Capacity of the Gaussian Arbitrarily Varying Channel

Imre Csiszár and Prakash Narayan, *Member, IEEE*

Abstract **—The Gaussian arbitrarily varying channel with input constraint Γ and state constraint Λ admits input sequences $x=(x_1,\cdots,x_n)$ of real numbers with $\sum x_i^2 \le n\Gamma$ and state sequences $s=(s_1,\cdots,s_n)$ of real numbers with $\sum s_i^2 \le n\Lambda$; the output sequence $x+s+V$, where $V=(V_1,\cdots,V_n)$ is a sequence of independent and identically distributed Gaussian random variables with mean 0 and variance σ^2. It is proved that the capacity of this arbitrarily varying channel for deterministic codes and the average probability of error criterion equals $\frac{1}{2}\log(1+\Gamma/(\Lambda+\sigma^2))$ if $\Lambda<\Gamma$ and is 0 otherwise.**

Index Terms **—Arbitrarily varying channel, Gaussian, capacity.**

I. Introduction

Arbitrarily varying channels (AVC's) were introduced by Blackwell *et al.* [5] to model communication channels with unknown parameters that may vary with time in an arbitrary and unknown manner during the transmission of a codeword. In this paper, attention is restricted to AVC's without memory; further, it is assumed that the sequence of channel states is selected arbitrarily subject to a constraint specified later, and possibly depending on the codebook but independently of the codeword actually sent.

AVC's exhibit various mathematical complexities even in the case of discrete alphabets (cf. Csiszár–Körner [6, Section 2.6]). In particular, their capacity may depend on whether or not random codes are permitted, and whether the average or maximum probability of error criterion is used. The random coding capacity admits a simple characterization as a min-max of mutual information, a result dating back to Blackwell *et al.* [5]. In contrast, the problem of capacity for deterministic codes is much harder. In particular, for the maximum probability of error criterion, a single-letter capacity formula is known only under certain conditions on the structure of the AVC (cf. Ahlswede [2] and Csiszár–Körner [7]).

Unless stated otherwise, the term capacity will hereafter always refer to capacity for *deterministic codes* and the *average probability of error criterion*. In the absence of state constraints, Ahlswede [1] proved that this capacity was either equal to the random coding capacity or otherwise to zero. The necessary and sufficient condition for positive capacity, as well as capacity under a state constraint, have been determined by Csiszár–Narayan [8]; it was further shown that Ahlswede's alternatives do not necessarily obtain under a state constraint.

Less attention has been bestowed in the literature on the capacity of AVC's with continuous alphabets. Presumably motivated by random coding capacity, there have been game-theoretic considerations concentrating on the min-max of mutual information (cf. McEliece [11]). Hughes–Narayan [10] have used a geometric approach to determine the random coding capacity of the Gaussian AVC defined formally in the following paragraph. Blachman [4] has provided lower and upper bounds on capacity in a communication situation differing from ours in that the interference (i.e., state sequence) could depend on the actual codeword transmitted. Our incomplete understanding of his paper seems to indicate that he, too, considered random coding capacity. To our knowledge, Ahlswede's [3] is the only paper treating the capacity of a continuous alphabet AVC for deterministic codes. His AVC (a Gaussian channel with the noise variance arbitrarily varying but not exceeding a given bound) allowed a very simple approach, which may not be extendable to other cases of interest.

In this paper, we determine the capacity of the Gaussian AVC formally defined as follows. Let the input and output alphabets, and the set of states, be the real line. For any input sequence $x=(x_1,\cdots,x_n)$ and state sequence $s=(s_1,\cdots,s_n)$, let the output be $x+s+V$, where $V=(V_1,\cdots,V_n)$ is a sequence of independent and identically distributed (i.i.d.) Gaussian random variables with mean 0 and variance σ^2. We adopt an input constraint Γ and state constraint Λ, namely the permissible input sequences of length n are those satisfying

$$\|x\|^2=\sum_{i=1}^{n} x_i^2 \le n\Gamma, \qquad (\Gamma>0) \tag{1.1}$$

and the permissible state sequences are those satisfying

$$\|s\|^2=\sum_{i=1}^{n} s_i^2 \le n\Lambda, \qquad (\Lambda>0). \tag{1.2}$$

A code of block-length n comprises a set of codewords $x_1,\cdots,x_N$, each in $\mathbb{R}^n$, and a decoder $\varphi\colon \mathbb{R}^n \to \{0,\cdots,N\}$. The average probability of error of this code, used on the Gaussian AVC as above when the state sequence is s, equals

$$\bar{e}(s)=\frac{1}{N}\sum_{i=1}^{N}\Pr\{\varphi(x_i+s+V)\neq i\}. \tag{1.3}$$

Manuscript received April 25, 1989; revised February 13, 1990. I. Csiszár was supported by the Hungarian National Foundation Scientific Research Grant No. 1806. P. Narayan was supported by the Systems Research Center at the University of Maryland under NSF Grant OIR-85-00108. This work was presented at the IEEE International Symposium on Information Theory, San Diego, CA, January 14–19, 1990.

I. Csiszár is with the Mathematical Institute of the Hungarian Academy of Sciences, H-1364 Budapest, POB 127, Hungary.

P. Narayan is with the Electrical Engineering Department and the Systems Research Center, University of Maryland, College Park, MD 20742.

IEEE Log Number 9038858.

0018-9448/91/0100-0018$01.00 © 1991 IEEE

The capacity C of the Gaussian AVC with input constraint Γ and state constraint Λ is the largest number with the property that for every $\delta>0$ and sufficiently large n, there exist codes with $N\ge\exp\{n(C-\delta)\}$ codewords, each satisfying (1.1), such that the supremum of $\bar{e}(s)$ subject to (1.2) converges to 0 as $n\to\infty$.

Our main result is the following.

Theorem 1: The capacity of the Gaussian AVC with input constraint Γ and state constraint Λ is

$$C=\begin{cases}\dfrac{1}{2}\log\left(1+\dfrac{\Gamma}{\Lambda+\sigma^2}\right), & \text{if } \Gamma>\Lambda\\ 0, & \text{if } \Gamma\le\Lambda\end{cases}.$$

According to Hughes–Narayan [10], the random coding capacity of the Gaussian AVC equals $\frac{1}{2}\log(1+\Gamma/(\Lambda+\sigma^2))$. Thus, in this case Ahlswede's alternatives do obtain. Yet a proof of the theorem above by the elegant "elimination technique" of Ahlswede [1] is not apparent. Rather, we shall use the straightforward but more computational method of Csiszár–Narayan [8]. Suitable approximation arguments would enable a derivation of our theorem directly from the results of [8]. Instead, we prefer to present a more transparent and direct proof, which will also serve to keep this paper self-contained.

We also determine the capacity of the noiseless additive AVC whose output is $x+s$ rather than $x+s+V$. The capacity of this AVC is defined similarly to that of the Gaussian AVC with the exception that (1.3) is now replaced by

$$\bar{e}(s)=\frac{1}{N}\left|\{i\colon \varphi(x_i+s)\ne i\}\right|. \tag{1.4}$$

Theorem 2: The capacity of the noiseless additive AVC with input constraint Γ and state constraint Λ is

$$C=\begin{cases}\dfrac{1}{2}\log\left(1+\dfrac{\Gamma}{\Lambda}\right), & \text{if } \Gamma>\Lambda\\ 0, & \text{if } \Gamma\le\Lambda.\end{cases}$$

Whereas this result is not a formal special case of Theorem 1, both theorems can be proved by the same method. We shall prove the simpler Theorem 2 first so that the reader may better understand the key ideas. Observe that Theorem 1 requires a separate proof only in the case $\Lambda+\sigma^2\ge\Gamma$. In fact, since (1.2) implies for an arbitrary $\epsilon>0$ that $\|s+V\|^2\le n(\Lambda+\sigma^2+\epsilon)$ with probability arbitrarily close to 1 if n is sufficiently large, in the case $\Lambda+\sigma^2<\Gamma$ the assertion of Theorem 1 follows immediately from that of Theorem 2.

Actually, we shall show that the capacity as claimed in Theorems 1 and 2 can be achieved using the minimum-distance decoder, namely

$$\phi(y)=\begin{cases} i, & \text{if } \|y-x_i\|^2<\|y-x_j\|^2, \quad \text{for } j\ne i\\ 0, & \text{if no such } 1\le i\le N \text{ exists.}\end{cases} \tag{1.5}$$

It is worth pointing out that the result of Theorem 2 with this decoder provides a solution to a weakened version of the unsolved sphere-packing problem. This problem seeks the exponential rate of the maximum number of nonintersecting spheres of radius $\sqrt{\Lambda}$ in $\mathbb{R}^n$ with centers in a sphere of radius $\sqrt{\Gamma}$. In our case, the spheres may intersect but for any given s in $\mathbb{R}^n$ of norm $\le\sqrt{\Lambda}$, only for a vanishingly small fraction of sphere centers x_i can x_i+s be closer to another sphere center than to x_i. The number C in Theorem 2 then gives the exponential rate of the maximum number of spheres satisfying this condition. A similar weakened version of the sphere-packing problem in Hamming space was solved in [8] as a special case of the coding theorem for the binary adder AVC.

