Quadratically Constrained Myopic Adversarial Channels

Yihan Zhang, Shashank Vatedka, *Member, IEEE*, Sidharth Jaggi, *Senior Member, IEEE*, and Anand D. Sarwate, *Senior Member, IEEE*

***Abstract*— We study communication in the presence of a jamming adversary where quadratic power constraints are imposed on the transmitter and the jammer. The jamming signal is allowed to be a function of the codebook, and a noncausal but noisy observation of the transmitted codeword. For a certain range of the noise-to-signal ratios (NSRs) of the transmitter and the jammer, we are able to characterize the capacity of this channel under deterministic encoding or stochastic encoding, i.e., with no common randomness between the encoder/decoder pair. For the remaining NSR regimes, we determine the capacity under the assumption of a small amount of common randomness (at most $2\log(n)$ bits in one sub-regime, and at most $\Omega(n)$ bits in the other sub-regime) available to the encoder-decoder pair. Our proof techniques involve a novel myopic list-decoding result for achievability, and a Plotkin-type push attack for the converse in a subregion of the NSRs, both of which may be of independent interest. We also give bounds on the strong secrecy capacity of this channel assuming that the jammer is simultaneously eavesdropping.**

***Index Terms*— Channel coding, communication channels, channel capacity, channel state information, Gaussian channels, multiuser channels, time-varying channels, communication system security.**

Manuscript received 5 January 2020; revised 4 May 2021; accepted 24 October 2021. Date of publication 14 April 2022; date of current version 13 July 2022. This work was supported in part by the University Grants Committee of the Hong Kong Special Administrative Region under Project AoE/E-02/08 and in part by the General Research Fund (GRF), Research Grants Council (RGC) under Grant 14208315 and Grant 14313116. The work of Yihan Zhang was supported by the European Union's Horizon 2020 Research and Innovation Programme under Grant 682203-ERC-[Inf-Speed-Tradeoff]. The work of Shashank Vatedka was supported in part by the Chinese University of Hong Kong (CUHK) Direct under Grant 4055039 and Grant 4055077. The work of Anand D. Sarwate was supported in part by the U.S. National Science Foundation under Award CCF-1909468. An earlier version of this paper was presented in part at the 2012 International Conference on Signal Processing and Communications [1] [DOI: 10.1109/ISIT.2018.8437457] and in part at the 2018 IEEE International Symposium on Information theory [2] [DOI: 10.1109/SPCOM.2012.6290241]. *(Corresponding author: Yihan Zhang.)*

Yihan Zhang was with the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong, and also with the Henry and Marilyn Taub Faculty of Computer Science, Technion—Israel Institute of Technology, Haifa 3200003, Israel. He is now with the Institute of Science and Technology Austria, 3400 Klosterneuburg, Austria (e-mail: yhzhang@link.cuhk.edu.hk).

Shashank Vatedka was with the Institute of Network Coding, The Chinese University of Hong Kong, Hong Kong, and also with the Department of Communications and Electronics, Télécom Paris (Institut Polytechnique de Paris), 91120 Palaiseau, France. He is now with the Department of Electrical Engineering, Indian Institute of Technology Hyderabad, Hyderabad 502285, India (e-mail: shashankvatedka@ee.iith.ac.in).

Sidharth Jaggi is with the School of Mathematics, University of Bristol, Bristol BS8 1TH, U.K. (e-mail: sid.jaggi@bristol.ac.uk).

Anand D. Sarwate is with the Department of Electrical and Computer Engineering, Rutgers, The State University of New Jersey, New Brunswick, NJ 08854 USA (e-mail: anand.sarwate@rutgers.edu).

Communicated by R. Safavi-Naini, Associate Editor for Cryptography.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIT.2022.3167554.

Digital Object Identifier 10.1109/TIT.2022.3167554

I. Introduction and Prior Work

CONSIDER a point-to-point communication system where a transmitter, Alice, wants to send a message to a receiver, Bob, through a channel distorted by additive noise. She does so by encoding the message to a length-n codeword, which is fed into the channel. Much of traditional communication and information theory has focused on the scenario where the noise is independent of the transmitted signal and the coding scheme. We study the case where communication takes place in the presence of a malicious jammer (whom we call James) who tries to ensure that Bob is unable to recover the transmitted message. The channel is a discrete-time, real-alphabet channel, and the codeword transmitted by Alice is required to satisfy a quadratic power constraint. It is assumed that the coding scheme is known to all three parties, and James also observes a noisy version of the transmitted signal (hence the term *myopic*). The jamming signal is required to satisfy a separate power constraint, but otherwise can be a noncausal function of the noisy observation and the coding scheme.

This problem is part of the general framework of arbitrarily varying channels (AVCs), introduced by Blackwell et al. [3]. The quadratically constrained AVC (also called the Gaussian AVC) was studied by Blachman [4], who gave upper and lower bounds on the capacity of the channel under the assumption that James observes a noiseless version of the transmitted codeword (a.k.a. the *omniscient* adversary). The lower bound used a sphere packing argument similar to the one used to prove the Gilbert-Varshamov (GV) bound for binary linear codes. The upper bound was based on Rankin's upper bound on the number of non-intersecting spherical caps that can be placed on a sphere [5]. The quadratically constrained AVC is closely related to the sphere-packing problem where the objective is to find the densest arrangement of identical n-dimensional balls of radius $\sqrt{nN}$ subject to the constraint that the center of each ball lies within a ball of radius $\sqrt{nP}$. An exact characterization of the capacity of this problem is not known, though inner [4] and outer bounds [6], [7] are known. At the other end of the spectrum, Hughes and Narayan [8], and later Csiszár and Narayan [9], studied the problem with an "oblivious" James, who knows the codebook, but does not see the transmitted codeword. They consider the regime when $P > N$ (it can be shown that no positive throughput is possible when $P < N$). They showed that under an average probability of error metric, the capacity of the

oblivious adversarial channel is equal to that of an additive white Gaussian noise (AWGN) channel whose noise variance is equal to the power constraint imposed on James. These omniscient and oblivious cases are two extreme instances of the general myopic adversary that we study in this paper.

The oblivious vector Gaussian AVC was studied by Hughes and Narayan [10], and later Thomas and Hughes [11] derived bounds on the error exponents for the oblivious Gaussian AVC. Sarwate and Gastpar [12] showed that the randomized coding capacity of the oblivious channel is the same under average and maximum error probability constraints.

