On AVCs with Quadratic Constraints

Farzin Haddadpour, Mahdi Jafari Siavoshani, Mayank Bakshi, Sidharth Jaggi

***Abstract*—"THIS PAPER IS ELIGIBLE FOR THE STUDENT PAPER AWARD". In this work we study an Arbitrarily Varying Channel (AVC) with quadratic power constraints on the transmitter and a so-called "oblivious" jammer (along with additional AWGN) under a *maximum probability of error* criterion, and no private randomness between the transmitter and the receiver. This is in contrast to similar AVC models under the *average probability of error* criterion considered in [1], and models wherein common randomness is allowed [2] – these distinctions are important in some communication scenarios outlined below.**

We consider the regime where the jammer's power constraint is smaller than the transmitter's power constraint (in the other regime it is known no positive rate is possible). For this regime we show the existence of stochastic codes (with *no common randomness* between the transmitter and receiver) that enables reliable communication at the same rate as when the jammer is replaced with AWGN with the same power constraint. This matches known information-theoretic outer bounds. In addition to being a stronger result than that in [1] (enabling recovery of the results therein), our proof techniques are also somewhat more direct, and hence may be of independent interest.

I. Introduction

Aerial Alice is flying in a surveillance plane high over Hostile Harry's territory. She wishes to relay her observations of Harry's troop movements back to Base-station Bob over n channel uses of an AWGN channel with variance σ^2. Harry obviously wishes to jam Alice's transmissions. However, both Alice's transmission energy and Harry's jamming energy are constrained – they have access to energy sources of nP and $n\Lambda$ Joules respectively.[1] Harry already knows *what message* Alice wants to transmit (after all, he knows the movements of his own troops), and also *roughly how* she'll transmit it (*i.e.*, her *communication protocol/code*, having recently captured another surveillance drone) but he doesn't know *exactly how* she'll transmit it (*i.e.*, her *codeword* – for instance, Alice could choose to focus her transmit power on some random subset of the n channel uses). Further, since Alice's transmissions are very quick, Harry has no time to tune his jamming strategy to Alice's actual codeword – he can only jam based on his prior knowledge of Alice's code, and her message.

Even in such an adverse jamming setting we demonstrate that Alice can communicate with Bob at a rate equalling $\frac{1}{2}\log\left(1+\frac{P}{\Lambda+\sigma^2}\right)$ as long as $P > \Lambda$. Note that this equals the capacity of an AWGN with noise parameter equal to $\Lambda + \sigma^2$ – this means that no "smarter" jamming strategy exists for Harry than simply behaving like AWGN with variance Λ. If $P < \Lambda$ no positive rate is possible since Harry can "spoof" by transmitting a fake message using the same strategy as Alice – Bob is unable to distinguish between the real and fake transmissions[2].

A. Relationship with prior work

The model considered in this work is essentially a special type of Arbitrarily Varying Channel (AVC) for which, to the best of our knowledge, the capacity has not been characterized before in the literature. The notion of AVCs was first introduced by Blackwell *et al.* [6], to capture communication models wherein channel have unknown parameters that may vary arbitrarily during the transmission of a codeword. The case when both the transmitter and the jammer operate under constraints (analogous to the quadratic constraints in this work) has also been considered [3]. For an extensive survey on AVCs the reader may refer to the excellent survey [5] and the references therein.

The class of AVCs over discrete alphabets has been studied in great detail in the literature [5]. However, less is known about AVCs with continuous alphabets. The bulk of the work on continuous alphabet AVCs (outlined below in this section) focuses on quadratically-constrained AVCs. This is also the focus of our work.

It is important to stress several features of the model considered in this work, and the differences with prior work:

- Stochastic encoding: To generate her codeword from her message, Alice is allowed to use private randomness (known only to her *a priori*, but not to Harry *or* Bob. This is in contrast to the *deterministic encoding* strategies often considered in the information theory/coding theory literature, wherein the codeword is a deterministic function of the message.
- Public code: Everything Bob knows about Alice's transmission *a priori*, Harry also knows.[3] This is in contrast to the *randomized encoding* model also considered in the literature (see for instance [2], [8]), in which it is critical that Alice and Bob share *common randomness* that is unknown to Harry.
- Message-aware jamming: The jammer already knows Alice's message. This is one important difference between our model, and that of [1].
- Oblivious adversary: The jammer has no extra knowledge of the codeword being transmitted than what he has already gleaned from his knowledge of Alice's code and her

[1] These are so-called *peak power* constraints – they must hold for *all* codewords, rather than averaged over all codewords *average power* constraints. If the peak power constraints are relaxed to average power constraints, for either Alice's transmissions, or Harry's jamming (or both), it is known [2] that standard capacity results do not hold – only "λ-capacities" exist.