II. Proof of the Main Result

The proof of the converse parts of Theorems 1 and 2 being standard, is relegated to the Appendix. The essential contribution of this paper consists in the direct part of coding Theorems 1 and 2.

Our goal is to show that, when $\Gamma>\Lambda$, for all sufficiently large n there exist $N=\exp(nR)$ codewords $x_1,\cdots,x_N$ in $\mathbb{R}^n$ satisfying $\|x_i\|^2\le n\Gamma$, $i=1,\cdots,N$, with R arbitrarily close to the asserted capacity value, such that for a suitable decoder φ the average probability of error $\bar{e}(s)$ is arbitrarily small uniformly subject to $\|s\|^2\le n\Lambda$.

Using the minimum distance decoder ϕ of (1.5) for the noiseless AVC, (1.4) becomes

$$\bar{e}(s)=\frac{1}{N}\left|\left\{i\colon \|x_i+s-x_j\|^2\le\|s\|^2, \text{ for some } j\ne i\right\}\right|, \tag{2.1}$$

and for the Gaussian case, (1.3) gives

$$\bar{e}(s)=\frac{1}{N}\sum_{i=1}^{N}\Pr\left\{\|x_i+s+V-x_j\|^2\le\|s+V\|^2, \text{ for some } j\ne i\right\}. \tag{2.2}$$

We can assume without any loss of generality that $\Gamma=1$, $0<\Lambda<1$. Further, (2.1) and (2.2) remain unchanged if all vectors are multiplied by $1/\sqrt{n}$. Hence it suffices to prove that for every sufficiently small $\delta>0$ and sufficiently large n there exist $N=\exp(nR)$ unit vectors $x_1,\cdots,x_N$ in $\mathbb{R}^n$ with $C-2\delta<R<C-\delta$ where $C=\frac{1}{2}\log(1+1/\Lambda)$ for the noiseless AVC and $C=\frac{1}{2}\log(1+1/(\Lambda+\sigma^2))$ for the Gaussian AVC, such that $\bar{e}'(s)$ is arbitrarily small, uniformly subject to $\|s\|^2\le\Lambda$, where

$$\bar{e}'(s)=\frac{1}{N}\left|\left\{i\colon \|x_i+s-x_j\|^2\le\|s\|^2, \text{ for some } j\ne i\right\}\right| \tag{2.3}$$

in the noiseless case, and

$$\bar{e}'(s)=\frac{1}{N}\sum_{i=1}^{N}\Pr\left\{\|x_i+s+V-x_j\|^2\le\|s+V\|^2, \text{ for some } j\ne i\right\} \tag{2.4}$$

in the Gaussian case where $V=(V_1,\cdots,V_n)$ is now a sequence of i.i.d. Gaussian random variables with mean 0 and variance σ^2/n.

We claim that the unit vectors $x_1,\cdots,x_N$ of the following Lemma 1 do possess the property above if η and ϵ are sufficiently small.

Lemma 1 (Codeword Properties): For every $\epsilon>0$, $8\sqrt{\epsilon}<\eta<1$, $K>2\epsilon$ and $N=\exp(nR)$ with $2\epsilon\le R\le K$, for $n\ge n_0(\epsilon,\eta,K)$ there exist unit vectors $x_1,\cdots,x_N$ in $\mathbb{R}^n$ such that for every unit vector u in $\mathbb{R}^n$ and constants α,β in $[0,1]$, we have

1) $\left|\{j: (\boldsymbol{x}_j, \boldsymbol{u}) \geq \alpha\}\right| \leq \exp\left\{n\left(|R + \tfrac{1}{2}\log(1-\alpha^2)|^+ + \epsilon\right)\right\}$

and, if $\alpha \geq \eta$, $\alpha^2 + \beta^2 > 1 + \eta - \exp(-2R)$,

2) $$\frac{1}{N}\left|\{i: |(\boldsymbol{x}_j, \boldsymbol{x}_i)| \geq \alpha,\, |(\boldsymbol{x}_j, \boldsymbol{u})| \geq \beta, \text{ for some } j \neq i\}\right| \leq \exp(-n\epsilon).$$

Here $(\cdot,\cdot)$ denotes inner product and $|\cdot|^+$ denotes "the positive part of." This lemma is an analog of the key Lemma 3 of [8], and can be proved similarly. The proof is in the Appendix.

Commencing with the noiseless case, in order to bound (2.3) for $\|\boldsymbol{s}\|^2 \leq \Lambda$, note that

$$\|\boldsymbol{x}_i + \boldsymbol{s} - \boldsymbol{x}_j\|^2 = \|\boldsymbol{x}_i\|^2 + \|\boldsymbol{s}\|^2 + \|\boldsymbol{x}_j\|^2 + 2(\boldsymbol{x}_i, \boldsymbol{s}) - 2(\boldsymbol{x}_i, \boldsymbol{x}_j) - 2(\boldsymbol{x}_j, \boldsymbol{s}). \quad (2.5)$$

Hence

$$\begin{aligned}\bar{e}'(\boldsymbol{s}) &= \frac{1}{N}\left|\{i: (\boldsymbol{x}_j, \boldsymbol{s}) + (\boldsymbol{x}_j, \boldsymbol{x}_i) \geq 1 + (\boldsymbol{x}_i, \boldsymbol{s}), \text{ for some } j \neq i\}\right| \\ &\leq \frac{1}{N}\left|\{i: (\boldsymbol{x}_i, \boldsymbol{s}) \leq -\eta\}\right| \\ &\quad + \frac{1}{N}\left|\{i: (\boldsymbol{x}_j, \boldsymbol{s}) + (\boldsymbol{x}_j, \boldsymbol{x}_i) > 1 - \eta, \text{ for some } j \neq i\}\right|. \end{aligned} \quad (2.6)$$

The first term of the sum in (2.6) can be bounded by Lemma 1(i). In fact, letting $\boldsymbol{u}$ be the unit vector such that

$$\boldsymbol{s} = \boldsymbol{u}\|\boldsymbol{s}\|, \quad (2.7)$$

$(\boldsymbol{x}_i, \boldsymbol{s}) \leq -\eta$ implies by the assumption $\Lambda < 1$ that $(\boldsymbol{x}_i, \boldsymbol{u}) \leq -\eta$. Thus if $R > -\frac{1}{2}\log(1-\eta^2)$, we get that

$$\begin{aligned}\frac{1}{N}\left|\{i: (\boldsymbol{x}_i, \boldsymbol{s}) \leq -\eta\}\right| &\leq \frac{1}{N}\left|\{i: |(\boldsymbol{x}_i, \boldsymbol{u})| \geq \eta\}\right| \\ &\leq \frac{1}{N}\exp\left\{n\left(R + \frac{1}{2}\log(1-\eta^2) + \epsilon\right)\right\} \\ &= \exp\left\{n\left(\frac{1}{2}\log(1-\eta^2) + \epsilon\right)\right\} \\ &\leq \exp\left\{n\left(\epsilon - \frac{\eta^2}{2}\right)\right\} \to 0, \\ &\qquad \text{as } n \to \infty. \end{aligned} \quad (2.8)$$

The second term of the sum in (2.6) can be bounded using 2) of Lemma 1 by suitably partitioning the set of possible values of the inner product $(\boldsymbol{x}_j, \boldsymbol{x}_i)$. To this end, let $\alpha_1 = 1 - \eta - \sqrt{\Lambda} < \alpha_2 < \cdots < \alpha_K = 1 - 2\eta$, with $\alpha_{k+1} - \alpha_k \leq \eta$, $k = 1, \cdots, K-1$. Then $(\boldsymbol{x}_j, \boldsymbol{s}) + (\boldsymbol{x}_j, \boldsymbol{x}_i) > 1 - \eta$ implies that $(\boldsymbol{x}_j, \boldsymbol{x}_i) \geq \alpha_1$, and if $\alpha_k \leq (\boldsymbol{x}_j, \boldsymbol{x}_i) \leq \alpha_{k+1}$ then necessarily $(\boldsymbol{x}_j, \boldsymbol{s}) > 1 - 2\eta - \alpha_k$. The latter, in turn, implies by (2.7) that $(\boldsymbol{x}_j, \boldsymbol{u}) > (1 - 2\eta - \alpha_k)/\sqrt{\Lambda}$. Hence

$$\begin{aligned}&\frac{1}{N}\left|\{i: (\boldsymbol{x}_j, \boldsymbol{s}) + (\boldsymbol{x}_j, \boldsymbol{x}_i) > 1 - \eta, \text{ for some } j \neq i\}\right| \\ &\quad \leq \sum_{k=1}^{K-1} \frac{1}{N}\left|\left\{i: (\boldsymbol{x}_j, \boldsymbol{x}_i) \geq \alpha_k, (\boldsymbol{x}_j, \boldsymbol{u}) \geq \frac{1 - 2\eta - \alpha_k}{\sqrt{\Lambda}}, \text{ for some } j \neq i\right\}\right| \\ &\qquad + \frac{1}{N}\left|\{i: (\boldsymbol{x}_j, \boldsymbol{x}_i) \geq 1 - 2\eta, \text{ for some } j \neq i\}\right|. \end{aligned}$$

