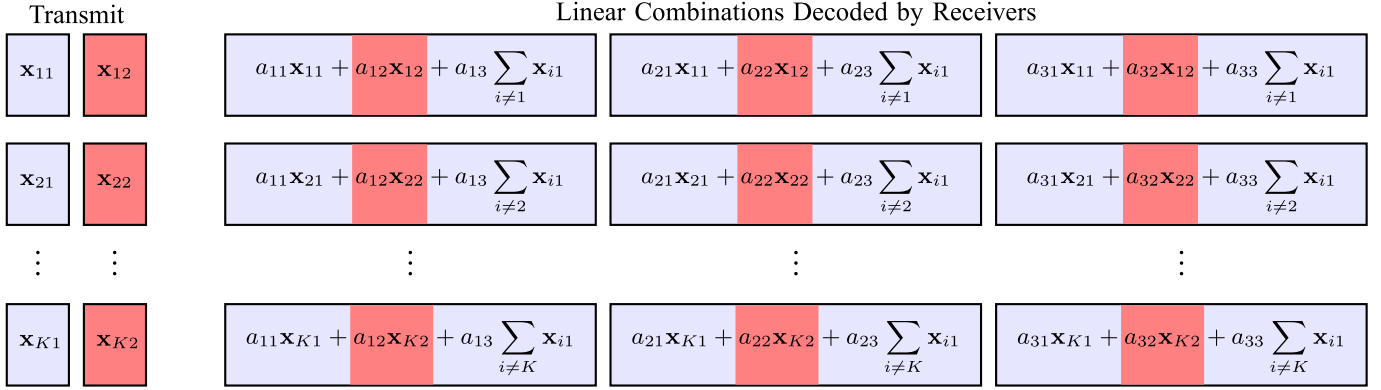


Fig. 12. Illustration of the lattice Han-Kobayashi scheme. Each transmitter sends a public (blue) and a private (red) lattice codeword. Each receiver decodes three linear combinations of the public codewords as well as its desired private codeword while treating the other private codewords as noise. From these linear combinations, the receivers can infer their desired public and private messages.

*Theorem 9:* Let  $R_{\text{comp},1} \geq R_{\text{comp},2}$  be the optimal computation rates for the effective MAC (48) induced by the symmetric  $K$ -user interference channel (2). Any symmetric rate  $R_{\text{SYM}} < R_{\text{comp},2}$  is achievable for the symmetric  $K$ -user interference channel (2).

*Proof:* Corollary 2 implies that for any symmetric rate  $R_{\text{SYM}} < R_{\text{comp},2}$  there exists a pair of nested lattices  $\Lambda \subseteq \Lambda_1$  such that both effective lattice points can be decoded at each receiver. Since the first effective user  $\mathbf{x}_{\text{eff},k1}$  carries all the desired information for the  $k$ th receiver, it follows that any  $R_{\text{SYM}} < R_{\text{comp},2}$  is achievable. ■

The next theorem gives an achievable rate region for the  $K$ -user interference channel when each receiver decodes only its desired codeword, while treating all other interfering codewords as noise. This theorem can be trivially proved using i.i.d. Gaussian codebooks. Nevertheless, we prove the theorem using nested lattice codebooks for completeness.

*Theorem 10:* Any symmetric rate satisfying

$$R_{\text{SYM}} < \frac{1}{2} \log \left( 1 + \frac{\text{SNR}}{1 + (K-1)g^2 \text{SNR}} \right)$$

is achievable for the symmetric  $K$ -user interference channel (2).

*Proof:* Decoding  $\mathbf{x}_k$  at the  $k$ th receiver of the symmetric  $K$ -user interference channel (2), while treating all other users as noise, is equivalent to decoding the linear combination with coefficient vector  $\mathbf{a} = [1 \ 0]^T$  in the effective two-user MAC (48). Therefore, any symmetric rate satisfying  $R_{\text{SYM}} < R_{\text{comp}}(\mathbf{g}, [1 \ 0]^T, \mathbf{B})$  is achievable. The effective noise variance for decoding this linear combination is found using (41) to be

$$\sigma_{\text{eff}}^2(\mathbf{g}, [1 \ 0]^T, \mathbf{B}) = \text{SNR} \left( 1 + \frac{\text{SNR}}{1 + (K-1)g^2 \text{SNR}} \right)^{-1},$$

which, using (43), implies that

$$R_{\text{comp}}(\mathbf{g}, [1 \ 0]^T, \mathbf{B}) = \frac{1}{2} \log \left( 1 + \frac{\text{SNR}}{1 + (K-1)g^2 \text{SNR}} \right).$$

For the two-user case, it is known that in the weak and moderately weak interference regimes each receiver should decode only part of the message transmitted by the other

user [5]. A natural extension of this Han-Kobayashi [15] approach to the  $K$ -user case is for each receiver to decode linear combinations that only include parts of the interfering messages. This is enabled by using a superposition of two lattice codewords at each transmitter, as we describe next. See Figure 12 for an illustration.

*Second scheme - Lattice Han-Kobayashi:* This scheme employs a chain of nested lattices  $\Lambda \subseteq \Lambda_2 \subseteq \Lambda_1$  to construct two codebooks  $\mathcal{L}_1$  and  $\mathcal{L}_2$  with rates  $R_1$  and  $R_2$ , respectively. Each user  $k$  splits its message  $w_k$  into two messages, a public message  $w_{k1}$  that is mapped into a codeword  $\mathbf{x}_{k1}$  from  $\mathcal{L}_1$  and a private message  $w_{k2}$  that is mapped into a codeword  $\mathbf{x}_{k2}$  from  $\mathcal{L}_2$ . It is convenient to treat each user  $k$  as two virtual users with codewords  $\mathbf{x}_{k1}$  and  $\mathbf{x}_{k2}$  that carry messages  $w_{k1}$  and  $w_{k2}$ , respectively. User  $k$  transmits a superposition of its virtual users' codewords,

$$\mathbf{x}_k = \sqrt{1 - \gamma^2} \mathbf{x}_{k1} + \gamma \mathbf{x}_{k2},$$

for  $\gamma \in [0, 1)$ . The signal seen by the  $k$ th receiver is

$$\mathbf{y}_k = \sqrt{1 - \gamma^2} \mathbf{x}_{k1} + \gamma \mathbf{x}_{k2} + g \sqrt{1 - \gamma^2} \sum_{i \neq k} \mathbf{x}_{i1} + g \gamma \sum_{i \neq k} \mathbf{x}_{i2} + \mathbf{z}_k, \quad (49)$$

which induces the effective four-user MAC

$$\mathbf{y}_k = \sqrt{1 - \gamma^2} \mathbf{x}_{\text{eff},k1} + \gamma \mathbf{x}_{\text{eff},k2} + g \sqrt{1 - \gamma^2} \mathbf{x}_{\text{eff},k3} + g \gamma \mathbf{x}_{\text{eff},k4} + \mathbf{z}_k \quad (50)$$

with effective users  $\mathbf{x}_{\text{eff},k1} = \mathbf{x}_{k1}$ ,  $\mathbf{x}_{\text{eff},k2} = \mathbf{x}_{k2}$ ,  $\mathbf{x}_{\text{eff},k3} = \sum_{i \neq k} \mathbf{x}_{i1}$  and  $\mathbf{x}_{\text{eff},k4} = \sum_{i \neq k} \mathbf{x}_{i2}$ . The effective users' weights are  $b_{\text{eff},1}^2 = 1$ ,  $b_{\text{eff},2}^2 = 1$ ,  $b_{\text{eff},3}^2 = K - 1$  and  $b_{\text{eff},4}^2 = K - 1$ , and

$$\mathbf{g} = \left[ \sqrt{1 - \gamma^2} \quad \gamma \quad g \sqrt{1 - \gamma^2} \quad g \gamma \right]^T$$

is the vector of effective channel gains.

The receiver aims to decode the effective codewords  $\mathbf{x}_{\text{eff},k1}$ ,  $\mathbf{x}_{\text{eff},k2}$  and  $\mathbf{x}_{\text{eff},k3}$  while treating the fourth effective codeword  $\mathbf{x}_{\text{eff},k4}$  as noise. The next lemma will be useful for the derivation of rates achieved by this scheme. Its proof is given in Appendix C.

*Lemma 1:* Consider the effective  $L$ -user MAC (33), where the decoder is only interested in the first  $L-1$  effective lattice points  $\mathbf{t}_{\text{eff},1}, \dots, \mathbf{t}_{\text{eff},L-1}$  and let  $\kappa = 1/\sqrt{1 + \text{SNR}g_L^2 b_{\text{eff},L}^2}$ . Any rate tuple achievable for decoding  $\mathbf{t}_{\text{eff},1}, \dots, \mathbf{t}_{\text{eff},L-1}$  over the effective  $(L-1)$ -user MAC

$$\sum_{\ell=1}^{L-1} \kappa g_{\ell} \mathbf{x}_{\text{eff},\ell} + \mathbf{z} \quad (51)$$

is also achievable for decoding the desired  $L-1$  lattice points over (33).

The next theorem gives the achievable rate region for the lattice Han-Kobayashi scheme.

*Theorem 11:* Let  $\kappa(\gamma) = 1/\sqrt{1 + \text{SNR}g^2\gamma^2(K-1)}$  and consider the effective MAC

$$\mathbf{y}_k = \kappa(\gamma)\sqrt{1-\gamma^2}\mathbf{x}_{\text{eff},k1} + \kappa(\gamma)\gamma\mathbf{x}_{\text{eff},k2} + \kappa(\gamma)g\sqrt{1-\gamma^2}\mathbf{x}_{\text{eff},k3} + \mathbf{z}_k, \quad (52)$$

with effective channel vector

$$\mathbf{g} = \left[ \kappa(\gamma)\sqrt{1-\gamma^2} \quad \kappa(\gamma)\gamma \quad \kappa(\gamma)g\sqrt{1-\gamma^2} \right]^T,$$

and effective users' weights  $b_{\text{eff},1}^2 = 1$ ,  $b_{\text{eff},2}^2 = 1$ , and  $b_{\text{eff},3}^2 = K-1$ . Let  $\{\mathbf{a}_1(\gamma), \mathbf{a}_2(\gamma), \mathbf{a}_3(\gamma)\}$  and  $R_{\text{comp},1}(\gamma) \geq R_{\text{comp},2}(\gamma) \geq R_{\text{comp},3}(\gamma)$  be the optimal coefficient vectors and computation rates, respectively. Any symmetric rate satisfying

$$R_{\text{SYM}} < \max_{\gamma \in [0,1]} R_{\text{comp},2}(\gamma) + R_{\text{comp},3}(\gamma)$$

is achievable for the symmetric  $K$ -user interference channel (2).

*Proof:* The receiver is only interested in the effective lattice points  $\mathbf{t}_{\text{eff},k1}, \mathbf{t}_{\text{eff},k2}$ . Nevertheless, we require that it decodes the three effective lattice points  $\mathbf{t}_{\text{eff},k1}, \mathbf{t}_{\text{eff},k2}$  and  $\mathbf{t}_{\text{eff},k3}$ . Due to Lemma 1, any rate tuple that is achievable over the effective channel (52) is also achievable for decoding  $\mathbf{t}_{\text{eff},k1}, \mathbf{t}_{\text{eff},k2}$  and  $\mathbf{t}_{\text{eff},k3}$  from the original effective channel (50) induced by the lattice Han-Kobayashi scheme.

Note that  $\mathbf{t}_{\text{eff},k1}$  and  $\mathbf{t}_{\text{eff},k3}$  are points from the same codebook  $\mathcal{L}_1$  with rate  $R_1$ , and  $\mathbf{t}_{\text{eff},k2}$  is a codeword from  $\mathcal{L}_2$  with rate  $R_2$ .

Consider a compute-and-forward transform coefficient matrix  $\mathbf{A}(\gamma) = [\mathbf{a}_1(\gamma) \quad \mathbf{a}_2(\gamma) \quad \mathbf{a}_3(\gamma)]^T$  for (52). For any full-rank matrix there exists at least one order of pseudo-triangularization. Therefore, there exists a pseudo-triangularization of  $\mathbf{A}(\gamma)$  with at least one permutation vector  $\pi$ .

Consider first the case where  $\pi(3) = 2$ , i.e., the effective lattice point  $\mathbf{t}_{\text{eff},2}$  is the last to be removed in the algebraic successive cancellation decoding procedure of the compute-and-forward transform. According to Theorem 8, for any  $R_1 < R_{\text{comp},2}(\gamma)$  and  $R_2 < R_{\text{comp},3}(\gamma)$  there exists a chain  $\Lambda \subseteq \Lambda_2 \subseteq \Lambda_1$  such that  $\mathbf{t}_{\text{eff},k1}, \mathbf{t}_{\text{eff},k2}$  and  $\mathbf{t}_{\text{eff},k3}$  can be decoded from the effective channel (52) via the compute-and-forward transform.

Otherwise,  $\pi(1) = 2$  or  $\pi(2) = 2$ , which means that the effective lattice point  $\mathbf{t}_{\text{eff},2}$  is either removed first or second from the preceding linear combinations in the algebraic successive cancellation decoding procedure of the compute-

and-forward transform. According to Theorem 8 for any  $R_1 < R_{\text{comp},3}(\gamma)$  and  $R_2 < R_{\text{comp},2}(\gamma)$  there exists a chain  $\Lambda \subseteq \Lambda_2 \subseteq \Lambda_1$  such that  $\mathbf{t}_{\text{eff},k1}, \mathbf{t}_{\text{eff},k2}$  and  $\mathbf{t}_{\text{eff},k3}$  can be decoded from the effective channel (52) via the compute-and-forward transform.

Since  $R_{\text{SYM}} = R_1 + R_2$ , and  $\gamma$  can be chosen such as to maximize  $R_{\text{SYM}}$ , the theorem is proved.  $\blacksquare$

The problem of optimizing the power allocation  $\gamma$  between the private and public codewords, played a major role in the approximation of the two-user interference channel capacity [5]. Here, we follow the approach of [5] and choose  $\gamma$  such that, at each unintended receiver, the received power of each private codeword is equal to that of the additive noise. Specifically, in the sequel we set  $\gamma^2 = 1/(g^2\text{SNR})$ . While this choice of  $\gamma$  may be sub-optimal, it suffices to develop our capacity approximations in closed form. The achievable symmetric rate for  $\gamma^2 = 1/(g^2\text{SNR})$  is given in the following corollary to Theorem 11.

*Corollary 3:* Assume  $g^2\text{SNR} > 1$  and consider the effective MAC

$$\mathbf{y}_k = \sqrt{\frac{g^2\text{SNR}-1}{K \cdot g^2\text{SNR}}}\mathbf{x}_{\text{eff},k1} + \sqrt{\frac{1}{K \cdot g^2\text{SNR}}}\mathbf{x}_{\text{eff},k2} + g\sqrt{\frac{g^2\text{SNR}-1}{K \cdot g^2\text{SNR}}}\mathbf{x}_{\text{eff},k3} + \mathbf{z}_k, \quad (53)$$

with effective channel vector

$$\mathbf{g} = \left[ \sqrt{\frac{g^2\text{SNR}-1}{K \cdot g^2\text{SNR}}} \quad \sqrt{\frac{1}{K \cdot g^2\text{SNR}}} \quad g\sqrt{\frac{g^2\text{SNR}-1}{K \cdot g^2\text{SNR}}} \right]^T, \quad (54)$$

and effective users' weights  $b_{\text{eff},1}^2 = 1$ ,  $b_{\text{eff},2}^2 = 1$ , and  $b_{\text{eff},3}^2 = K-1$ . Let  $\{\mathbf{a}_1^{\text{HK}}, \mathbf{a}_2^{\text{HK}}, \mathbf{a}_3^{\text{HK}}\}$  and  $R_{\text{comp},1}^{\text{HK}} \geq R_{\text{comp},2}^{\text{HK}} \geq R_{\text{comp},3}^{\text{HK}}$  be the optimal coefficient vectors and computation rates for this effective MAC. Any symmetric rate

$$R_{\text{SYM}} < R_{\text{comp},2}^{\text{HK}} + R_{\text{comp},3}^{\text{HK}}$$

is achievable for the symmetric  $K$ -user interference channel (2).

Computing the achievable rates given by Theorem 9 and Corollary 3 requires finding the optimal computation rates for the effective MACs (48) and (53), which involves solving an integer least-squares optimization problem (see Section III-D). In the remainder of this section, we derive lower bounds on these achievable rates that depend only on the values of SNR and INR and can therefore be directly compared to the upper bounds (47). To simplify the exposition, we assume  $g > 0$  in the sequel, although all results easily follow for  $g < 0$  as well.

