Fifty-fourth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 27 - 30, 2016

Diophantine Approximation for Network Information Theory: A Survey of Old and New Results

Bobak Nazer
Boston University
bobak@bu.edu

Or Ordentlich
MIT
ordent@mit.edu

***Abstract*—Recent developments in network information theory, such as interference alignment and compute-and-forward, have highlighted connections to the branch of mathematics known as Diophantine approximation. At a high level, Diophantine approximation is concerned with how well the reals can be approximated by rationals. This paper surveys some classical and modern Diophantine results and demonstrates their applications in establishing degrees-of-freedom and capacity bounds.**

I. Introduction

It is by now well-known that codes with algebraic structure (e.g., lattice codes, linear codes) are quite useful for establishing achievability theorems in network information theory. Scenarios where random structured codebooks are useful include distributed source coding [1], [2], relaying [3]–[5], interference alignment [6]–[10], dirty-paper multiple-access [11]–[13], and physical-layer secrecy [14]–[16]. See the recent textbook of Zamir for a comprehensive treatment of lattice codes and their applications [17]. It has also been noted that these achievability results are often quite sensitive to how close the source or channel parameters are to integers or rationals. This stands in contrast to the behavior of achievability results proven with random i.i.d. codebooks, which are often monotonic functions of the parameters.

At a high level, Diophantine approximation is a branch of number theory that studies how well real numbers can be approximated by rationals (with a bounded denominator) [18], [19]. Classical and modern results from this field have been employed to establish the degrees-of-freedom for several structured coding strategies. For instance, the real interference alignment strategy of Motahari *et al.* [7] uses approximation results for manifolds [20], [21] to establish that $K/2$ degrees-of-freedom are achievable K-user Gaussian interference channel for almost every channel matrix.

In 2014 and 2016, the York Workshop on Interactions between Number Theory and Wireless Communication provided an opportunity for number theorists and information theorists to exchange ideas and discuss open problems of mutual interest. For instance, most Diophantine approximation results are stated as "zero-one laws," which seems to limit their applicability to degrees-of-freedom bounds, i.e., taking the signal-to-noise ratio (SNR) to infinity. Thus, it is of interest to develop approximation results that are amenable to finite SNR bounds. Below, we provide a (necessarily incomplete) survey of classical and recent results from Diophantine approximation. Using compute-and-forward as a case study, we will discuss how these approximation theorems can be used to obtain simple upper and lower bounds on achievable rate expressions.

The work of B. Nazer was supported by NSF grant CCF-1253918.

II. Classic Results in Diophantine Approximation

The most basic question in Diophantine approximation is how well can a real number $h \in \mathbb{R}$ be approximated by a rational number whose denominator is smaller than some positive integer Q. The simplest estimate on the approximation error was obtained by Dirichlet [22].

Theorem 1 (Dirichlet 1842): For any $h \in \mathbb{R}$ and $Q \in \mathbb{N}$, there exist integers p and q such that $1 \leq q \leq Q$ and

$$\left| h - \frac{p}{q} \right| < \frac{1}{qQ}.$$

Proof. For a real number x denote the floor operation by $\lfloor x \rfloor$. Consider the following $Q+1$ numbers

$$0, 1, h - \lfloor h \rfloor, 2h - \lfloor 2h \rfloor, \ldots, Qh - \lfloor Qh \rfloor \tag{1}$$

and the Q intervals

$$\left[0, \frac{1}{Q}\right), \left[\frac{1}{Q}, \frac{2}{Q}\right), \ldots, \left[\frac{Q-1}{Q}, 1\right).$$

By the pigeonhole principle, at least one interval contains two or more numbers from (1). Hence, there are integers q_1, q_2, p_1, p_2 with $0 \leq q_1 < q_2 \leq Q$ such that

$$|(q_2 h - p_2) - (q_1 h - p_1)| < \frac{1}{Q}.$$

Thus, taking $q = q_2 - q_1 \leq Q$ and $p = p_2 - p_1$ we have that

$$|qh - p| < \frac{1}{Q}.$$

Dividing both sides by q establishes the theorem. ■

In particular, Dirichlet's theorem shows that for any $h \in \mathbb{R}$, there are infinitely many solutions $(q,p) \in \mathbb{N} \times \mathbb{Z}$ to the equation

$$|qh - p| < \psi(q),$$

with $\psi(q) = q^{-1}$.

Definition 1 (Approximating Function): We say that $\psi : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is an *approximating function* if it is monotonically decreasing and satisfies $\psi(r) \to 0$ as $r \to \infty$.

Note that the monotonicity assumption can often be removed. See [23] for more details.

We can now define the set $\mathcal{W}(\psi)$ of *ψ-well approximable numbers* as

$$\mathcal{W}(\psi) \triangleq \Big\{ h \in [0,1) \ : \ |qh - p| < \psi(q) \quad \text{for i.m. } (q,p) \in \mathbb{N} \times \mathbb{Z} \Big\}, \qquad (2)$$

where 'i.m.' reads 'infinitely many'.[1] Denote the Lebesgue measure of a subset $\mathcal{A} \subset \mathbb{R}$ by $\mu(\mathcal{A})$. Consider the following basic question in metric Diophantine approximation: How does $\mu\left(\mathcal{W}(\psi)\right)$ depend on ψ? The answer is given by Khintchine's Theorem [24].