To complete the proof of the direct part of Theorem 2, it suffices to check for every $(\alpha, \beta) = (1 - 2\eta, 0)$ and $(\alpha_k, 1 - 2\eta - \alpha_k)/\sqrt{\Lambda})$, $k = 1, \cdots, K$, the condition $\alpha^2 + \beta^2 > 1 + \eta - \exp(-2R)$ of 2) of Lemma 1. (The condition $\alpha \geq \eta$ is clearly satisfied provided $\eta < \min\{\frac{1}{3}, (1 - \sqrt{\Lambda})/2\}$.) Differentiation shows that $\alpha^2 + (1 - 2\eta - \alpha)^2/\Lambda$ is minimized by $\alpha = (1 - 2\eta)/1 + \Lambda$, and the minimum equals $(1 - 2\eta)^2/1 + \Lambda$. Thus, the condition to be satisfied is

$$\frac{(1 - 2\eta)^2}{1 + \Lambda} > 1 + \eta - \exp(-2R)$$

or,

$$R < -\frac{1}{2}\log\left[1 + \eta - \frac{(1 - 2\eta)^2}{1 + \Lambda}\right]. \quad (2.9)$$

Obviously, if $C - 2\delta < R < C - \delta$ for any fixed $\delta > 0$, where

$$C = \frac{1}{2}\log\left(1 + \frac{1}{\Lambda}\right) = -\frac{1}{2}\log\left(1 - \frac{1}{1 + \Lambda}\right),$$

the inequality (2.9) will be satisfied if η is sufficiently small.

The proof for the Gaussian case (Theorem 1) is similar but bounding (2.4) is not as easy. We first present two simple technical lemmas.

Lemma 2: Let the r.v. $\boldsymbol{U}$ be uniformly distributed on the unit n-sphere. Then for every vector $\boldsymbol{u}$ on this sphere and any $0 < \alpha < 1$, we have

$$\Pr\{|(\boldsymbol{U}, \boldsymbol{u})| \geq \alpha\} \leq 2(1 - \alpha^2)^{(n-1)/2}, \qquad \text{if } \alpha \geq \frac{1}{\sqrt{2\pi n}}.$$

Proof: Denote the angle between the unit vectors $\boldsymbol{U}$ and $\boldsymbol{u}$ by $\Theta(\boldsymbol{U}, \boldsymbol{u})$. Then by Shannon [12, (28)],

$$\Pr\{\Theta(\boldsymbol{U}, \boldsymbol{u}) \leq \theta\} \leq \frac{1}{\sqrt{2\pi n}} \frac{\sin^{n-1}\theta}{\cos\theta}, \qquad \text{if } 0 \leq \theta < \pi/2.$$

With $\alpha = \cos\theta$, it follows that

$$\Pr\{(\boldsymbol{U}, \boldsymbol{u}) \geq \alpha\} \leq \frac{1}{\sqrt{2\pi n}} \frac{(1 - \alpha^2)^{(n-1)/2}}{\alpha} \leq (1 - \alpha^2)^{(n-1)/2}, \qquad \text{if } \alpha \geq \frac{1}{\sqrt{2\pi n}}.$$

The proof is completed by observing that $\Pr\{(\boldsymbol{U}, \boldsymbol{u}) \leq -\alpha\} = \Pr\{(\boldsymbol{U}, \boldsymbol{u}) \geq \alpha\}$. □

Lemma 3: Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be unit vectors with $|(\boldsymbol{u}, \boldsymbol{v})| \leq \eta$. Then for any unit vector $\boldsymbol{x}$, the component $\boldsymbol{x}^{\perp}$ of $\boldsymbol{x}$ orthogonal to $\text{span}\{\boldsymbol{u}, \boldsymbol{v}\}$ has norm

$$\|\boldsymbol{x}^{\perp}\|^2 \leq 1 - (\boldsymbol{u}, \boldsymbol{x})^2 - (\boldsymbol{v}, \boldsymbol{x})^2 + 4\eta. \quad (2.10)$$

Further, for any pair of constants α, β,

$$\|\alpha \boldsymbol{u} + \beta \boldsymbol{v}\|^2 \leq (\alpha^2 + \beta^2)(1+\eta). \tag{2.11}$$

Proof: Let $\boldsymbol{v}' = (\boldsymbol{v} - (\boldsymbol{u},\boldsymbol{v})\boldsymbol{u}) / \|\boldsymbol{v} - (\boldsymbol{u},\boldsymbol{v})\boldsymbol{u}\|$ be the unit vector orthogonal to $\boldsymbol{u}$ such that $\operatorname{span}\{\boldsymbol{u},\boldsymbol{v}'\} = \operatorname{span}\{\boldsymbol{u},\boldsymbol{v}\}$. Then

$$\begin{aligned}|(\boldsymbol{v}',\boldsymbol{x})| &= \frac{|(\boldsymbol{v},\boldsymbol{x}) - (\boldsymbol{u},\boldsymbol{v})(\boldsymbol{u},\boldsymbol{x})|}{\|\boldsymbol{v} - (\boldsymbol{u},\boldsymbol{v})\boldsymbol{u}\|} \geq \frac{|(\boldsymbol{v},\boldsymbol{x})| - \eta}{1+\eta}\\ &\geq (|(\boldsymbol{v},\boldsymbol{x})| - \eta)(1-\eta)\\ &\geq |(\boldsymbol{v},\boldsymbol{x})| - 2\eta.\end{aligned}$$

Since $\|\boldsymbol{x}^{\perp}\|^2 = 1 - (\boldsymbol{u},\boldsymbol{x})^2 - (\boldsymbol{v}',\boldsymbol{x})^2$, this implies (2.10). Finally

$$\begin{aligned}\|\alpha\boldsymbol{u} + \beta\boldsymbol{v}\|^2 &= \alpha^2 + \beta^2 + 2\alpha\beta(\boldsymbol{u},\boldsymbol{v})\\ &= (\alpha^2+\beta^2)\left(1 + \frac{2\alpha\beta}{\alpha^2+\beta^2}(\boldsymbol{u},\boldsymbol{v})\right)\\ &\leq (\alpha^2+\beta^2)(1+\eta),\end{aligned}$$

as $|2\alpha\beta/(\alpha^2+\beta^2)| \leq 1$, thereby proving (2.11).

Continuing with the proof of the direct part of Theorem 1, note that on account of (2.8) it suffices to consider only those terms in (2.4) for which $|(\boldsymbol{x}_i,\boldsymbol{u})| \leq \eta$, where $\boldsymbol{u}$ is a unit vector satisfying $\boldsymbol{s} = \boldsymbol{u}\|\boldsymbol{s}\|$. We shall bound these terms using

$$\begin{aligned}\|\boldsymbol{x}_i + \boldsymbol{s} + \boldsymbol{V} - \boldsymbol{x}_j\|^2 = \|\boldsymbol{x}_i\|^2 + \|\boldsymbol{s}+\boldsymbol{V}\|^2 + \|\boldsymbol{x}_j\|^2&\\ + 2(\boldsymbol{x}_i,\boldsymbol{s}) + 2(\boldsymbol{x}_i,\boldsymbol{V})&\\ - 2(\boldsymbol{x}_j,\boldsymbol{x}_i) - 2(\boldsymbol{x}_j,\boldsymbol{s}) - 2(\boldsymbol{x}_j,\boldsymbol{V}).&\end{aligned} \tag{2.12}$$

Decomposing $\boldsymbol{x}_j$ and $\boldsymbol{V}$ into components in $M_{i,\boldsymbol{u}} = \operatorname{span}\{\boldsymbol{x}_i,\boldsymbol{u}\}$ and in $M_{i,\boldsymbol{u}}^{\perp}$, we have

$$\begin{aligned}(\boldsymbol{x}_j,\boldsymbol{V}) &= \left(\boldsymbol{x}_j^{M_{i,u}}, \boldsymbol{V}^{M_{i,u}}\right) + \left(\boldsymbol{x}_j^{M_{i,u}^{\perp}}, \boldsymbol{V}^{M_{i,u}^{\perp}}\right)\\ &= \left(\boldsymbol{x}_j, \boldsymbol{V}^{M_{i,u}}\right) + \left(\boldsymbol{x}_j^{M_{i,u}^{\perp}}, \boldsymbol{V}\right).\end{aligned} \tag{2.13}$$