#### A. Very Strong Interference Regime

The very strong interference regime corresponds to  $g^2 \geq \text{SNR}$ . The sum capacity for

$$g^2 \geq \frac{(\text{SNR} + 1)^2}{\text{SNR}}, \quad (55)$$

which covers almost all of this regime was characterized exactly by Sridharan *et al.* [20] using a lattice encoding scheme very similar to the one used in Theorem 9. The key difference

$$\begin{aligned} \sigma_{\text{eff}}^2(\mathbf{g}, [0 \ 1]^T, \mathbf{B}) &= \text{SNR} \cdot [0 \ 1] \left( \begin{bmatrix} 1 & 0 \\ 0 & K-1 \end{bmatrix} - \frac{\text{SNR} \begin{bmatrix} 1 & 0 \\ 0 & K-1 \end{bmatrix} \begin{bmatrix} 1 \\ g \end{bmatrix} \begin{bmatrix} 1 & g \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & K-1 \end{bmatrix}}{1 + \text{SNR} \begin{bmatrix} 1 & 0 \\ 0 & K-1 \end{bmatrix} \begin{bmatrix} 1 \\ g \end{bmatrix}} \right) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \text{SNR} \cdot \frac{(K-1)(1 + \text{SNR})}{1 + \text{SNR} + (K-1)g^2 \text{SNR}} \end{aligned} \quad (57)$$

is that in [20] each receiver decodes successively: it first decodes the sum of interfering codewords and then subtracts it in order to get a clean view of the desired signal. Recall that in our scheme, each receiver decodes two linear combinations of its signal and the interference.

A slight modification of the scheme given in [20] suffices to achieve the interference-free capacity to within a (small) constant gap for all  $g^2 > \text{SNR}$ .<sup>7</sup> Nevertheless, rather than using the results of [20], we now proceed to lower bound the achievable rate of Theorem 9 for the case  $\alpha \geq 2$ , i.e.,  $g^2 \geq \text{SNR}$ . We do this in order to show that our lattice encoding and decoding framework suffices to achieve the approximate capacity in all regimes.

Using the single-layer scheme presented above, the channel seen by each receiver is converted to an effective two-user MAC (48). Let  $R_{\text{comp},1} \geq R_{\text{comp},2}$  be the optimal computation rates for this effective channel. Theorem 9 implies that any  $R_{\text{SYM}} < R_{\text{comp},2}$  is achievable, and hence, it suffices to lower bound  $R_{\text{comp},2}$ . We have

$$R_{\text{comp},2} = R_{\text{comp},1} + R_{\text{comp},2} - R_{\text{comp},1}.$$

Applying Theorem 7 to the effective MAC (48), we find that the sum of the optimal computation rates is lower bounded by

$$R_{\text{comp},1} + R_{\text{comp},2} \geq \frac{1}{2} \log \left( \frac{1 + \text{SNR}(1 + g^2(K-1))}{K-1} \right) - 1.$$

Therefore

$$R_{\text{comp},2} \geq \frac{1}{2} \log \left( \frac{1 + \text{SNR}(1 + g^2(K-1))}{K-1} \right) - 1 - R_{\text{comp},1}, \quad (56)$$

and it suffices to upper bound  $R_{\text{comp},1}$ .

Let  $R_{\text{comp}}(\mathbf{g}, [0 \ 1]^T, \mathbf{B})$  be the computation rate for decoding the linear combination with coefficient vector  $\mathbf{a} = [0 \ 1]^T$  over the effective MAC (48) with  $\mathbf{g} = [1 \ g]^T$  and  $\mathbf{B} = \text{diag}(1, K-1)$ . The effective noise variance for the coefficient vector  $\mathbf{a} = [0 \ 1]^T$ , which is calculated using (41), is given in (57), shown at the top of the page. Substituting  $\sigma_{\text{eff}}^2(\mathbf{g}, [0 \ 1]^T, \mathbf{B})$  into (40) gives

$$R_{\text{comp}}(\mathbf{g}, [0 \ 1]^T, \mathbf{B}) = \frac{1}{2} \log \left( \frac{1 + \text{SNR}(1 + g^2(K-1))}{(K-1)(1 + \text{SNR})} \right). \quad (58)$$

The coefficient vector  $\mathbf{a} = [0 \ 1]^T$  either gives the highest computation rate or not. If it does, i.e., if  $R_{\text{comp},1} =$

$R_{\text{comp}}(\mathbf{g}, [0 \ 1]^T, \mathbf{B})$ , substituting (58) into (56) gives

$$R_{\text{comp},2} \geq \frac{1}{2} \log(1 + \text{SNR}) - 1. \quad (59)$$

It follows from (18) and (20) that  $[0 \ 1]^T$  yields the highest computation rate among all integer coefficient vectors that are linearly dependent with it. Thus, if  $R_{\text{comp},1} \neq R_{\text{comp}}(\mathbf{g}, [0 \ 1]^T, \mathbf{B})$ , any coefficient vector that attains  $R_{\text{comp},1}$  must be linearly independent of  $[0 \ 1]^T$ . It follows that

$$\begin{aligned} R_{\text{comp},2} &\geq R_{\text{comp}}(\mathbf{g}, [0 \ 1]^T, \mathbf{B}) \\ &> \frac{1}{2} \log \left( g^2 \frac{\text{SNR}}{1 + \text{SNR}} \right). \end{aligned} \quad (60)$$

Taking the minimum of the two bounds (59) and (60), and using the fact that  $g^2 \geq \text{SNR}$  we obtain

$$\begin{aligned} R_{\text{comp},2} &\geq \min \left( \frac{1}{2} \log(1 + \text{SNR}) - 1, \frac{1}{2} \log \left( \frac{\text{SNR}^2}{1 + \text{SNR}} \right) \right)^+ \\ &\geq \frac{1}{2} \log(1 + \text{SNR}) - 1. \end{aligned}$$

Thus, in the very strong regime, any symmetric rate satisfying

$$R_{\text{SYM}} < \frac{1}{2} \log(1 + \text{SNR}) - 1 \quad (61)$$

is achievable, which is within 1 bit of the outer bound (47).

### B. Strong Interference Regime

The strong interference regime corresponds to  $1 \leq \alpha < 2$ , or equivalently  $1 \leq g^2 < \text{SNR}$ . As in the previous subsection, we lower bound  $R_{\text{comp},2}$  in order to obtain a closed-form expression for the achievable symmetric rate. In contrast to the very strong interference regime, where the lower bound on  $R_{\text{comp},2}$  is valid for any  $g^2 \geq \text{SNR}$ , here we must exclude certain channel gains in order to get a constant gap from the outer bound (47). That is, the lower bounds we derive for the strong interference regime are only valid for a predefined subset of the interval  $g^2 \in [1, \text{SNR})$ . As we increase the measure of this subset, our approximation gap worsens. This somewhat strange behavior is to be expected from the existing literature. The results of [8] and [21] show that for the  $K$ -user interference channel the DoF are discontinuous at the rationals. The notion of DoF corresponds to  $\alpha \approx 1$ . Since the strong interference regime contains values of  $\alpha$  near 1, we cannot expect to achieve rates which are a constant gap from the upper bounds of [5] for all values of  $g$ . Instead, we show that these upper bounds can be approached up to a constant gap for all  $1 \leq g^2 < \text{SNR}$  except for some outage set whose measure can be controlled at the price of increasing the gap. We will

<sup>7</sup>Namely, if  $\text{SNR} \leq g^2 < \text{SNR} + 2 + 1/\text{SNR}$  all transmitters can reduce their transmission power by a small factor such that the very strong interference condition from [20] is satisfied. This power reduction results in a constant rate-loss.

see a similar phenomenon when we analyze the moderately weak interference regime.

From (56), we have

$$\begin{aligned} R_{\text{comp},2} &\geq \frac{1}{2} \log \left( \frac{1 + \text{SNR}(1 + g^2(K-1))}{K-1} \right) - 1 - R_{\text{comp},1} \\ &> \frac{1}{2} \log (g^2 \text{SNR}) - 1 - R_{\text{comp},1} \end{aligned} \quad (62)$$

The optimal computation rate for the effective MAC (48) can be written, by substituting  $\mathbf{g} = [1 \ g]^T$  and  $\mathbf{B} = \text{diag}(1, K-1)$  into (38), as

$$\begin{aligned} R_{\text{comp},1} &= \frac{1}{2} \log(\text{SNR}) - \frac{1}{2} \log(\sigma_g^2) \\ \sigma_g^2 &= \min_{\beta, a_1, a_2} \left( (\beta - a_1)^2 \text{SNR} + (\beta g - a_2)^2 (K-1) \text{SNR} + \beta^2 \right) \end{aligned} \quad (63)$$

where  $\sigma_g^2$  is the effective noise variance and the minimization is over  $\beta \in \mathbb{R}$ , and  $\mathbf{a} = [a_1 \ a_2]^T \in \mathbb{Z}^2 \setminus \mathbf{0}$ . Substituting (63) into (62) and applying Theorem 9, we see that any symmetric rate satisfying

$$R_{\text{SYM}} < \frac{1}{2} \log(g^2) + \frac{1}{2} \log(\sigma_g^2) - 1, \quad (64)$$

is achievable over the  $K$ -user interference channel. Thus, in order to obtain a lower bound on  $C_{\text{SYM}}$  it suffices to lower bound  $\sigma_g^2$ .

*Remark 13:* It may at first seem counterintuitive that the symmetric rate expression in (64) is an increasing function of the effective noise variance  $\sigma_g^2$  for the highest computation rate  $R_{\text{comp},1}$ . However, as discussed in Section I-A, when  $\sigma_g^2$  is small, the desired signal and the interference are aligned. From another perspective, if the channel vector  $\mathbf{g}$  is very close to the integer vector  $\mathbf{a}$  (after scaling by  $\beta$ ), then it must be far from the integer coefficient vector that determines  $R_{\text{comp},2}$ , which in turn determines  $R_{\text{SYM}}$ . Thus, the best performance is attained when the channel vector is hard to approximate with an integer vector. Building on this idea, the lower bound derived below connects our problem to a Diophantine approximation problem<sup>8</sup> and characterizes the outage set in terms of channel gains that are well-approximated by rationals.

The effective noise  $\sigma_g^2$  can be bounded as

$$\sigma_g^2 \geq \min_{\beta, a_1, a_2} \left( ((\beta - a_1)^2 + (\beta g - a_2)^2) \text{SNR} + \beta^2 \right). \quad (65)$$

We first hold  $\beta$  constant and minimize over  $a_1, a_2$ . If  $|\beta| \geq 1/(2g)$ , the optimal choices for the integers  $a_1, a_2$  are

$$a_1 = \lfloor \beta \rfloor, \quad a_2 = \lfloor \beta g \rfloor. \quad (66)$$

If  $|\beta| < 1/(2g)$ , rounding the gains will set both  $a_1$  and  $a_2$  to zero, which is not allowed. Since  $g \geq 1$ , the optimal choice is

$$a_1 = 0, \quad a_2 = \text{sign}(\beta). \quad (67)$$

Now, we are left with the problem of minimizing (65) over  $\beta$ . Rather than explicitly solving this minimization problem, we give a lower bound on its solution. We do this by splitting

the real line into three intervals, and lower bounding  $\sigma_g^2$  for all values of  $\beta$  within each one. Then, we take the minimum over these three bounds.

**Interval 1:**  $0 < |\beta| \leq 1/(2g)$

In this interval it is optimal to set  $a_2 = \text{sign}(\beta)$ . Moreover,  $|\beta g| \leq 1/2$ , and therefore  $|\beta g - a_2| > 1/2$ . Combining this with (65) gives

$$\sigma_g^2 \geq \frac{\text{SNR}}{4}. \quad (68)$$

**Interval 2:**  $1/(2g) < |\beta| \leq 1/2$

Here, it is optimal to set  $a_1 = \lfloor \beta \rfloor = 0$ . Substituting  $a_1 = 0$  in (65) gives

$$\sigma_g^2 \geq \beta^2 \text{SNR} > \frac{\text{SNR}}{4g^2} > \frac{\text{SNR}^{1/2}}{4\sqrt{g^2}}, \quad (69)$$

where the last inequality follows since  $g^2 < \text{SNR}$  in the strong interference regime.

**Interval 3:**  $1/2 < |\beta|$

Since  $|\beta| > 1/2$ , we can write  $\beta = q + \varphi$  where  $q$  is a nonzero integer and  $\varphi \in [-1/2, 1/2)$ . Substituting into (65), we get

$$\begin{aligned} \sigma_g^2 &\geq \min_{\varphi, q, a_1, a_2} \left( (\varphi + q - a_1)^2 \text{SNR} \right. \\ &\quad \left. + (qg - a_2 + \varphi g)^2 \text{SNR} + (\varphi + q)^2 \right) \\ &\geq \min_{\varphi, q, a_2} \left( (\varphi^2 + (qg - a_2 + \varphi g)^2) \text{SNR} + \frac{q^2}{4} \right). \end{aligned} \quad (70)$$

The minimization of (70) with respect to  $\varphi$  (where the constraint  $\varphi \in [-1/2, 1/2)$  is ignored) can be obtained by differentiation. The minimizing value of  $\varphi$  is

$$\varphi^* = -\frac{g}{1+g^2}(qg - a_2).$$

Substituting  $\varphi^*$  into (70) gives

$$\sigma_g^2 \geq \min_{q, a_2} \left( \frac{1}{1+g^2} (qg - a_2)^2 \text{SNR} + \frac{q^2}{4} \right), \quad (71)$$

which, using the fact that  $g^2 \geq 1$ , can be further bounded by

$$\sigma_g^2 \geq \frac{1}{4} \min_{q, a_2} \max \left( \frac{1}{g^2} (qg - a_2)^2 \text{SNR}, q^2 \right). \quad (72)$$

We would like to obtain a lower bound on  $\sigma_g^2$  that is valid for all  $g \notin \mathcal{S}$ , where  $\mathcal{S}$  is an outage set with bounded measure. Consider first the interval  $[b, b+1)$  for some integer  $1 \leq b < \sqrt{\text{SNR}}$ . Define

$$q_{\max, b} \triangleq \frac{1}{\sqrt{b+1/2}} \text{SNR}^{1/4-\delta/2}, \quad (73)$$

for some  $\delta > 0$  to be specified later, and note that  $q_{\max, b}$  is not necessarily an integer. Also, define

$$\Phi_b \triangleq \sqrt{b+1/2} \text{SNR}^{-1/4-\delta/2} \quad (74)$$

and let  $\mathcal{S}_b$  be the set of all values of  $g \in [b, b+1)$  such that the inequality

$$|qg - a_2| < \Phi_b \quad (75)$$

<sup>8</sup>Diophantine approximation refers to the branch of number theory that studies how well real numbers can be approximated by rational numbers.

has at least one solution with integers  $q$  and  $a_2$ , where  $q$  is in the range  $0 < q \leq q_{\max,b}$ . Let  $\bar{\mathcal{S}}_b = [b, b+1] \setminus \mathcal{S}_b$ . By (72), (73), and (74), we have that for all  $g \in \bar{\mathcal{S}}_b$

$$\begin{aligned} \sigma_g^2 &\geq \frac{1}{4} \min \left( \min_{0 < q \leq \lfloor q_{\max,b} \rfloor, a_2} \max \left( \frac{1}{g^2} (qg - a_2)^2 \text{SNR}, q^2 \right), \right. \\ &\quad \left. \min_{\lceil q_{\max,b} \rceil \leq q, a_2} \max \left( \frac{1}{g^2} (qg - a_2)^2 \text{SNR}, q^2 \right) \right) \\ &\geq \frac{1}{4} \min \left( \frac{1}{g^2} \Phi_b^2 \text{SNR}, q_{\max,b}^2 \right) \\ &= \frac{1}{4} \min \left( \frac{b+1/2}{g^2} \text{SNR}^{1/2-\delta}, \frac{1}{b+1/2} \text{SNR}^{1/2-\delta} \right). \end{aligned} \quad (76)$$

Since  $b \geq 1$ , we have that

$$\frac{g}{2} < b + \frac{1}{2} < 2g.$$

Thus, (76) can be further bounded by

$$\sigma_g^2 \geq \frac{1}{8|g|} \text{SNR}^{1/2-\delta}. \quad (77)$$

We now turn to upper bound the Lebesgue measure of the set  $\mathcal{S}_b$ . Our derivation is quite similar to the proof of the convergent part of Khinchine's Theorem [61]. Let  $\mathcal{I} = [-1, 1]$  and define the set

$$\mathcal{T}_b(q) = \left[ \left\{ b, b + \frac{1}{q}, \dots, b + \frac{q-1}{q} \right\} + \frac{\Phi_b}{q} \mathcal{I} \right] \bmod [b, b+1],$$

where the sum of the two sets is a Minkowski sum. Writing the Diophantine approximation problem (75) as

$$\left| g - \frac{a_2}{q} \right| < \frac{\Phi_b}{q}, \quad (78)$$

we see that for a given  $q$  and  $g \in [b, b+1]$  the inequality admits a solution if and only if  $g \in \mathcal{T}_b(q)$ . It follows that

$$\mathcal{S}_b = \bigcup_{q=1}^{\lfloor q_{\max,b} \rfloor} \mathcal{T}_b(q). \quad (79)$$

See Figure 13 for an illustration of the sets  $\mathcal{T}_b(q)$  and  $\mathcal{S}_b$ . Thus, the Lebesgue measure of  $\mathcal{S}_b$  can be upper bounded by

$$\begin{aligned} \mu(\mathcal{S}_b) &= \text{Vol}(\mathcal{S}_b) \\ &\leq \sum_{q=1}^{\lfloor q_{\max,b} \rfloor} \text{Vol}(\mathcal{T}_b(q)) \\ &\leq \sum_{q=1}^{\lfloor q_{\max,b} \rfloor} q \cdot 2 \frac{\Phi_b}{q} \\ &\leq 2q_{\max,b} \Phi_b \\ &= 2\text{SNR}^{-\delta}. \end{aligned} \quad (80)$$

Setting  $\delta = (c+1)/\log(\text{SNR})$  and substituting into (77) and (80) gives

$$\sigma_g^2 \geq \frac{2^{-c}}{16\sqrt{g^2}} \text{SNR}^{1/2} \quad (81)$$

for all  $g \in [b, b+1]$  up to an outage set  $\mathcal{S}_b$  of measure not greater than  $2^{-c}$ .