Theorem 2 (Khintchine 1924): Let ψ be an approximating function. Then

$$\mu\left(\mathcal{W}(\psi)\right) = \begin{cases} 0 & \text{if } \sum_{q=1}^{\infty} \psi(q) \leq \infty \\ 1 & \text{if } \sum_{q=1}^{\infty} \psi(q) = \infty \end{cases}.$$

It is instructive to consider the family of approximating functions $\psi(r) = r^{-m}$. From Dirichlet's Theorem, we see that for all $m \leq 1$ it holds that $\mathcal{W}(\psi) = [0,1)$ and thus $\mu\left(\mathcal{W}(\psi)\right) = 1$. (The latter statement can also be obtained directly from the convergent part of Khintchine's Theorem.) The divergent part of Khintchine's Theorem shows that for any $m > 1$ we have that $\mu\left(\mathcal{W}(\psi)\right) = 0$. Thus, $m = 1$ is the *critical exponent*.

Both Dirichlet's and Khintchine's Theorems admit generalizations for matrices of arbitrary dimensions. For brevity, we will only mention the generalization of the latter, due to Groshev [25], and referred to as the Khintchine-Groshev Theorem. Let

$$\mathcal{W}_{N,M}(\psi) \triangleq \Big\{ \mathbf{H} \in [0,1)^{N \times M} \ : \ \|\mathbf{Hq} - \mathbf{p}\|_\infty < \psi(\|\mathbf{q}\|_\infty) \quad \text{for i.m. } (\mathbf{q},\mathbf{p}) \in \mathbb{Z}^M \times \mathbb{Z}^N \Big\},$$

and let μ be the MN-dimensional Lebesgue measure.

Theorem 3 (Khintchine-Groshev): Let $N, M \in \mathbb{N}$ and let ψ be an approximating function. Then

$$\mu\left(\mathcal{W}_{N,M}(\psi)\right) = \begin{cases} 0 & \text{if } \sum_{q=1}^{\infty} q^{M-1}\psi(q)^N < \infty \\ 1 & \text{if } \sum_{q=1}^{\infty} q^{M-1}\psi(q)^N = \infty \end{cases}.$$

[1]Note that the restriction to the unit interval entails no loss of generality, since h is ψ-well approximable if and only if $h + z$ is, for all $z \in \mathbb{Z}$.

The following simple corollary of the convergent part of Theorem 3 plays an important role in the applications of Diophantine approximation to network information theory.

Corollary 1 ([23, Corollary 1]): Let $N, M \in \mathbb{N}$ and let ψ be an approximating function satisfying $\sum_{q=1}^{\infty} q^{M-1}\psi(q)^N < \infty$. Then for almost every $\mathbf{H} \in \mathbb{R}^{N \times M}$ (w.r.t. to the MN-dimensional Lebesgue measure) there exist a constant $\kappa(\mathbf{H}) > 0$ such that

$$\|\mathbf{Hq} - \mathbf{p}\|_\infty > \kappa(\mathbf{H})\psi(\|\mathbf{q}\|_\infty) \ \ \forall \ \mathbf{q} \in \mathbb{Z}^M \setminus \{\mathbf{0}\}, \mathbf{p} \in \mathbb{Z}^N.$$

The study of lattices, and more generally of the "geometry of numbers", initiated by Minkowski at the end of the 19th century, turned out to provide a powerful framework for proving Dirichlet-like theorems for matrices of arbitrary dimensions. Let $\mathbf{F} \in \mathbb{R}^{K \times K}$ be a full-rank matrix, and let

$$\Lambda(\mathbf{F}) \triangleq \left\{ \mathbf{Fa} \ : \ \mathbf{a} \in \mathbb{Z}^K \right\},$$

be the lattice generated by the matrix $\mathbf{F}$.

Definition 2 (Successive Minima): For $k = 1, \ldots, K$, we define the kth successive minimum of $\Lambda(\mathbf{F})$ as

$$\lambda_k(\mathbf{F}) \triangleq \inf \left\{ r \ : \ \dim\left(\mathrm{span}\left(\Lambda(\mathbf{F}) \bigcap \mathcal{B}(\mathbf{0}, r) \right) \right) \geq k \right\}$$

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

Theorem 4 (Minkowski's First Theorem): For any full-rank matrix $\mathbf{F} \in \mathbb{R}^{K \times K}$

$$\lambda_1^2(\mathbf{F}) \leq K \left|\det(\mathbf{F})\right|^{\frac{2}{K}}.$$

Theorem 5 (Minkowski's Second Theorem): For any full-rank matrix $\mathbf{F} \in \mathbb{R}^{K \times K}$

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

Note that these two statements are a specialization of Minkowski's original results [18, p.156] into a form with explicit constants [26].

III. Some Applications in Network Information Theory

We now demonstrate some applications of the classical results discussed above to network information theory. In particular, we will focus on the compute-and-forward strategy and the symmetric Gaussian interference channel.

A. Multiple-Access Channels and Compute-and-Forward

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

$$\mathbf{y} = \sum_{k=1}^{K} h_k \mathbf{x}_k + \mathbf{z}, \qquad (3)$$

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

$$\|\mathbf{x}_k\|^2 \le n\mathsf{snr}, \ k = 1, \ldots, K.$$

The compute-and-forward framework, introduced in [5], provides a communication scheme based on nested lattice codes for decoding integer-linear combinations of the transmitted codewords. In particular, it was shown in [5], [27] that for n large enough and any $\mathbf{a} = [a_1 \ \cdots \ a_K] \in \mathbb{Z}^K$, it is possible to decode the integer-linear combination

$$\mathbf{v} = \sum_{k=1}^{K} a_k \mathbf{x}_k$$

with a vanishing error probability, as long as all K users transmit from the same nested lattice codebook of rate

$$R < R_{\text{comp}}(\mathbf{h}, \mathbf{a}, \mathsf{snr}) \triangleq -\frac{1}{2} \log \|\mathbf{F}\mathbf{a}\|^2, \tag{4}$$

where $\mathbf{F} \in \mathbb{R}^{K \times K}$ satisfies the equation $\mathbf{F}^T\mathbf{F} = \left(\mathbf{I} + \mathsf{snr}\mathbf{h}^T\mathbf{h}\right)^{-1}$.