Since $\boldsymbol{V} = (V_1,\cdots,V_n)$ is a sequence of i.i.d. Gaussian random variables with mean 0 and variance σ^2/n, we have as $n \to \infty$ that

$$\Pr\{|(\boldsymbol{x}_i,\boldsymbol{V})| > \eta\} \to 0, \qquad \Pr\{\|\boldsymbol{V}^{M_{i,u}}\| > \eta\} \to 0$$

uniformly in i and $\boldsymbol{u}$. This along with (2.12), (2.13), implies that

$$\begin{aligned}&\Pr\{\|\boldsymbol{x}_i + \boldsymbol{s} + \boldsymbol{V} - \boldsymbol{x}_j\|^2 \leq \|\boldsymbol{s}+\boldsymbol{V}\|^2, \text{ for some } j \neq i\}\\ &\quad = \Pr\Big\{(\boldsymbol{x}_j,\boldsymbol{x}_i) + (\boldsymbol{x}_j,\boldsymbol{s}) + \left(\boldsymbol{x}_j^{M_{i,u}^{\perp}},\boldsymbol{V}\right) > 1 + (\boldsymbol{x}_i,\boldsymbol{s})\\ &\quad + (\boldsymbol{x}_i,\boldsymbol{V}) - \left(\boldsymbol{x}_j^{M_{i,u}},\boldsymbol{V}\right), \text{ for some } j \neq i\Big\}\\ &\quad \leq \Pr\Big\{\left(\boldsymbol{x}_j^{M_{i,u}^{\perp}},\boldsymbol{V}\right) > 1 - 3\eta - |(\boldsymbol{x}_j,\boldsymbol{x}_i)| - |(\boldsymbol{x}_j,\boldsymbol{s})|,\\ &\qquad\qquad \text{for some } j \neq i\Big\} + \epsilon,\end{aligned} \tag{2.14}$$

for all sufficiently large n, whenever $|(\boldsymbol{x}_i,\boldsymbol{u})| \leq \eta$.

Hence, in order to show that $\bar{e}'(\boldsymbol{s})$ in (2.4) goes to 0 uniformly subject to $\|\boldsymbol{s}\|^2 \leq \Lambda$, it suffices to prove that

$$\begin{aligned}\frac{1}{N}\sum_{i:\,|(\boldsymbol{x}_i,\boldsymbol{u})| \leq \eta} \Pr\Big\{\left(\boldsymbol{x}_j^{M_{i,u}^{\perp}},\boldsymbol{V}\right) > 1 - 3\eta - |(\boldsymbol{x}_j,\boldsymbol{x}_i)|&\\ - |(\boldsymbol{x}_j,\boldsymbol{u})|\sqrt{\Lambda}, \text{ for some } j \neq i\Big\}&\end{aligned} \tag{2.15}$$

converges to 0 uniformly for unit vectors $\boldsymbol{u} \in \mathbb{R}^n$, as $n \to \infty$.

To this end, we partition the set of all possible values of the inner products $(\boldsymbol{x}_j,\boldsymbol{x}_i)$ and $(\boldsymbol{x}_j,\boldsymbol{u})$. Let $\alpha_1 = 0 < \alpha_2 < \cdots < \alpha_K = 1$ and $\beta_1 = 0 < \beta_2 < \cdots < \beta_L = 1$. with $\alpha_{k+1} - \alpha_k \leq \eta$, $k = 1,\cdots,K-1$, and $\beta_{l+1} - \beta_l \leq \eta$, $l = 1,\cdots,L-1$. Further let

$$\begin{aligned}F_{ikl} = \Big\{j\colon j \neq i,\ \alpha_k \leq |(\boldsymbol{x}_j,\boldsymbol{x}_i)| \leq \alpha_{k+1},&\\ \beta_l \leq |(\boldsymbol{x}_j,\boldsymbol{u})| \leq \beta_{l+1}\Big\}&\end{aligned}$$

and

$$\begin{aligned}G = \Big\{(k,l)\colon 1 \leq k \leq K,\ 1 \leq l \leq L,\ \alpha_k \geq \eta,&\\ \alpha_k^2 + \beta_l^2 > 1 + \eta - \exp(-2R)\Big\}.&\end{aligned}$$

Then the expression (2.15) is

$$\begin{aligned}&\leq \sum_{(k,l)\in G}\frac{1}{N}|\{i\colon F_{ikl} \neq \phi\}|\\ &\quad + \frac{1}{N}\sum_{i:\,|(\boldsymbol{x}_i,\boldsymbol{u})| \leq \eta}\ \sum_{(k,l)\notin G}\Pr\Big\{\bigcup_{j\in F_{ikl}}\Big\{\left(\boldsymbol{x}_j^{M_{i,u}^{\perp}},\boldsymbol{V}\right)\\ &> 1 - 5\eta - \alpha_k - \beta_l\sqrt{\Lambda}\Big\}\Big\}.\end{aligned}$$

As the first term above goes to zero uniformly in $\boldsymbol{u}$ as $n \to \infty$ by 2) of Lemma 1, it remains to consider only the second term. Recalling again that $\boldsymbol{V} = (V_1,\cdots,V_n)$ is an i.i.d. sequence of $N(0,\sigma^2/n)$ random variables, we note that $\Pr\{\|\boldsymbol{V}\|^2 > \sigma^2 + \eta\} \to 0$ as $n \to \infty$. Therefore, it suffices to prove that

$$\begin{aligned}\frac{1}{N}\sum_{i:\,|(\boldsymbol{x}_i,\boldsymbol{u})| \leq \eta}\ \sum_{(k,l)\notin G}\Pr\Big\{\bigcup_{j\in F_{ikl}}\Big\{\|\boldsymbol{V}\|^2 \leq \sigma^2 + \eta,&\\ \left(\boldsymbol{x}_j^{M_{i,u}^{\perp}},\boldsymbol{V}\right) > 1 - 5\eta - \alpha_k - \beta_l\sqrt{\Lambda}\Big\}\Big\}&\end{aligned} \tag{2.16}$$

goes to 0 uniformly in $\boldsymbol{u}$.

Now, observe that

$$\begin{aligned}|F_{ikl}| &\leq |\{j\colon |(\boldsymbol{x}_j,\boldsymbol{x}_i)| \geq \alpha_k,\ |(\boldsymbol{x}_j,\boldsymbol{u})| \geq \beta_l\}|\\ &\leq |\{j\colon |(\boldsymbol{x}_j,\alpha_k\boldsymbol{x}_i + \beta_l\boldsymbol{u})| \geq \alpha_k^2 + \beta_l^2\}|\\ &\leq \left|\left\{j\colon |(\boldsymbol{x}_j,\boldsymbol{w})| \geq \frac{\alpha_k^2+\beta_l^2}{\|\alpha_k\boldsymbol{x}_i + \beta_l\boldsymbol{u}\|}\right\}\right|\end{aligned}$$

where $\boldsymbol{w} = (\alpha_k\boldsymbol{x}_i + \beta_l\boldsymbol{u})/\|\alpha_k\boldsymbol{x}_i + \beta_l\boldsymbol{u}\|$. By Lemma 3 (for $|(\boldsymbol{x}_i,\boldsymbol{u})| \leq \eta$), $\|\alpha_k\boldsymbol{x}_i + \beta_l\boldsymbol{u}\| \leq \sqrt{(\alpha_k^2+\beta_l^2)(1+\eta)}$, so that

$$\begin{aligned}|F_{ikl}| &\leq \left|\left\{j\colon |(\boldsymbol{x}_j,\boldsymbol{w})| \geq \sqrt{\frac{\alpha_k^2+\beta_l^2}{1+\eta}}\right\}\right|\\ &\leq \exp\left\{n\left(\left|R + \frac{1}{2}\log\left(1 - \frac{\alpha_k^2+\beta_l^2}{1+\eta}\right)\right|^{+} + \epsilon\right)\right\}\\ &\qquad\qquad \text{by Lemma } 1(i)\\ &\leq \exp\left\{n\left(\left|R + \frac{1}{2}\log(1 - \alpha_k^2 - \beta_l^2 + \eta)\right|^{+} + \epsilon\right)\right\}\end{aligned} \tag{2.17}$$

for all n sufficiently large, where we can assume that

$$\alpha_k^2 + \beta_l^2 < 1 + \eta \tag{2.18}$$

(as otherwise $F_{ikl} = \phi$).