[2] Such a jamming strategy is equivalent to the more general *symmetrizability* condition in the AVC literature (see, for instance [3], [4], and [5]).

[3] This requirement is an analogue for communication of Kerckhoffs' Principle [7] in cryptography, which states that in a secure system, everything about the system is public knowledge, except possibly Alice's *private* randomness.

message. This is in contrast to the *omniscient adversary* often considered in the coding theory literature.

These model assumptions are equivalent to requiring public stochastic codes with small maximum error of probability against an oblivious adversary. Several papers also operate under *some* of these assumptions, but as far as we know, none examines the scenario where *all* these constraints are active.

The literature on *sphere packing* focuses on an AVC model wherein zero-error probability of decoding is required (or, equivalently, when the probability (over Alice's codeword and Harry's jamming actions) of Bob's decoding error is required to equal zero). Inner and outer bounds were obtained by Blachman [9]. Like several other zero-error communication problems (including Shannon's classic work [10]) characterization of the optimal throughput possible is challenging, and in general still an open problem.[4]

Other related models include:

- The *vector Gaussian AVC* [13]. As in the "usual" vector Gaussian channels, optimal code designs require "water-filling".
- The *per-sequence/universal* coding schemes in [14].
- The *correlated/myopic* jammers in [15], [16], wherein jammers obtain a noisy version of Alice's transmission and base their jamming strategy on this.
- The *joint source-channel coding, and coding with feedback* models considered by Başar [17], [18].
- Several other AVC variants, including *dirty paper coding*, in [19].

We summarize some of the above in Table I.

II. Notation and Problem Statement

Throughout the paper, we use capital letters to denote random variables and random vectors, and corresponding lowercase letters to denote their realizations. Moreover, bold letters are reserved for vectors and calligraphic symbols denote sets. Random sets are represented by an extra star as superscripts. Some constants are also denoted by capital letters.

We use $N(a,\sigma^2)$ to denote for a Gaussian random variable with mean a and variance σ^2. To denote a ball in an n-dimensional real space of radius r which centered at the point $\boldsymbol{c} \in \mathbb{R}^n$, we write $\mathcal{B}_n(\boldsymbol{c},r)$. In Table II, we summarize the notation used in this paper.

A. Problem Statement

In this paper we study the capacity of a quadratic constrained AVC with stochastic encoder under the attack of a malicious adversary who knows the transmitted message but is oblivious to the actual transmitted codewords.

Let the input and output of the channel are denoted by the random variables X and Y where $X, Y \in \mathbb{R}$. Then, formally, the channel is defined as follows

$$Y = X + S + V, \tag{1}$$

Symbol	Meaning
$\Psi(i)$	Stochastic encoder applied to the message i
$\phi(Y)$	Deterministic decoder
$e(\boldsymbol{s}, i)$	Error probability (over the stochastic encoder and the channel noise) for a fixed message i and jamming vector $\boldsymbol{s}$
$e_{\max}(\boldsymbol{s})$	Maximum (over messages) error probability for a fixed jamming vector $\boldsymbol{s}$
$N(a,\sigma^2)$	Gaussian random variable with mean a and variance σ^2
$\mathcal{B}_n(\boldsymbol{c}, r)$	A ball of radius r in $\mathbb{R}^n$ which centered at $\boldsymbol{c} \in \mathbb{R}^n$

Table II
Commonly used symbols.

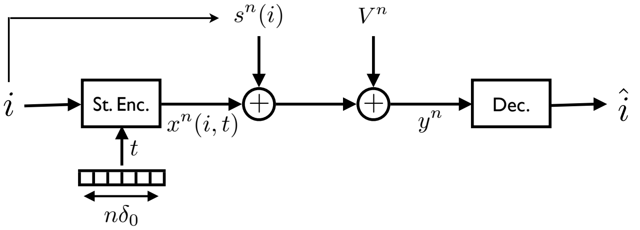

Figure 1. A power constraint AVC with stochastic encoder. Here we assume that the adversary has access to the transmitted message i but not to the transmitted codeword $x^n(i,t)$.

where $S \in \mathbb{R}$ is the channel state chosen by a malicious adversary and $V \sim N(0,\sigma^2)$ is Gaussian random variable. Here we assume that the noise V is independent over different uses of channel (1). The channel input is subjected to a peak power constraint as follows

$$\|\boldsymbol{x}\|^2 = \sum_{i=1}^{n} x_i^2 \leq nP, \tag{2}$$

and the permissible state sequences are those satisfying

$$\|\boldsymbol{s}\|^2 = \sum_{i=1}^{n} s_i^2 \leq n\Lambda. \tag{3}$$

The problem setup is depicted pictorially in Figure 1.