This work builds on [1], which characterized the capacity of this channel under the assumption that James knows a noisy version of the transmitted signal, but Alice's codebook is shared only with Bob. This can be interpreted as a myopic channel with an unlimited amount of common randomness (or shared secret key, CR) between Alice and Bob. A related model was studied by Haddadpour et al. [13], who assumed that James knows the message, but not the exact codeword transmitted by Alice. In this setup, Alice has access to private randomness which is crucially used to pick a codeword for a given message. However, Alice and Bob do not share any common randomness. Game-theoretic versions of the problems have also been considered in the literature, notably by Médard [14], Shafiee and Ulukus [15] and Baker and Chao [16]. Shafiee and Ulukus [15] considered a more general two-sender scenario, while Baker and Chao [16] studied a multiple antenna version of the problem. More recently, Hosseinigoki and Kosut [17] derived the list-decoding capacity of the Gaussian AVC with an oblivious adversary. Zhang and Vatedka [18] derived bounds on achievable list-sizes for random spherical and lattice codes. Pereg and Steinberg [19] have analyzed a relay channel where the observation of the destination is corrupted by a power-constrained oblivious adversary. Beemer et al. [20] studied a related problem of authentication against a myopic adversary, where the goal of the decoder is to correctly either decode the message or detect adversarial interference. Zhang et al. [21] also studied a quadratically constrained two-way interference channel with a jamming adversary, where proof techniques similar to ours were used to obtain upper and lower bounds on the capacity. Budkuley et al. [22] gave an improved symmetrization (known as CP-*symmetrization* where CP is for *completely positive*) bound for myopic AVCs over *discrete* alphabets. The result expands the parameter region where the capacity (without common randomness) is zero. The proof is based on a significant generalization of the Plotkin bound in classical coding theory which is proved by Wang et al. [23]. Dey et al. [24] studied, among others, the binary erasure-erasure myopic adversarial channels and gave nontrivial achievability schemes beating the Gilbert–Varshamov bound in the *insufficiently* myopic regime.

Communication in the presence of a myopic jammer has also received considerable attention in the discrete-alphabet case (see [25] and references therein, and the recent [26], [27] on covert communication with myopic jammers). We would like to draw connections to the bit-flip adversarial problem where communication takes place over a binary channel, and James observes the codeword through a binary symmetric channel (BSC) with crossover probability q. He is allowed to flip at most np bits, where $0 < p < 1/2$ can be interpreted as his "jamming power." Dey et al. [25] showed that when James is sufficiently myopic, i.e., $q > p$, the capacity is equal to $1 - H(p)$. In other words, he can do no more damage than an oblivious adversary. As we will see in the present article, this is not true for the quadratically constrained case. We will show that as long as the *omniscient list-decoding capacity* for Bob is greater than the *AWGN channel capacity* for James, the capacity is equal to a certain *myopic list-decoding capacity* for Bob. In this regime, James cannot uniquely determine the transmitted codeword among exponentially many. As a result no attack strategy by James that "pushes" the transmitted codeword to the nearest other codeword is as bad as in the omniscient case since the nearest codeword in general will be different for different choices of the transmitted codeword.

Recent works have also considered communication with simultaneous active and passive attacks [28]–[34]. However, in these works, the eavesdropper and jammer are assumed to be independent entities and the jammer is assumed to be an oblivious adversary. In this work, we derive lower bounds on the capacity of the myopic adversarial channel with an additional wiretap secrecy [35] constraint, treating the jammer as an eavesdropper at the same time.

Let us now describe the problem we address in this paper. The setup is illustrated in Fig. 1. Alice wants to send a message $\mathbf{m}$ to Bob. The message is assumed to be uniformly chosen from $\{0,1\}^{nR}$, where $R > 0$ is a parameter called the *rate*. Alice and Bob additionally have n_{key} bits of shared secret key, $\mathbf{k}$ (n_{key} could be zero — indeed, some of the major results in this work derive AVC capacity for some NSR regimes when $n_{\text{key}} = 0$). This key is kept private from James. Alice encodes the message $\mathbf{m}$ (using $\mathbf{k}$) to a codeword $\underline{\mathbf{x}} \in \mathbb{R}^n$, which is transmitted across the channel. Let $\mathcal{C}$ denote the set of all possible codewords (the codebook). In this work, we study three types of encoding:

- *Deterministic encoding:* $n_{\text{key}} = 0$ and $\underline{\mathbf{x}}$ is a deterministic function of $\mathbf{m}$
- *Stochastic encoding:* $n_{\text{key}} = 0$, but $\underline{\mathbf{x}}$ is a function of $\mathbf{m}$ and private random bits known only to Alice
- *Randomized encoding:* $n_{\text{key}} > 0$, and $\underline{\mathbf{x}}$ can be a function of the shared key $\mathbf{k}$ and random bits known only to Alice.

If the code is non-deterministic, then the *codebook rate* $R_{\text{code}} := \frac{1}{n}\log|\mathcal{C}|$ could be different from the *message rate* R (which we sometimes simply refer to as the rate). The codebook must satisfy a power constraint of $P > 0$, i.e. $\|\underline{x}\|_2 \le \sqrt{nP}$ for all $\underline{x} \in \mathcal{C}$. James sees $\underline{\mathbf{z}} = \underline{\mathbf{x}} + \underline{\mathbf{s}}_z$, where $\underline{\mathbf{s}}_z$ is an AWGN with mean zero and variance σ^2. He chooses a jamming vector $\underline{\mathbf{s}} \in \mathbb{R}^n$ as a noncausal function of $\underline{\mathbf{z}}$, the codebook $\mathcal{C}$, and his private randomness. The jamming vector is also subject to a power constraint: $\|\underline{s}\|_2 \le \sqrt{nN}$ for some $N > 0$. Bob obtains $\underline{\mathbf{y}} = \underline{\mathbf{x}} + \underline{\mathbf{s}}$, and decodes this to a message $\widehat{\mathbf{m}}$. The message is said to have been conveyed reliably if $\widehat{\mathbf{m}} = \mathbf{m}$. The probability of error, P_e, is defined as the probability that $\widehat{\mathbf{m}} \neq \mathbf{m}$, where the randomness is over the message $\mathbf{m}$, the private randomness that Alice uses,