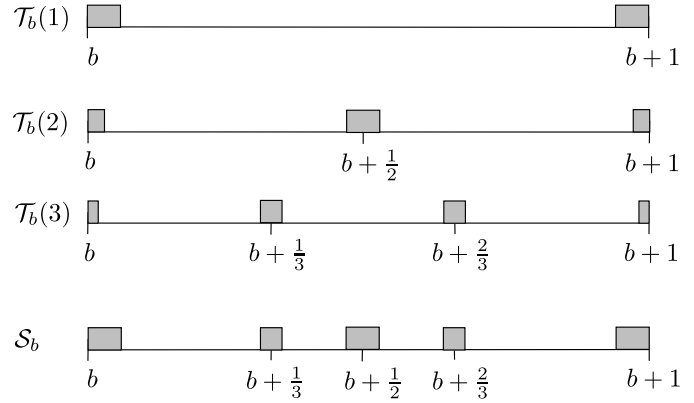


Fig. 13. An illustration of the sets  $\mathcal{T}_b(1)$ ,  $\mathcal{T}_b(2)$ ,  $\mathcal{T}_b(3)$  and their union  $\mathcal{S}_b$ . In this illustration,  $q_{\max,b} = 3$  and  $\Phi_b = 1/16$ .

Combining the three bounds (68), (69), and (81) yields

$$\begin{aligned} \sigma_g^2 &\geq \min \left( \frac{\text{SNR}}{4}, \frac{1}{4\sqrt{g^2}} \text{SNR}^{1/2}, \frac{2^{-c}}{16\sqrt{g^2}} \text{SNR}^{1/2} \right) \\ &\geq \frac{2^{-c}}{16\sqrt{g^2}} \text{SNR}^{1/2} \end{aligned} \quad (82)$$

for all  $g \in [b, b+1]$  up to an outage set  $\mathcal{S}_b$  with measure at most  $2^{-c}$ .

Combining (64) and (82) we see that, for all  $g \in [b, b+1]$  up to an outage set  $\mathcal{S}_b$  of measure not greater than  $2^{-c}$ , any symmetric rate satisfying

$$\begin{aligned} R_{\text{SYM}} &< \frac{1}{4} \log(g^2 \text{SNR}) - \frac{c}{2} - 3 \\ &= \frac{1}{4} \log(\text{INR}) - \frac{c}{2} - 3 \end{aligned} \quad (83)$$

is achievable. We conclude that the symmetric rate (83) is achievable for all channel gains in the strong interference regime except for an outage set whose measure is a fraction of  $2^{-c}$  of the interval  $1 \leq |g| < \sqrt{\text{SNR}}$ , for any  $c > 0$ .

*Remark 14:* In the high SNR limit, the total DoF of the symmetric  $K$ -user IC drops below  $K/2$  when the channel gain  $g$  is rational [8], [21]. At finite SNR, we observe that the channel gains in the interval  $g \in [b, b+1]$  that fall within the outage set are the ones close to rational numbers with denominator smaller than  $q_{\max,b}$ . Since  $q_{\max,b}$  scales as  $\text{SNR}^{1/4}/\sqrt{|g|}$ , only channel gains close to rational numbers with denominator smaller than  $\text{SNR}^{1/4}$  may result in outage. Moreover, the sensitivity of the achievable rate to the ‘‘rationality’’ of  $g$  decreases as  $g$  increases.

*Remark 15:* We note that for any  $c > 0$  the set of channel coefficients that fall in the outage set

$$\mathcal{S} = \bigcup_{b=1}^{\lfloor \sqrt{\text{SNR}} \rfloor} \mathcal{S}_b,$$

can be easily determined by setting  $\delta = (c+1)/\log(\text{SNR})$  in (73), (74), and applying (79).

### C. Moderately Weak Interference Regime

The moderately weak interference regime is characterized by  $2/3 \leq \alpha < 1$ , or equivalently,  $\text{SNR}^{-1/3} \leq g^2 < 1$ . As in the strong interference regime, we show the achievability of symmetric rates which are a constant gap from the upper bound for a certain fraction of the channel gains. As opposed to the very strong and strong interference regimes, where a single-layered lattice scheme suffices to achieve the approximate capacity, here we will need the second scheme, which employs two layers of lattice codes at each transmitter.

We will set the power of the private lattice codewords so that they are perceived at noise level at the unintended receivers. Let  $g_1$ ,  $g_2$  and  $g_3$  be the channel gains in the effective three-user MAC (53) from Corollary 3, and recall that, for this effective channel, the effective weight matrix is  $\mathbf{B} = \text{diag}(1, 1, K-1)$ . Let  $R_{\text{comp},1}^{\text{HK}} \geq R_{\text{comp},2}^{\text{HK}} \geq R_{\text{comp},3}^{\text{HK}}$  be the three optimal computation rates for the effective channel (53). Corollary 3 states that any symmetric rate satisfying  $R_{\text{SYM}} < R_{\text{comp},2}^{\text{HK}} + R_{\text{comp},3}^{\text{HK}}$  is achievable, and we now turn to lower bounding this achievable rate in closed form. First, note that

$$R_{\text{comp},2}^{\text{HK}} + R_{\text{comp},3}^{\text{HK}} = \sum_{m=1}^3 R_{\text{comp},m}^{\text{HK}} - R_{\text{comp},1}^{\text{HK}}. \quad (84)$$

By applying Theorem 7 to the effective channel (53), we obtain the following lower bound on the sum of optimal computation rates,

$$\begin{aligned} & \sum_{m=1}^3 R_{\text{comp},m}^{\text{HK}} \\ & \geq \frac{1}{2} \log \left( \frac{1 + \text{SNR}(g_1^2 + g_2^2 + (K-1)g_3^2)}{K-1} \right) - \frac{3}{2} \log(3) \\ & > \frac{1}{2} \log(\text{SNR}(g_1^2 + g_2^2)) - \frac{1}{2} \log(3^3(K-1)) \\ & = \frac{1}{2} \log(\text{SNR}) - \frac{1}{2} \log(27K(K-1)), \end{aligned} \quad (85)$$

where we have used the fact that  $g_1^2 + g_2^2 = 1/K$  in the last equality. The highest computation rate can be written as

$$R_{\text{comp},1}^{\text{HK}} = \frac{1}{2} \log(\text{SNR}) - \frac{1}{2} \log(\sigma_{\text{HK}}^2) \quad (86)$$

where  $\sigma_{\text{HK}}^2$  is given in (87), shown at the bottom of the page. The minimization in (87) is performed over all  $\beta \in \mathbb{R}$  and  $\mathbf{a}_1^{\text{HK}} = [a_1 \ a_2 \ a_3]^T \in \mathbb{Z}^3 \setminus \mathbf{0}$ . Combining (84), (86), and (87) and applying Corollary 3, we see that any symmetric rate satisfying

$$R_{\text{SYM}} < \frac{1}{2} \log(\sigma_{\text{HK}}^2) - \frac{1}{2} \log(27K(K-1)) \quad (88)$$

is achievable for the  $K$ -user interference channel. Therefore, it suffices to lower bound the effective noise variance  $\sigma_{\text{HK}}^2$ .

Substituting

$$\beta = \sqrt{\frac{Kg^2\text{SNR}}{g^2\text{SNR}-1}} \tilde{\beta}$$

in (87), which is allowed since  $\beta$  can take any value in  $\mathbb{R}$ , gives

$$\sigma_{\text{HK}}^2 = \min_{\tilde{\beta}, a_1, a_2, a_3} \left( \tilde{\beta}^2 \cdot \frac{Kg^2\text{SNR}}{g^2\text{SNR}-1} + (\tilde{\beta} - a_1)^2 \text{SNR} + \left( \frac{\tilde{\beta}}{\sqrt{g^2\text{SNR}-1}} - a_2 \right)^2 \text{SNR} + (K-1)(\tilde{\beta}g - a_3)^2 \text{SNR} \right). \quad (89)$$

In the sequel, we assume<sup>9</sup>  $\text{SNR} > 4$ . With this assumption,  $\sqrt{g^2\text{SNR}-1} > 1$  for all  $g^2 \geq \text{SNR}^{-1/3}$ , i.e., for all values of  $g$  in the moderately weak interference regime. We will also use the fact that the inequality  $\sqrt{g^2\text{SNR}-1} > 1$  continues to hold for all  $g^2 \geq \text{SNR}^{-1/2}$ , i.e., for all values of  $g$  in the weak interference regime. This implies that  $g^2\text{SNR}/(g^2\text{SNR}-1) > 1$  and hence (89) can be lower bounded as

$$\begin{aligned} \sigma_{\text{HK}}^2 & \geq \min_{\tilde{\beta}, a_1, a_2, a_3} \left( K\tilde{\beta}^2 + (\tilde{\beta} - a_1)^2 \text{SNR} + \left( \frac{\tilde{\beta}}{\sqrt{g^2\text{SNR}-1}} - a_2 \right)^2 \text{SNR} + (K-1)(\tilde{\beta}g - a_3)^2 \text{SNR} \right). \end{aligned} \quad (90)$$

We first hold  $\tilde{\beta}$  constant, and minimize over  $a_1, a_2, a_3$ . If  $|\tilde{\beta}| \geq 1/2$ , the optimal choices for the integers  $a_1, a_2, a_3$  are

$$a_1 = \lfloor \tilde{\beta} \rfloor, \quad a_2 = \lfloor \tilde{\beta} / \sqrt{g^2\text{SNR}-1} \rfloor, \quad a_3 = \lfloor \tilde{\beta}g \rfloor. \quad (91)$$

If  $|\tilde{\beta}| < 1/2$ , all three integers  $a_1, a_2, a_3$  from (91) are zero, which is not permitted. Therefore, for these values of  $\tilde{\beta}$  one of the integers must take the value 1 or  $-1$ . Since for  $\text{SNR} > 4$  and  $\text{SNR}^{-1/2} \leq g^2 < 1$  we have

$$\max \left( |\tilde{\beta}|, \left| \tilde{\beta} / \sqrt{g^2\text{SNR}-1} \right|, |\tilde{\beta}g| \right) = |\tilde{\beta}|,$$

the optimal choices of  $a_1, a_2, a_3$  for values of  $|\tilde{\beta}| < 1/2$  are

$$a_1 = \text{sign}(\tilde{\beta}), \quad a_2 = 0, \quad a_3 = 0. \quad (92)$$

Now, the problem of lower bounding  $\sigma_{\text{HK}}^2$  reduces to minimizing (90) over  $\tilde{\beta}$ . Rather than solving this cumbersome

<sup>9</sup>This assumption is valid, since for  $\text{SNR} \leq 4$  the symmetric capacity is upper bounded by  $1/2 \log(1+4) = 1.161$  bits. Our capacity approximations in this subsection, and also in the next subsection, exhibit a constant gap greater than  $7/2$  bits, and therefore hold for  $\text{SNR} < 4$ .

$$\sigma_{\text{HK}}^2 = \min_{\beta, \mathbf{a}_1^{\text{HK}}} \left( \beta^2 + \left( \left( \beta \sqrt{\frac{g^2\text{SNR}-1}{K \cdot g^2\text{SNR}}} - a_1 \right)^2 + \left( \beta \sqrt{\frac{1}{K \cdot g^2\text{SNR}}} - a_2 \right)^2 + (K-1) \left( \beta g \sqrt{\frac{g^2\text{SNR}-1}{K \cdot g^2\text{SNR}}} - a_3 \right)^2 \right) \text{SNR} \right) \quad (87)$$

minimization problem, we split the real line into four intervals, and lower bound  $\sigma_{\text{HK}}^2$  for all values of  $\tilde{\beta}$  within each one. Then, we take the minimum over these four lower bounds. In a similar manner to the previous subsection, we define  $\delta = (2c + 8)/\log(\text{SNR})$ , where  $c > 0$  is some constant. The lower bounds below are derived in Appendix D-A.

**Interval 1:**  $0 < |\tilde{\beta}| \leq 1/2$

$$\sigma_{\text{HK}}^2 \geq \frac{\text{SNR}}{4}. \quad (93)$$

**Interval 2:**  $1/2 < |\tilde{\beta}| \leq \sqrt{|g|}\text{SNR}^{1/4-\delta/2}/2$ <sup>10</sup>

For all values of  $\text{SNR}^{-1/3} < |g| \leq 1$  except for an outage set with measure not greater than  $2^{-c}$  we have

$$\sigma_{\text{HK}}^2 > \frac{2^{-2c}}{4 \cdot 2^8} \frac{\text{SNR}^{1/2}}{\sqrt{g^2}}. \quad (94)$$

**Interval 3:**  $\sqrt{|g|}\text{SNR}^{1/4-\delta/2}/2 < |\tilde{\beta}| \leq \text{SNR}^{1/4}/\sqrt{8|g|}$

$$\sigma_{\text{HK}}^2 \geq \frac{2^{-2c}}{4 \cdot 2^8} \frac{\text{SNR}^{1/2}}{\sqrt{g^2}}. \quad (95)$$

**Interval 4:**  $\text{SNR}^{1/4}/\sqrt{8|g|} < |\tilde{\beta}|$

$$\sigma_{\text{HK}}^2 \geq \frac{1}{8} \frac{\text{SNR}^{1/2}}{\sqrt{g^2}}. \quad (96)$$

Combining the four lower bounds (93), (94), (95), and (96), we have

$$\begin{aligned} \sigma_{\text{HK}}^2 &\geq \min \left( \frac{1}{4} \text{SNR}, \frac{2^{-2c}}{2^{10}} \frac{\text{SNR}^{1/2}}{\sqrt{g^2}}, \frac{1}{8} \frac{\text{SNR}^{1/2}}{\sqrt{g^2}} \right) \\ &= \frac{2^{-2c}}{2^{10}} \frac{\text{SNR}^{1/2}}{\sqrt{g^2}} \end{aligned}$$

for all  $\text{SNR}^{-1/3} \leq g^2 < 1$  up to an outage set of measure not greater than  $2^{-c}$ . Thus, substituting our lower bound for  $\sigma_{\text{HK}}^2$  into (88), we find that any symmetric rate satisfying

$$R_{\text{SYM}} < \frac{1}{2} \log \left( \frac{\text{SNR}^{1/2}}{\sqrt{g^2}} \right) - c - 5 - \frac{1}{2} \log(27) - \frac{1}{2} \log(K^2)$$

is achievable over the symmetric  $K$ -user interference channel for all  $\text{SNR}^{-1/3} \leq g^2 < 1$  up to an outage set of measure not greater than  $2^{-c}$ . Since

$$\begin{aligned} &\frac{1}{2} \log \left( \frac{\text{SNR}^{1/2}}{\sqrt{g^2}} \right) - c - 5 - \frac{1}{2} \log(27) - \frac{1}{2} \log(K^2) \\ &> \frac{1}{2} \log \left( \frac{\text{SNR}^{1/2}}{\sqrt{g^2}} \right) - c - 8 - \log(K) \\ &= \frac{1}{2} \log \left( \frac{\text{SNR}}{\sqrt{\text{INR}}} \right) - c - 8 - \log(K), \end{aligned}$$

any symmetric rate satisfying

$$R_{\text{SYM}} < \frac{1}{2} \log \left( \frac{\text{SNR}}{\sqrt{\text{INR}}} \right) - c - 8 - \log(K) \quad (97)$$

is achievable.

<sup>10</sup>If  $\sqrt{|g|}\text{SNR}^{1/4-\delta/2}/2 < 1/2$  this interval is empty, and we skip to interval 3.

*Remark 16:* It follows from the derivation in Appendix D-A that, as in the strong interference regime, the channel gains that fall within the outage set are the ones close to rational numbers with denominator smaller than  $\text{SNR}^{1/4}$ . Here, the sensitivity of the achievable rate to the ‘‘rationality’’ of  $g$  increases as  $g$  approaches 1.