A *code with stochastic encoder* (Ψ, ϕ) of block-length n consists of a set of encoders that are denoted by a random variable $\Psi : \ \{1, \ldots, M\} \mapsto \mathbb{R}^n$ and a deterministic decoder $\phi : \ \mathbb{R}^n \mapsto \{0, \ldots, M\}$ where 0 denote for an error and $M \triangleq e^{nR}$ is the number of messages[5]. Each encoder ψ is constructed by a set of codewords $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\}$ from $\mathbb{R}^n$.

In this paper we focus on the *maximum probability of error*. For a fixed jamming vector $\boldsymbol{s}$, let us define the probability of error given that the message i has been sent as follows

$$e(\boldsymbol{s}, i) \triangleq \mathbb{P}_{\Psi,V}\left[\phi\left(\Psi(i) + \boldsymbol{s} + \boldsymbol{V}\right) \neq i\right]. \tag{4}$$

The maximum probability of error for a fixed $\boldsymbol{s}$ is defined as

$$e_{\max}(\boldsymbol{s}) \triangleq \max_{i \in \{1,\ldots,M\}} e(\boldsymbol{s}, i). \tag{5}$$

The *capacity* for this channel is now defined as:

Definition 1. *The capacity C of an AVC with stochastic encoder under the quadratic transmit constraint P and jamming constraint Λ is the supremum over the set of real numbers*

[4] The literature on *Spherical Codes* (see [11] and [12] for some relatively recent work) looks at the related problem of packing unit hyperspheres on the *surface* of a hypershere. This corresponds to design of codes where each codeword meets the quadratic power constraint with equality, rather than allowing for an inequality.

[5] For notational convenience we assume that e^{nR} is an integer.

	Error Criterion	**Capacity**
Blachman [9]	$\sup_{\boldsymbol{s}} \sup_{\psi} \left[\phi(\psi(i)+\boldsymbol{s}+\boldsymbol{V}) \neq i\right] = 0$	upper and lower bounds for the capacity
Hughes & Narayan [2]	$\sup_{\boldsymbol{s}} \max_i \mathbb{P}_{(\Phi,\Psi),V} \left[\Phi(\Psi(i)+\boldsymbol{s}+\boldsymbol{V}) \neq i\right] \leq \epsilon$	$C = \frac{1}{2}\log(1+\frac{P}{\Lambda+\sigma^2})$
Csiszar & Narayan [1]	$\sup_{\boldsymbol{s}} \frac{1}{M}\sum_{i=1}^{M} \mathbb{P}_V \left[\phi(\psi(i)+\boldsymbol{s}+\boldsymbol{V}) \neq i\right] \leq \epsilon$	$C = \begin{cases} \frac{1}{2}\log(1+\frac{P}{\Lambda+\sigma^2}) & \text{if } P > \Lambda \\ 0 & \text{Otherwise.} \end{cases}$
Our Setup	$\sup_{\boldsymbol{s}} \max_i \mathbb{P}_{\Psi,V} \left[\phi(\Psi(i)+\boldsymbol{s}+\boldsymbol{V}) \neq i\right] \leq \epsilon$	$C = \begin{cases} \frac{1}{2}\log(1+\frac{P}{\Lambda+\sigma^2}) & \text{if } P > \Lambda \\ 0 & \text{Otherwise.} \end{cases}$

Table I
Comparison of existing results on Quadratic-constrained AVCs with AWGN.

such that for every $\delta > 0$ and sufficiently large n there exist codes with stochastic encoder (Ψ, ϕ) that satisfies the following conditions. First, for the number of messages M encoded by the code we have $M > \exp(n(C-\delta))$. Moreover, each codeword satisfies the quadratic constraint (2) *and finally for the code we have $\lim_{n\to\infty} \sup_{\boldsymbol{s}:\|\boldsymbol{s}\|^2 \leq n\Lambda} e_{\max}(\boldsymbol{s}) = 0$.*

III. Main Results

Theorem 1. *The capacity of a quadratic-constrained AVC channel under the maximum probability of error criterion with transmit constraint P and jamming constraint Λ and additive Gaussian noise of power σ^2 is given by*

$$C = \begin{cases} \frac{1}{2}\log(1+\frac{P}{\Lambda+\sigma^2}) & \text{if } P > \Lambda, \\ 0 & \text{Otherwise.} \end{cases}$$

Remark 1. *The result of Theorem 1 matches the result of stochastic encoder over discrete alphabets [20], [5, Theorem 7], in which it is shown that for the* average *probability of error criterion, using a stochastic encoder doesn't increase the capacity. Because the number of possible adversarial actions here is uncountably large, the technique of [20], which relies on taking a union bound over at most exponential-sized set of possible adversarial actions, does not work.*

Corollary 1. *The capacity of a quadratic-constrained AVC under the maximum probability of error criterion with transmit constraint P and jamming constraint Λ is given by*

$$C = \begin{cases} \frac{1}{2}\log(1+\frac{P}{\Lambda}) & \text{if } P > \Lambda, \\ 0 & \text{Otherwise.} \end{cases}$$

IV. Proof of Main Results

In this section, we present the proof of Theorem 1 and its corollary. Due to space constrain the proof of the converse parts of Theorem 1 is relegated to [21, Section IV-B].