Further, by Lemma 3, if $\boldsymbol{x}_j' = \boldsymbol{x}_j'(i, \boldsymbol{u})$ represents the unit vector in the direction of $\boldsymbol{x}_j^{M_{i,u}}$, for $j \in F_{ikl}$ we have $\|\boldsymbol{x}_j^{M_{i,u}}\| \leq \sqrt{1-(\boldsymbol{x}_j,\boldsymbol{x}_i)^2-(\boldsymbol{x}_j,\boldsymbol{u})^2+4\eta} \leq \sqrt{1-\alpha_k^2-\beta_l^2+4\eta}$ if $|(\boldsymbol{x}_i,\boldsymbol{u})| \leq \eta$. Hence

$$\Pr\left\{\|V\|^2 \leq \sigma^2+\eta,\left(\boldsymbol{x}_j^{M_{i,u}},V\right) > 1-5\eta-\alpha_k-\beta_l\sqrt{\Lambda}\right\}$$

$$\leq \Pr\left\{\|V\|^2 \leq \sigma^2+\eta,(\boldsymbol{x}_j',V) > \frac{1-\alpha_k-\beta_l\sqrt{\Lambda}-5\eta}{\sqrt{1-\alpha_k^2-\beta_l^2+4\eta}}\right\}$$

$$= \int_0^{\sigma^2+\eta}\Pr\left\{(\boldsymbol{x}_j',V) > \frac{1-\alpha_k-\beta_l\sqrt{\Lambda}-5\eta}{\sqrt{1-\alpha_k^2-\beta_l^2+4\eta}}\,\middle|\,\|V\|^2 = r\right\} \cdot dF(r),$$

$$\text{where } F(r) = \Pr\{\|V\|^2 \leq r\}$$

$$= \int_0^{\sigma^2+\eta}\Pr\left\{\left(\boldsymbol{x}_j',\frac{V}{\|V\|}\right) > \frac{1-\alpha_k-\beta_l\sqrt{\Lambda}-5\eta}{\sqrt{r\left(1-\alpha_k^2-\beta_l^2+4\eta\right)}}\,\middle|\,\|V\|^2 = r\right\} dF(r)$$

$$\leq \Pr\left\{(\boldsymbol{u}',U) > \frac{1-\alpha_k-\beta_l\sqrt{\Lambda}-5\eta}{\sqrt{(\sigma^2+\eta)\left(1-\alpha_k^2-\beta_l^2+4\eta\right)}}\right\} \tag{2.19}$$

where U is a r.v. distributed uniformly on the unit n-sphere and $\boldsymbol{u}'$ is any fixed unit vector in $\mathbb{R}^n$. Together with (2.17), this implies that (2.16) is overbounded by $\Sigma_{(k,l)\notin G}A_{kl}^{(n)}$, where

$$A_{kl}^{(n)} = \exp\left\{n\left(\left|R+\frac{1}{2}\log\left(1-\alpha_k^2-\beta_l^2+\eta\right)\right|^+ + \epsilon\right)\right\}$$

$$\cdot\Pr\left\{(\boldsymbol{u}',U) > \frac{1-\alpha_k-\beta_l\sqrt{\Lambda}-5\eta}{\sqrt{(\sigma^2+\eta)\left(1-\alpha_k^2-\beta_l^2+4\eta\right)}}\right\}. \tag{2.20}$$

Hence it suffices to show that $A_{kl}^{(n)} \to 0$ as $n \to \infty$ for every $(k,l) \notin G$.

Since $(k,l) \notin G$, there are two cases to consider:

a) $\alpha_k \leq \eta, \quad \alpha_k^2+\beta_l^2 > 1+\eta-\exp(-2R),$

and

b) $\alpha_k^2+\beta_l^2 \leq 1+\eta-\exp(-2R).$

We first observe that in both cases

$$1-\alpha_k-\beta_l\sqrt{\Lambda}-5\eta > 0 \tag{2.21}$$

provided that η is chosen sufficiently small. Indeed, in case a), the expression in (2.21) is $\geq 1-\sqrt{\Lambda}-6\eta$. In case b), the assumption $R < C-\delta = \frac{1}{2}\log(1+1/(\Lambda+\sigma^2))-\delta$ implies that

$$\alpha_k^2+\beta_l^2 \leq 1+\eta-\frac{1}{1+\dfrac{1}{\Lambda+\sigma^2}}\exp(2\delta)$$

$$< 1+\eta-\frac{\Lambda}{1+\Lambda}\exp(2\delta).$$

Since

$$\alpha_k+\beta_l\sqrt{\Lambda} \leq \sqrt{\left(\alpha_k^2+\beta_l^2\right)(1+\Lambda)}$$

(as can be directly verified by squaring both sides), this yields

$$1-\alpha_k-\beta_l\sqrt{\Lambda}-5\eta > 1-5\eta$$
$$-\sqrt{1-\Lambda(\exp(2\delta)-1)+\eta(1+\Lambda)} > 0$$

if η is sufficiently small.

Now, in case a) we obtain, using Lemma 2, that

$$A_{kl}^{(n)} \leq \exp(n\epsilon)\left[1-\frac{\left(1-\alpha_k-\beta_l\sqrt{\Lambda}-5\eta\right)^2}{(\sigma^2+\eta)\left(1-\alpha_k^2-\beta_l^2+4\eta\right)}\right]^{(n-1)/2}$$

$$\leq \exp(n\epsilon)\left[1-\frac{\left(1-\sqrt{\Lambda}-6\eta\right)^2}{(\sigma^2+\eta)(1+4\eta)}\right]^{(n-1)/2} \to 0$$

if ϵ and η are chosen small enough.

In case b), we have $R+\frac{1}{2}\log(1-\alpha_k^2-\beta_l^2+\eta) > 0$. Then, using Lemma 2 we obtain from (2.20) that

$$A_{kl}^{(n)} \leq \exp\left\{n\left(R+\frac{1}{2}\log\left(1-\alpha_k^2-\beta_l^2+\eta\right)+\epsilon\right\}\right.$$

$$\cdot\left[1-\frac{\left(1-\alpha_k-\beta_l\sqrt{\Lambda}-5\eta\right)^2}{(\sigma^2+\eta)\left(1-\alpha_k^2-\beta_l^2+4\eta\right)}\right]^{(n-1)/2}$$

$$= \exp\left\{R+\frac{1}{2}\log\left(1-\alpha_k^2-\beta_l^2+\eta\right)+n\epsilon\right\}$$

$$\cdot\exp\left\{(n-1)\left(R+\frac{1}{2}\log\left(1-\alpha_k^2-\beta_l^2+\eta\right)\right.\right.$$

$$\left.\left.+\frac{1}{2}\log\left[1-\frac{\left(1-\alpha_k-\beta_l\sqrt{\Lambda}-5\eta\right)^2}{(\sigma^2+\eta)\left(1-\alpha_k^2-\beta_l^2+4\eta\right)}\right]\right)\right\}$$

$$\leq \exp\left(R+\frac{1}{2}\log(1+\eta)+\epsilon\right)$$

$$\cdot\exp\left\{(n-1)\left(R+\epsilon+\frac{1}{2}\log\left[1-\alpha_k^2-\beta_l^2+4\eta\right.\right.\right.$$

$$\left.\left.\left.-\frac{\left(1-\alpha_k-\beta_l\sqrt{\Lambda}-5\eta\right)^2}{\sigma^2+\eta}\right]\right)\right\}.$$

Evaluating the maximum of

$$\gamma(\alpha,\beta) = 1-\alpha^2-\beta^2+4\eta-\frac{\left(1-\alpha-\beta\sqrt{\Lambda}-5\eta\right)^2}{\sigma^2+\eta},$$

we obtain by differentiation that the maximum is attained at

$$\alpha = \frac{1-5\eta}{1+\Lambda+\sigma^2+\eta}, \qquad \beta = \alpha\sqrt{\Lambda}$$

and the value of the maximum is

$$1+4\eta-\frac{(1-5\eta)^2}{1+\Lambda+\sigma^2+\eta}.$$

Thus, in case b), $A_{kl}^{(n)} \to 0$ if

$$R < -\frac{1}{2}\log\left\{1+4\eta-\frac{(1-5\eta)^2}{1+\Lambda+\sigma^2+\eta}\right\}-\epsilon. \quad (2.22)$$

Obviously, if $C-2\delta < R < C-\delta$ for any fixed $\delta > 0$, where

$$C=\frac{1}{2}\log\left(1+\frac{1}{\Lambda+\sigma^2}\right)=-\frac{1}{2}\log\left(1-\frac{1}{1+\Lambda+\sigma^2}\right),$$

the inequality (2.22) will be satisfied if η and ϵ are sufficiently small.

This completes the proof. □

III. Discussion

We have established the capacity of the Gaussian AVC with input constraint Γ, state constraint Λ, and noise power σ^2, for deterministic codes and the average probability of error criterion. It is 0 if $\Lambda \geq \Gamma$, and equals the capacity of an ordinary memoryless channel with additive Gaussian noise of power $\Lambda+\sigma^2$, for the same input constraint Γ, if $\Lambda < \Gamma$. The limiting value of this capacity as $\sigma^2 \to 0$ is, as expected, the capacity of the noiseless AVC with input constraint Γ and state constraint Λ. The previous result solves a weakened version of the problem of determining the exponential rate of the tightest sphere-packing in $\mathbb{R}^n$ as $n \to \infty$.

A comparison of the Gaussian AVC with the discrete case treated in [8], [9] indicates that the former is simpler in that it does not call for a complex decoding rule. Indeed, simple minimum-distance decoding suffices to achieve capacity. On the other hand, since the powerful and intuitive method of types is unavailable, the computations are less lucid and appear to rely, to a degree, on analytical artifices.