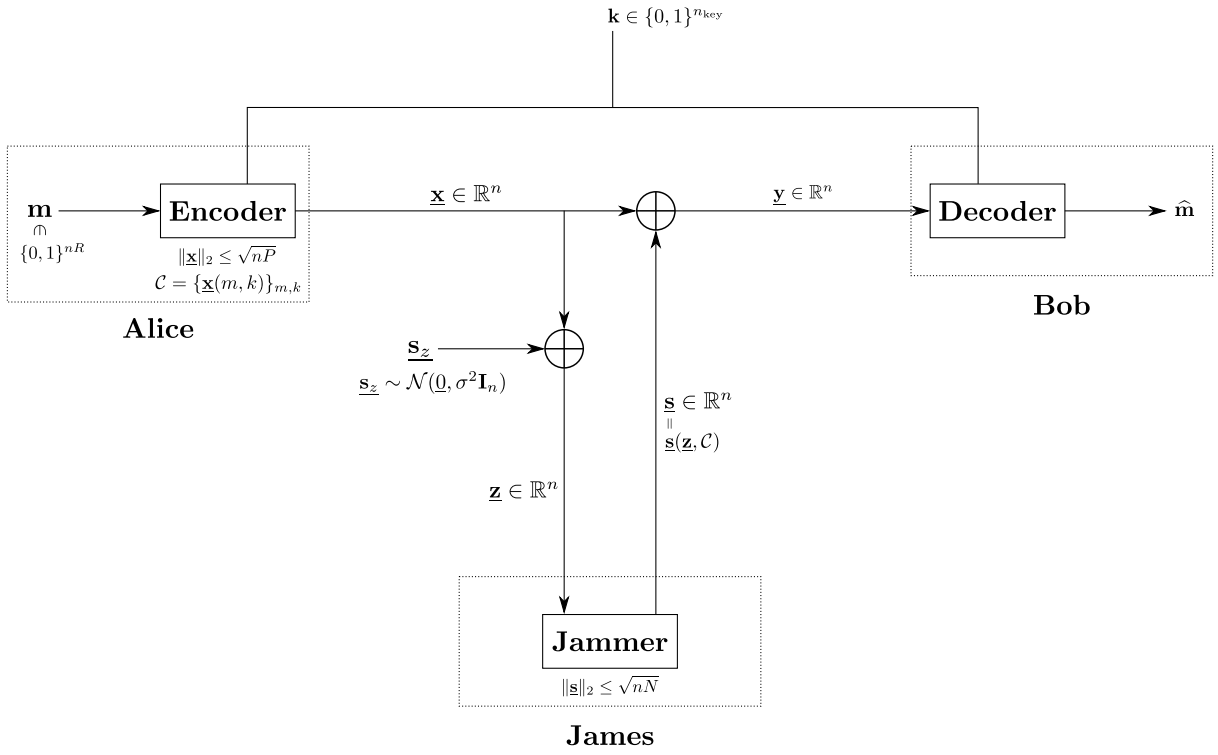


Fig. 1. The setup studied in this paper: Alice wants to transmit nR-bit message $\mathbf{m}$ to Bob. Across the channel, she transmits a codeword $\underline{\mathbf{x}}$, which is a function of $\mathbf{m}$ (and potentially a shared key $\mathbf{k}$ of n_{key} bits, though we also study the scenario when $n_{\text{key}} = 0$). The collection of codewords $\mathcal{C}$ is called the codebook, and every codeword in the codebook must satisfy a power constraint of $\sqrt{nP}$. The jammer James observes $\underline{\mathbf{z}}$ corresponding to the output of an AWGN channe with variance σ^2. He then chooses a jamming/state sequence $\underline{\mathbf{s}}$ (satisfying a power constraint of $\sqrt{nN}$) as a noncausal function of $\underline{\mathbf{s}_z}$ and $\mathcal{C}$. On observing $\underline{\mathbf{y}} = \underline{\mathbf{x}} + \underline{\mathbf{s}}$, Bob must output his estimate $\hat{\mathbf{m}}$ of the message $\mathbf{m}$ such that the probability of error (averaged over $\mathbf{m}$ and $\underline{\mathbf{s}_z}$) vanishes.

the random noise $\underline{\mathbf{s}_z}$, the key $\mathbf{k}$, and the private random bits available to James.[1] In all our code constructions,[2] we will assume that Alice and Bob may share a secret key, but the mapping from $(\mathbf{m}, \mathbf{k})$ to $\underline{\mathbf{x}}$ is deterministic. In other words, Alice does not possess any source of additional private randomness. Conversely, all our impossibility results are robust to the presence of private randomness at the encoder (since in some AVC scenarios, private randomness is known to boost capacity — e.g. [36]) We study the problem with different amounts of common randomness shared by Alice and Bob but unknown to James, and present results in each case.

We say that a rate $R > 0$ is achievable if there exists a sequence (in increasing n) of codebooks for which the probability of error[3] goes to zero as $n \to \infty$. The supremum of all achievable rates is called the capacity of the channel.

We say that a rate $R > 0$ is achievable with (wiretap) secrecy if there exists a sequence (in increasing n) of codebooks for which the probability of error and the mutual information $I(\mathbf{m}; \underline{\mathbf{z}})$ both go to zero as $n \to \infty$. This is commonly referred to as the *strong secrecy* requirement in the literature. The supremum of all achievable rates is called the secrecy capacity of the channel.

[1]An averaging argument shows that the rate cannot be improved even if Bob uses additional private random bits for randomized decoding.

[2]An exception is Appendix C, where we show that private randomness does not increase the capacity of the omniscient adversarial channel. However, we have reason to believe (albeit unsupported by formal proof) that additional private randomness *may* increase the achievable rate — this is part of our ongoing investigation.

[3]See Sec. V for a formal definition.

A. Organization of the Paper

We give a summary of our results and proof techniques in Sec. II. The formal statements of the results are presented in Sec. VI. The main results are also compactly summarized in Table I and the results with secrecy are tabulated in Table II. We then discuss the connection between our work and several closely related prior works in Sec. III. Notation and preliminaries are described in Sec. IV and Sec. V, respectively. This, as mentioned, is followed by ideas and details of the proof techniques in Sec. VI. In Sec. VII, we describe the results for infinite common randomness and give a formal proof of the converse. Sec. VIII contains the main ideas required to prove our results with linear and logarithmic amounts of common randomness. Our results on list-decoding are described in Sec. VIII-A, with Theorem 14 giving the main result. Coming to the no-common randomness regime, we present a technical yet high-level proof sketch of the achievability and a full proof of the symmetrization converse in Sec. IX. Sec. X contains a detailed proof of Theorem 11, and Sec. XI gives the proof of Theorem 14. Appendix C has a note on why private randomness does not improve the capacity if James is omniscient. We transcribe a rigorous proof of a folklore theorem regarding list-decoding in Euclidean space against an omniscient adversary in Sec. D. Some of the technical

details of proofs appear in the other appendices (specifically Appendix B and Appendices E–H). Frequently used notation is summarized in Table III in Appendix A. Fig. 21 is a flowchart outlining steps involved in the proof.

II. Overview of Results and Proof Techniques

A. Overview of Results

We now briefly describe our results and proof techniques. It is helpful to visualize our results in terms of the noise-to-signal ratio (NSR), using a N/P (adversarial NSR to Bob) versus σ^2/P (random NSR to James) plot similar to the one shown in Fig. 8.[4] In [1], it was shown that with an infinite amount of common randomness, the capacity is $R_{\mathrm{LD}} := \frac{1}{2}\log\frac{P}{N}$ in the red region, and $R_{\mathrm{LD,myop}} := \frac{1}{2}\log\left(\frac{(P+\sigma^2)(P+N)-2P\sqrt{N(P+\sigma^2)}}{N\sigma^2}\right)$ in the blue region. The capacity is zero in the grey region.

In this article, while the major results are for the case when $n_{\mathrm{key}} = 0$, along the way we prove anciliary results for the regimes where $n_{\mathrm{key}} = \Theta(n)$ and $n_{\mathrm{key}} = \Theta(\log n)$.