For the achievability part of Theorem 1, we claim that the same *minimum distance decoder* proposed in [1] to achieve the capacity for the average probability of error criterion, which is given by

$$\phi(\boldsymbol{y}) = \begin{cases} i & \text{if } \|\boldsymbol{y}-\boldsymbol{x}_i\|^2 < \|\boldsymbol{y}-\boldsymbol{x}_j\|^2, \quad \text{for } j \neq i, \\ 0 & \text{if no such } i: 1 \leq i \leq M \text{ exists,} \end{cases} \tag{6}$$

also achieves the capacity for the maximum probability of error criterion.

Note that in order to show the suprimum over $\boldsymbol{s}$ subject to (3) of $e_{\max}(\boldsymbol{s})$ goes to zero it is sufficient to show that for every message i the suprimum over $\boldsymbol{s}$ subject to (3) of $e(\boldsymbol{s}, i)$ goes to zero.

To communicate, Alice (the transmitter) randomly picks a codebook $\mathcal{C}$ and fixes it. The codebook $\mathcal{C}$ comprises $e^{n(\delta_0+R)}$ codewords $\boldsymbol{x}(i,t)$, $1 \leq i \leq e^{nR}$ and $1 \leq t \leq e^{n\delta_0}$, each chosen uniformly at random and independently from a sphere of radius $\sqrt{nP}$. Then, the ith row of the codebook, i.e., $\{\boldsymbol{x}(i,1), \ldots, \boldsymbol{x}(i, e^{n\delta_0})\}$, is assigned to the ith message. In order to transmit the message i, the encoder randomly picks a codeword from the ith row of the codebook and sends it over the channel.

Now, given that the message i has been transmitted, the error probability $e(\boldsymbol{s}, i)$ of an stochastic code used over a quadratic-constrained AVC under the use of the minimum distance decoder (defined by (6)) equals

$$\begin{aligned} e(\boldsymbol{s}, i) =& \mathbb{P}_{\Psi,V}\left[\phi\left(\Psi(i)+\boldsymbol{s}+\boldsymbol{V}\right) \neq i\right] \\ =& \mathbb{P}_T\mathbb{P}_V\Big[\|\boldsymbol{x}(i,T)+\boldsymbol{s}+\boldsymbol{V}-\boldsymbol{x}(j,t')\|^2 \\ & \quad \leq \|\boldsymbol{s}+\boldsymbol{V}\|^2 \text{ for some } i \neq j \text{ and } t'\Big] \\ =& \mathbb{P}_T\mathbb{P}_V\Big[\langle \boldsymbol{x}(j,t'), \boldsymbol{x}(i,T)+\boldsymbol{s}+\boldsymbol{V}\rangle \geq nP \\ & \quad + \langle \boldsymbol{x}(i,T), \boldsymbol{s}+\boldsymbol{V}\rangle \text{ for some } j \neq i \text{ and } t'\Big]. \end{aligned} \tag{7}$$

where T is a uniformly distributed random variable defined over the set $\{1, \ldots, e^{n\delta_0}\}$.

A. Achievability proof of Theorem 1

The main step in proving the achievability part of Theorem 1 consists in asserting the doubly exponential probability bounds which is stated in Lemma 1.

Lemma 1. *Let $\mathcal{C}^* = \{\boldsymbol{X}(i,t)\}$ in which $1 \leq i \leq \exp(nR)$ and $1 \leq t \leq \exp(n\delta_0)$ be a random codebook comprises of independent random vectors $\boldsymbol{X}(i,t)$ each uniformly distributed on the n-dimensional sphere of radius $\sqrt{nP}$. First, fix a vector $\boldsymbol{s} \in \mathcal{B}_n(0, \sqrt{n\Lambda})$. Then for every $\delta_0 > \delta_1 > 0$ and for sufficiently large n if $R < \frac{1}{2}\log\left(1+\frac{P}{\sigma^2+\Lambda}\right)$ we have*

$$\begin{aligned} &\mathbb{P}_{\mathcal{C}^*}\Bigg[\mathbb{P}_T\mathbb{P}_V\Big[\langle \boldsymbol{X}(j,t'), \boldsymbol{X}(i,T)+\boldsymbol{s}+\boldsymbol{V}\rangle \geq nP \\ &\quad + \langle \boldsymbol{X}(i,T), \boldsymbol{s}+\boldsymbol{V}\rangle \text{ for some j} \neq \text{i and } t'\Big] \geq Ke^{-n\delta_1}\Bigg] \\ &\leq \exp\Big(-(K\log 2 - 10)\exp((\delta_0-\delta_1)n)\Big). \end{aligned}$$

Proof sketch of Lemma 1: For the complete proof of Lemma 1 refer to [21].