Typically, the receiver is only interested in decoding L linearly independent linear combinations, but does not care about the particular coefficients. Therefore, we can use the L linearly independent integer vectors $\mathbf{a}_1, \ldots, \mathbf{a}_L$ that yield the highest rates $R_{\text{comp}}(\mathbf{h}, \mathbf{a}_1, \mathsf{snr}) \ge \cdots \ge R_{\text{comp}}(\mathbf{h}, \mathbf{a}_L, \mathsf{snr})$. Accordingly, we define the kth computation rate $R_{\text{comp},k}(\mathbf{h}, \mathsf{snr}) \triangleq R(\mathbf{h}, \mathbf{a}_k, \mathsf{snr})$ to be the rate associated with decoding the kth best integer coefficient vector, linearly independent of $\{\mathbf{a}_1, \ldots, \mathbf{a}_{k-1}\}$. It follows by definition that

$$R_{\text{comp},k}(\mathbf{h}, \mathsf{snr}) = -\frac{1}{2} \log \lambda_k^2(\mathbf{F}).$$

By Sylvester's Theorem [28], we have that

$$\begin{aligned} |\det(\mathbf{F})|^2 &= \det(\mathbf{I} + \mathsf{snr}\mathbf{h}^T\mathbf{h})^{-1} \\ &= (1 + \mathsf{snr}\|\mathbf{h}\|^2)^{-1}. \end{aligned}$$

Combining this identity with Minkowski's First and Second Theorems yields the following results, respectively.

Theorem 6: For any $\mathbf{h} \in \mathbb{R}^K$ and $\mathsf{snr} > 0$, we have that

$$R_{\text{comp},1}(\mathbf{h}, \mathsf{snr}) \ge \frac{1}{K} \cdot \frac{1}{2} \log(1 + \mathsf{snr}\|\mathbf{h}\|^2) - \frac{1}{2} \log K.$$

Theorem 7 ([9, Theorem 3]): For any $\mathbf{h} \in \mathbb{R}^K$ and $\mathsf{snr} > 0$, we have that

$$\begin{aligned} \frac{1}{2} \log(1 + \mathsf{snr}\|\mathbf{h}\|^2) - \frac{K}{2} \log K &\le \sum_{k=1}^{K} R_{\text{comp},k}(\mathbf{h}, \mathsf{snr}) \\ &\le \frac{1}{2} \log(1 + \mathsf{snr}\|\mathbf{h}\|^2). \end{aligned}$$

Recalling that $\frac{1}{2} \log(1 + \mathsf{snr}\|\mathbf{h}\|^2)$ is the sum-capacity of the MAC (3), the interpretation of Theorem 6 is that the rate for decoding the best equation is never much smaller than the symmetric capacity of the channel, whereas Theorem 7 tells us that the sum of the rates for decoding the best K independent equations is never much smaller than the channel's sum-capacity. Note that decoding K linearly independent equations is equivalent to decoding all K codewords. Consequently, Theorem 7 was used in [9] to show that a low-complexity scheme based on nested lattice codes can achieve the sum-capacity of the Gaussian MAC (3) to within a gap of at most $\frac{K}{2} \log K$ bits.

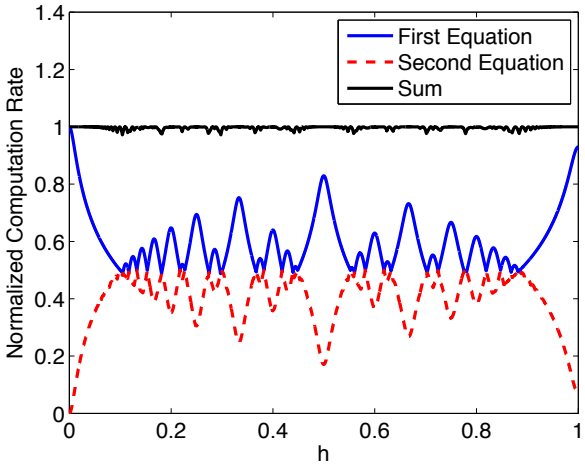

Fig. 1. Computation rates for the best two linearly independent integer linear combinations vs. h for the channel $\mathbf{y} = \mathbf{x}_1 + h\mathbf{x}_2 + \mathbf{z}$ at $\mathsf{snr} = 40$dB. The sum of these computation rates is nearly equal to the multiple-access sum capacity. All rates are normalized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

Theorem 7 shows that the sum of computation rates depends of $\mathbf{h}$ mainly through its Euclidean norm, and is not too sensitive to perturbations in its individual entries. However, while the sum is near constant, the way it is distributed between $R_{\text{comp},1}(\mathbf{h}, \mathsf{snr}), \ldots, R_{\text{comp},K}(\mathbf{h}, \mathsf{snr})$ can be highly sensitive to the channel vector $\mathbf{h}$. See Figure 1. Theorem 6 provides a universal lower bound on $R_{\text{comp},1}(\mathbf{h}, \mathsf{snr})$, but the only universal upper bound we can obtain is

$$R_{\text{comp},1}(\mathbf{h}, \mathsf{snr}) \le \frac{1}{2} \log(1 + \mathsf{snr} \max_k |h_k|^2),$$

which is attained with equality when $\mathbf{h}$ contains only one nonzero entry [5, Theorem 14]. Nevertheless, in the limit of high snr, it is possible to obtain informative upper bounds on $R_{\text{comp},1}(\mathbf{h}, \mathsf{snr})$ that hold for almost every $\mathbf{h} \in \mathbb{R}^K$. First, let us define the number of degrees-of-freedom (DoF) associated with the kth computation rate as

$$d_{\text{comp},k}(\mathbf{h}) \triangleq \lim_{\mathsf{snr} \to \infty} \frac{R_{\text{comp},k}(\mathbf{h}, \mathsf{snr})}{\frac{1}{2} \log(1 + \mathsf{snr})}.$$