- *List-Decoding:* We prove a general result for list-decoding in the presence of a myopic adversary. For an omniscient adversary, the list-decoding capacity is $R_{\mathrm{LD}} = \frac{1}{2}\log\frac{P}{N}$. This is a folklore result, but we give a proof of this statement in Appendix D for completeness. When the adversary is myopic, and the encoder-decoder pair shares $\mathcal{O}(n)$ bits of common randomness, we give achievable rates for list-decoding. This is equal to R_{LD} for $\frac{\sigma^2}{P} \leq \frac{P}{N} - 1$, and is larger than R_{LD} in a certain regime (depending on the amount of common randomness) where $\frac{\sigma^2}{P} > \frac{P}{N} - 1$. The achievable rates are illustrated in Fig. 7. With no common randomness, we can achieve R_{LD} and $R_{\mathrm{LD,myop}}$ in the red and blue regions of Fig. 7a respectively. If Alice and Bob share n_{key} bits, then $R_{\mathrm{LD,myop}}$ is achievable in a larger region. For instance, if $n_{\mathrm{key}} = 0.2n$, then the blue region can be expanded to give Fig. 7b.
- *Linear CR:* When common randomness is present, we combine our list-decoding result with [37, Lemma 13] to give achievable rates over the myopic adversarial channel. Let us first discuss the case the amount of common randomness is linear in n, i.e., $n_{\mathrm{key}} = nR_{\mathrm{key}}$ for some $R_{\mathrm{key}} > 0$. If $R_{\mathrm{key}} \geq \frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right) - R_{\mathrm{LD,myop}}$, then we are able to give a complete characterization of the capacity of the channel for all values of the NSRs. We can achieve everything in Fig. 8. If $R_{\mathrm{key}} < \frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right) - R_{\mathrm{LD,myop}}$, then we are able to characterize the capacity in only a sub-region of the NSRs — This is illustrated in Fig. 6c and Fig. 6d for different values of R_{key}. In the dotted regions, we only have nonmatching upper and lower bounds. It is worth pointing out that no fixed R_{key} will let us achieve $R_{\mathrm{LD,myop}}$ in the entire blue region of Fig. 8. However, for every point in the blue region, there exists a finite value of R_{key} such that $R_{\mathrm{LD,myop}}$ is achievable at that point. In other words, an $n_{\mathrm{key}} = \Omega(n)$ is sufficient to achieve $R_{\mathrm{LD,myop}}$ at every point in the interior of the blue region in Fig. 8.
- *Logarithmic CR:* For the $n_{\mathrm{key}} = \Theta(\log n)$ case, we are able to find the capacity in the red and blue regions in Fig. 6b. In the dotted regions, we have nonmatching upper and lower bounds.
- *No CR:* For $n_{\mathrm{key}} = 0$, we require a more involved approach to find the capacity. We use some of the results on myopic list-decoding in our bounds for the probability of error. We find the capacity in the red, blue and grey regions in Fig. 6a, but only have nonmatching upper and lower bounds in the dotted green and white regions.
- *Sufficiency of Deterministic Encoding Against Omniscient Adversaries:* We show that if James is omniscient, then private randomness at the encoder does not help improve the capacity. This is based on a similar observation made by Dey et al. [36] for the bit-flip adversarial channel. See Appendix C for details.
- *Wiretap Secrecy:* We use the above results to derive achievable rates under strong secrecy constraints. Specifically, we want to ensure that the mutual information to James, $I(\mathbf{m};\underline{\mathbf{z}}) = o(1)$ in addition to Bob being able to decode $\mathbf{m}$ reliably. Since the proof of reliability uses a random spherical code construction, we are able to obtain strong secrecy using random binning by ensuring that the code corresponding to each bin is a good resolvability code [38] for the AWGN channel from Alice to James.

The variation of the regions of the noise-to-signal ratios (NSR) where we can obtain achievable rates is illustrated in Fig. 9. As seen in the figure, even $\Theta(\log n)$ bits of common randomness is sufficient to ensure that the red and blue regions are expanded. An additional $\Theta(n)$ bits can be used to expand the blue region even further, eventually achieving everything in Fig. 8. The rates derived in this paper are compared with prior work in Table I.[5]

B. Proof Techniques for Converse Results

We begin by outlining the proof techniques used in our converse results. At first sight, it might seem that geometric/sphere packing bounds such as in [6] may be used when Bob's NSR N/P is higher than James's NSR σ^2/P, since whenever Bob can hope to decode Alice's message, so can James. If Alice's encoder is deterministic, James can therefore infer Alice's transmitted codeword, and thereby "push" it to the nearest codeword. However, such a reasoning applies only to deterministic codes, i.e., when Alice does not use any private or common randomness. We therefore highlight two converse techniques that apply even when Alice's encoder is *not* deterministic.

[4] Our parameterization makes the parameter regions of interest *compact* and concentrated in a bounded region around the origin (rather than scattered or shooting infinitely far away) in the two-dimensional plane spanned by σ^2/P and N/P.

[5] Many prior works (for example [1], [14] etc.) also consider additional random noise, independent of the jamming noise introduced by James, on the channel from Alice to Bob. In principle the techniques in this paper carry over directly even to that setting, but for ease of exposition we choose not to present those results.

1) Scale-and-Babble: The scale-and-babble attack is a strategy that reduces the channel from Alice to Bob into an AWGN channel. James expends a certain amount of power in cancelling the transmitted signal, and the rest in adding independent Gaussian noise. Since the capacity of the AWGN channel cannot be increased using common randomness, the scale-and-babble attack gives an upper bound that is valid for all values of n_{key}. This technique gives us the rate region illustrated in Fig. 10. The capacity is upper bounded by R_{LD} in the red region, $R_{\text{LD,myop}}$ in the blue region, and is zero in the grey region.

We remark that the scale-and-babble attack is not an original idea of this work. This proof was suggested by Sarwate [1], and is an application of a more general technique proposed by Csiszár and Narayan [39] to convert an AVC into a discrete memoryless channel. Nevertheless we give a complete proof to keep the paper self-contained.