To derive the doubly exponential bound stated in the lemma, we use Lemma 2. To this end let us define the functions f_t for $1 \leq t \leq e^{n\delta_0}$ as follows

$$f_t\left(\boldsymbol{X}(i,1),\ldots,\boldsymbol{X}(i,t)\right) \triangleq \mathbb{P}_V\Big[\langle \boldsymbol{X}(j,t'), \boldsymbol{X}(i,t)+\boldsymbol{s}+\boldsymbol{V}\rangle \\ \geq nP + \langle \boldsymbol{X}(i,t), \boldsymbol{s}+\boldsymbol{V}\rangle \text{ for some } j \neq i \text{ and } t'\Big].$$

Now, by using the functions f_t, the probability expression in the statement of lemma can be written as follows

$$\mathbb{P}_{\mathcal{C}^*}\left[\frac{1}{e^{n\delta_0}}\sum_{t=1}^{e^{n\delta_0}} f_t\left(\boldsymbol{X}(i,1),\ldots,\boldsymbol{X}(i,t)\right) \geq Ke^{-n\delta_1}\right]. \quad (8)$$

In order to bound (8) we use Lemma 2 where we have to bound the expected values of the functions f_t. To this end we can first show that the vectors $\boldsymbol{s}$, $\boldsymbol{V}$, and $\boldsymbol{X}(i,t)$ are almost orthogonal with high probability. By properly bounding the terms in the expectation and using the above orthogonality conditions we can finally derive the following bound for the expectation of functions f_t

$$\mathbb{E}_{\mathcal{C}^*}\left[f_t(\boldsymbol{X}(i,1),\ldots,\boldsymbol{X}(i,t))|\boldsymbol{X}(i,1),\ldots,\boldsymbol{X}(i,t-1)\right] \\ \leq 2\exp\left(-\frac{n-1}{2}\frac{{\delta_2}^2}{\|\boldsymbol{s}\|^2+\sigma^2+\delta_2}\right) + 2e^{-n\xi} + e^{-\frac{n\eta^2}{2\sigma^2\Lambda}} \\ + 2e^{n(R+\delta_0)+\frac{n-1}{2}\log\left(1-\frac{P-\delta_2'}{P+\Lambda+\sigma^2-\delta_2}\right)},$$

where $\xi = \frac{1}{2}\left[1+\frac{\delta_2-2\eta}{\sigma^2}-\sqrt{1+2\frac{\delta_2-2\eta}{\sigma^2}}\right]$, $\delta_2' = 2\sqrt{P}\delta_2-\delta_2^2$, and $\delta_2 > 2\eta > 0$. Now, as it is shown in [21] in more details, by introducing δ_1 and imposing the conditions $\delta_2 \leq \|\boldsymbol{s}\|^2+\sigma^2$, $\delta_2 > 2\eta+4\sigma^2\sqrt{\delta_1}$, $\eta > \sqrt{2\Lambda\sigma^2\delta_1}$, $\delta_2 > \sqrt{\frac{4P(\Lambda+\sigma^2)\delta_1}{1-1/n}}$, and by choosing

$$R < \frac{1-1/n}{2}\log\left(1+\frac{P-\delta_2'}{\Lambda+\sigma^2-\delta_2+\delta_2'}\right)-\delta_0-\delta_1,$$

we can show that

$$\mathbb{E}_{\mathcal{C}^*}\left[f_t(\boldsymbol{X}(i,1),\ldots,\boldsymbol{X}(i,t))|\boldsymbol{X}(i,1),\ldots,\boldsymbol{X}(i,t-1)\right] \\ \leq 10\exp\left(-n\delta_1\right).$$

Then by applying Lemma 2 and choosing $a = 10e^{-n\delta_1}$ and $\tau = Ke^{-n\delta_1}$ we have

$$\mathbb{P}_{\mathcal{C}^*}\Bigg[\frac{1}{e^{n\delta_0}}\sum_{t=1}^{e^{n\delta_0}}\mathbb{P}_V\big[\langle \boldsymbol{X}(j,t'),\boldsymbol{X}(i,t)+\boldsymbol{s}+\boldsymbol{V}\rangle \geq nP \\ + \langle \boldsymbol{X}(i,t),\boldsymbol{s}+\boldsymbol{V}\rangle \text{ for some j} \neq \text{i and } t'\big]\Bigg] \geq Ke^{-n\delta_1}\Bigg] \\ \leq \exp\Big(-\exp(n\delta_0)\Big(K\log 2\exp(-n\delta_1)-10\exp(-n\delta_1)\Big)\Big) \\ = \exp\Big(-(K\log 2-10)\exp(n(\delta_0-\delta_1))\Big).$$

By assuming $\delta_0 > \delta_1 > 0$ we obtain the desired doubly exponential bound, hence we are done. ■

Our proof requires the following "martingale concentration lemma" proven in [1, Lemma A1].