Theorem 6 and Theorem 7 imply that for any $\mathbf{h} \in \mathbb{R}^K$

$$d_{\text{comp},1}(\mathbf{h}) \ge \frac{1}{K} \tag{5}$$

and

$$\sum_{k=1}^{K} d_{\text{comp},k}(\mathbf{h}) = 1, \tag{6}$$

respectively. The following result uses the Khintchine-Groshev Theorem to show that the lower bound (5) is tight for almost every $\mathbf{h} \in \mathbb{R}^K$.

Theorem 8 ([9, Corollary 4]): For almost every $\mathbf{h} \in \mathbb{R}^K$ (w.r.t. Lebesgue measure)

$$d_{\text{comp},1}(\mathbf{h}) = \frac{1}{K}.$$

Theorem 8 is a simple consequence of the following result.

Lemma 1: Let $\mathbf{h} \in \mathbb{R}^K \setminus \{\mathbf{0}\}$, and let $\vec{\mathbf{h}} \in \mathbb{R}^K \setminus \{\mathbf{0}\}$ be its permuted version such $\bar{h} = |\vec{h}_1| \geq \cdots \geq |\vec{h}_K|$. Further define $\tilde{h}_k = \vec{h}_k/\bar{h}$ for $k = 1, \ldots, K$ and

$$c_0(\mathbf{h}) \triangleq \min\left(\frac{1}{4\bar{h}^2}, \max_{k=2,\ldots,K} \frac{1}{1+\tilde{h}_k^2}\right).$$

For any $\epsilon > 0$, let $0 \leq \kappa_\epsilon(\mathbf{h}) \leq 1/2$ be the largest number for which

$$\max_{k=2,\ldots,K} |q\tilde{h}_k - a_k| \geq \kappa_\epsilon(\mathbf{h}) \cdot |q|^{-(1+\epsilon)/(K-1)} \quad \forall q \in \mathbb{N}, (a_2, \ldots, a_K) \in \mathbb{Z}^{K-1}. \tag{7}$$

Then,

$$R_{\text{comp},1}(\mathbf{h}, \mathsf{snr}) \leq \frac{1+\epsilon}{K+\epsilon} \cdot \frac{1}{2}\log(\mathsf{snr}) - \frac{1}{2}\log(c_0(\mathbf{h})) - \log(\kappa_\epsilon(\mathbf{h})). \tag{8}$$

Proof of Theorem 8. Clearly, $c_0(\mathbf{h}) > 0$ for every $\mathbf{h} \in \mathbb{R}^K \setminus \{\mathbf{0}\}$. Let $\psi(r) = r^{-\frac{1+\epsilon}{K-1}}$ and note that $\sum_{q=1}^{\infty}(\psi(q))^{K-1}$ converges. Corollary 1 therefore implies that $\kappa(\mathbf{h}) > 0$ for almost every $\mathbf{h} \in \mathbb{R}^K$. The result now follows by applying (8) and taking the double limit

$$\lim_{\mathsf{snr}\to\infty} \lim_{\epsilon\to 0} \frac{\frac{1+\epsilon}{K+\epsilon} \cdot \frac{1}{2}\log(\mathsf{snr}) - \frac{1}{2}\log(c_0(\mathbf{h})) - \log(\kappa_\epsilon(\mathbf{h}))}{\frac{1}{2}\log(1+\mathsf{snr})}.$$

■

Proof of Lemma 1. From [9], an alternative expression for $R_{\text{comp},1}(\mathbf{h}, \mathsf{snr})$ is

$$R_{\text{comp},1}(\mathbf{h}, \mathsf{snr}) = \frac{1}{2}\log\left(\frac{\mathsf{snr}}{\min_{\alpha\in\mathbb{R}, \mathbf{a}\in\mathbb{Z}^K\setminus\{\mathbf{0}\}} \sigma^2(\alpha, \mathbf{a}, \mathbf{h}, \mathsf{snr})}\right), \tag{9}$$

where

$$\sigma^2(\alpha, \mathbf{a}, \mathbf{h}, \mathsf{snr}) \triangleq \alpha^2 + \mathsf{snr}\sum_{k=1}^{K}(\alpha h_k - a_k)^2.$$

Our goal is to lower bound $\min_{\alpha\in\mathbb{R}, \mathbf{a}\in\mathbb{Z}^K\setminus\{\mathbf{0}\}} \sigma^2(\alpha, \mathbf{a}, \mathbf{h}, \mathsf{snr})$. Since $R_{\text{comp},1}(\mathbf{h}, \mathsf{snr})$ is invariant to permutations on the entries of $\mathbf{h}$, we can assume w.l.o.g. that $\mathbf{h} = \vec{\mathbf{h}}$, such that $\bar{h} = |h_1|$. It is easy to see that

$$\sigma^2(\alpha, \mathbf{a}, \mathbf{h}, \mathsf{snr}) \geq \frac{\mathsf{snr}}{4} \geq \kappa_\epsilon^2(\mathbf{h})\mathsf{snr} \tag{10}$$

for all $|\alpha| < \frac{1}{2}, \mathbf{a} \in \mathbb{Z}^K \setminus \{\mathbf{0}\}$.