#### D. Weak Interference Regime

This regime is characterized by  $1/2 \leq \alpha < 2/3$ , or equivalently,  $\text{SNR}^{-1/2} \leq g^2 < \text{SNR}^{-1/3}$ . As in the moderately weak interference regime, we develop a closed-form lower bound on the achievable symmetric rate of Corollary 3. A key difference is that the bound derived here is valid for all channel gains, rather than up to an outage set.

We first note that equations (88) and (90) continue to hold in this regime as in the moderately weak interference regime, and the optimal choices of  $a_1, a_2, a_3$  are also as in (91) and (92). As before, we divide the real line into four intervals, give lower bounds on  $\sigma_{\text{HK}}^2$  which hold for all values of  $\tilde{\beta}$  in each one, and conclude that  $\sigma_{\text{HK}}^2$  is lower bounded by the minimum of these four bounds. The lower bounds below are derived in Appendix D-B

**Interval 1:**  $0 < |\tilde{\beta}| \leq 1/2$

$$\sigma_{\text{HK}}^2 \geq \frac{\text{SNR}}{4}. \quad (98)$$

**Interval 2:**  $1/2 < |\tilde{\beta}| \leq 1/(2|g|)$

$$\sigma_{\text{HK}}^2 \geq \frac{g^2 \text{SNR}}{4}. \quad (99)$$

**Interval 3:**  $1/(2|g|) < |\tilde{\beta}| \leq \sqrt{g^2 \text{SNR}/8}$

$$\sigma_{\text{HK}}^2 \geq \frac{1}{4g^4}. \quad (100)$$

**Interval 4:**  $\sqrt{g^2 \text{SNR}/8} < |\tilde{\beta}|$

$$\sigma_{\text{HK}}^2 > \frac{g^2 \text{SNR}}{4}. \quad (101)$$

Combining the four lower bounds (98), (99), (100), and (101), we have

$$\begin{aligned} \sigma_{\text{HK}}^2 &\geq \frac{1}{4} \min \left( \text{SNR}, g^2 \text{SNR}, g^{-4} \right) \\ &= \frac{g^2 \text{SNR}}{4}, \end{aligned} \quad (102)$$

where (102) is true since  $\text{SNR}^{-1/2} \leq g^2 < \text{SNR}^{-1/3}$ . It follows by substituting (102) into (88) that any symmetric rate satisfying

$$R_{\text{SYM}} < \frac{1}{2} \log \left( \frac{g^2 \text{SNR}}{4} \right) - \frac{1}{2} \log(27) - \frac{1}{2} \log(K^2)$$

is achievable for the symmetric  $K$ -user interference channel with  $\text{SNR}^{-1/2} \leq g^2 < \text{SNR}^{-1/3}$ . Since

$$\begin{aligned} &\frac{1}{2} \log \left( \frac{g^2 \text{SNR}}{4} \right) - \frac{1}{2} \log(27) - \frac{1}{2} \log(K^2) \\ &> \frac{1}{2} \log(g^2 \text{SNR}) - \frac{7}{2} - \log(K) \\ &= \frac{1}{2} \log(\text{INR}) - \frac{7}{2} - \log(K) \end{aligned}$$

any symmetric rate satisfying

$$R_{\text{SYM}} < \frac{1}{2} \log(\text{INR}) - \frac{7}{2} - \log(K) \quad (103)$$

is achievable.

### E. Noisy Interference Regime

The noisy interference regime is characterized by  $\alpha < 1/2$ , or equivalently  $g^2 < \text{SNR}^{-1/2}$ . In this regime, each receiver decodes its desired codeword while treating all interfering codewords as noise. Lattice codes are not necessary in this regime in order to approximate the symmetric capacity: random i.i.d. Gaussian codebooks suffice. Nevertheless, the same performance can be achieved with lattice codes as shown in Theorem 10 which states that any symmetric rate

$$R_{\text{SYM}} < \frac{1}{2} \log \left( 1 + \frac{\text{SNR}}{1 + (K-1)g^2\text{SNR}} \right)$$

is achievable. It follows that any symmetric rate satisfying

$$\begin{aligned} R_{\text{SYM}} &< \frac{1}{2} \log \left( 1 + \frac{\text{SNR}}{1 + g^2\text{SNR}} \right) - \frac{1}{2} \log(K-1) \\ &= \frac{1}{2} \log \left( 1 + \frac{\text{SNR}}{1 + \text{INR}} \right) - \frac{1}{2} \log(K-1) \end{aligned} \quad (104)$$

is achievable.

## VII. DEGREES-OF-FREEDOM

In the previous section, we have shown that the compute-and-forward transform can approximate the capacity of the symmetric  $K$ -user interference channel up to a constant gap for all channel gains outside a small outage set. Ideally, we would like to use a similar approach to approximate the capacity of the general (non-symmetric) interference channel. In contrast to the symmetric case, where all interferers are automatically aligned (if they all use the same lattice codebook), in a general interference channel the interferers will be observed through different channel gains. A linear combination of lattice codewords is always a codeword only if all of the coefficients are integers. Thus, in order to induce alignment, all of the interfering gains should be steered towards integers, which is an overconstrained problem.

The compute-and-forward transform proposed in this paper is quite general, in that its performance can be evaluated for any Gaussian interference network, and it can be combined with precoding schemes that induce alignment. For instance, consider the class of *real interference alignment* precoding schemes that transform the channel seen by each receiver in a non-symmetric interference channel to an effective MAC where some of the interfering users are aligned. A remarkable example of such a scheme is that of [8], which is used to prove that the DoF offered by almost every Gaussian  $K$ -user interference channel is  $K/2$ . To date, essentially all real interference alignment schemes utilized a scalar lattice constellation (e.g.,  $p$ -ary pulse amplitude modulation), concatenated with a random i.i.d. outer code. Potentially, replacing this construction with AWGN capacity achieving  $n$ -dimensional lattice codes can improve the performance of

such schemes and may eventually lead to achievable rate regions that outperform TDMA at reasonable values of SNR. Here, we take a first step and verify that the compute-and-forward transform can attain the same high SNR asymptotics.

Specifically, we show that for almost every  $K$ -user MAC, each user can achieve  $1/K$  DoF using the compute-and-forward transform. In [8], it is shown that the same is true using a scalar lattice concatenated with a random i.i.d. outer code and maximum likelihood decoding. This result is then used as a building block for the interference alignment scheme. Since real interference alignment schemes often induce effective multiple-access channels whose coefficients are dependent [8], our analysis assumes that the channel coefficients belong to a *manifold*, and our results apply for a set of full Lebesgue measure with respect to the considered manifold.

Theorem 3 in Section IV guarantees that the sum of the optimal computation rates is close to the sum capacity of the MAC. However, the theorem does not tell us how the sum rate is divided between the  $K$  rates. We now show that, in a DoF sense, the sum is equally split between all  $K$  rates for almost every channel realization. Recall the definition for DoF:

$$d_{\text{comp},k} = \lim_{\text{SNR} \rightarrow \infty} \frac{R_{\text{comp},k}(\text{SNR})}{\frac{1}{2} \log(1 + \text{SNR})}. \quad (105)$$

First, we upper bound  $d_{\text{comp},1}$ , the DoF provided by the highest computation rate.

*Theorem 12:* Let  $f_1, \dots, f_K$  be functions from  $\mathbb{R}^m$  to  $\mathbb{R}$  satisfying

- 1)  $f_k$  for  $k = 1, \dots, K$  is analytic in  $\mathbb{R}^m$ ,
- 2)  $1, f_1, \dots, f_K$  are linearly independent over  $\mathbb{R}$ ,

and define the manifold

$$\mathcal{M} = \left\{ [f_1(\tilde{\mathbf{h}}) \cdots f_K(\tilde{\mathbf{h}})] : \tilde{\mathbf{h}} \in \mathbb{R}^m \right\}. \quad (106)$$

For almost every  $\mathbf{h} \in \mathcal{M}$ , the DoF offered by the highest computation rate is upper bounded by

$$d_{\text{comp},1} \leq \frac{1}{K}. \quad (107)$$

The proof is given in Appendix E, and is based on showing that restricting the scaling coefficient  $\beta$  from (9) to the form  $\beta = q/h_1$  for  $q \in \mathbb{Z}$  (almost surely) incurs no loss from a DoF point of view. This way, the first coefficient of  $\beta\mathbf{h}$  is an integer. Then, a result from the field of Diophantine approximation which is due to Kleinbock and Margulis [62] is used in order to lower bound the error in approximating the remaining  $K-1$  channel gains with integers.

As a special case of Theorem 12 we may choose the manifold  $\mathcal{M}$  as  $\mathbb{R}^K$  which implies the following corollary.

*Corollary 4:* For almost every  $\mathbf{h} \in \mathbb{R}^K$  the DoF offered by the highest computation rate is upper bounded by

$$d_{\text{comp},1} \leq \frac{1}{K}.$$

*Remark 17:* Niesen and Whiting [63] studied the DoF offered by the highest computation rate and showed that

$$d_{\text{comp},1} \leq \begin{cases} 1/2 & K = 2 \\ 2/(K+1) & K > 2 \end{cases} \quad (108)$$



for almost every  $\mathbf{h} \in \mathbb{R}^K$ . Our bound therefore agrees with that of [63] for  $K = 2$  and improves it for  $K > 2$ .

The next corollary shows that all  $K$  optimal computation rates offer  $1/K$  DoF for almost every  $\mathbf{h}$  satisfying mild conditions.

*Corollary 5:* Let  $\mathcal{M}$  be a manifold satisfying the conditions of Theorem 12. For almost every  $\mathbf{h} \in \mathcal{M}$  the DoF provided by each of the  $K$  optimal computation rates is  $d_{\text{comp},k} = 1/K$ .

*Proof:* Theorem 3 implies that  $\sum_{k=1}^K d_{\text{comp},k} \geq 1$ . Using the fact that  $d_{\text{comp},k}$  is monotonically decreasing in  $k$  and that  $d_{\text{comp},1} \leq 1/K$  for almost every  $\mathbf{h} \in \mathcal{M}$ , the corollary follows. ■

The corollary above implies that, in the limit of very high SNR, not only is the sum of computation rates close to the sum capacity of the MAC, but each computation rate scales like the symmetric capacity of the MAC for almost all channel gains. Note that our analysis (as well as that of [63]) is within the context of the achievable computation rates stemming from Theorem 2.

The next corollary follows from Corollary 5 and Theorem 5.

*Corollary 6:* Let  $\mathcal{M}$  be a manifold satisfying the conditions of Theorem 12. The DoF attained by each user in the  $K$ -user MAC under the compute-and-forward transform is  $1/K$  for almost every  $\mathbf{h} \in \mathcal{M}$ . In particular, the DoF attained by each user in the  $K$ -user MAC under the compute-and-forward transform is  $1/K$  for almost every  $\mathbf{h} \in \mathbb{R}^K$ .

The next theorem shows that for almost every effective  $L$ -user multiple access channel of the form introduced in Section IV-C each of the effective users achieves  $1/L$  degree of freedom. The proof is given in Appendix F.

*Theorem 13:* Let  $f_1, \dots, f_L$  be functions from  $\mathbb{R}^m$  to  $\mathbb{R}$  satisfying

- 1)  $f_\ell$  for  $\ell = 1, \dots, L$  is analytic in  $\mathbb{R}^m$ ,
- 2)  $1, f_1, \dots, f_L$  are linearly independent over  $\mathbb{R}$ ,

and define the manifold

$$\mathcal{M} = \left\{ [f_1(\tilde{\mathbf{g}}) \cdots f_L(\tilde{\mathbf{g}})] : \tilde{\mathbf{g}} \in \mathbb{R}^m \right\}.$$

For almost every  $\mathbf{g} \in \mathcal{M}$  the DoF offered by each of the  $L$  optimal computation rates for the effective MAC (33) is

$$d_{\text{comp},\ell} = \lim_{\text{SNR} \rightarrow \infty} \frac{R_{\text{comp},\ell}(\text{SNR})}{\frac{1}{2} \log(1 + \text{SNR})} = \frac{1}{L}. \quad (109)$$

*Remark 18:* The manifold  $\mathcal{M}$  is the same manifold used in [8]. This manifold was general enough to allow the derivation of the DoF characterization of the  $K$ -user interference channel in [8]. Thus, the DoF results from [8] can be re-derived using the same real interference alignment scheme from [8] with  $n$ -dimensional lattice codes instead of 1-D integer constellations concatenated with outer codes.

## VIII. DISCUSSION

In this paper, we have developed a new decoding framework for lattice-based interference alignment. We used this framework as a building block for two lattice-based interference alignment schemes for the symmetric real Gaussian  $K$ -user interference channel. These schemes perform well starting from the moderate SNR regime, and are within a constant gap

from the upper bounds on the capacity for all channel gains outside of some outage set whose measure can be controlled.

A natural question for future research is how to extend the results above to the general Gaussian  $K$ -user interference channel. The main problem is that, in the general case, the interfering lattice codewords are not naturally aligned, as their gains are not integer-valued. Therefore, in order to successfully apply lattice interference alignment, some form of precoding, aimed towards forcing the cross channel gains to be integers, is required. Unfortunately, simple power-backoff strategies do not suffice, even in the three-user case.

One option for overcoming this problem is to use many layers at each transmitter, as in [8], and create partial alignment between interfering layers. While this achieves the optimal DoF, it performs poorly at reasonable values of SNR, as there will be a rate loss for each additional layer. As a result the rate region obtained by combining the compute-and-forward transform with the precoding scheme of [8] is inferior to that obtained by time-sharing, for values of SNR of practical interest. Another option is to precode not only using power-backoff, but also over time, which may partially compensate for the lack of sufficient free parameters. An example for such a precoding scheme is the power-time code introduced in [23].

A positive feature of the compute-and-forward framework is that it does not require perfect alignment of the lattice points participating in the integer linear combinations. Namely, the effect of not perfectly equalizing the channel gains to integers is an enhanced effective noise. For the general interference channel, this suggests that it may suffice to find precoding schemes that only approximately force the cross-channel gains to integers.

## APPENDIX A PROOF OF THEOREM 5

We begin with two lemmas which will be useful for the proof of Theorem 5.

*Lemma 2:* Let  $\mathbf{A}$  be a  $K \times K$  matrix with integer entries of magnitudes bounded from above by some constant  $a_{\text{max}}$ . If there exists a real-valued  $K \times K$  lower triangular matrix  $\mathbf{L}$  with unit diagonal such that  $\tilde{\mathbf{A}} = \mathbf{L}\mathbf{A}$  is upper triangular up to column permutation  $\pi$ , then for any prime  $p > K(K!)^2(Ka_{\text{max}})^{2K}a_{\text{max}}$  there also exists a lower triangular matrix  $\mathbf{L}^{(p)}$  with elements from  $\{0, 1, \dots, p-1\}$  and unit diagonal such that  $\tilde{\mathbf{A}}^{(p)} = [\mathbf{L}^{(p)}\mathbf{A}] \bmod p$  is upper triangular up to column permutation  $\pi$ .