One generalization of Theorem 1 is immediate. Namely, if in the representation $x+s+V$ of the channel output, the variances of the independent, zero-mean Gaussian components of V are allowed to vary arbitrarily subject to $\sigma_i \leq \sigma$, $i=1,\cdots,N$, the capacity remains unaltered. Indeed the only change necessitated in the error-bounding is the replacement of $\|V\|^2 \leq \sigma^2+\eta$ in (2.15) and (2.18) by $\|\tilde{V}\|^2 \leq \sigma^2+\eta$, where $\tilde{V}=(V_1\sigma/\sigma_1,\cdots,V_n\sigma/\sigma_n)$ has i.i.d. components, followed by the observation that $(x_j', V)=(\tilde{x}_j', \tilde{V})$, where $\|\tilde{x}_j'\| < 1$ since $\tilde{x}_j'$ is obtained by multiplying the components of the unit vector x_j' by $\sigma_i/\sigma \leq 1$, $i=1,\cdots,n$.

A further generalization with arbitrarily varying noise variances subject only to $(1/n)\sum_{i=1}^n \sigma_i^2 \leq \sigma^2$, when we believe the capacity to yet remain unchanged, does not yield to such a simple artifice. On the contrary, it apparently requires more complex calculations, including a generalization of Lemma 1. Indeed for more general AVC models with continuous alphabets, the direct approach may well become unmanageable, thus necessitating recourse to the method of approximations by discrete AVC's.

Appendix

We shall first prove the converse parts of Theorems 1 and 2, followed by the proof of Lemma 1 (cf. Section II).

The fact that $\Gamma \leq \Lambda$ implies $C=0$ follows by a well-known argument of Blackwell *et al.* [5]. Namely, let $x_1,\cdots,x_N$, $N \geq 2$, be arbitrary codewords in $\mathbb{R}^n$ satisfying (1.1), and—assuming $\Gamma \leq \Lambda$—consider the state sequences $s_1 = x_1,\cdots,s_N=x_N$. Then for any decoder φ,

$$\begin{aligned}\Pr\{\varphi(x_i+s_j+V)\neq i\} &= \Pr\{\varphi(x_j+s_i+V)\neq i\}\\ &\geq 1-\Pr\{\varphi(x_j+s_i+V)\neq j\}\end{aligned}$$

whenever $i \neq j$. Hence

$$\begin{aligned}\frac{1}{N}\sum_{j-1}^{N}\bar{e}(s_j) &= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Pr\{\varphi(x_i+s_j+V)\neq i\}\\ &\geq \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{i-1}\big[\Pr\{\varphi(x_i+s_j+V)\neq i\}\\ &\qquad +\Pr\{\varphi(x_j+s_i+V)\neq j\}\big]\\ &\geq \frac{1}{N^2}\cdot\frac{N(N-1)}{2}\geq\frac{1}{4},\end{aligned}$$

and consequently $\bar{e}(s_j) \geq \frac{1}{4}$ for at least one $j \in \{1,\cdots,N\}$.

Next, to prove that $C \leq \frac{1}{2}\log(1+\Gamma/(\sigma^2+\Lambda))$, consider an i.i.d. sequence $S=(S_1,\cdots,S_n)$ of Gaussian random variables (also independent of V) with mean 0 and variance $\Lambda' < \Lambda$. Given any code with codewords $x_1,\cdots,x_N$, and decoder φ, the expectation of $\bar{e}(S)$ (cf. (1.3)) equals

$$E\bar{e}(S)=\frac{1}{N}\sum_{i=1}^{N}\Pr\{\varphi(x_i+S+V)\neq i\}$$

while

$$E\bar{e}(S) \leq \max_{s:\,\|s\|^2\leq n\Lambda}\bar{e}(s)+\Pr\{\|S\|^2>n\Lambda\}.$$

Since $\operatorname{var} S_i = \Lambda' < \Lambda$, it holds that $\Pr\{\|S\|^2 > n\Lambda\} \to 0$ and, therefore, if $\max_{\|s\|^2\leq n\Lambda}\bar{e}(s) < \epsilon$, we have

$$\frac{1}{N}\sum_{i=1}^{N}\Pr\{\varphi(x_i+S+V)\neq i\}\leq 2\epsilon,$$

for $n \geq n_0(\epsilon,\Lambda')$.

Here the left side is the average probability of error of the given code on the "ordinary" memoryless channel with additive Gaussian noise of variance $\Lambda'+\sigma^2$. Hence, by the converse to the coding theorem for such channels, it follows that

$$C \leq C' = \frac{1}{2}\log\left(1+\frac{\Gamma}{\Lambda'+\sigma^2}\right).$$

Since $\Lambda' < \Lambda$ was arbitrary, this completes the proof of the converse part of Theorem 1 (and, with the obvious changes, also that of Theorem 2).

We now prove Lemma 1. We shall show that $N=\exp(nR)$ randomly selected unit vectors will possess, with probability arbitrarily close to 1, all the properties states in Lemma 1. Its proof entails Chernoff bounding applied to dependent random variables, which is provided by the following

Lemma A1: Let $\mathbf{Z}_1,\cdots,\mathbf{Z}_N$ be arbitrary r.v.'s and $f_i(\mathbf{Z}_1,\cdots,\mathbf{Z}_i)$ be arbitrary with $0 \leq f_i \leq 1$, $i=1,\cdots,N$. Then the condition

$$E[f_i(\mathbf{Z}_1,\cdots,\mathbf{Z}_i)|\mathbf{Z}_1,\cdots,\mathbf{Z}_{i-1}] \leq a \text{ a.s.}, \qquad i=1,\cdots,N,$$

implies that

$$\Pr\left\{\frac{1}{N}\sum_{i=1}^{N}f_i(\mathbf{Z}_1,\cdots,\mathbf{Z}_i)>t\right\}\leq\exp\{-N(t\log 2-a)\}.$$

Proof: This lemma is the same as Lemma A1 of [8], with the exponentials and logarithms to the base 2 in the latter replaced by natural exponentials and logarithms.

Proof of Lemma 1: Throughout this proof "for large n" will mean "for all n larger than some threshold n_0 depending only on ϵ, η, and K."

Let $\mathbf{Z}_1,\cdots,\mathbf{Z}_N$ be independent r.v.'s each uniformly distributed on the unit n-sphere. First, fix a unit vector $\boldsymbol{u}$ in $\mathbb{R}^n$ and α,β in $[0,1]$. The main step of the proof consists in asserting the doubly exponential probability bounds that for large n

$$\Pr\left\{\left|\{j: |(\mathbf{Z}_j,\boldsymbol{u})| \geq \alpha\}\right| > \exp\left\{n\left(\left|R+\frac{1}{2}\log(1-\alpha^2)\right|^+ + \frac{\epsilon}{2}\right)\right\}\right\} \leq \exp\left(-\frac{1}{2}\exp\left(\frac{n\epsilon}{2}\right)\right) \quad \text{(A1)}$$

and if

$$\alpha \geq \frac{\eta}{2}, \qquad \alpha^2+\beta^2 > 1+\frac{3\eta}{4}-\exp(-2R), \quad \text{(A2)}$$

$$\Pr\left\{\frac{1}{N}\left|\{i: |(\mathbf{Z}_j,\mathbf{Z}_i)| \geq \alpha,\ |(\mathbf{Z}_j,\boldsymbol{u})| \geq \beta,\ \text{for some } j \neq i\}\right| > \exp(-n\epsilon)\right\}$$
$$\leq 4\exp\left(-\frac{1}{10}\exp\left(\frac{n\epsilon}{2}\right)\right). \quad \text{(A3)}$$

To establish (A1), (A3), we shall apply Lemma A1 to $\mathbf{Z}_1,\cdots,\mathbf{Z}_N$ for two different choices of the functions f_i.

Observe that (A1) holds trivially if $\frac{1}{2}\log(1-\alpha^2)+\epsilon/2 > 0$, i.e., if $\alpha < \sqrt{1-\exp(-\epsilon)}$. Hence, restricting attention to $\alpha \geq \sqrt{1-\exp(-\epsilon)} > 0$, let

$$f_j(\mathbf{Z}_1,\cdots,\mathbf{Z}_j) = \begin{cases} 1, & \text{if } |(\mathbf{Z}_j,\boldsymbol{u})| \geq \alpha \\ 0, & \text{otherwise} \end{cases}.$$

The hypothesis of Lemma A1 is then satisfied by Lemma 2 with $a = 2(1-\alpha^2)^{(n-1)/2}$ for large n. Thus with $t = (1/N)\exp\{n(|R+\frac{1}{2}\log(1-\alpha^2)|^+ + \epsilon/2)\}$ in Lemma A1, we get

$$\Pr\left\{\frac{1}{N}\left|\{j: |(\mathbf{Z}_j,\boldsymbol{u})| \geq \alpha\}\right| > \frac{1}{N}\exp\left\{n\left(\left|R+\frac{1}{2}\log(1-\alpha^2)\right|^+ + \frac{\epsilon}{2}\right)\right\}\right\}$$
$$\leq \exp\left\{-\left[N\left(\frac{1}{N}\exp\left\{n\left(\left|R+\frac{1}{2}\log(1-\alpha^2)\right|^+ + \frac{\epsilon}{2}\right)\right\}\cdot\log 2 - 2(1-\alpha^2)^{(n-1)/2}\right)\right]\right\}.$$