For $|\alpha| > 1/2$ we can write $\alpha = (q+\varphi)/h_1$ where $q \in \mathbb{Z} \setminus \{0\}$ and $\varphi \in [-1/2, 1/2)$, which gives rise to

$$\begin{aligned}
\sigma^2(\alpha, \mathbf{a}, \mathbf{h}, \mathsf{snr}) &= \sigma^2(\varphi, q, \mathbf{a}, \mathbf{h}, \mathsf{snr}) \\
&= \frac{(q+\varphi)^2}{h_1^2} + \mathsf{snr}\sum_{k=1}^{K}\Big((q+\varphi)\tilde{h}_k - a_k\Big)^2 \\
&= \frac{(q+\varphi)^2}{h_1^2} + \mathsf{snr}\Big((q+\varphi-a_1)^2 \\
&\qquad + \max_{k=2,\ldots,K}\big((q+\varphi)\tilde{h}_k - a_k\big)^2\Big) \qquad (11) \\
&\geq \frac{(q/2)^2}{h_1^2} + \mathsf{snr}\Big(\varphi^2 + \max_{k=2,\ldots,K}\big((q+\varphi)\tilde{h}_k - a_k\big)^2\Big) \qquad (12)
\end{aligned}$$

where in (12) we used the fact that for every $q, a_1 \in \mathbb{Z}$ and every $\varphi \in [-1/2, 1/2)$ it holds that $|q+\varphi| \geq |q/2|$ and $|q - a_1 + \varphi| \geq |\varphi|$. Next, we note that for every choice of $a_k, q, \tilde{h}_k$ the function $g_k(\varphi) = \varphi^2 + \big((q+\varphi)\tilde{h}_k - a_k\big)^2$ is convex and attains its minimum at

$$\varphi_k^* = \frac{-\tilde{h}_k}{1+\tilde{h}_k^2}(q\tilde{h}_k - a_k)$$

and this minimum value is equal to

$$g_k(\varphi_k^*) = \frac{1}{1+\tilde{h}_k^2}(q\tilde{h}_k - a_k)^2.$$

We therefore have that

$$\begin{aligned}
\sigma^2(\varphi, q, \mathbf{a}, \mathbf{h}, \mathsf{snr}) &\geq \frac{q^2}{4h_1^2} + \mathsf{snr}\max_{k=2,\ldots,K}\frac{1}{1+\tilde{h}_k^2}(q\tilde{h}_k - a_k)^2 \\
&\geq c_0(\mathbf{h})\left(\max_{k=2,\ldots,K}|q\tilde{h}_k - a_k|^2\mathsf{snr} + q^2\right) \\
&= c_0(\mathbf{h})\Gamma(q, (a_2, \ldots, a_K), \mathbf{h}, \mathsf{snr}) \qquad (13)
\end{aligned}$$

where

$$\Gamma(q, (a_2, \ldots, a_K), \mathbf{h}, \mathsf{snr}) \triangleq \max_{k=2,\ldots,K}|q\tilde{h}_k - a_k|^2\mathsf{snr} + q^2.$$

Since in (12) we set $a_1 = q$, the restriction that $\mathbf{a} \in \mathbb{Z}^K \setminus \{\mathbf{0}\}$ translates to the restriction that $(q, a_2, \ldots, a_K) \in \mathbb{Z}^K \setminus \{\mathbf{0}\}$. By definition of $\kappa_\epsilon(\mathbf{h})$, for any choice of $(q, a_2, \ldots, a_K) \in \mathbb{Z}^K \setminus \{\mathbf{0}\}$ with $q \neq 0$, we have that

$$\begin{aligned}
\Gamma(q, (a_2, \ldots, a_K), \mathbf{h}, \mathsf{snr}) &\geq \kappa_\epsilon^2(\mathbf{h})\Big(|q|^{-\frac{2(1+\epsilon)}{K-1}}\mathsf{snr} + q^2\Big) \\
&\geq \kappa_\epsilon^2(\mathbf{h})\mathsf{snr}^{\frac{K-1}{K+\epsilon}}, \qquad (14)
\end{aligned}$$

where in the last inequality we have used the fact that $x^{-2\frac{1+\epsilon}{K-1}}\mathsf{snr} + x^2 \geq \mathsf{snr}^{\frac{K-1}{K+\epsilon}}$ for all $x > 0$. Combining (14) and (10) we obtained that for any $(q, a_2, \ldots, a_K) \in \mathbb{Z}^K \setminus \{\mathbf{0}\}$ it holds that

$$\Gamma(q, (a_2, \ldots, a_K), \mathbf{h}, \mathsf{snr}) \geq \kappa_\epsilon^2(\mathbf{h})\mathsf{snr}^{\frac{K-1}{K+\epsilon}}. \tag{15}$$

Substituting this into (13) gives that for any $\alpha \in \mathbb{R}$ and any $\mathbf{a} \in \mathbb{Z}^K \setminus \{\mathbf{0}\}$ we have that

$$\sigma^2(\alpha, \mathbf{a}, \mathbf{h}, \mathsf{snr}) \geq c_0(\mathbf{h})\kappa_\epsilon^2(\mathbf{h})\mathsf{snr}^{\frac{K-1}{K+\epsilon}}. \tag{16}$$

Combining this with (9) establishes the desired result. ■

Remark 1: Niesen and Whiting [29] studied the DoF offered by the highest computation rate and showed that

$$d_{\text{comp},1} \leq \begin{cases} 1/2 & K = 2 \\ 2/(K+1) & K > 2 \end{cases} \tag{17}$$

for almost every $\mathbf{h} \in \mathbb{R}^K$. Thus, Theorem 8 improves upon [29] for $K > 2$.