*Proof:* Assume that there exists a lower triangular matrix  $\mathbf{L}$  with unit diagonal such that  $\tilde{\mathbf{A}} = \mathbf{L}\mathbf{A}$  is upper triangular up to column permutation  $\pi$ . We begin by showing that all elements in the  $i$ th ( $i > 1$ ) row of  $\mathbf{L}$  can be written as rational numbers with the same denominator  $1 \leq q_i \leq K!(Ka_{\text{max}}^2)^K$ . To see this note that if  $\tilde{\mathbf{A}}$  is triangular up to column permutation vector  $\pi$ , then its  $i$ th row contains at least  $i-1$  zeros, namely  $\tilde{a}_{ij} = 0$  for  $j = \pi(1), \dots, \pi(i-1)$ . Since  $\mathbf{L}$  is lower triangular, the following equations must hold

$$\tilde{a}_{ij} = \sum_{m=1}^i \ell_{im} a_{mj} = 0, \text{ for } j = \pi(1), \dots, \pi(i-1). \quad (110)$$

By definition  $\ell_{ii} = 1$ , therefore (110) can be written as

$$\sum_{m=1}^{i-1} \ell_{im} a_{mj} = -a_{ij}, \text{ for } j = \pi(1), \dots, \pi(i-1). \quad (111)$$

Define the vectors  $\boldsymbol{\ell}^{(i)} = [\ell_{i1} \ \dots \ \ell_{i,i-1}]^T$ ,  $\mathbf{a}^{(i\pi)} = -[a_{i,\pi(1)} \ \dots \ a_{i,\pi(i-1)}]^T$  and the matrix

$$\mathbf{A}^{(i,\pi)} = \begin{pmatrix} a_{1\pi(1)} & \dots & a_{i-1\pi(1)} \\ \vdots & \ddots & \vdots \\ a_{1\pi(i-1)} & \dots & a_{i-1\pi(i-1)} \end{pmatrix}.$$

We have,

$$\mathbf{A}^{(i,\pi)} \boldsymbol{\ell}^{(i)} = \mathbf{a}^{(i,\pi)}. \quad (112)$$

From the fact that  $\mathbf{A}$  can be pseudo-triangularized with permutation vector  $\pi$ , we know that the system of equations (112) has a solution. Assume that

$$\text{rank}(\mathbf{A}^{(i,\pi)}) = u \leq i-1.$$

It follows that there are  $u$  linearly independent columns in  $\mathbf{A}^{(i,\pi)}$ . Let  $\mathcal{U} \subseteq \{1, \dots, i-1\}$  be a set of indices corresponding to  $u$  such linearly independent columns, and  $\bar{\mathcal{U}}$  be its complement. Let  $\mathbf{A}_{\mathcal{U}}^{(i,\pi)} \in \mathbb{Z}^{i-1 \times u}$  be the matrix obtained by taking the columns of  $\mathbf{A}^{(i,\pi)}$  with indices in  $\mathcal{U}$ . Since (112) has a solution, we have  $\mathbf{a}^{(i,\pi)} \in \text{span}(\mathbf{A}_{\mathcal{U}}^{(i,\pi)})$ . Thus, we can set  $\ell^{(i)}(k) = 0$  for all  $k \in \bar{\mathcal{U}}$ , and (112) will still have a solution. Letting  $\boldsymbol{\ell}_{\mathcal{U}}^{(i)} \in \mathbb{R}^{u \times 1}$  be the vector obtained by taking from  $\boldsymbol{\ell}^{(i)}$  only the entries with indices in  $\mathcal{U}$ , it follows that

$$\mathbf{A}_{\mathcal{U}}^{(i,\pi)} \boldsymbol{\ell}_{\mathcal{U}}^{(i)} = \mathbf{a}^{(i,\pi)} \quad (113)$$

has a solution. Now, multiplying both sides of (113) by  $(\mathbf{A}_{\mathcal{U}}^{(i,\pi)})^T$  gives

$$\mathbf{A}'^{(i,\pi)} \boldsymbol{\ell}_{\mathcal{U}}^{(i)} = \mathbf{a}'^{(i,\pi)}, \quad (114)$$

where  $\mathbf{A}'^{(i,\pi)} = (\mathbf{A}_{\mathcal{U}}^{(i,\pi)})^T \mathbf{A}_{\mathcal{U}}^{(i,\pi)} \in \mathbb{Z}^{u \times u}$  is a full-rank matrix and  $\mathbf{a}'^{(i,\pi)} = (\mathbf{A}_{\mathcal{U}}^{(i,\pi)})^T \mathbf{a}^{(i,\pi)} \in \mathbb{Z}^{u \times 1}$ . Note that all entries of  $\mathbf{A}'^{(i,\pi)}$  as well as all entries of  $\mathbf{a}'^{(i,\pi)}$  have magnitude bounded from above by  $\tilde{a}_{\max} \triangleq u a_{\max}^2$ . Cramer's rule for solving a system of linear equations (see [50]) implies that all elements of  $\boldsymbol{\ell}_{\mathcal{U}}^{(i)}$  can be expressed as rational numbers with denominator  $q_i \triangleq |\det(\mathbf{A}'^{(i,\pi)})|$ . Recall the Leibnitz formula (see [50]) for the determinant of an  $n \times n$  matrix  $\mathbf{G}$

$$\det(\mathbf{G}) = \sum_{\sigma \in \mathcal{S}_n} \text{sign}(\sigma) \prod_{i=1}^n G_{i,\sigma_i}, \quad (115)$$

where  $\mathcal{S}_n$  is the set of all permutations of  $\{1, \dots, n\}$ . It follows that  $\det(\mathbf{A}'^{(i,\pi)})$  must be an integer and in addition  $1 \leq |\det(\mathbf{A}'^{(i,\pi)})| \leq u! (\tilde{a}_{\max})^u$ . Thus,  $1 \leq q_i \leq u! (\tilde{a}_{\max})^u$ . Moreover, Cramer's rule also implies that the numerator of each element in  $\boldsymbol{\ell}_{\mathcal{U}}^{(i)}$  is an integer not greater than  $u! (\tilde{a}_{\max})^u$  in magnitude. Since  $u \leq K$ , and since each element of  $\boldsymbol{\ell}^{(i)}$  is either zero or corresponds to an element in  $\boldsymbol{\ell}_{\mathcal{U}}^{(i)}$ , each element  $\ell_{ij}$ ,  $j \leq i$  of  $\mathbf{L}$  can be written as a rational

number  $\ell_{ij} = m_{ij}/q_i$  with  $1 \leq q_i \leq K!(K a_{\max}^2)^K$  and  $|m_{ij}| \leq K!(K a_{\max}^2)^K$  for  $i = 1, \dots, K$ .

Now, define the matrix  $\tilde{\mathbf{L}} = \text{diag}(q_1, \dots, q_K) \mathbf{L}$  and note that  $\tilde{\mathbf{L}} \in \mathbb{Z}^{K \times K}$  due to the above. Let  $\tilde{\mathbf{A}}^{(p)} = [\tilde{\mathbf{L}} \mathbf{A}] \bmod p$ . Since multiplying a row in a matrix by a constant leaves its zero entries unchanged, the entries of the matrix

$$\begin{aligned} \tilde{\mathbf{A}}^{(p)} &= [\tilde{\mathbf{L}} \mathbf{A}] \bmod p \\ &= [\text{diag}(q_1, \dots, q_K) \mathbf{L} \mathbf{A}] \bmod p \\ &= [\text{diag}(q_1, \dots, q_K) \tilde{\mathbf{A}}] \bmod p, \end{aligned} \quad (116)$$

are zero whenever the entries of  $\tilde{\mathbf{A}}$  are equal to zero. Moreover, since all elements of  $\mathbf{L}$  are bounded in magnitude by  $K!(K a_{\max}^2)^K$  and all elements of  $\mathbf{A}$  are bounded in magnitude by  $a_{\max}$ , all elements of  $\tilde{\mathbf{A}} = \mathbf{L} \mathbf{A}$  are bounded in magnitude by  $K K!(K a_{\max}^2)^K a_{\max}$ . Combining with the fact that  $1 \leq q_i \leq K!(K a_{\max}^2)^K$ , we have  $|\tilde{a}_{ij}^{(p)}| \leq K(K!)^2 (K a_{\max})^{2K} a_{\max}$  for all  $i = 1, \dots, K$ ,  $j = 1, \dots, K$ . Therefore, for a prime number  $p > K(K!)^2 (K a_{\max})^{2K} a_{\max}$  the modulo reduction in (116) does not change any of the non-zero entries of  $\text{diag}(q_1, \dots, q_K) \tilde{\mathbf{A}}$  to zero.

Recall that if  $\mathbf{A}$  can be pseudo-triangularized with a matrix  $\mathbf{L}$  and permutation vector  $\pi$  then  $\tilde{a}_{i,\pi(i)} \neq 0$ , and hence also  $\tilde{a}_{i,\pi(i)}^{(p)} \neq 0$  for  $i = 1, \dots, K$ . We have therefore shown that for  $p$  large enough there exists a lower-triangular matrix  $\tilde{\mathbf{L}}^{(p)} = [\text{diag}(q_1, \dots, q_K) \mathbf{L}] \bmod p$  with elements from  $\{0, 1, \dots, p-1\}$  such that  $\tilde{\mathbf{A}}^{(p)} = [\tilde{\mathbf{L}}^{(p)} \mathbf{A}] \bmod p$  is upper-triangular up to column permutation  $\pi$ . In order to complete the proof, it is left to transform  $\tilde{\mathbf{L}}^{(p)}$  to a lower-triangular matrix with elements from  $\{0, 1, \dots, p-1\}$  and unit diagonal. Let  $(q_i)^{-1}$  be an integer that satisfies  $[(q_i)^{-1} q_i] \bmod p = 1$ . Such an integer always exists since  $q_i$  is an integer different than zero, and  $p$  is prime. It is easy to verify that the matrix  $\mathbf{L}^{(p)} = [\text{diag}((q_1)^{-1}, \dots, (q_K)^{-1}) \tilde{\mathbf{L}}^{(p)}] \bmod p$  is a lower-triangular matrix with elements from  $\{0, 1, \dots, p-1\}$  and unit diagonal, and  $\tilde{\mathbf{A}}^{(p)} = [\mathbf{L}^{(p)} \mathbf{A}] \bmod p$  is upper triangular up to column permutation  $\pi$ . ■

*Lemma 3:* Let  $\mathbf{t}_1, \dots, \mathbf{t}_k$  be lattice points from a chain of nested lattices satisfying the conditions of Theorem 2. Let  $\mathbf{v} = [\sum_{k=1}^K a_k \mathbf{t}_k] \bmod \Lambda$  and  $\mathbf{u} = [\sum_{k=1}^K b_k \mathbf{t}_k] \bmod \Lambda$  be integer linear combinations of these points. Then

$$[\mathbf{v} + \mathbf{u}] \bmod \Lambda = \left[ \sum_{k=1}^K ((a_k + b_k) \bmod p) \mathbf{t}_k \right] \bmod \Lambda.$$

*Proof:* Due to the distributive property of the modulo operation we have

$$\begin{aligned} [\mathbf{v} + \mathbf{u}] \bmod \Lambda &= \left[ \sum_{k=1}^K (a_k + b_k) \mathbf{t}_k \right] \bmod \Lambda \\ &= \left[ \sum_{k=1}^K [(a_k + b_k) \bmod p + M_k \cdot p] \mathbf{t}_k \right] \bmod \Lambda \\ &= \left[ \sum_{k=1}^K ((a_k + b_k) \bmod p) \mathbf{t}_k \right. \\ &\quad \left. + \sum_{k=1}^K M_k \cdot [p \cdot \mathbf{t}_k] \bmod \Lambda \right] \bmod \Lambda \end{aligned} \quad (117)$$

where  $\{M_k\}_{k=1}^K$  are some integers. Utilizing the fact that  $[p \cdot \mathbf{t}_k] \bmod \Lambda = \mathbf{0}$  for all lattice points in the chain, which follows from Theorem 2(c), the lemma is established. ■

We are now ready to prove Theorem 5.

*Proof of Theorem 5:* Let  $\mathbf{T} = [\mathbf{t}_1^T \cdots \mathbf{t}_K^T]^T$  and  $\mathbf{V} = [\mathbf{v}_1^T \cdots \mathbf{v}_K^T]^T = [\mathbf{A}\mathbf{T}] \bmod \Lambda$ . The compute-and-forward transform of the MAC (4) can be written as

$$\begin{aligned} \mathbf{S} &= \left[ \mathbf{A} \begin{pmatrix} \mathbf{t}_1 \\ \vdots \\ \mathbf{t}_K \end{pmatrix} + \mathbf{Z}_{\text{eff}} \right] \bmod \Lambda \\ &= [\mathbf{A}\mathbf{T} + \mathbf{Z}_{\text{eff}}] \bmod \Lambda \\ &= [\mathbf{V} + \mathbf{Z}_{\text{eff}}] \bmod \Lambda. \end{aligned}$$

Assume there exists a pseudo-triangularization of  $\mathbf{A}$  with permutation vector  $\pi$ , i.e., there exists a lower triangular matrix  $\mathbf{L}$  with unit diagonal such that  $\tilde{\mathbf{A}} = \mathbf{L}\mathbf{A}$  is upper triangular up to column permutation  $\pi$ . Lemma 2 implies that there exists a lower triangular matrix  $\mathbf{L}^{(p)}$  with elements from  $\{0, 1, \dots, p-1\}$  and unit diagonal such that  $\tilde{\mathbf{A}}^{(p)} = [\mathbf{L}^{(p)}\mathbf{A}] \bmod p$  is upper triangular up to column permutation  $\pi$ . Since  $\mathbf{L}^{(p)}$  has a unit diagonal it can be written as  $\mathbf{L}^{(p)} = \mathbf{I} + \mathbf{R}$  where  $\mathbf{I}$  is the identity matrix and  $\mathbf{R}$  has non-zero entries only below the main diagonal.

Assume the receiver has access to the side information  $\mathbf{v}_1, \dots, \mathbf{v}_{K-1}$ . As the entries of  $\mathbf{R}$  are non-zero only below the main diagonal, the receiver could compute  $\mathbf{R} \cdot \mathbf{V}$ , add it to  $\mathbf{S}$  and reduce modulo  $\Lambda$ , giving rise to

$$\begin{aligned} \mathbf{S}^{\text{SI}} &= [\mathbf{S} + \mathbf{R} \cdot \mathbf{V}] \bmod \Lambda \\ &= [\mathbf{A}\mathbf{T} + \mathbf{R}\mathbf{A}\mathbf{T} + \mathbf{Z}_{\text{eff}}] \bmod \Lambda \\ &= [(\mathbf{I} + \mathbf{R})\mathbf{A}\mathbf{T} + \mathbf{Z}_{\text{eff}}] \bmod \Lambda \\ &= [\mathbf{L}^{(p)}\mathbf{A}\mathbf{T} + \mathbf{Z}_{\text{eff}}] \bmod \Lambda \\ &= [[\mathbf{L}^{(p)}\mathbf{A}] \bmod p \cdot \mathbf{T} + \mathbf{Z}_{\text{eff}}] \bmod \Lambda \quad (118) \\ &= \left[ \tilde{\mathbf{A}}^{(p)} \begin{pmatrix} \mathbf{t}_1 \\ \vdots \\ \mathbf{t}_K \end{pmatrix} + \mathbf{Z}_{\text{eff}} \right] \bmod \Lambda \end{aligned}$$

where (118) follows from Lemma 3. Let  $\tilde{\mathbf{V}} = [\tilde{\mathbf{A}}^{(p)}\mathbf{T}] \bmod \Lambda$  and recall that  $\tilde{\mathbf{A}}^{(p)}$  is upper-triangular up to column permutation  $\pi$ , thus  $\tilde{a}_{j,\pi(m)}^{(p)} = 0$  for all  $j = \pi(m) + 1, \dots, K$ . It follows that for any  $m < K$  the lattice point  $\mathbf{t}_{\pi(m)}$  does not participate in any of the linear combinations  $\tilde{\mathbf{v}}_{m+1}, \dots, \tilde{\mathbf{v}}_K$ .

Assume the mapping function between users and lattices is chosen as  $\theta(k) = \pi^{-1}(k)$ , i.e., each user  $k$  employs the codebook  $\mathcal{L}_k = \Lambda_{\pi^{-1}(k)} \cap \Lambda$ . In this case, the densest lattice participating in linear combination  $\tilde{\mathbf{v}}_m$  is  $\Lambda_m$ . The decoder uses  $\mathbf{s}_m^{\text{SI}}$  in order to produce an estimate

$$\hat{\tilde{\mathbf{v}}}_m = [Q_{\Lambda_m}(\mathbf{s}_m^{\text{SI}})] \bmod \Lambda \quad (119)$$

for each one of the linear combinations  $\tilde{\mathbf{v}}_m$ ,  $m = 1, \dots, K$ . It follows from Theorem 2 that there exists a chain of nested lattices  $\Lambda \subseteq \Lambda_K \subseteq \dots \subseteq \Lambda_1$  forming the set of codebooks  $\mathcal{L}_1, \dots, \mathcal{L}_K$  with rates  $R_1, \dots, R_K$  such that all linear combinations  $\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_K$  can be decoded with a vanishing

error probability as long as the rates of all users satisfy the constraints of (29).

We have shown that if the receiver has access to  $\mathbf{v}_1, \dots, \mathbf{v}_{K-1}$  it can decode the set of linear combinations  $\tilde{\mathbf{V}}$ . We now show a sequential decoding procedure which guarantees that the receiver has the right amount of side information at each step. First, note that

$$\mathbf{s}_m^{\text{SI}} = \left[ \mathbf{s}_m + \sum_{\ell=1}^{m-1} r_{m\ell} \mathbf{v}_\ell \right] \bmod \Lambda, \quad (120)$$

thus the necessary side information for decoding  $\tilde{\mathbf{v}}_m$  is only  $\mathbf{v}_1, \dots, \mathbf{v}_{m-1}$ . In particular,  $\mathbf{s}_1^{\text{SI}} = \mathbf{s}_1$  and hence  $\mathbf{v}_1$  can be decoded with a vanishing error probability with no side information. After decoding  $\mathbf{v}_1$  the receiver has it as side information, and can therefore compute  $\mathbf{s}_2^{\text{SI}}$  and decode  $\tilde{\mathbf{v}}_2$ . As  $\tilde{\mathbf{v}}_2 = [r_{21}\mathbf{v}_1 + \mathbf{v}_2] \bmod \Lambda$  and the receiver knows  $\mathbf{v}_1$ , it can use it in order to recover  $\mathbf{v}_2$ . Now, the receiver has  $\mathbf{v}_1$  and  $\mathbf{v}_2$  as side information and can use it to compute  $\mathbf{s}_3^{\text{SI}}$ . The process continues sequentially until all linear combinations  $\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_K$  are decoded.