The inequality (A1) would then follow if we showed that the term within the square brackets, denoted $h(n,R,\epsilon,\alpha)$, was bounded below by $\frac{1}{2}\exp(n\epsilon/2)$ for large n. There are two cases to consider.

a) If

$$R > -\frac{1}{2}\log(1-\alpha^2), \quad \text{(A4)}$$

observe that

$$\begin{aligned} h(n,R,\epsilon,\alpha) &= \exp\left\{n\left(R+\frac{1}{2}\log(1-\alpha^2)+\frac{\epsilon}{2}\right)\right\}\log 2 \\ &\quad - 2\exp(nR)\exp\left\{\frac{n-1}{2}\log(1-\alpha^2)\right\} \\ &= \exp\left\{n\left(R+\frac{1}{2}\log(1-\alpha^2)\right)\right\} \\ &\quad \cdot\left\{\exp\left(\frac{n\epsilon}{2}\right)\cdot\log 2 - 2\exp\left\{-\frac{1}{2}\log(1-\alpha^2)\right\}\right\} \\ &\geq \exp\left(\frac{n\epsilon}{2}\right)\cdot\log 2 - 2\exp R \qquad \text{(by (A4))} \\ &\geq \frac{1}{2}\exp\left(\frac{n\epsilon}{2}\right) \end{aligned}$$

for large n.

b) If

$$R \leq -\frac{1}{2}\log(1-\alpha^2) \quad \text{(A5)}$$

then

$$\begin{aligned} h(n,R,\epsilon,\alpha) &= \exp\left(\frac{n\epsilon}{2}\right)\cdot\log 2 \\ &\quad - 2\exp(nR)\exp\left\{\frac{n-1}{2}\log(1-\alpha^2)\right\} \\ &\geq \exp\left(\frac{n\epsilon}{2}\right)\cdot\log 2 - 2\exp R \qquad \text{(by A5)} \end{aligned}$$

with the bounding completed as in case a).

Turning to (A3) next, define

$$A_i = \left\{j: j < i,\ \left|(\mathbf{Z}_j,\boldsymbol{u})\right| \geq \beta\right\}$$

and

$$\tilde{A}_i = \begin{cases} A_i, & \text{if } |A_i| \leq \exp\left\{n\left(\left|R+\frac{1}{2}\log(1-\beta^2)\right|^+ + \frac{\epsilon}{2}\right)\right\} \\ \phi, & \text{otherwise.} \end{cases}$$

First note that $\{\tilde{A}_i \neq A_i \text{ for some } i\} \subset \{|A_N| > \exp\{n(|R+\frac{1}{2}\log(1-\beta^2)|^+ + \epsilon/2)\}\}$, which by (A1) has probability less than $\exp(-\frac{1}{2}\exp(n\epsilon/2))$ for large n. Next, let

$$f_i(\mathbf{Z}_1,\cdots,\mathbf{Z}_i) = \begin{cases} 1, & \text{if } \left|(\mathbf{Z}_j,\mathbf{Z}_i)\right| \geq \alpha \text{ for some } j \in \tilde{A}_i \\ 0, & \text{otherwise.} \end{cases}$$

Then for large n

$$\Pr\left\{\frac{1}{N}\left|\{i: |(\mathbf{Z}_j,\mathbf{Z}_i)| \geq \alpha,\ |(\mathbf{Z}_j,\boldsymbol{u})| \geq \beta,\ \text{for some } j < i\}\right| > \exp(-n\epsilon)\right\}$$
$$\leq \exp\left(-\frac{1}{2}\exp\left(\frac{n\epsilon}{2}\right)\right) + \Pr\left\{\frac{1}{N}\sum_{i=1}^{N} f_i(\mathbf{Z}_1,\cdots,\mathbf{Z}_i) > \exp(-n\epsilon)\right\}. \quad \text{(A6)}$$

The second term on the right side of (A6) will be bounded using Lemma A1. To this end, we introduce the event $\{|(\mathbf{Z}_i, \boldsymbol{u})| \geq \eta/4\}$ and note from Lemma 2 that its probability is less than $2(1-\eta^2/16)^{(n-1)/2}$ for large n. Also, writing $\mathbf{Z}_i = (\mathbf{Z}_i, \boldsymbol{u})\boldsymbol{u} + \mathbf{Z}_i^{\perp}$, we see that $(\mathbf{Z}_j, \mathbf{Z}_i) = (\mathbf{Z}_i, \boldsymbol{u})(\mathbf{Z}_j, \boldsymbol{u}) + (\mathbf{Z}_j, \mathbf{Z}_i^{\perp})$. Hence

$$E[f_i(\mathbf{Z}_1, \cdots, \mathbf{Z}_i)|\mathbf{Z}_1, \cdots, \mathbf{Z}_{i-1}]$$

$$= \Pr\left\{ \bigcup_{j \in \bar{A}_i} \{|(\mathbf{Z}_j, \mathbf{Z}_i)| \geq \alpha\} | \mathbf{Z}_1, \cdots, \mathbf{Z}_{i-1} \right\}$$

$$\leq \Pr\left\{ |(\mathbf{Z}_i, \boldsymbol{u})| \geq \frac{\eta}{4} \right\}$$

$$+ \sum_{j \in \bar{A}_i} \Pr\left\{ |(\mathbf{Z}_i, \boldsymbol{u})| < \frac{\eta}{4}, |(\mathbf{Z}_j, \mathbf{Z}_i)| \geq \alpha | \mathbf{Z}_1, \cdots, \mathbf{Z}_{i-1} \right\}$$

$$\leq 2\left(1 - \frac{\eta^2}{16}\right)^{(n-1)/2} + \sum_{j \in \bar{A}_i} \Pr\left\{ \left|(\mathbf{Z}_j, \mathbf{Z}_i^{\perp})\right| \geq \alpha - \frac{\eta}{4} \middle| \mathbf{Z}_j = 1\boldsymbol{z}_j \right\}$$

$$= 2\left(1 - \frac{\eta^2}{16}\right)^{(n-1)/2} + \sum_{j \in \bar{A}_i} \Pr\left\{ \left|(\boldsymbol{z}_j, \mathbf{Z}_i^{\perp})\right| \geq \alpha - \frac{\eta}{4} \right\}. \tag{A7}$$

Now, $(\boldsymbol{z}_j, \mathbf{Z}_i^{\perp}) = (\boldsymbol{z}_j^{\perp}, \mathbf{Z}_i)$, where $\|\boldsymbol{z}_j^{\perp}\|^2 = 1 - (\boldsymbol{z}_j, \boldsymbol{u})^2 \leq 1 - \beta^2$. Then if $\tilde{\boldsymbol{u}}$ is any fixed unit vector in $\mathbb{R}^n$, we obtain for large n that

$$\Pr\left\{ \left|(\boldsymbol{z}_j, \mathbf{Z}_i^{\perp})\right| \geq \alpha - \frac{\eta}{4} \right\} \leq \Pr\left\{ |(\tilde{\boldsymbol{u}}, \mathbf{Z}_i)| \geq \frac{\alpha - \frac{\eta}{4}}{\sqrt{1-\beta^2}} \right\}$$

$$\leq 2\left(1 - \frac{\left(\alpha - \frac{\eta}{4}\right)^2}{1-\beta^2}\right)^{(n-1)/2} \tag{A8}$$

by Lemma 2 if $\alpha - \eta/4 < \sqrt{1-\beta^2}$; otherwise the probability is trivially zero. Since $|\bar{A}_i| \leq \exp\{n(|R + \frac{1}{2}\log(1-\beta^2)|^+ + \epsilon/2)\}$, we obtain from (A7) and (A8) that the hypothesis of Lemma A1 is satisfied with

$$a = 2\left(1 - \frac{\eta^2}{16}\right)^{(n-1)/2} + 2\exp\left\{ n\left(\left| R + \frac{1}{2}\log(1-\beta^2) \right|^+ + \frac{\epsilon}{2} \right) \right\}$$

$$\cdot \exp\left\{ \frac{n-1}{2} \log\left(1 - \frac{(\alpha - \eta/4)^2}{1-\beta^2} \right) \right\} \tag{A9}$$

if $\alpha - \eta/4 < \sqrt{1-\beta^2}$, and otherwise with

$$a = 2\left(1 - \frac{\eta^2}{16}\right)^{(n-1)/2}. \tag{A10}$$

Thus, with $t = \exp(-n\epsilon)$ in Lemma A1, we get

$$\Pr\left\{ \frac{1}{N} \sum_{i=1}^{N} f_i(\mathbf{Z}_1, \cdots, \mathbf{Z}_i) > \exp(-n\epsilon) \right\}$$

$$\leq \exp\{-N[\exp(-n\epsilon)\log 2 - a]\}. \tag{A11}$$

We claim that for large n

$$a < \frac{1}{2}\exp(-n\epsilon). \tag{A12}$$

Observe that for large n (using the hypothesis $\eta > 8\sqrt{\epsilon}$)

$$2\left(1 - \frac{\eta^2}{16}\right)^{(n-1)/2} \leq 2\exp\left\{ -(n-1)\frac{\eta^2}{32} \right\} \leq \frac{1}{4}\exp(-n\epsilon) \tag{A13}$$

establishing (A12) when $\alpha - \eta/4 \geq \sqrt{1-\beta^2}$ (cf. (A10)). If $\alpha - \eta/4 < \sqrt{1-\beta^2}$, the second term in (A9) is, for large n,

$$\leq 2\exp\left\{ (n-1)\left[\left| R + \frac{1}{2}\log(1-\beta^2) \right|^+ + \frac{3\epsilon}{4} + \frac{1}{2}\log\left(1 - \frac{\left(\alpha - \frac{\eta}{4}\right)^2}{1-\beta^2} \right) \right] \right\}. \tag{A14}$$