Corollary 2 ([9, Corollary 5]): For almost every $\mathbf{h} \in \mathbb{R}^K$ (w.r.t. Lebesgue measure)

$$d_{\text{comp},1}(\mathbf{h}) = \cdots = d_{\text{comp},K}(\mathbf{h}) = \frac{1}{K}.$$

Proof. By (6), we have that $\sum_{k=1}^K d_{\text{comp},k} = 1$. Using the fact that $d_{\text{comp},k}$ is monotonically decreasing in k and that $d_{\text{comp},1} = 1/K$ for almost every $\mathbf{h} \in \mathbb{R}^K$, the corollary follows. ■

Remark 2: In [30], Theorem 8 was extended to multiple access channels where the receiver is equipped with $N > 1$ antennas, and it was shown that $d_{\text{comp},1}(\mathbf{H}) = \min(1, N/K)$ for almost every $\mathbf{H} \in \mathbb{R}^{N\times K}$ (w.r.t. Lebesgue measure), where H_{ij} is the channel gain from the jth user to the ith receive antenna.

B. Interference Channels and Lattice Alignment

We now consider the K-user symmetric Gaussian interference channel, which consists of K transmitter-receiver pairs. The kth receiver observes

$$\mathbf{y}_k = \mathbf{x}_k + h\sum_{\ell \neq k} \mathbf{x}_\ell + \mathbf{z}_k$$

where $\mathbf{x}_\ell \in \mathbb{R}^n$ is the codeword from transmitter ℓ, $h \in \mathbb{R}$ is the interfering channel gain and $\mathbf{z}_k$ is i.i.d. AWGN with zero mean and unit variance. The goal of receiver k is to recover $\mathbf{x}_k$ from $\mathbf{y}_k$ and we are interested in determining the highest symmetric rate R_{sym} (i.e., all users have the same rate). As before, we assume all users face the same power constraint $\|\mathbf{x}_\ell\|^2 \leq n\mathsf{snr}$.

If all transmitters employ the same lattice codebook, then the interference will automatically be aligned from the perspective of each receiver, owing to the fact that lattices are closed under integer-linear combinations. Therefore, from the perspective of each receiver, it observes an effective two-user multiple-access channel, and can recover its desired codeword from only two integer-linear combinations of the codewords. It follows from the analysis above that $d_{\text{comp},2} = 1/2$ and thus each user can attain $1/2$ degrees-of-freedom, for almost every $h \in \mathbb{R}$. For the general K-user Gaussian interference channel, we need to employ the signaling strategy from [7], which leads to effective channel gains that are drawn from a certain manifold. As a result, we require more sophisticated Diophantine approximation theorems [20], [21], [31] in order to show that that $1/2$ degrees-of-freedom are achievable for almost all channel gains.

We now consider the symmetric rate at finite SNR. As shown in [9], all receivers can decode so long as

$$R < -\frac{1}{2}\log(\lambda_2^2(\mathbf{F})) \tag{18}$$

where $\mathbf{F} \in \mathbb{R}^{2\times 2}$ satisfies $\mathbf{F}^T\mathbf{F} = (\mathbf{B}^{-1} + \mathsf{snr}\mathbf{g}^T\mathbf{g})^{-1}$ where

$$\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & (K-1) \end{bmatrix} \qquad \mathbf{g} = [1 \; h].$$

While this rate expression can be evaluated numerically, it cannot be directly compared to available upper bounds in this form. Consider the strong regime, $h \in [1, \sqrt{\mathsf{snr}}]$. In [9], it was shown that, in the strong regime, for any $\gamma > 0$

$$\frac{1}{4}\log(h^2\mathsf{snr}) - \frac{\gamma}{2} - 3 \leq R_{\text{sym}} \leq \frac{1}{4}\log(h^2\mathsf{snr}) + 1 \tag{19}$$

up to an outage set of measure $2^{-\gamma}$ with respect to the channel gains in the interval $[1, \sqrt{\mathsf{snr}}]$. Similar results can also be found for the X channel in [8]. The proof closely follows the proof of the convergent part of Khintchine's Theorem, but does not take the limit that leads to the zero-one law in the usual theorem statement. In the next section, we discuss recent results in Diophantine approximation that capture a general form of this result.

IV. Recent Developments

Let ψ be an approximating function, $\kappa > 0$ be a positive real number, and define the set

$$\mathcal{I}_{N,M}(\psi, \kappa) \triangleq \Big\{ \mathbf{H} \in \mathbb{R}^{N\times M} \; : \|\mathbf{H}\mathbf{q} - \mathbf{p}\|_\infty \geq \kappa\psi(\|\mathbf{q}\|_\infty) \\ \forall \mathbf{q} \in \mathbb{Z}^M \setminus \{\mathbf{0}\}, \mathbf{p} \in \mathbb{Z}^N \Big\}.$$

Assuming the matrix $\mathbf{H}$ is randomly drawn from some given distribution, it is often of interest to calculate the quantity

$$\Pr\left(\mathbf{H} \notin \mathcal{I}_{N,M}(\psi, \kappa)\right).$$

This quantity has been the subject of a recent work [23], which studied it under various assumptions on the underlying probability distribution and the approximating function. We reproduce one of the main results below.