Conditioned on correct decoding, we obtain  $K$  noiseless linear combinations

$$\begin{pmatrix} \tilde{\mathbf{v}}_1 \\ \vdots \\ \tilde{\mathbf{v}}_K \end{pmatrix} = \left[ \tilde{\mathbf{A}}^{(p)} \begin{pmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \\ \vdots \\ \mathbf{t}_K \end{pmatrix} \right] \bmod \Lambda. \quad (121)$$

Since  $\tilde{\mathbf{A}}^{(p)}$  is upper-triangular up to column permutation, and in particular full-rank modulo  $p$ , the original lattice points  $\mathbf{t}_1, \dots, \mathbf{t}_K$  each user transmitted can be recovered. ■

## APPENDIX B PROOF OF THEOREMS 7 AND 8

*Proof of Theorem 7:* The proof is identical to that of Theorem 3 with  $\mathbf{F} = (\text{SNR}^{-1}\mathbf{B}^{-1} + \mathbf{g}\mathbf{g}^T)^{-1/2}$ . ■

*Proof of Theorem 8:* Let

$$\mathbf{S} = \left[ \mathbf{A} \begin{pmatrix} \mathbf{t}_{\text{eff},1} \\ \vdots \\ \mathbf{t}_{\text{eff},L} \end{pmatrix} + \mathbf{Z}_{\text{eff}} \right] \bmod \Lambda$$

be the compute-and-forward transform of the effective  $L$ -user MAC, and assume that  $\mathbf{A}$  can be pseudo-triangularized with permutation vector  $\pi$ . Repeating the proof of Theorem 5 it is easy to see that, for any set of rates

$$R_\ell < R_{\text{comp},\pi^{-1}(\ell)}, \quad \ell = 1, \dots, L,$$

there exists a chain of nested lattices  $\Lambda \subseteq \Lambda_L \subseteq \dots \subseteq \Lambda_1$  inducing the codebooks  $\mathcal{L}_\ell = \Lambda_{\pi^{-1}(\ell)} \cap \mathcal{V}$  with rates  $R_\ell$ , such that if  $\mathbf{t}_{\text{eff},\ell} \in \mathcal{L}_\ell$  for all  $\ell = 1, \dots, L$ , all effective lattice points can be decoded from  $\mathbf{S}$ .

If each of the users  $i \in \mathcal{K}_\ell$  that comprise effective user  $\ell$  uses the lattice codebook  $\mathcal{L}_\ell$  (or any codebook nested in  $\mathcal{L}_\ell$ ), then  $\mathbf{t}_{\text{eff},\ell} \in \mathcal{L}_\ell$  and all effective lattice points can be decoded. ■

APPENDIX C  
PROOF OF LEMMA 1

In order to decode the desired effective lattice points, it suffices to decode  $L - 1$  linearly independent integer linear combinations of them, in which  $\mathbf{t}_{\text{eff},L}$  does not participate. Let  $\bar{\mathbf{a}} = [\bar{a}_1 \cdots \bar{a}_{L-1} 0]^T$  be some coefficient vector for such a linear combination. The effective rate for computing the linear combination  $\bar{\mathbf{v}} = [\sum_{\ell=1}^{L-1} \bar{a}_\ell \mathbf{t}_{\text{eff},\ell}] \bmod \Lambda$  with the coefficient vector  $\bar{\mathbf{a}}$  over the channel (33) is

$$R_{\text{comp}}(\mathbf{g}, \bar{\mathbf{a}}, \mathbf{B}) = \frac{1}{2} \log \left( \frac{\text{SNR}}{\sigma_{\text{eff}}^2(\mathbf{g}, \bar{\mathbf{a}}, \mathbf{B})} \right), \quad (122)$$

where

$$\begin{aligned} \sigma_{\text{eff}}^2(\mathbf{g}, \bar{\mathbf{a}}, \mathbf{B}) &= \min_{\tilde{\beta} \in \mathbb{R}} \text{SNR} \sum_{\ell=1}^{L-1} (\tilde{\beta} g_\ell - \bar{a}_\ell)^2 b_{\text{eff},\ell}^2 \\ &\quad + \tilde{\beta}^2 (1 + \text{SNR} g_L^2 b_{\text{eff},L}^2) \\ &= \min_{\tilde{\beta} \in \mathbb{R}} \text{SNR} \sum_{\ell=1}^{L-1} (\beta \kappa g_\ell - \bar{a}_\ell)^2 b_{\text{eff},\ell}^2 + \beta^2, \end{aligned} \quad (123)$$

where (123) follows by substituting  $\tilde{\beta} = \beta \kappa$ . The effective noise variance and computation rate for decoding a linear combination with coefficient vector  $\bar{\mathbf{a}} = [\bar{a}_1 \cdots \bar{a}_{L-1} 0]$  over the effective channel (33) are therefore the same as those of decoding a linear combination with  $\mathbf{a} = [\bar{a}_1 \cdots \bar{a}_{L-1}]$  over the effective channel (51). Thus, for purposes of decoding integer linear combinations of effective lattice points  $\mathbf{t}_{\text{eff},1}, \dots, \mathbf{t}_{\text{eff},L-1}$  the two channels are equivalent. Since this is all we need in order to decode  $\mathbf{t}_{\text{eff},1}, \dots, \mathbf{t}_{\text{eff},L-1}$ , the lemma follows.

APPENDIX D

DERIVATION OF THE UPPER BOUNDS ON  $\sigma_{\text{HK}}^2$  WITHIN  
THE DIFFERENT INTERVALS

*A. Moderately Weak Interference Regime*

We upper bound  $\sigma_{\text{HK}}^2$  for all values of  $\tilde{\beta}$  within each of the four intervals. Recall that in the moderately weak interference regime  $\text{SNR}^{-1/3} \leq g^2 \leq 1$ . Define  $\delta = (2c + 8) / \log(\text{SNR})$ , where  $c > 0$  is some constant.

**Interval 1:**  $0 < |\tilde{\beta}| \leq 1/2$

In this interval the choice  $a_1 = \text{sign}(\tilde{\beta})$  is optimal due to (92). Therefore, for all  $|\tilde{\beta}| \leq 1/2$  we have  $(\tilde{\beta} - a_1)^2 \geq 1/4$ . Thus,

$$\sigma_{\text{HK}}^2 \geq \frac{\text{SNR}}{4}.$$

**Interval 2:**  $1/2 < |\tilde{\beta}| \leq \sqrt{|g|} \text{SNR}^{1/4 - \delta/2} / 2$

Since  $|\tilde{\beta}| > 1/2$  we can express it as  $\tilde{\beta} = q + \varphi$  with  $q \in \mathbb{Z} \setminus 0$  and  $\varphi \in [-1/2, 1/2)$ . We can further lower bound  $\sigma_{\text{HK}}^2$  as

$$\begin{aligned} \sigma_{\text{HK}}^2 &> \min_{\varphi, q, a_1, a_3} \left( ((\varphi + q - a_1)^2 + (qg - a_3 + \varphi g)^2) \text{SNR} \right) \\ &= \min_{\varphi, q, a_3} \left( (\varphi^2 + (qg - a_3 + \varphi g)^2) \text{SNR} \right). \end{aligned} \quad (124)$$

Ignoring the constraint  $\varphi \in [-1/2, 1/2)$ , the minimizing value of  $\varphi$  is found to be

$$\varphi^* = -\frac{g}{1+g^2}(qg - a_3).$$

Substituting  $\varphi^*$  into (124) gives

$$\begin{aligned} \sigma_{\text{HK}}^2 &> \min_{q, a_3} \left( \frac{1}{1+g^2} (qg - a_3)^2 \text{SNR} \right) \\ &\geq \frac{1}{2} \min_{q, a_3} \left( (qg - a_3)^2 \text{SNR} \right). \end{aligned} \quad (125)$$

For  $b = 1, 2, \dots, \lceil 1/6 \log(\text{SNR}) \rceil$  we define the sets

$$\mathcal{G}_b = \{g : g \in [2^{-b}, 2^{-b+1})\}, \quad (126)$$

and the quantities

$$\begin{aligned} q_{\max,b} &\triangleq \sqrt{2^{-b+1}} \text{SNR}^{1/4 - \delta/2}, \\ \Phi_b &\triangleq \frac{1}{\sqrt{2^{-b+1}}} \text{SNR}^{-1/4 - \delta/2}. \end{aligned}$$

Let  $\mathcal{S}_b$  be the set of all values of  $g \in \mathcal{G}_b$  such that the inequality

$$|qg - a_3| < \Phi_b \quad (127)$$

has at least one solution with  $0 < |q| \leq q_{\max,b}$  and  $a_3 \in \mathbb{Z}$ . Note that since  $q = \tilde{\beta} - \varphi$  and we assume in this interval that  $1/2 < |\tilde{\beta}| \leq \sqrt{|g|} \text{SNR}^{1/4 - \delta/2} / 2$ , we have

$$|q| < \sqrt{|g|} \text{SNR}^{1/4 - \delta/2}.$$

Thus, for all  $g \in \mathcal{G}_b$  and  $\tilde{\beta}$  in the considered interval, we have

$$|q| < q_{\max,b}.$$

Let  $\tilde{\mathcal{S}}_b = \mathcal{G}_b \setminus \mathcal{S}_b$ . Using (125), we have that for all  $g \in \tilde{\mathcal{S}}_b$  and  $\tilde{\beta}$  in the considered interval

$$\begin{aligned} \sigma_{\text{HK}}^2 &\geq \frac{1}{2} \Phi_b^2 \text{SNR} \\ &\geq \frac{1}{2} \frac{\text{SNR}^{1/2 - \delta}}{2^{-b+1}} \\ &\geq \frac{1}{4} \frac{\text{SNR}^{1/2 - \delta}}{\sqrt{g^2}}. \end{aligned} \quad (128)$$

The condition (127), which defines the set  $\mathcal{S}_b$ , can be written equivalently as

$$|q \cdot 2^b g - 2^b a_3| < 2^b \Phi_b. \quad (129)$$

Define  $\tilde{g} = 2^b \cdot g$ , and note that for all  $g \in \mathcal{G}_b$  we have  $\tilde{g} \in [1, 2)$ . With this notation, (129) becomes

$$\left| \tilde{g} - \frac{2^b a_3}{q} \right| < 2^b \frac{\Phi_b}{q}. \quad (130)$$

Define

$$\begin{aligned} \mathcal{I}_b(q) &= \left[ \left\{ \frac{0}{q}, \frac{1 \cdot 2^b}{q}, \frac{2 \cdot 2^b}{q}, \dots, \frac{\lfloor \frac{2q-1}{2^b} \rfloor \cdot 2^b}{q} \right\} \right. \\ &\quad \left. + 2^b \frac{\Phi_b}{q} \mathcal{I} \right] \bmod [0, 2), \end{aligned} \quad (131)$$

where  $\mathcal{I} = [-1, 1)$  and the sum in (131) is a Minkowski sum.

It is easy to verify that

$$\mathcal{S}_b \subseteq 2^{-b} \bigcup_{q=1}^{\lfloor q_{\max} \rfloor} \mathcal{T}_b(q). \quad (132)$$

Since  $\lfloor \frac{2q-1}{2^b} \rfloor = 0$  for all  $0 < q < 2^{b-1}$ , for all values of  $q$  in this range we have

$$\begin{aligned} \mathcal{T}_b(q) &= \left[ 2^b \frac{\Phi_b \mathcal{I}}{q} \right] \bmod [0, 2) \\ &\subseteq [2^b \Phi_b \mathcal{I}] \bmod [0, 2) \\ &= \mathcal{T}_b(1). \end{aligned} \quad (133)$$

Therefore,

$$\begin{aligned} \mathcal{S}_b &\subseteq 2^{-b} \left( \left( \bigcup_{q=1}^{2^{b-1}-1} \mathcal{T}_b(q) \right) \cup \left( \bigcup_{q=2^{b-1}}^{\lfloor q_{\max} \rfloor} \mathcal{T}_b(q) \right) \right) \\ &= 2^{-b} \left( \mathcal{T}_b(1) \cup \left( \bigcup_{q=2^{b-1}}^{\lfloor q_{\max} \rfloor} \mathcal{T}_b(q) \right) \right). \end{aligned} \quad (134)$$

The Lebesgue measure of  $\mathcal{S}_b$  is bounded by

$$\begin{aligned} \mu(\mathcal{S}_b) &= \text{Vol}(\mathcal{S}_b) \\ &\leq 2^{-b} \left( \text{Vol}(\mathcal{T}_b(1)) + \sum_{q=2^{b-1}}^{\lfloor q_{\max, b} \rfloor} \text{Vol}(\mathcal{T}_b(q)) \right) \\ &\leq 2^{-b} \left( 2 \cdot 2^b \Phi_b + \sum_{q=2^{b-1}}^{\lfloor q_{\max, b} \rfloor} \left\lceil \frac{2q}{2^b} \right\rceil \cdot 2 \cdot 2^b \frac{\Phi_b}{q} \right) \\ &\leq 2\Phi_b + 2\Phi_b \sum_{q=2^{b-1}}^{\lfloor q_{\max, b} \rfloor} 2 \frac{2q}{2^b} \frac{1}{q} \\ &\leq 2\Phi_b + 8 \cdot 2^{-b} \Phi_b q_{\max, b} \\ &= 2\Phi_b + 8 \cdot 2^{-b} \text{SNR}^{-\delta} \\ &= \sqrt{2} \cdot 2^{b/2} \text{SNR}^{-1/4-\delta/2} + 8 \cdot 2^{-b} \text{SNR}^{-\delta}. \end{aligned} \quad (135)$$

We can now upper bound the measure of the outage set

$$\mathcal{S} = \bigcup_{b=1}^{\lceil 1/6 \log(\text{SNR}) \rceil} \mathcal{S}_b,$$

of all values of  $\text{SNR}^{-1/6} \leq g < 1$  for which (128) does not necessarily hold, as

$$\begin{aligned} \mu(\mathcal{S}) &= \sum_{b=1}^{\lceil 1/6 \log(\text{SNR}) \rceil} \mu(\mathcal{S}_b) \\ &< \sqrt{2} \text{SNR}^{-1/4-\delta/2} \sum_{b=1}^{\lceil 1/6 \log(\text{SNR}) \rceil + 1} (\sqrt{2})^b \\ &\quad + 8 \text{SNR}^{-\delta} \sum_{b=1}^{\lceil 1/6 \log(\text{SNR}) \rceil + 1} 2^{-b}. \end{aligned}$$

Using the identity

$$\sum_{b=1}^B \rho^b = \frac{\rho}{\rho-1} (\rho^B - 1),$$

which is valid for all  $\rho \neq 1$ , and the fact that  $\sum_{b=1}^{\infty} 2^{-b} < 1$ , we have

$$\begin{aligned} \mu(\mathcal{S}) &< \sqrt{2} \text{SNR}^{-1/4-\delta/2} \frac{\sqrt{2}}{\sqrt{2}-1} \sqrt{2} \text{SNR}^{1/12} + 8 \text{SNR}^{-\delta} \\ &< 7 \text{SNR}^{-\delta/2} + 8 \text{SNR}^{-\delta} \\ &< 16 \text{SNR}^{-\delta/2}. \end{aligned} \quad (136)$$

Substituting  $\delta = (2c+8)/\log(\text{SNR})$  into (128) and (136), we see that in the interval  $1/2 < |\tilde{\beta}| \leq \sqrt{|g|} \text{SNR}^{1/4-\delta/2}/2$  for all values of  $\text{SNR}^{-1/3} < |g| \leq 1$  except for an outage set with measure not greater than  $2^{-c}$  we have

$$\sigma_{\text{HK}}^2 > \frac{2^{-2c} \text{SNR}^{1/2}}{4 \cdot 2^8 \sqrt{g^2}}.$$

**Interval 3:**  $\sqrt{|g|} \text{SNR}^{1/4-\delta/2}/2 < |\tilde{\beta}| \leq \text{SNR}^{1/4}/\sqrt{8|g|}$   
Since  $\text{SNR}^{-1/3} \leq g^2 < 1$  and we assumed  $\text{SNR} > 4$ , we have

$$g^2 \text{SNR} - 1 > \frac{g^2 \text{SNR}}{2}. \quad (137)$$

Note that (137) continues to hold for all  $g^2 > \text{SNR}^{-1/2}$ . This will be useful in the weak interference regime. For all values of  $|\tilde{\beta}|$  in this interval

$$\begin{aligned} \left| \frac{\tilde{\beta}}{\sqrt{g^2 \text{SNR} - 1}} \right| &\leq \frac{\text{SNR}^{1/4}/\sqrt{8|g|}}{\sqrt{g^2 \text{SNR} - 1}} \\ &< \frac{\text{SNR}^{1/4}}{\sqrt{8|g|} \cdot \frac{g^2 \text{SNR}}{2}} \\ &\leq \frac{1}{2} |g|^{-3/2} \text{SNR}^{-1/4} \\ &\leq \frac{1}{2}, \end{aligned}$$

and hence, using (91), the optimal value of  $a_2$  is

$$a_2 = \left\lfloor \tilde{\beta} / \sqrt{g^2 \text{SNR} - 1} \right\rfloor = 0.$$

Therefore, using the fact that  $\delta = (2c+8)/\log(\text{SNR})$ , we can upper bound (90) as

$$\sigma_{\text{HK}}^2 \geq \frac{\tilde{\beta}^2 \text{SNR}}{g^2 \text{SNR} - 1} \geq \frac{2^{-2c} \text{SNR}^{1/2}}{4 \cdot 2^8 \sqrt{g^2}}.$$

**Interval 4:**  $\text{SNR}^{1/4}/\sqrt{8|g|} < |\tilde{\beta}|$

In this interval,

$$\sigma_{\text{HK}}^2 \geq K \tilde{\beta}^2 \geq \frac{1}{8} \frac{\text{SNR}^{1/2}}{\sqrt{g^2}}.$$

## B. Weak Interference Regime

We upper bound  $\sigma_{\text{HK}}^2$  for all values of  $\tilde{\beta}$  within each of the four intervals. Recall that in this regime  $\text{SNR}^{-1/2} \leq g^2 < \text{SNR}^{-1/3}$ .