2) Symmetrization Attacks: Symmetrization attacks give us upper bounds on the throughput when Alice and Bob do not share a secret key, but hold regardless of whether Alice's encoder uses private randomness or not. We give two attacks for James:

- *A $\underline{\mathbf{z}}$-aware symmetrization attack:* James picks a codeword $\underline{\mathbf{x}}'$ from Alice's codebook uniformly at random and independently of $\underline{\mathbf{z}}$. He transmits $(\underline{\mathbf{x}}' - \underline{\mathbf{z}})/2$ — since $\underline{\mathbf{z}} = \underline{\mathbf{x}} + \underline{\mathbf{s}}_z$ for some vector $\underline{\mathbf{s}}_z$ with $\mathcal{N}(0, \sigma^2)$ components, therefore Bob receives $(\underline{\mathbf{x}} + \underline{\mathbf{x}}' - \underline{\mathbf{s}}_z)/2$. If $\underline{\mathbf{x}} \neq \underline{\mathbf{x}}'$, then Bob makes a decoding error with nonvanishing probability. This attack is inspired by a technique used to prove the Plotkin bound for bit-flip channels. The symmetrization attack lets us prove that the capacity is zero when $\frac{\sigma^2}{P} \leq \frac{1}{1-N/P} - 2$ (Fig. 11c). The $\underline{\mathbf{z}}$-aware attack is novel in the context of myopic channels, but is also inspired by similar ideas in [40].
- *A $\underline{\mathbf{z}}$-agnostic symmetrization argument:* This lets us show that the capacity is zero for $N > P$ (Fig. 11b). James picks a codeword $\underline{\mathbf{x}}'$ as before but instead transmits $\underline{\mathbf{s}} = \underline{\mathbf{x}}'$. Bob receives $\underline{\mathbf{x}} + \underline{\mathbf{x}}'$ and we can show that the probability of error is nonvanishing. The $\underline{\mathbf{z}}$-agnostic symmetrization attack was used by Csiszár and Narayan [9] to show that the capacity of the oblivious adversarial channel is zero for $N > P$.

The scale-and-babble attack holds for all values of n_{key} since it involves reducing the channel into an equivalent AWGN channel, and the capacity of the AWGN channel cannot be increased using common randomness. On the other hand, the symmetrization arguments are not valid when $n_{\text{key}} > 0$. Indeed, we will show that strictly positive rates can be achieved in the symmetrizable regions with even $\Omega(\log n)$ bits of common randomness.

Combining the three techniques give us the upper bounds in Fig. 12.

C. Proof Techniques for Achievability Results

The achievability proofs for the three regimes of n_{key} outlined above involve some common techniques. We now give a high-level description of some of the ideas. Fundamental to the achievability proofs is the concept of list-decoding. In all the achievability proofs, we use random spherical codes $\mathcal{C} = \{\underline{x}(m,k) : 1 \leq m \leq 2^{nR}, 1 \leq k \leq 2^{n_{\text{key}}}\}$, where each $\underline{x}(m,k)$ is sampled independently and uniformly from the sphere $\mathcal{S}^{n-1}(0, \sqrt{nP})$ in $\mathbb{R}^n$ centred at 0 and comprising of vectors of magnitude $\sqrt{nP}$.

1) Myopic List-Decoding: This is a central idea in our proofs, and a novel contribution of this work. The broad idea is to use myopia to ensure that James is unable to uniquely recover the transmitted codeword. We show that if the codebook rate is sufficiently large, then there are exponentially many codewords that from James's perspective Alice could plausibly have transmitted. Due to this confusion, no attack strategy (by pushing the transmitted $\underline{\mathbf{x}}$ in the direction of the nearest other codeword $\underline{\mathbf{x}}'$, since the nearest codeword will in general be different directions for different $\underline{\mathbf{x}}$) by James is as bad as the one he could instantiate in the omniscient case. We study the list-decoding problem, where instead of recovering the transmitted message uniquely, Bob tries to output a poly(n) sized list that includes the transmitted codeword. Since James is myopic, we could hope to achieve rates greater than the omniscient list-decoding capacity R_{LD}. Even with $n_{\text{key}} = 0$, we can achieve a higher rate, equal to $R_{\text{LD,myop}}$, in the blue region in Fig. 7a. The blue region can be expanded with a larger amount of common randomness, as seen in Fig. 7b. We will in fact show that the list-decoding capacity is equal to $C_{\text{myop,rand}}$ (see Eqn. (VI.1) for its definition) if n_{key} is large enough.

Let us briefly outline the proof techniques. We show that conditioned on $\underline{\mathbf{z}}$, the transmitted codeword lies in a strip (informally denoted by $\mathcal{S}tr$ for now) approximately at a distance $\sqrt{n\sigma^2}$ to $\underline{\mathbf{z}}$. See Fig. 2 for an illustration. If the codebook rate exceeds $\frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right)$, then this strip will contain exponentially many codewords. All these codewords are roughly at the same distance to $\underline{\mathbf{z}}$ and are therefore nearly indistinguishable from the one actually transmitted. We operate under the assumption of a more powerful adversary who has, in addition to $\underline{\mathbf{z}}$, access to an oracle. This is a matter of convenience and will greatly simplify our proofs. The oracle reveals an exponential sized subset of the codewords (that includes $\underline{\mathbf{x}}$) from the strip. We call this the oracle-given set (OGS). We prove that for most codewords in the OGS, no attack vector $\underline{\mathbf{s}}$ can eventually force a list-size greater than poly(n) as long as the rate is less than $C_{\text{myop,rand}}$. To prove this result, we obtain a bound on the *typical* area of the decoding region $\mathcal{S}^{n-1}(0, \sqrt{nP}) \cap \mathcal{B}^n(\underline{\mathbf{x}} + \underline{\mathbf{s}}, \sqrt{nN})$. We will show that for certain regimes of the noise-to-signal ratios (NSRs), the volume of the decoding region is typically much less than the worst-case scenario (i.e., had James known $\underline{\mathbf{x}}$). This gives us an improvement over the omniscient list-decoding capacity.

2) Reverse List-Decoding: This technique, along with myopic list-decoding outlined above, is used to obtain achievable rates in the case where Alice and Bob do not share any common randomness. Given an attack vector $\underline{s}$, we say that $\underline{x}'$ confuses $\underline{x}$ if $\underline{x}'$ lies within $\mathcal{B}^n(\underline{x} + \underline{s}, \sqrt{nN})$. In list-decoding, we attempt to find the number of codewords that could potentially confuse the transmitted codeword. Our goal in the list-decoding problem is to keep the number of confusable

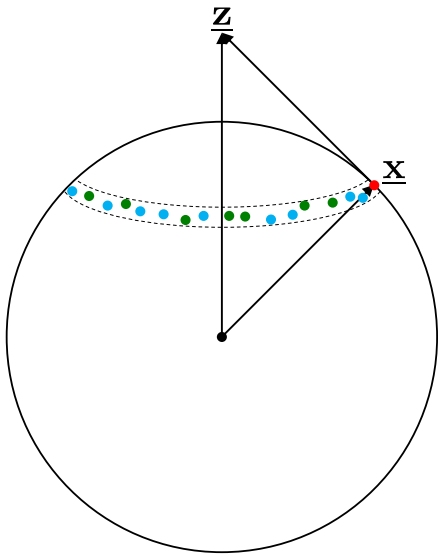


Fig. 2. Illustration of the strip and the oracle-given set (OGS). All codewords in the strip (the collection of red, blue and green points) are roughly at the same distance from $\underline{\mathbf{z}}$, and hence approximately have the same likelihood of being transmitted. The OGS (illustrated as red and green points) is a randomly chosen subset of the codewords in this strip.

codewords to a minimum. In reverse list-decoding, we ask the opposite question: Given a potentially confusing codeword, how many codewords in the strip could this confuse?