Lemma 2 ([1, Lemma A1])**.** *Let $X_1,\ldots,X_L$ be arbitrary r.v.'s and $f_i(X_1,\ldots,X_L)$ be arbitrary function with $0 \leq f_i \leq 1$, $i = 1,\ldots,L$. Then if the conditions $\mathbb{E}\left[f_i(X_1,\ldots,X_L)|X_1,\ldots,X_{i-1}\right] \leq a$ hold almost surely for $i = 1,\ldots,L$, we have*

$$\mathbb{P}\left[\frac{1}{L}\sum_{i=1}^{L} f_i(X_1,\ldots,X_i) > \tau\right] \leq \exp\left(-L(\tau\log 2-a)\right).$$

Lemma 3 (Quantizing Adversarial Vector)**.** *For a fixed jamming vector $\boldsymbol{s}$, for sufficiently small $\varepsilon > 0$, and for every $\delta_0 > \delta_1 > 0$, there exists a codebook $\mathcal{C} = \{\boldsymbol{x}(i,t)\}$ of rate $R \leq \frac{1}{2}\log(1+\frac{P}{\Lambda+\sigma^2})$ comprises of vectors $\boldsymbol{x}(i,t) \in \mathbb{R}^n$ of size $\sqrt{nP}$ with $1 \leq i \leq e^{nR}$ and $1 \leq t \leq e^{n\delta_0}$ which performs well over the AVC defined in Section II for all $\boldsymbol{s}' \in \mathcal{B}_n(\boldsymbol{s},\varepsilon)$, i.e., it satisfies*

$$e(\boldsymbol{s},i) = \mathbb{P}_T\mathbb{P}_V\Big[\langle \boldsymbol{x}(j,t'),\boldsymbol{x}(i,T)+\boldsymbol{s}+\boldsymbol{V}\rangle \\ \geq nP + \langle \boldsymbol{x}(i,T),\boldsymbol{s}+\boldsymbol{V}\rangle \text{ for some } j \neq i \text{ and } t'\Big] \\ < K\exp(-n\delta_1) \quad (9)$$

for all $\boldsymbol{s}' \in \mathcal{B}_n(\boldsymbol{s},\varepsilon)$.

Proof: For a particular $\boldsymbol{s}$, instead of (9), let us assume that the code $\mathcal{C}$ satisfies a stronger condition

$$\mathbb{P}_T\mathbb{P}_V\Big[\langle \boldsymbol{x}(j,t'),\boldsymbol{x}(i,T)+\boldsymbol{s}+\boldsymbol{V}\rangle \geq nP \\ - 2\varepsilon\sqrt{nP} + \langle \boldsymbol{x}(i,T),\boldsymbol{s}+\boldsymbol{V}\rangle \text{ for some } j \neq i \text{ and } t'\Big] \\ < K\exp(-n\delta_1). \quad (10)$$

Then it can be verified that for all $\boldsymbol{s}' \in \mathcal{B}_n(\boldsymbol{s},\varepsilon)$ the code $\mathcal{C}$ satisfies (9) where $\boldsymbol{s}$ is replaced by $\boldsymbol{s}'$. To show this let $\boldsymbol{s}' = \boldsymbol{s}+\rho\boldsymbol{u}$ where $\boldsymbol{u}$ is an arbitrary unit vector and $\rho \in [-\varepsilon,\varepsilon]$. Hence for all $\boldsymbol{s}' \in \mathcal{B}_n(\boldsymbol{s},\varepsilon)$ we can write

$$e(\boldsymbol{s}',i) = \mathbb{P}_T\mathbb{P}_V\Big[\langle \boldsymbol{x}(j,t'),\boldsymbol{x}(i,T)+\boldsymbol{s}+\boldsymbol{V}\rangle + \rho\langle \boldsymbol{x}(j,t'),\boldsymbol{u}\rangle \\ \geq nP + \langle \boldsymbol{x}(i,T),\boldsymbol{s}+\boldsymbol{V}\rangle \\ + \rho\langle \boldsymbol{x}(i,T),\boldsymbol{u}\rangle \text{ for some } j \neq i \text{ and } t'\Big] \\ \leq \mathbb{P}_T\mathbb{P}_V\Big[\langle \boldsymbol{x}(j,t'),\boldsymbol{x}(i,T)+\boldsymbol{s}+\boldsymbol{V}\rangle + \varepsilon\sqrt{nP} \\ \geq nP + \langle \boldsymbol{x}(i,T),\boldsymbol{s}+\boldsymbol{V}\rangle \\ - \varepsilon\sqrt{nP} \text{ for some } j \neq i \text{ and } t'\Big] \\ \overset{(a)}{\leq} K\exp(-n\delta_1),$$

where (a) follows from (10).