**Interval 1:**  $0 < |\tilde{\beta}| \leq 1/2$

As  $a_1 \neq 0$ , in this interval  $(\tilde{\beta} - a_1)^2 > 1/4$ . Thus,

$$\sigma_{\text{HK}}^2 \geq \frac{\text{SNR}}{4}.$$

**Interval 2:**  $1/2 < |\tilde{\beta}| \leq 1/(2|g|)$

In this interval  $a_3 = \lfloor \tilde{\beta}g \rfloor = 0$ . Thus,

$$\sigma_{\text{HK}}^2 \geq (\tilde{\beta}g)^2 \text{SNR} \geq \frac{g^2 \text{SNR}}{4}.$$

**Interval 3:**  $1/(2|g|) < |\tilde{\beta}| \leq \sqrt{g^2 \text{SNR}/8}$

Under our assumption that  $\text{SNR} > 4$ , for all values of  $|\tilde{\beta}|$  in this interval we have

$$\begin{aligned} \left| \frac{\tilde{\beta}}{\sqrt{g^2 \text{SNR} - 1}} \right| &\leq \frac{\sqrt{g^2 \text{SNR}/8}}{\sqrt{g^2 \text{SNR} - 1}} \\ &< \frac{\sqrt{g^2 \text{SNR}/8}}{\sqrt{g^2 \text{SNR}/2}} \\ &\leq \frac{1}{2}, \end{aligned}$$

where the second inequality follows from (137). Thus, the optimal choice for  $a_2$  is

$$a_2 = \left\lfloor \tilde{\beta} / \sqrt{g^2 \text{SNR} - 1} \right\rfloor = 0. \quad (138)$$

Therefore, (90) can be lower bounded by

$$\sigma_{\text{HK}}^2 \geq \frac{\tilde{\beta}^2}{g^2 \text{SNR}} \text{SNR} \geq \frac{1}{4g^4}.$$

**Interval 4:**  $\sqrt{g^2 \text{SNR}/8} < |\tilde{\beta}|$

In this interval,

$$\sigma_{\text{HK}}^2 \geq K \tilde{\beta}^2 > \frac{g^2 \text{SNR}}{4}.$$

#### APPENDIX E PROOF OF THEOREM 12

For the proof we will need a key result from the field of metric Diophantine approximation which is due to Kleinbock and Margulis. The following theorem is a special case of [62, Th. A].

*Theorem 14:* Let  $\mathcal{U}$  be a domain in  $\mathbb{R}^m$  and let  $f_1, f_2, \dots, f_K$  be real analytic functions in  $\tilde{\mathbf{h}} \in \mathcal{U}$ , which together with 1 are linearly independent over  $\mathbb{R}$ , and define the manifold

$$\mathcal{M} = \left\{ [f_1(\tilde{\mathbf{h}}) \ \dots \ f_K(\tilde{\mathbf{h}})] : \tilde{\mathbf{h}} \in \mathcal{U} \right\}.$$

For almost every  $\mathbf{h} \in \mathcal{M}$  and any  $\delta > 0$ , the inequality

$$\max_{\ell=1, \dots, K} |qh_\ell - a_\ell| \leq |q|^{-\frac{1}{K} - \delta} \quad (139)$$

has at most finitely many solutions  $(q, \mathbf{a}) \in \mathbb{Z} \times \mathbb{Z}^K$ .

For the proof of Theorem 12 we will need a corollary of Theorem 14.

*Corollary 7:* Let  $f_1, f_2, \dots, f_K$  be functions from  $\mathbb{R}^m$  to  $\mathbb{R}$  satisfying the following conditions:

- 1)  $f_i$  for  $i = 1, \dots, K$  is analytic in  $\mathbb{R}^m$ ,
- 2)  $1, f_1, \dots, f_K$  are linearly independent over  $\mathbb{R}$ .

Let  $\mathcal{D} = \{\tilde{\mathbf{h}} \in \mathbb{R}^m : f_1(\tilde{\mathbf{h}}) = 0\}$  and  $\mathcal{D}(\epsilon) = \mathcal{D} + \mathcal{B}(\mathbf{0}, \epsilon)$ , where the sum is a Minkowski sum and  $\mathcal{B}(\mathbf{0}, \epsilon)$  is an  $m$ -dimensional closed ball with some radius  $\epsilon > 0$ . Define the set  $\mathcal{U}(\epsilon) = \mathbb{R}^m \setminus \mathcal{D}(\epsilon)$ , the set of functions  $\tilde{f}_k(\tilde{\mathbf{h}}) = f_k(\tilde{\mathbf{h}})/f_1(\tilde{\mathbf{h}})$

from  $\mathcal{U}(\epsilon)$  to  $\mathbb{R}$  for  $k = 2, \dots, K$ , and the manifold

$$\tilde{\mathcal{M}}(\epsilon) = \left\{ [\tilde{f}_2(\tilde{\mathbf{h}}) \ \dots \ \tilde{f}_K(\tilde{\mathbf{h}})] : \tilde{\mathbf{h}} \in \mathcal{U}(\epsilon) \right\}. \quad (140)$$

For all  $\epsilon > 0$ , almost every  $\tilde{\mathbf{h}} \in \tilde{\mathcal{M}}(\epsilon)$ , and any  $\delta > 0$  the inequality

$$\max_{\ell=1, \dots, K-1} |q\tilde{h}_\ell - a_\ell| \leq |q|^{-\frac{1}{K-1} - \delta} \quad (141)$$

has at most finitely many solutions  $(q, \mathbf{a}) \in \mathbb{Z} \times \mathbb{Z}^{K-1}$ .

*Proof of Corollary 7:* We would like to apply Theorem 14 for the set of functions  $\tilde{f}_2, \dots, \tilde{f}_K$  from  $\mathcal{U}(\epsilon)$  to  $\mathbb{R}$ . To that end we have to show that for all  $\epsilon > 0$  the functions  $\tilde{f}_2, \dots, \tilde{f}_K$  are analytic in  $\mathcal{U}(\epsilon)$  and together with 1 are linearly independent over  $\mathbb{R}$ .

The reciprocal of an analytic function that is nowhere zero is analytic. Thus, for any  $\epsilon > 0$ , the function  $1/f_1(\tilde{\mathbf{h}})$  is analytic in  $\mathcal{U}(\epsilon)$ . Furthermore, the product of two analytic functions is analytic. Therefore, for any  $\epsilon > 0$ , the functions  $\tilde{f}_k = f_k(\tilde{\mathbf{h}}) \cdot (1/f_1(\tilde{\mathbf{h}}))$  are analytic in  $\mathcal{U}(\epsilon)$  for  $k = 2, \dots, K$ .

We show that the functions  $1, \tilde{f}_2, \dots, \tilde{f}_K$  from  $\mathcal{U}(\epsilon)$  to  $\mathbb{R}$  are linearly independent for all  $\epsilon > 0$  by contradiction. Assume they are linearly dependent. Thus, there exists a measurable set  $\mathcal{S} \in \mathcal{U}(\epsilon)$  and a set of coefficients  $\{t_1(\epsilon), \dots, t_K(\epsilon)\} \in \mathbb{R}$  not all zero such that  $\forall \tilde{\mathbf{h}} \in \mathcal{S}$

$$t_1(\epsilon) \cdot 1 + t_2(\epsilon) \cdot \frac{f_2(\tilde{\mathbf{h}})}{f_1(\tilde{\mathbf{h}})} + \dots + t_K(\epsilon) \cdot \frac{f_K(\tilde{\mathbf{h}})}{f_1(\tilde{\mathbf{h}})} = 0.$$

This implies that  $\forall \tilde{\mathbf{h}} \in \mathcal{S}$

$$0 \cdot 1 + t_1(\epsilon) \cdot f_1(\tilde{\mathbf{h}}) + t_2(\epsilon) \cdot f_2(\tilde{\mathbf{h}}) + \dots + t_K(\epsilon) \cdot f_K(\tilde{\mathbf{h}}) = 0,$$

in contradiction to the assumption that the functions  $1, f_1, \dots, f_K$  from  $\mathbb{R}^m$  to  $\mathbb{R}$  are linearly independent over  $\mathbb{R}$ .

We can therefore apply Theorem 14 to the set of functions  $\tilde{f}_2, \dots, \tilde{f}_K$  from  $\mathcal{U}(\epsilon)$  to  $\mathbb{R}$  for all  $\epsilon > 0$ , and the corollary follows. ■

We are now ready to prove Theorem 12. Define the sets  $\mathcal{D}$ ,  $\mathcal{D}(\epsilon)$  and  $\mathcal{U}(\epsilon)$  as in Corollary 7, and the manifold

$$\tilde{\mathcal{M}}(\epsilon) = \left\{ [f_1(\tilde{\mathbf{h}}) \ \dots \ f_K(\tilde{\mathbf{h}})] : \tilde{\mathbf{h}} \in \mathcal{U}(\epsilon) \right\}.$$

We begin by showing that, for any  $\epsilon > 0$ , the DoF offered by the first computation rate is upper bounded by

$$d_{\text{comp},1} \leq \frac{1}{K} \quad (142)$$

for almost every  $\mathbf{h} \in \tilde{\mathcal{M}}(\epsilon)$ . Then we take  $\epsilon$  to zero in order to show that this holds for almost every

$$\mathbf{h} \in \mathcal{M} = \left\{ [f_1(\tilde{\mathbf{h}}) \ \dots \ f_K(\tilde{\mathbf{h}})] : \tilde{\mathbf{h}} \in \mathbb{R}^m \right\}.$$

Consider the manifold  $\tilde{\mathcal{M}}(\epsilon)$  for some  $\epsilon > 0$ . Note that  $h_1 = f_1(\tilde{\mathbf{h}}) \neq 0$  for any  $\tilde{\mathbf{h}} \in \tilde{\mathcal{M}}(\epsilon)$ , and we can therefore define  $\bar{\mathbf{h}} = \mathbf{h}/h_1$ . We have  $\bar{h}_1 = 1$  and  $[\bar{h}_2 \ \dots \ \bar{h}_K] \in \tilde{\mathcal{M}}(\epsilon)$ , where  $\tilde{\mathcal{M}}(\epsilon)$  is the manifold from (140) in Corollary 7.

The channel (4) is equivalent to the channel

$$\bar{\mathbf{y}} = \frac{1}{h_1} \mathbf{y} = \mathbf{x}_1 + \sum_{k \neq 1} \bar{h}_k \mathbf{x}_k + \frac{1}{h_1} \mathbf{z}. \quad (143)$$

Let  $\mathbf{a}$  be a vector of integer coefficients, and  $\beta$  be the scaling factor used by the receiver in order to decode the linear combination  $\mathbf{v} = [\sum_{k=1}^K a_k \mathbf{t}_k] \bmod \Lambda$ , see Section III. The effective noise encountered in decoding the linear combination  $\mathbf{v}$  with coefficient vector  $\mathbf{a}$  is

$$\mathbf{z}_{\text{eff}}(\mathbf{h}, \mathbf{a}, \beta) = (\beta - a_1)\mathbf{x}_1 + \sum_{k \neq 1} (\beta \bar{h}_k - a_k)\mathbf{x}_k + \frac{\beta}{h_1}\mathbf{z},$$

and its effective variance is given by

$$\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, \beta) = (\beta - a_1)^2 \text{SNR} + \sum_{k \neq 1} (\beta \bar{h}_k - a_k)^2 \text{SNR} + \frac{\beta^2}{|h_1|^2}. \quad (144)$$

Recall that

$$\begin{aligned} R_{\text{comp},1} &= \max_{\mathbf{a}, \beta} \frac{1}{2} \log \left( \frac{\text{SNR}}{\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, \beta)} \right) \\ &= \frac{1}{2} \log(\text{SNR}) - \frac{1}{2} \log \left( \min_{\mathbf{a}, \beta} \sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, \beta) \right). \end{aligned} \quad (145)$$

Thus, in order to obtain an upper bound on  $R_{\text{comp},1}$  we need to lower bound  $\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, \beta)$  for all values of  $\beta \in \mathbb{R}$  and  $\mathbf{a} \in \mathbb{Z}^K \setminus \mathbf{0}$ . Let

$$\bar{h}^* = \max_{k=1, \dots, K} \bar{h}_k,$$

and

$$k^* = \arg \max_{k=1, \dots, K} \bar{h}_k.$$

Note that if  $|\beta| < 1/(2\bar{h}^*)$  the minimizing corresponding choice of integers  $a_1, \dots, a_K$  in (144) is  $a_{k^*} = \text{sign}(\beta)$ , and  $a_k = 0$  for all  $k \neq k^*$ . This in turn, implies that for  $|\beta| < 1/(2\bar{h}^*)$  we have

$$\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, \beta) > (\beta \bar{h}^* - \text{sign}(\beta))^2 \text{SNR} > \frac{\text{SNR}}{4}, \quad (146)$$

which means  $d_{\text{comp},1} = 0$ . Thus, in order to obtain a positive DoF,  $|\beta|$  must be greater than  $1/(2\bar{h}^*)$ .

If  $1/(2\bar{h}^*) \leq |\beta| \leq 1/2$ , then the minimizing value of  $a_1$  in (144) is  $a_1 = 0$ . This implies that for all values of  $1/(2\bar{h}^*) \leq |\beta| \leq 1/2$  we have

$$\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, \beta) > \beta^2 \text{SNR} > \frac{\text{SNR}}{4(\bar{h}^*)^2}, \quad (147)$$

which also means  $d_{\text{comp},1} = 0$ . Thus, in order to obtain a positive DoF,  $|\beta|$  must be greater than  $1/2$ .

Hence, in order to lower bound (144) in the limit of very high SNR, it suffices to limit the optimization space of  $\beta$  to  $|\beta| > 1/2$ . For such values,  $\beta$  can be written in the form  $\beta = q + \varphi$  where  $\varphi \in [-1/2, 1/2)$ , and  $q \in \mathbb{Z} \setminus 0$  is the nearest integer to  $\beta$ .

For any  $|\varphi| < 1/2$ ,  $q \in \mathbb{Z} \setminus 0$  and  $\mathbf{a} \in \mathbb{Z}^K \setminus \mathbf{0}$  we have

$$\begin{aligned} &\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, q, \varphi) \\ &\geq (\varphi + q - a_1)^2 \text{SNR} \\ &\quad + \max_{k \neq 1} (q \bar{h}_k - a_k + \varphi \bar{h}_k)^2 \text{SNR} + \frac{(q/2)^2}{|h_1|^2} \\ &\geq \varphi^2 \text{SNR} + \max_{k \neq 1} (q \bar{h}_k - a_k + \varphi \bar{h}_k)^2 \text{SNR} + \frac{(q/2)^2}{|h_1|^2} \end{aligned}$$

$$= \max_{k \neq 1} \left( (\varphi^2 + (q \bar{h}_k - a_k + \varphi \bar{h}_k)^2) \text{SNR} + \frac{q^2}{|2h_1|^2} \right) \quad (148)$$

We further bound (148) by substituting the minimizing value of  $\varphi$  for each  $k \neq 1$ . It follows by simple differentiation that for each  $k \neq 1$  the minimum occurs at

$$\varphi^*(k) = \frac{-\bar{h}_k}{1+\bar{h}_k^2} (q \bar{h}_k - a_k).$$

Substituting  $\varphi^*(k)$  into (148) yields

$$\begin{aligned} &\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, q, \varphi) \\ &\geq \max_{k \neq 1} \left( \frac{1}{1+\bar{h}_k^2} (q \bar{h}_k - a_k)^2 \text{SNR} + \frac{q^2}{|2h_1|^2} \right) \\ &> \frac{1}{1+\max_{k \neq 1} \bar{h}_k^2} \cdot \max_{k \neq 1} |q \bar{h}_k - a_k|^2 \text{SNR} + \frac{q^2}{|2h_1|^2} \\ &\geq c_0(\mathbf{h}) \cdot \left( \max_{k \neq 1} |q \bar{h}_k - a_k|^2 \text{SNR} + q^2 \right), \end{aligned} \quad (149)$$

where  $c_0(\mathbf{h}) > 0$  is some constant independent of the SNR.