For every codeword $\underline{x}'$ and attack vector $\underline{s}$, we could define the reverse list-size as $|\{\underline{x} \in \mathcal{S}tr : \|\underline{x} + \underline{s} - \underline{x}'\|_2 \leq \sqrt{nN}\}|$. In other words, this is the number of codewords in the strip which when translated by $\underline{s}$, are confusable with $\underline{x}'$. Our goal is to keep this as small as possible (in fact, poly(n)) for all possible attack vectors $\underline{s}$. We show that small reverse list-sizes are guaranteed as long as $\frac{1}{2}\log\frac{P}{N} > \frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right)$, i.e. in the red and blue regions of Fig. 13.

3) Going From List-Decoding to Unique Decoding: Obtaining results for unique decoding uses two different approaches that depends on the amount of common randomness.

- *Linear/logarithmic amount of common randomness*: Langberg [41] gave a combinatorial technique to convert any list-decodable code (with no common randomness) for a binary channel into a uniquely decodable code of the same rate with $\Omega(\log n)$ bits of common randomness. This was later generalized by Sarwate [37] to arbitrary AVCs, and [42] recently showed that only $(1+\varepsilon)\log n$ bits suffices (where ε denotes the difference between the list-decoding capacity and the transmission rate). This combined with our result on myopic list-decoding will give us an achievable rate for reliable communication over the myopic channel.
- *No common randomness*: The ideas in myopic list-decoding can be used to show that there are at most poly(n) codewords that can potentially confuse the exponentially many codewords in the OGS. Using reverse list-decoding, we can conclude that each codeword outside the OGS can confuse at most poly(n) codewords in the OGS. Using this, and a "grid argument" along the lines of [25], we can show that the probability of decoding error is vanishingly small.

4) Two Uses of the Common Randomness: When Alice and Bob have only $\mathcal{O}(\log n)$ bits of common randomness, $\mathbf{k}$ is only used to find the true message from the list using the approach proposed in [37], [41]. However, when Alice and Bob have $\Omega(n)$ bits of shared secret key, there are two different uses for $\mathbf{k}$.

Recall from our discussion above that to obtain an improvement over omniscient list-decoding, we must ensure that James is sufficiently confused about the true codeword. If he can recover $\underline{\mathbf{x}}$ with high probability (w.h.p.), then there would be no hope of achieving rates greater than R_{LD}. When Alice and Bob share linear amounts of common randomness, all but $(1+\varepsilon)\log n$ bits is used to ensure that the strip contains exponentially many codewords. This is done by generating $2^{n_{\mathrm{key}}-(1+\varepsilon)\log n}$ independent codebooks, each containing 2^{nR} codewords. Based on the realization of the key, Alice picks the appropriate codebook for encoding. Since Bob knows the key, he can use this codebook for decoding. However, James does not have access to the shared secret key and to him, all the $2^{nR+n_{\mathrm{key}}-(1+\varepsilon)\log n}$ codewords are equally likely to have been transmitted. This ensures that James is sufficiently confused about the true codeword. The remaining $(1+\varepsilon)\log n$ bits are then used to disambiguate the list at the decoder.

III. Comparison With Other Works

We now describe the similarities and differences with three closely related works.

Sarwate [1] derived the capacity of the myopic adversarial channel with unlimited common randomness. The present paper is in some sense a continuation of [1], with a restriction on the amount of shared secret key. In [1], the codebook was privately shared by Alice and Bob. A minimum angle decoder was used, and the achievable rate was the solution of an optimization problem identical to what we obtain in our analysis of myopic list-decoding. The converse used the scale-and-babble attack that we describe in a subsequent section. When Alice and Bob share $\Omega(n)$ bits of common randomness, we find that the upper bound is indeed optimal. We were unable to obtain an improved upper bound for smaller amounts of common randomness. However, we do give an improved converse for certain NSRs using symmetrization when there is no common randomness shared by the encoder-decoder pair.

Dey et al. [25] studied the discrete myopic adversarial channel. We borrow several proof techniques from their paper to obtain results when $n_{\mathrm{key}} = 0$. This includes the ideas of blob list-decoding, reverse list-decoding, and the grid argument to prove that the probability of error is vanishingly small even when transmitting at rates above R_{GV}, despite the fact that the jamming pattern may be correlated with the transmission. However, there are several differences between our work and the discrete case. A key difference between this work and [25] is our use of myopic list-decoding. A direct extension of the techniques in [25] would only let us achieve rates up to the omniscient list-decoding capacity $\frac{1}{2}\log\frac{P}{N}$. The study of myopic list-decoding is novel, and is one of the main

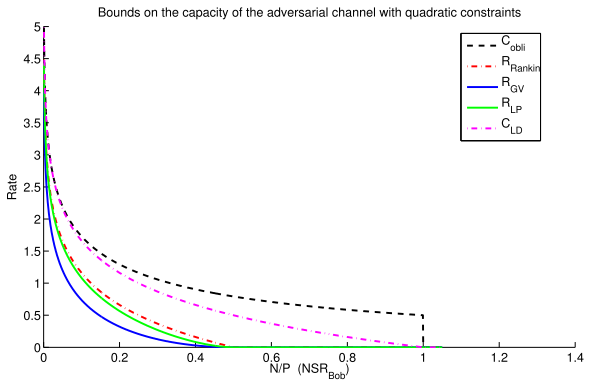


Fig. 3. Achievable rates for the quadratically constrained adversarial channel — prior work.

contributions of this work. Furthermore, the random variables involved in this work are continuous which introduces several challenges. Several arguments involving union bounds and fixed distances do not go through directly from the discrete setting. We overcome some of these by quantizing the relevant random variables, a standard trick to approximate the continuous variables by discrete ones. In [25], the oracle-given set (OGS) was chosen to be a subset of codewords all at the same distance to the vector received by James. All codewords in the OGS would then have the same posterior probability given James's observation and the OGS. As $\underline{\mathbf{s}}_z$ is Gaussian in our case, such an argument cannot be used. Instead, we choose the OGS from a thin strip. Given $\underline{\mathbf{z}}$ and the OGS, the codewords are not uniformly distributed. However, we carefully choose the thickness of the strip to be small enough to ensure that this posterior distribution is close to being uniform. Due to the random variables being continuous, dealing with the quadratically constrained case requires a more careful analysis involving many more slackness parameters than the discrete analogue.

The symmetrization argument we use in Section IX-B is inspired by the "scaled babble-and-push" attack studied by Li et al. [40]. This work studies the attack where James generates an independent codeword $\underline{\mathbf{x}}'$ and transmits $(\underline{\mathbf{x}}'-\underline{\mathbf{z}})/2$. A similar idea with optimized parameters allows us to prove that the capacity is zero in the regime where $\frac{\sigma^2}{P} < \frac{1}{1-N/P} - 2$.