We distinguish between two cases.

a) If $R > -\frac{1}{2}\log(1-\beta^2)$, then the bound in (A14) is

$$= 2\exp\left\{ (n-1)\left[R + \frac{3\epsilon}{4} + \frac{1}{2}\log\left(1 - \beta^2 - \left(\alpha - \frac{\eta}{4}\right)^2 \right) \right] \right\}$$

$$\leq 2\exp\left\{ (n-1)\left[R + \frac{3\epsilon}{4} + \frac{1}{2}\log\left(1 - \alpha^2 - \beta^2 + \frac{\eta}{2} \right) \right] \right\}$$

$$\leq 2\exp\left\{ (n-1)\left[\frac{1}{2}\log\left(\frac{1 - \alpha^2 - \beta^2 + \eta/2}{1 - \alpha^2 - \beta^2 + 3\eta/4} \right) + \frac{3\epsilon}{4} \right] \right\}, \quad \text{by (A2)}$$

$$\leq 2\exp\left\{ (n-1)\left[\frac{-\eta/2}{4 + 3\eta} + \frac{3\epsilon}{4} \right] \right\}, \quad \text{using } \log x \leq x - 1$$

$$\leq 2\exp\left\{ -(n-1)\frac{5\epsilon}{4} \right\}, \quad \text{by the hypothesis } \eta > 8\sqrt{\epsilon}$$

$$\leq \frac{1}{4}\exp(-n\epsilon)$$

for large n.

b) If $R \leq -\frac{1}{2}\log(1-\beta^2)$, the bound in (A14) is

$$= 2\exp\left\{ (n-1)\left[\frac{3\epsilon}{4} + \frac{1}{2}\log\left(1 - \frac{\left(\alpha - \frac{\eta}{4}\right)^2}{1-\beta^2} \right) \right] \right\}$$

$$\leq 2\exp\left\{ (n-1)\left[\frac{3\epsilon}{4} + \frac{1}{2}\log\left(1 - \frac{\eta^2}{16} \right) \right] \right\}, \quad \text{by (A2)}$$

$$\leq \frac{1}{4}\exp(-n\epsilon)$$

for large n, using the condition of the lemma that $\eta > 8\sqrt{\epsilon}$.

Thus, in both cases, the term in (A14) is less than $\frac{1}{4}\exp(-n\epsilon)$. This, together with (A13), establishes (A12).

Then from (A6) and (A11), we obtain that for large n

$$\begin{aligned}\Pr\left\{\frac{1}{N}\left|\{i: |(\mathbf{Z}_j,\mathbf{Z}_i)| \geq \alpha,\ |(\mathbf{Z}_j,\boldsymbol{u})| \geq \beta, \text{ for some } j<i\}\right|\right.\\ \left. > \exp(-n\epsilon)\right\}\\ \leq \exp\left(-\frac{1}{2}\exp\left(\frac{n\epsilon}{2}\right)\right)\\ +\exp\left\{-N\left[\exp(-n\epsilon)\log 2-\frac{1}{2}\exp(-n\epsilon)\right]\right\}\\ \leq 2\exp\left(-\frac{1}{10}\exp\left(\frac{n\epsilon}{2}\right)\right),\end{aligned}$$

where we use $N=\exp(nR)>\exp(2n\epsilon)$.

By symmetry, the same bound holds if "for some $j<i$" is replaced by "for some $j>i$," thereby validating the claim in (A3).

The doubly exponential bounds in (A1), (A3) imply that for any finite set $\mathcal{S}_n$ of unit vectors in $\mathbb{R}^n$ with $|\mathcal{S}_n|$ increasing exponentially in n, and any finite subsets A and B of $[0,1]$, the probability of the joint occurrence of the events

$$\left|\{j: |(\mathbf{Z}_j,\boldsymbol{u})| \geq \alpha\}\right| \leq \exp\left\{n\left(\left|R+\frac{1}{2}\log(1-\alpha^2)\right|^{+}+\frac{\epsilon}{2}\right)\right\}$$

for all $\boldsymbol{u}\in\mathcal{S}^n$ and $\alpha\in A$, and

$$\frac{1}{N}\left|\{i: |(\mathbf{Z}_j,\mathbf{Z}_i)| \geq \alpha,\ |(\mathbf{Z}_j,\boldsymbol{u})| \geq \beta, \text{ for some } j\neq i\}\right| \leq \exp(-n\epsilon)$$

for all $\boldsymbol{u}\in\mathcal{S}_n$ and $\alpha\in A$, $\beta\in B$ satisfying (A2) will be arbitrarily close to 1. We complete the proof of the lemma by observing that for an appropriate choice of $\mathcal{S}_n$, A and B as above, if $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_N$ are unit vectors satisfying

$$\left|\{j: |(\boldsymbol{x}_j,\boldsymbol{u})| \geq \alpha\}\right| \leq \exp\left\{n\Bigg(\left|R+\frac{1}{2}\log(1-\alpha^2)\right|^{-}+\frac{\epsilon}{2}\right\} \tag{A15}$$

for all $\boldsymbol{u}\in\mathcal{S}^n$ and $\alpha\in A$, and

$$\frac{1}{N}\left|\{i: |(\boldsymbol{x}_j,\boldsymbol{x}_i)| \geq \alpha,\ |(\boldsymbol{x}_j,\boldsymbol{u})| \geq \beta, \text{ for some } j\neq i\}\right| \leq \exp(-n\epsilon), \tag{A16}$$

for all $\boldsymbol{u}\in\mathcal{S}^n$ and $\alpha\in A$, $\beta\in B$ satisfying (A2), then $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_N$ will satisfy (A15) with $\epsilon/2$ replaced by ϵ for every unit vector $\boldsymbol{u}\in\mathbb{R}^n$ and every $\alpha\in[0,1]$, and satisfy (A16) for every unit vector $\boldsymbol{u}\in\mathbb{R}^n$ and every $\alpha\in[0,1]$, $\beta\in[0,1]$ satisfying $\alpha\geq\eta$, $\alpha^2+\beta^2>1+\eta-\exp(-2R)$. Indeed, it suffices to choose $\mathcal{S}^n$ as a ν-dense subset of the unit sphere in $\mathbb{R}^n$, and A and B as ν-dense subsets of $[0,1]$, with $\nu>0$ sufficiently small.

References

[1] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels," *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, vol. 44, pp. 159–175, 1978.

[2] ——, "A method of coding and an application to arbitrarily varying channels," *J. Combinat., Inform. Syst. Sci.*, vol. 5, pp. 10–35, 1980.

[3] ——, "The capacity of a channel with arbitrarily varying additive Gaussian channel probability functions," *Trans. Sixth Prague Conf. Inform. Theory, Statistical Decision Functions, Random Processes*, pp. 13–21, Sept. 1971.

[4] N. M. Blachman, "On the capacity of a band-limited channel perturbed by statistically dependent interference," *IRE Trans. Inform. Theory*, vol. IT-8, no. 1, pp. 48–55, Jan. 1962.

[5] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacities of certain channel classes under random coding," *Ann. Math. Statist.*, vol. 31, pp. 558–567, 1960.

[6] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic Press, 1981.

[7] ——, "On the capacity of the arbitrarily varying channel for maximum probability of error," *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, vol. 57, pp. 87–101, 1981.

[8] I. Csiszár and P. Narayan, "The capacity of the arbitrarily varying channel revisited: Positivity, constraints," *IEEE Trans. Inform. Theory*, vol. 34, no. 2, pp. 181–193, Mar. 1988.

[9] ——, "Capacity and decoding rules for classes of arbitrarily varying channels," *IEEE Trans. Inform. Theory*, vol. 35, no. 4, pp. 752–769, July 1989.

[10] B. Hughes and P. Narayan, "Gaussian arbitrarily varying channels," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 267–284, Mar. 1987.

[11] R. J. McEliece, "Communication in the presence of jamming—An information theory approach," in *Secure Digital Communications, CISM Courses and Lectures*, No. 279, G. Longo, Ed. New York: Springer Verlag, 1983.

[12] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, no. 3, pp. 611–656, May 1959.