Consider the limit of  $\text{SNR} \rightarrow \infty$ , and assume  $|q|$  is upper bounded by some finite integer  $q_0 > 0$ . Then, for almost every  $\mathbf{h} \in \tilde{\mathcal{M}}(\epsilon)$ , there exists a constant  $c_1(\mathbf{h}, q_0) > 0$ , independent of the SNR, for which

$$\max_{k \neq 1} |q \bar{h}_k - a_k| > c_1(\mathbf{h}, q_0) \quad (150)$$

for all  $0 < |q| \leq q_0$  and  $\mathbf{a} \in \mathbb{Z}^{K-1}$ . Note that  $\mathbf{h}$  does not satisfy (150) only if all elements of  $\bar{\mathbf{h}}$  are rational. Substituting (150) into (149) gives  $\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, q, \varphi) > c_2(\mathbf{h}, q_0) \text{SNR}$  which means that the DoF is zero. Therefore, in order to get a positive DoF,  $q$  must tend to infinity when the SNR tends to infinity.

Any positive integer  $|q|$  can be expressed as  $|q| = \text{SNR}^\gamma$  for some  $\gamma > 0$ . From Corollary 7 we know that for any  $\epsilon, \delta > 0$ , almost every  $\mathbf{h} \in \tilde{\mathcal{M}}(\epsilon)$ , and  $q$  large enough, we have

$$\max_{k \neq 1} |q \bar{h}_k - a_k| > |q|^{-\frac{1}{K-1} - \delta} = \text{SNR}^{-\frac{\gamma}{K-1} - \gamma\delta}. \quad (151)$$

Thus, for  $|q|$  large enough and almost every  $\mathbf{h} \in \tilde{\mathcal{M}}(\epsilon)$ , we have

$$\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, q, \varphi) \geq c_0^2(\mathbf{h}) \cdot \max \left( \text{SNR}^{1 - \frac{2\gamma}{K-1} - 2\gamma\delta}, \text{SNR}^{2\gamma} \right). \quad (152)$$

Minimizing (152) with respect to  $\gamma$  gives

$$\gamma = \frac{K-1}{2(K + \delta K - \delta)}.$$

Hence, for all  $q \in \mathbb{Z}$ ,  $\varphi \in [-1/2, 1/2)$ ,  $\mathbf{a} \in \mathbb{Z}^K \setminus \mathbf{0}$  and almost every  $\mathbf{h} \in \tilde{\mathcal{M}}(\epsilon)$

$$\sigma_{\text{eff}}^2(\mathbf{h}, \mathbf{a}, q, \varphi) > c_3(\mathbf{h}) \text{SNR}^{\frac{K-1}{K+\delta(K-1)}}, \quad (153)$$

where  $c_3(\mathbf{h}) > 0$  is also a constant independent of the SNR. Substituting into (145) gives

$$\begin{aligned} R_{\text{comp},1} &< \frac{1 + \delta(K-1)}{K + \delta(K-1)} \cdot \frac{1}{2} \log(\text{SNR}) \\ &\quad - \frac{1}{2} \log(c_3(\mathbf{h})) \end{aligned} \quad (154)$$

for any  $\delta > 0$ . Taking  $\delta \rightarrow 0$ , it follows that the DoF the highest computation rate offers is upper bounded by

$$\lim_{\text{SNR} \rightarrow \infty} \frac{R_{\text{comp},1}}{\frac{1}{2} \log(1 + \text{SNR})} \leq \frac{1}{K}, \quad (155)$$

for almost every  $\mathbf{h} \in \tilde{\mathcal{M}}(\epsilon)$ . Since this holds for all  $\epsilon > 0$ , we can now take  $\epsilon$  to zero (note that the bound does not depend on  $\epsilon$ ). The set  $\mathcal{D}$  has measure zero since  $f_1$  is analytic on  $\mathbb{R}^m$  and is not identically zero (otherwise, the set of functions  $1, f_1, \dots, f_K$  is not linearly independent). Note that the measure of  $\mathcal{D}(\epsilon)$  goes to zero as  $\epsilon \rightarrow 0$ , and furthermore  $\mathcal{D} = \bigcap_{\epsilon > 0} \mathcal{D}(\epsilon)$ . Therefore, the claim holds for almost every  $\mathbf{h} \in \mathcal{M}$ .

#### APPENDIX F PROOF OF THEOREM 13

Consider the reference  $L$ -user MAC

$$\mathbf{y}_{\text{ref}} = \sum_{\ell=1}^L g_{\ell} \mathbf{x}_{\ell} + \mathbf{z}, \quad (156)$$

where  $\mathbf{z}$  is AWGN with zero mean and unit variance and all users are subject to the power constraint  $\|\mathbf{x}_{\ell}\|^2 \leq n\text{SNR}$ . Applying Corollary 5 to this channel implies that, for almost every  $\mathbf{g} \in \mathcal{M}$ , the DoF that each optimal computation rate offers is  $1/L$ . Let  $R_{\text{comp}}^{\text{ref}}(\mathbf{g}, \mathbf{a})$  be the computation rate corresponding to the coefficient vector  $\mathbf{a}$  over the reference MAC (156). We now show the computation rate of the same coefficient vector  $R_{\text{comp}}(\mathbf{g}, \mathbf{a}, \mathbf{B})$  over the effective MAC (33) is within a constant number of bits from  $R_{\text{comp}}^{\text{ref}}(\mathbf{g}, \mathbf{a})$ .

For the reference channel (156) the effective noise variance for a given  $\mathbf{a}$  and  $\beta$  is

$$\sigma_{\text{ref}}^2(\mathbf{g}, \mathbf{a}, \beta) = \text{SNR} \|\beta \mathbf{g} - \mathbf{a}\|^2 + \beta^2,$$

while for the effective  $L$ -user MAC (33) the effective variance for the same  $\mathbf{a}$  and  $\beta$  is

$$\sigma_{\text{eff}}^2(\mathbf{g}, \mathbf{a}, \beta, \mathbf{B}) = \text{SNR} \sum_{\ell=1}^L (\beta g_{\ell} - a_{\ell})^2 b_{\text{eff},\ell}^2 + \beta^2.$$

Letting  $b^* = \max_{\ell=1, \dots, L} b_{\text{eff},\ell}^2$  and noting that  $b^* \geq 1$  gives

$$\sigma_{\text{ref}}^2(\mathbf{g}, \mathbf{a}, \beta) \leq \sigma_{\text{eff}}^2(\mathbf{g}, \mathbf{a}, \beta, \mathbf{B}) \leq b^* \sigma_{\text{ref}}^2(\mathbf{g}, \mathbf{a}, \beta).$$

Since the above inequalities are valid for any value of  $\beta$ , in particular they hold true for the optimal value of  $\beta$  and it follows that

$$R_{\text{comp}}^{\text{ref}}(\mathbf{g}, \mathbf{a}) - \frac{1}{2} \log(b^*) \leq R_{\text{comp}}(\mathbf{g}, \mathbf{a}, \mathbf{B}) \leq R_{\text{comp}}^{\text{ref}}(\mathbf{g}, \mathbf{a}).$$

As  $b^*$  is independent of the SNR, it follows that the DoF offered by each computation rate for the reference and effective MACs (156) and (33) are equal. In particular, this is the case for the optimal computation rates, thus the theorem follows.

#### REFERENCES

- [1] M. A. Maddah-Ali, A. S. Motahari, and A. K. 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.
- [2] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom for the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.
- [3] B. Nazer, M. Gastpar, S. A. Jafar, and S. Vishwanath, "Ergodic interference alignment," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6355–6371, Oct. 2012.
- [4] A. B. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 1, pp. 60–70, Jan. 1978.
- [5] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.
- [6] S. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.
- [7] G. Bresler, A. Parekh, and D. 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.
- [8] A. S. Motahari, S. O. Gharan, M.-A. Maddah-Ali, and A. K. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," *IEEE Trans. Inf. Theory*, accepted [Online]. Available: <http://arxiv.org/abs/0908.2282>
- [9] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1250–1276, Jun. 2002.
- [10] U. Erez and R. Zamir, "Achieving  $\frac{1}{2} \log(1 + \text{SNR})$  on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.
- [11] R. Zamir, "Lattices are everywhere," in *Proc. 4th Annu. Workshop ITA, La Jolla, CA, USA*, Feb. 2009, pp. 392–421.
- [12] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 620–643, Feb. 2009.
- [13] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 689–699, Feb. 2009.
- [14] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3032–3050, Jul. 2009.
- [15] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.
- [16] H. Sato, "The capacity of the Gaussian interference channel under strong interference (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 27, no. 6, pp. 786–788, Nov. 1981.
- [17] A. B. Carleial, "A case where interference does not reduce capacity (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 21, no. 5, pp. 569–570, Sep. 1975.
- [18] 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.
- [19] T. M. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.
- [20] S. Sridharan, A. Jafarian, S. Vishwanath, and S. A. Jafar, "Capacity of symmetric K-user Gaussian very strong interference channels," in *Proc. IEEE Global Commun. Conf.*, New Orleans, LA, USA, Dec. 2008, pp. 1–5.
- [21] 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.
- [22] Y. Wu, S. Shamai, and S. Verdú, "Degrees of freedom of the interference channel: A general formula," in *Proc. IEEE ISIT*, St. Petersburg, Russia, Aug. 2011, pp. 1362–1366.
- [23] O. Ordentlich and U. Erez, "On the robustness of lattice interference alignment," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2735–2759, May 2013.
- [24] U. Niesen and M. A. Maddah-Ali, "Interference alignment: From degrees-of-freedom to constant-gap capacity approximations," *IEEE Trans. Inf. Theory*, vol. 59, no. 8, pp. 4855–4888, Aug. 2013.
- [25] A. G. Dimakis, K. Ramchandran, Y. Wu, and C. Suh, "A survey on network codes for distributed storage," *Proc. IEEE*, vol. 99, no. 3, pp. 476–489, Mar. 2011.
- [26] S. A. Jafar, "Interference alignment—A new look at signal dimensions in a communication network," *Found. Trends Commun. Inf. Theory*, vol. 7, no. 1, pp. 1–134, 2011.



- [27] G. Bresler and D. N. C. Tse, "3-user interference channel: Degrees of freedom as a function of channel diversity," in *Proc. 47th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Sep. 2009, pp. 265–271.
- [28] C. Wang, T. Gou, and S. A. Jafar, "Subspace alignment chains and the degrees of freedom of the three-user MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2432–2479, May 2014.
- [29] S. Sridharan, A. Jafarian, S. Vishwanath, S. A. Jafar, and S. Shamai, "A layered lattice coding scheme for a class of three-user Gaussian interference channels," in *Proc. 46th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Sep. 2008, pp. 531–538.
- [30] 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.
- [31] S. Saha and R. Berry, "Sum-capacity of a class of K-user Gaussian interference channels within  $O(K)$  bits," in *Proc. 49th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Sep. 2011, pp. 847–854.
- [32] L. Zhou and W. Yu, "On the capacity of the K-user cyclic Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 154–165, Jan. 2013.
- [33] B. Bandemer and A. El Gamal, "Interference decoding for deterministic channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2966–2975, May 2011.
- [34] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for canceling known interference," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3820–3833, Nov. 2005.
- [35] T. Philosof, R. Zamir, U. Erez, and A. J. Khisti, "Lattice strategies for the dirty multiple access channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5006–5035, Aug. 2011.
- [36] D. Krithivasan and S. S. Pradhan, "Lattices for distributed source coding: Jointly Gaussian sources and reconstruction of a linear function," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5268–5651, Dec. 2009.
- [37] A. B. Wagner, "On distributed compression of linear functions," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 79–94, Jan. 2011.
- [38] D. N. C. Tse and M. A. Maddah-Ali, "Interference neutralization in distributed lossy source coding," in *Proc. IEEE ISIT*, Austin, TX, USA, Jun. 2010, pp. 166–170.
- [39] B. Nazer, A. Sanderovich, M. Gastpar, and S. Shamai, "Structured superposition for backhaul constrained cellular uplink," in *Proc. IEEE ISIT*, Seoul, South Korea, Jun. 2009, pp. 1530–1534.
- [40] S.-N. Hong and G. Caire, "Compute-and-forward strategies for cooperative distributed antenna systems," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5227–5243, Sep. 2013.
- [41] B. Nazer and M. Gastpar, "Reliable physical layer network coding," *Proc. IEEE*, vol. 99, no. 3, pp. 438–460, Mar. 2011.
- [42] 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. 5641–5654, Nov. 2010.
- [43] W. Nam, S.-Y. Chung, and Y. H. Lee, "Capacity of the Gaussian two-way relay channel to within  $1/2$  bit," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5488–5494, Nov. 2010.
- [44] C. Feng, D. Silva, and F. Kschischang, "An algebraic approach to physical-layer network coding," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7576–7596, Nov. 2013.
- [45] R. Zamir, *Lattice Coding for Signals and Networks*. Cambridge Univ. Press, 2014.
- [46] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 219–221, Mar. 1979.
- [47] S. A. Jafar and S. Vishwanath, "Generalized degrees of freedom of the symmetric Gaussian K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3297–3303, Jul. 2010.
- [48] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.
- [49] O. Ordentlich and U. Erez, "A simple proof for the existence of 'good' pairs of nested lattices," in *Proc. 27th Conv. IEEEI*, Nov. 2012, pp. 1–12.
- [50] D. A. Harville, *Matrix Algebra From a Statistician's Perspective*. New York, NY, USA: Springer-Verlag, 1997.
- [51] B. Hern and K. Narayanan, "Multilevel coding schemes for compute-and-forward with flexible decoding," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7613–7631, Nov. 2013.
- [52] O. Ordentlich, J. Zhan, U. Erez, B. Nazer, and M. Gastpar, "Practical code design for compute-and-forward," in *Proc. IEEE ISIT*, St. Petersburg, Russia, Aug. 2011, pp. 1876–1880.
- [53] N. E. Tunalı and K. R. Narayanan, "Concatenated signal codes with applications to compute and forward," in *Proc. IEEE Global Commun. Conf.*, Houston, TX, USA, Dec. 2011, pp. 1–5.
- [54] J.-C. Belfiore and C. Ling, "The flatness factor in lattice network coding: Design criterion and decoding algorithm," in *Proc. IZS Commun.*, Zürich, Switzerland, Mar. 2012.
- [55] Y.-C. Huang, K. R. Narayanan, and N. E. Tunalı. (2014, Jan.). Multistage compute-and-forward with multilevel lattice codes based on product constructions. *IEEE Trans. Inf. Theory*, submitted for publication [Online]. Available: <http://arxiv.org/abs/1401.2228>
- [56] A. K. Lenstra, H. W. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," *Math. Ann.*, vol. 261, no. 4, pp. 515–534, 1982.
- [57] D. Wubben, R. Bohnke, V. Kuhn, and K. D. Kammeyer, "Near-maximum-likelihood detection of MIMO systems using MMSE-based lattice reduction," in *Proc. IEEE Int. Conf. Commun.*, vol. 2. Paris, France, Jun. 2004, pp. 798–802.
- [58] D. Micciancio and S. Goldwasser, *Complexity of Lattice Problems: A Cryptographic Perspective* (Engineering and Computer Science), vol. 671. Cambridge, U.K.: Kluwer, 2002.
- [59] B. Nazer, "Successive compute-and-forward," in *Proc. IZS Commun.*, Zürich, Switzerland, Mar. 2012.
- [60] O. Ordentlich, U. Erez, and B. Nazer, "Successive integer-forcing and its sum-rate optimality," in *Proc. 51th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Oct. 2013.
- [61] W. M. Schmidt, *Diophantine Approximation*. New York, NY, USA: Springer-Verlag, 1980.
- [62] D. Y. Kleinbock and G. A. Margulis, "Flows on homogeneous spaces and Diophantine approximation on manifolds," *Ann. Math.*, vol. 148, no. 1, pp. 339–360, Jul. 1998.
- [63] U. Niesen and P. Whiting, "The degrees of freedom of compute-and-forward," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5214–5232, Aug. 2012.

**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 Thalheimer Scholarship for graduate students, the Advanced Communication Center (ACC) Feder Family award for outstanding research work in the field of communication technologies (2011,2014), and the Weinstein Prize for research in signal processing (2011,2013).

**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.

**Bobak Nazer** (S'02–M'09) received the B.S.E.E. degree from Rice University, Houston, TX, in 2003, the M.S. degree from the University of California, Berkeley, CA, in 2005, and the Ph.D degree from the University of California, Berkeley, CA, in 2009, all in electrical engineering.

He is currently an Assistant Professor in the Department of Electrical and Computer Engineering at Boston University, Boston, MA. From 2009 to 2010, he was a postdoctoral associate in the Department of Electrical and Computer Engineering at the University of Wisconsin, Madison, WI. His research interests include network information theory and wireless communications.

Dr. Nazer received the Eli Jury Award from the EECS Department at UC Berkeley in 2009, the Dean's Catalyst Award from the College of Engineering at BU in 2011, the NSF CAREER Award in 2013, and the IEEE Communications Society and Information Theory Society Joint Paper Award in 2013.