Now, in Lemma 1 we can use the stronger error requirement (10) to show that there exists a code which satisfies (10). This stronger requirement results in a rate loss, but as ε goes to zero the rate loss due to that vanishes. By the above argument, we know that this code satisfies (9) for all $\boldsymbol{s}' \in \mathcal{B}_n(\boldsymbol{s},\varepsilon)$ and we are done. ■

Finally, Lemma 4 shows the existence of a good codebook for the quadratic constrained AVC problem with stochastic encoder which have been introduced in Section II-A and hence completes the proof of Theorem 1.

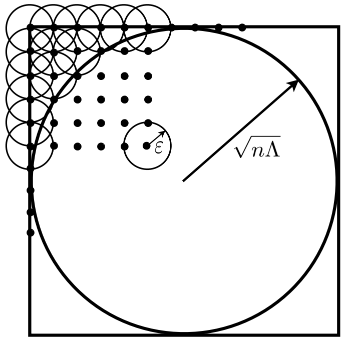

Figure 2. This figure shows that how the whole sphere $\mathcal{B}(0, \sqrt{n\Lambda})$ can be covered by ε-dense subsets χ_n. Here the set χ_n comprises of points from a hyper-cubical lattice.

Lemma 4 (Codebook Existence)**.** *For every $\delta_0 > \delta_1 > 0$ and $n \geq n_0(\delta_0, \delta_1)$ there exist a codebook $\mathcal{C} = \{\boldsymbol{x}(i,t)\}$ of rate $R \leq \frac{1}{2}\log(1 + \frac{P}{\sigma^2+\Lambda})$ comprises of vectors $\boldsymbol{x}(i,t) \in \mathbb{R}^n$ of size $\sqrt{nP}$ with $1 \leq i \leq e^{nR}$ and $1 \leq t \leq e^{n\delta_0}$ such that for every vector $\boldsymbol{s}$ and every transmitted message i we have*

$$\begin{aligned} e(\boldsymbol{s}, i) = \mathbb{P}_T\mathbb{P}_V \Big[&\langle \boldsymbol{x}(j,t'), \boldsymbol{x}(i,T) + \boldsymbol{s} + \boldsymbol{V}\rangle \\ &\geq nP + \langle \boldsymbol{x}(i,T), \boldsymbol{s} + \boldsymbol{V}\rangle \text{ for some } j \neq i \text{ and } t'\Big] \\ &< K\exp(-n\delta_1). \end{aligned} \tag{11}$$

Proof: For any fixed codebook $\mathcal{C} = \{\boldsymbol{x}(i,t)\}$, let us explicitly mention to the dependency of the error probability on $\mathcal{C}$ by defining $e_{\mathcal{C}}(\boldsymbol{s}, i) \triangleq e(\boldsymbol{s}, i)$. Then in order to prove the assertion of lemma we can equivalently show that

$$\liminf_{n\to\infty} \mathbb{P}_{\mathcal{C}^*}\left[\forall \boldsymbol{s}, \forall i \;\; e_{\mathcal{C}^*}(\boldsymbol{s}, i) < Ke^{-n\delta_1}\right] > 0.$$

However, by using Lemma 3, it is not necessary to check for all $\boldsymbol{s}$ but only for those belonging to an ε-net[6] χ_n that covers $\mathcal{B}_n(0, \sqrt{n\Lambda})$. Hence, we can write

$$\begin{aligned} \mathbb{P}_{\mathcal{C}^*}&\left[\forall \boldsymbol{s} \in \chi_n, \forall i \;\; e_{\mathcal{C}^*}(\boldsymbol{s}, i) < Ke^{-n\delta_1}\right] \\ &= 1 - \mathbb{P}_{\mathcal{C}^*}\left[\exists \boldsymbol{s} \in \chi_n, \exists i \;\; e_{\mathcal{C}^*}(\boldsymbol{s}, i) \geq Ke^{-n\delta_1}\right] \\ &\overset{\text{(a)}}{\geq} 1 - \sum_{\boldsymbol{s}\in\chi_n}\sum_{i=1}^{e^{nR}} \mathbb{P}_{\mathcal{C}^*}\left[e_{\mathcal{C}^*}(\boldsymbol{s}, i) \geq Ke^{-n\delta_1}\right], \end{aligned}$$

where (a) follows from the union bound.