We could potentially extend this idea, along the lines of the scaled babble-and-push attack [40] in the following manner. In a subset of codeword indices, James uses the scale-and-babble attack. In the remaining indices i, he transmits $(\mathbf{x}'_i - \mathbf{z}_i)/2$. We could hope to get an improved converse in the regime $1 \le \frac{P}{N} \le 1 + \frac{\sigma^2}{P} \le \frac{1}{N/P-1}$. We were unsuccessful in analyzing this and it has been left as future work.

IV. Notation

Random variables are denoted by lower case Latin letters in boldface, e.g., $\mathbf{m}$. Their realizations are denoted by corresponding letters in plain typeface, e.g., m. Vectors of length n, where n is the block-length, are denoted by lower case Latin letters with an underline, e.g., $\underline{\mathbf{x}}, \underline{\mathbf{s}}, \underline{x}, \underline{s}$, etc. The ith entry of a vector is denoted by a subscript i, e.g., $\underline{\mathbf{x}}_i, \underline{\mathbf{s}}_i, \underline{x}_i, \underline{s}_i$, etc. Matrices are denoted by capital Latin/Greek letters in boldface, e.g., $\mathbf{I}, \mathbf{\Sigma}$, etc.

Sets are denoted by capital Latin letters in calligraphic typeface, e.g., $\mathcal{C}, \mathcal{I}$, etc. In particular, an $(n-1)$-dimensional sphere in n-dimensional Euclidean space centered at $\underline{x}$ of radius r is denoted by

$$\mathcal{S}^{n-1}(\underline{x}, r) = \{\underline{y} \in \mathbb{R}^n : \|\underline{y}\|_2 = r\}.$$

An n-dimensional ball in Euclidean space centered at $\underline{x}$ of radius r is denoted by

$$\mathcal{B}^n(\underline{x}, r) = \{\underline{y} \in \mathbb{R}^n : \|\underline{y}\|_2 \le r\}.$$

As shown in Figure 4a, an $(n-1)$-dimensional cap centered at $\underline{x}$ of radius r living on an $(n-1)$-dimensional sphere of radius r' is denoted by

$$\begin{aligned}\mathcal{C}ap^{n-1}(\underline{x}, r, r') &= \{\underline{y} \in \mathcal{S}^{n-1}(O, r') : \|\underline{y} - \underline{x}\|_2 \le r\} \\ &= \mathcal{B}^n(\underline{x}, r) \cap \mathcal{S}^{n-1}(O, r').\end{aligned}$$

As shown in Figure 4b, an $(n-1)$-dimensional strip centered at $\underline{x_-}$ and $\underline{x_+}$ of radii r_- and r_+ is denoted by

$$\begin{aligned}&\mathcal{S}tr^{n-1}(\underline{x_-}, \underline{x_+}, r_-, r_+) \\ =&\{\underline{x} \in \mathcal{S}^{n-1}(r) : \|\underline{x} - \underline{x_-}\|_2 \ge r_-,\ \|\underline{x} - \underline{x_+}\|_2 \le r_+\} \\ =&\mathcal{B}^n(\underline{x_-}, r_-)^c \cap \mathcal{B}^n(\underline{x_+}, r_+) \cap \mathcal{S}^{n-1}(O, r),\end{aligned}$$

where r satisfies $\sqrt{r^2 - r_-^2} - \sqrt{r^2 - r_+^2} = \|\underline{x_-} - \underline{x_+}\|$. An n-dimensional shell centered at $\underline{x}$ of inner radius r_{in} and outer radius r_{out}, where $r_{\text{out}} > r_{\text{in}}$, is denoted by

$$\begin{aligned}\mathcal{S}h^n(\underline{x}, r_{\text{in}}, r_{\text{out}}) =& \mathcal{S}h^n\left(\underline{x}, \frac{r_{\text{in}} + r_{\text{out}}}{2} \pm \frac{r_{\text{out}} - r_{\text{in}}}{2}\right) \\ =& \mathcal{B}^n(\underline{x}, r_{\text{out}}) \backslash \mathcal{B}^n(\underline{x}, r_{\text{in}}).\end{aligned}$$

Let $\text{Vol}(\cdot)$ denote the Lebesgue volume of a Euclidean body and let $\text{Area}(\cdot)$ denote its Lebesgue area of an Euclidean surface. For $M \in \mathbb{Z}_{>0}$, we let $[M]$ denote the set of first M positive integers $\{1, 2, \ldots, M\}$.

The probability mass function (p.m.f.) of a discrete random variable $\mathbf{x}$ or a random vector $\underline{\mathbf{x}}$ is denoted by $p_{\mathbf{x}}$ or $p_{\underline{\mathbf{x}}}$. Here with a slight abuse of notation, we use the same to denote the probability density function (p.d.f.) of $\mathbf{x}$ or $\underline{\mathbf{x}}$ if they are continuous. If every entry of $\underline{\mathbf{x}}$ is independently and identically distributed (i.i.d.) according to $p_{\mathbf{x}}$, then we write $\underline{\mathbf{x}} \sim p_{\mathbf{x}}^{\otimes n}$. In other words,

$$p_{\underline{\mathbf{x}}}(\underline{x}) = p_{\mathbf{x}}^{\otimes n}(\underline{x}) := \prod_{i=1}^{n} p_{\mathbf{x}}(\underline{x}_i).$$

Let $\text{Unif}(\Omega)$ denote the uniform distribution over some probability space Ω. Let $\mathcal{N}(\underline{\mu}, \mathbf{\Sigma})$ denote the n-dimensional Gaussian distribution with mean vector $\underline{\mu}$ and covariance matrix $\mathbf{\Sigma}$.

The indicator function is defined as, for any $\mathcal{A} \subseteq \Omega$ and $x \in \Omega$,

$$\mathbb{1}_{\mathcal{A}}(x) = \begin{cases} 1, & x \in \mathcal{A} \\ 0, & x \notin \mathcal{A}. \end{cases}$$

At times, we will slightly abuse notation by saying that $\mathbb{1}_{\mathsf{A}}$ is 1 when relation A is satisfied and zero otherwise. We use

TABLE I

SUMMARY OF RESULTS FOR THE ADVERSARIAL CHANNEL WITH QUADRATIC CONSTRAINTS. HERE, $\mathbb{1}_{\{P\geq N\}}$ IS 1 WHEN $P \geq N$ AND ZERO OTHERWISE. ALSO, $C_J := \frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right)$. RATES KNOWN IN PRIOR WORK ARE PLOTTED IN FIG. 3. NOTE THAT IN ANY ENTRY IN THE THIRD COLUMN "*Level of Myopia*", ALL REGIMES OF σ^2/P AS A FUNCTION OF N/P ARE DISJOINT. IN PARTICULAR, IN THE LAST FOUR ENTRIES OF THAT COLUMN, ONLY ONE OF THE CASES $\sigma^2/P \leq 1/(N/P) - 1$ AND $\sigma^2/P \leq N/P - 1$ CAN OCCUR FOR ANY FIXED VALUE OF N/P