Theorem 9 ([23, Corollary 2]): Choose $0 < \delta < 1$. Let $\mathbf{H} \in \mathbb{R}^{N\times M}$ be a random matrix and ψ an approximating function. Define $S_\psi = \sup_{q\in\mathbb{N}} \psi(q)$ and $\Sigma_\psi = \sum_{q=1}^\infty q^{M-1}(\psi(q))^N$. Furthermore, assume that the pdf of $\mathbf{H}$ is upper bounded by a constant, $f_{\mathbf{H}}(\mathbf{H}) \leq c_{\max}$ and let T be the smallest positive integer such that $\Pr\left(\mathbf{H} \in [-T, T]^{N\times M}\right) \geq 1 - \delta/2$. Then,

$$\Pr\left(\mathbf{H} \notin \mathcal{I}_{N,M}(\psi, \kappa)\right) \leq \delta$$

where

$$\kappa = \frac{1}{2}\min\left\{\frac{1}{S_\psi}, \; \left(\left(\frac{\delta}{2c_{\max}(2T)^{MN}\Sigma_\psi}\right)^{1/M}\right\}.$$

We now demonstrate how this result can be used to prove the lower bound in (19). Using a variation on Theorem 7 from [9, Theorem 7], it can be shown that

$$\begin{aligned} R_{\text{sym}} &> \frac{1}{2}\log(h^2\mathsf{snr}) - \frac{1}{2}\log\left(\frac{\mathsf{snr}}{\sigma^2_{\text{eff}}}\right) - 1 \\ &> \frac{1}{2}\log(h^2) + \frac{1}{2}\log(\sigma^2_{\text{eff}}) - 1 \end{aligned} \tag{20}$$

where

$$\sigma^2_{\text{eff}} = \min_{\substack{\alpha\in\mathbb{R} \\ [a_1\ a_2]\in\mathbb{Z}^2\setminus\{\mathbf{0}\}}} \alpha^2 + \mathsf{snr}\Big((\alpha - a_1)^2 + (K-1)(\alpha h - a_2)^2\Big).$$

Following the steps in [9, p.3469], it follows that if $|\alpha| < 1/2$, then $\sigma^2_{\text{eff}} > \mathsf{snr}^{1/2}/(4h)$. For $|\alpha| \geq 1/2$, we get the lower bound

$$\sigma^2_{\text{eff}} \geq \frac{1}{4}\min_{q,p}\max\left\{q^2,\ \frac{\mathsf{snr}}{h^2}(qh-p)^2\right\}$$

In order to apply Theorem 9, we assume first that h is drawn uniformly from the interval $[1,2)$. Thus, we can have $c_{\max} = 1$ and $T = 2$. Select the approximating function

$$\psi(q) = \begin{cases} \frac{1}{Q} & 1 \leq q \leq \lfloor Q \rfloor, \\ 0 & \text{otherwise.} \end{cases}$$

which yields $S_\psi = \frac{1}{Q}$ and $\Sigma_\psi \leq 1$.

It follows from Theorem 9 that

$$\sigma^2_{\text{eff}} \geq \frac{1}{4}\min_{q,p}\max\left\{q^2,\ \frac{\mathsf{snr}}{h^2}(\kappa\psi(q))^2\right\}$$

for any q for all but a set of channel gains in $[1,2)$ with measure δ where $\kappa = \min\{\frac{Q}{2}, \frac{\delta}{16}\}$. Setting $Q = \mathsf{snr}^{1/4}\big(\frac{\delta}{16\lceil h\rceil}\big)^{1/2}$, we get that

$$\sigma^2_{\text{eff}} \geq \mathsf{snr}^{1/2}\frac{\delta}{64h}\ .$$

for all but a set of channel gains in $[1,2)$ with measure δ so long as $\mathsf{snr} > 1$.

Plugging this into (20), we get that

$$\begin{aligned} R_{\text{sym}} &> \frac{1}{2}\log(h^2) + \frac{1}{2}\log\left(\frac{\mathsf{snr}^{1/2}\delta}{64h}\right) - 1 \\ &\geq \frac{1}{4}\log(h^2\mathsf{snr}) + \frac{1}{2}\log(\delta) - 4 \\ &= \frac{1}{4}\log(h^2\mathsf{snr}) - \frac{\gamma}{2} - 4 \end{aligned}$$

for all but a set of channel gains in $[1,2)$ with measure $2^{-\gamma}$ where in the last step we have set $\delta = 2^{-\gamma}$. Since the ψ-approximability of h is equivalent to ψ-approximability of $h+b$ for any $b \in \mathbb{Z}$, the result holds for any $h \in [b, b+1)$.

One can similarly derive bounds for other distributions, such as Gaussian fading. More generally, the results in [23] can be used to derive outage-type approximations for higher dimensions as well as effective channel gains that lie on manifolds.

V. Conclusions

As we have seen, Diophantine approximation provides useful techniques and bounds for characterizing the performance of lattice-based coding strategies. In this brief survey, we have focused on the applications of the Khintchine-Groshev Theorem and its generalizations for bounding the performance of compute-and-forward. Two very recent papers have proposed alternative novel techniques for upper bounding the outage probability for the Kth computation rate in the more general MIMO setting. In particular, [32] upper bounded the outage probability under random unitary precoding for the compound MIMO channel with a given white-input mutual information, whereas in [33] the outage probability was bounded under various statistical assumptions on the distributions of the channel matrix.

References

[1] D. Krithivasan and S. S. Pradhan, "Lattices for distributed source coding: Jointly Gaussian sources and reconstruction of a linear function," *IEEE Transactions on Information Theory*, vol. 55, no. 12, pp. 5268–5651, December 2009.