Now, note that to bound $|\chi_n|$ one might cover $\mathcal{B}_n(0, \sqrt{n\Lambda})$ by a hypercube of edge size $2\sqrt{n\Lambda}$; see Figure 2. So we can write $|\chi_n| \leq \left(\frac{2\sqrt{n\Lambda}}{\varepsilon}\right)^n$. Then, by using Lemma 1 we have

$$\begin{aligned} \mathbb{P}_{\mathcal{C}^*}&\left[\forall \boldsymbol{s} \in \chi_n, \forall i \;\; e_{\mathcal{C}^*}(\boldsymbol{s}, i) < Ke^{-n\delta_1}\right] \\ &\geq 1 - \left(\frac{2\sqrt{n\Lambda}}{\varepsilon}\right)^n \times e^{nR} \times \exp\left(-K'e^{n(\delta_0-\delta_1)}\right), \end{aligned}$$

where, assuming $\delta_0 > \delta_1$, the right hand side goes to 1 as n goes to infinity and this completes the proof of lemma. ■

[6] An ε-net is a set of points in a metric space such that each point of the space is within distance ε of some point in the set.

V. Acknowledgments

This work is partially supported by a grant from the University Grants Committee of the Hong Kong Special Administrative Region, China (Project No. AoE/E-02/08), the CUHK MoE-Microsoft Key Laboratory for Human-centric Computing & Interface Technologies, and the Project MMT-p7-11 of the Shun Hing Institute of Advanced Engineering, The Chinese University of Hong Kong.

References

[1] I. Csiszár and P. Narayan, "Capacity of the gaussian arbitrarily varying channel," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 18–26, 1991.

[2] B. Hughes and P. Narayan, "Gaussian arbitrarily varying channels," *IEEE Transactions on Information Theory*, vol. 33, no. 2, pp. 267–284, 1987.

[3] I. Csiszar and P. Narayan, "Arbitrarily varying channels with constrained inputs and states," *IEEE Transactions on Information Theory*, vol. 34, no. 1, pp. 27–34, 1988.

[4] ——, "The capacity of the arbitrarily varying channel revisited: Positivity, constraints," *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 181–193, 1988.

[5] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2148–2177, 1998.

[6] D. Blackwell, L. Breiman, and A. Thomasian, "The capacity of a class of channels," *The Annals of Mathematical Statistics*, pp. 1229–1241, 1959.

[7] A. Kerckhoffs, "La cryptographie militaire," *Journal des sciences militaires IX*, vol. 5, no. 38, pp. 161–191, 1883.

[8] M. Agarwal, A. Sahai, and S. Mitter, "Coding into a source: A direct inverse rate-distortion theorem," *arXiv preprint cs/0610142*, 2006.

[9] N. Blachman, "On the capacity of a band-limited channel perturbed by statistically dependent interference," *IRE Transactions on Information Theory*, vol. 8, no. 1, pp. 48–55, 1962.

[10] C. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, 1956.

[11] A. D. Wyner, "Random packings and coverings of the unit n-sphere," *Bell Systems Technical Journal*, vol. 46, pp. 2111–2118, 1967.

[12] J. Hamkins and K. Zeger, "Asymptotically dense spherical codes. i. wrapped spherical codes," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1774–1785, 1997.

[13] B. Hughes and P. Narayan, "The capacity of a vector gaussian arbitrarily varying channel," *IEEE Transactions on Information Theory*, vol. 34, no. 5, pp. 995–1003, 1988.

[14] Y. Lomnitz and M. Feder, "Communication over individual channels," *IEEE Transactions on Information Theory*, vol. 57, no. 11, pp. 7333–7358, 2011.

[15] M. Médard, "Capacity of correlated jamming channels," in *Proceedings of the Annual Allerton Conference on Communications, Control and Computing*, vol. 35, 1997, pp. 1043–1052.

[16] A. Sarwate, "An avc perspective on correlated jamming," in *International Conference on Signal Processing and Communications (SPCOM)*. IEEE, 2012, pp. 1–5.

[17] T. Basar, "The gaussian test channel with an intelligent jammer," *IEEE Transactions on Information Theory*, vol. 29, no. 1, pp. 152–157, 1983.

[18] ——, "Optimum linear causal coding schemes for gaussian stochastic processes in the presence of correlated jamming," *IEEE Transactions on Information Theory*, vol. 35, no. 1, pp. 199–202, 1989.

[19] A. Sarwate and M. Gastpar, "Relaxing the gaussian avc," *arXiv preprint arXiv:1209.2755*, 2012.

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

[21] F. Haddadpour, M. Jafari Siavoshani, M. Bakshi, and S. Jaggi, "On avcs with quadratic constraints," *arXiv preprint*, 2013.