[2] D. Krithivasan and S. Pradhan, "Distributed source coding using Abelian group codes," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1495–1519, March 2011.

[3] M. P. Wilson, K. Narayanan, H. Pfister, and A. Sprintson, "Joint physical layer coding and network coding for bidirectional relaying," *IEEE Transactions on Information Theory*, vol. 11, no. 56, pp. 5641–5654, November 2010.

[4] W. Nam, S. Chung, and Y. Lee, "Nested lattice codes for Gaussian relay networks with interference," *IEEE Transactions on Information Theory*, vol. 57, no. 12, pp. 7733–7745, December 2011.

[5] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6463–6486, October 2011.

[6] G. Bresler, A. Parekh, and D. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4566–4592, September 2010.

[7] A. S. Motahari, S. Oveis-Gharan, M.-A. Maddah-Ali, and A. K. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 4799–4810, August 2014.

[8] U. Niesen and M. A. Maddah-Ali, "Interference alignment: From degrees-of-freedom to constant-gap capacity approximations," *IEEE Transactions on Information Theory*, vol. 59, no. 8, pp. 4855–4888, August 2013.

[9] O. Ordentlich, U. Erez, and B. Nazer, "The approximate sum capacity of the symmetric K-user Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 60, no. 6, pp. 3450–3482, June 2014.

[10] I. Shomorony and S. Avestimehr, "Degrees of freedom of two-hop wireless networks: Everyone gets the entire cake," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2417–2431, May 2014.

[11] T. Philosof and R. Zamir, "On the loss of single-letter characterization: The dirty multiple access channel," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2442–2454, June 2009.

[12] T. Philosof, R. Zamir, U. Erez, and A. J. Khisti, "Lattice strategies for the dirty multiple access channel," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5006–5035, August 2011.

[13] I.-H. Wang, "Approximate capacity of the dirty multiple-access channel with partial state information at the encoders," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2781–2787, May 2012.

[14] X. He and A. Yener, "Providing secrecy with structured codes: Tools and applications to two-user Gaussian channels," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2121–2138, April 2014.

[15] J. Xie and S. Ulukus, "Secure degrees of freedom of one-hop wireless networks," *IEEE Transactions on Information Theory*, vol. 60, no. 6, pp. 3359–3378, June 2014.

[16] S. Vatedka, N. Kashyap, and A. Thangaraj, "Secure compute-and-forward in a bidirectional relay," *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2531–2556, May 2015.
[17] R. Zamir, *Lattice Coding for Signals and Networks*. Cambridge University Press, 2014.
[18] J. W. S. Cassels, *An Introduction to Diophantine Approximations*. Cambridge University Press, 1957.
[19] V. I. Bernik and M. M. Dodson, *Metric Diophantine Approximation on Manifolds*. Cambridge University Press, 1991.
[20] V. Bernik, D. Kleinbock, and G. A. Margulis, "Khintchine-type theorems on manifolds: The convergence case for standard and multiplicative versions," *International Mathematics Research Notes*, vol. 9, pp. 453–486, 2001.
[21] V. Beresnevich, "A Groshev type theorem for convergence on manifolds," *Acta Mathematica Hungarica*, vol. 94, no. 1-2, pp. 99–130, 2002.
[22] L. Dirichlet, "Verallgemeinerung eines Satzes aus der Lehre von den Kettenbrüchen nebst einigen Anwendungen auf die Theorie der Zahlen," *SB Preuss. Akad. Wiss*, pp. 93–95, 1842.
[23] F. Adiceam, V. Beresnevich, J. Levesley, S. Velani, and E. Zorin, "Diophantine approximation and applications in interference alignment," *arXiv preprint arXiv:1506.03688*, 2015.
[24] A. Khintchine, "Einige sätze über kettenbrüche, mit anwendungen auf die theorie der diophantischen approximationen," *Mathematische Annalen*, vol. 92, no. 1, pp. 115–125, 1924.
[25] A. Groshev, "Une théorème sur les systèmes des formes linéaires," *Dokl. Akad. Nauk SSSR*, vol. 9, pp. 151–152, 1938.
[26] D. Micciancio and S. Goldwasser, *Complexity of Lattice Problems: A Cryptographic Perspective*. Cambridge, UK: Kluwer Academic Publishers, 2002, vol. 671 of The Kluwer International International Series in Engineering and Computer Science.
[27] B. Nazer, "Successive compute-and-forward," in *Proceedings of the International Zurich Seminar on Communications (IZS 2012)*, Zurich, Switzerland, March 2012.
[28] J. J. Sylvester, "On the relation between the minor determinants of linearly equivalent quadratic functions," *Philosophical Magazine*, vol. 1, no. 4, pp. 295–305, 1851.
[29] U. Niesen and P. Whiting, "The degrees-of-freedom of compute-and-forward," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5214–5232, August 2012.
[30] O. Ordentlich and U. Erez, "Precoded integer-forcing universally achieves the MIMO capacity to within a constant gap," *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 323–340, January 2015.
[31] D. Y. Kleinbock and G. A. Margulis, "Flows on homogeneous spaces and Diophantine approximation on manifolds," *Annals of Mathematics*, vol. 148, no. 1, pp. 339–360, July 1998.
[32] E. Domanovitz and U. Erez, "Outage behavior of randomly precoded integer forcing over MIMO channels," *arXiv preprint arXiv:1608.01588*, 2016.
[33] F. Adiceam and E. Zorin, "On the minimum of a positive definite quadratic form over non–zero lattice points. theory and applications," *arXiv preprint arXiv:1607.04467*, 2016.