Reference	**Rate**	**Level of myopia**	**Common randomness**
Folklore (Appendix D)	$R_{\mathrm{LD}} := \frac{1}{2}\log\frac{P}{N}$ for list-decoding	$\sigma^2 = 0$	0
Hughes-Narayan [8]	$C_{\mathrm{obli,rand}} := C_{\mathrm{AWGN}} := \frac{1}{2}\log\left(1+\frac{P}{N}\right)$	$\sigma^2 = \infty$	∞
Csiszár-Narayan [9]	$C_{\mathrm{obli}} := C_{\mathrm{AWGN}}\mathbb{1}_{\{P\geqslant N\}}$	$\sigma^2 = \infty$	0
Blachman [4]	$C_{\mathrm{omni}} \leqslant R_{\mathrm{Rankin}} := \frac{1}{2}\log\left(\frac{P}{2N}\right)\mathbb{1}_{\{P\geqslant 2N\}}$ $C_{\mathrm{omni}} \geqslant R_{\mathrm{GV}} := \frac{1}{2}\log\left(\frac{P^2}{4N(P-N)}\right)\mathbb{1}_{\{P\geqslant 2N\}}$	$\sigma^2 = 0$	0
Kabatiansky-Levenshtein [6]	$C_{\mathrm{omni}} \leqslant R_{\mathrm{LP}} := (\alpha\log\alpha - \beta\log\beta)\mathbb{1}_{\{P\geqslant 2N\}}$, where $\alpha := \frac{P+2\sqrt{N(P-N)}}{4\sqrt{N(P-N)}}$, $\beta := \frac{P-2\sqrt{N(P-N)}}{4\sqrt{N(P-N)}}$	$\sigma^2 = 0$	0
Sarwate [1]	$C_{\mathrm{myop,rand}} = R_{\mathrm{LD}} := \frac{1}{2}\log\frac{P}{N}$ $C_{\mathrm{myop,rand}} = R_{\mathrm{LD,myop}} := \frac{1}{2}\log\left(\frac{(P+\sigma^2)(P+N)-2P\sqrt{N(P+\sigma^2)}}{N\sigma^2}\right)$ $C_{\mathrm{myop,rand}} = 0$	$\frac{\sigma^2}{P} \leqslant \frac{1}{N/P} - 1$ $\frac{\sigma^2}{P} \geqslant \max\left\{\frac{1}{N/P}-1, \frac{N}{P}-1\right\}$ $\frac{\sigma^2}{P} \leqslant \frac{N}{P} - 1$	∞
This work (Theorem 14)	$C_{\mathrm{myop}} = R_{\mathrm{LD}}$ $C_{\mathrm{myop}} = R_{\mathrm{LD,myop}}$ $R_{\mathrm{GV}} \leqslant C_{\mathrm{myop}} \leqslant R_{\mathrm{LD}}$ $R_{\mathrm{GV}} \leqslant C_{\mathrm{myop}} \leqslant R_{\mathrm{LD,myop}}$ $C_{\mathrm{myop}} = 0$	$\frac{1}{1-N/P} - 1 \leqslant \frac{\sigma^2}{P} \leqslant \frac{1}{N/P} - 1$ $\frac{\sigma^2}{P} \geqslant \max\left\{\frac{1}{N/P}-1, \frac{1}{1-N/P}-1\right\}$ $\frac{1}{1-N/P} - 2 \leqslant \frac{\sigma^2}{P} \leqslant \min\left\{\frac{1}{N/P}-1, \frac{1}{1-N/P}-1\right\}$ $\max\left\{\frac{1}{N/P}-1, \frac{1}{1-N/P}-2\right\} \leqslant \frac{\sigma^2}{P} \leqslant \frac{1}{1-N/P}-1, \frac{N}{P} \leqslant 1$ $\frac{\sigma^2}{P} \leqslant \frac{1}{1-N/P} - 2$ or $\frac{N}{P} \geqslant 1$	0
This work (Lemma 13)	$C_{\mathrm{myop}} = R_{\mathrm{LD}}$ $C_{\mathrm{myop}} = R_{\mathrm{LD,myop}}$ $R_{\mathrm{LD}} \leqslant C_{\mathrm{myop}} \leqslant R_{\mathrm{LD,myop}}$ $C_{\mathrm{myop}} = 0$	$\frac{\sigma^2}{P} \leqslant \frac{1}{N/P} - 1$ $\frac{\sigma^2}{P} \geqslant \max\left\{\frac{1}{N/P}-1, 4\frac{N}{P}-1\right\}$ $\max\left\{\frac{1}{N/P}-1, \frac{N}{P}-1\right\} \leqslant \frac{\sigma^2}{P} \leqslant 4\frac{N}{P}-1$ $\frac{\sigma^2}{P} \leqslant \frac{N}{P} - 1$	$\Theta(\log n)$
This work (Lemma 12)	$C_{\mathrm{myop}} = R_{\mathrm{LD}}$ $C_{\mathrm{myop}} = R_{\mathrm{LD,myop}}$ $R_{\mathrm{LD}} \leqslant C_{\mathrm{myop}} \leqslant R_{\mathrm{LD,myop}}$ $C_{\mathrm{myop}} = 0$	$\frac{\sigma^2}{P} \leqslant \frac{1}{N/P} - 1$ $\frac{\sigma^2}{P} \geqslant \max\left\{\frac{1}{N/P}-1, \frac{N}{P}-1\right\}$, and $R_{\mathrm{key}} > \frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right) - R_{\mathrm{LD,myop}}$ $\frac{\sigma^2}{P} \geqslant \max\left\{\frac{1}{N/P}-1, \frac{N}{P}-1\right\}$, and $R_{\mathrm{key}} < \frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right) - R_{\mathrm{LD,myop}}$ $\frac{\sigma^2}{P} \leqslant \frac{N}{P} - 1$	nR_{key}

standard Bachmann-Landau (Big-Oh) notation for asymptotic functions. All logarithms are to the base two.

We use $H(\cdot)$ to denote interchangeably Shannon entropy and differential entropy; the exact meaning will usually be clear from context.

V. PRELIMINARIES

A. Arbitrarily Varying Channel (AVC)

A channel with a state controlled by an adversarial jammer is called an AVC in the literature. James's jamming strategy is a (potentially probabilistic) map which, based on his observation, constructs an *attack vector* $\underline{s}$ satisfying his *maximum power constraint* $\|\underline{s}\|_2 \leq \sqrt{nN}$,

$$\begin{array}{rccc} \mathrm{Jam}: & \mathbb{R}^n & \to & \mathbb{R}^n \\ & \underline{z} & \mapsto & \underline{s} \end{array}.$$

B. Code

A *deterministic encoder* is a deterministic map which encodes a *message* to a *codeword* of length n, where n is called *block-length* or the *number of channel uses*, satisfying Alice's *maximum power constraint* $\|\underline{x}(m)\|_2 \leq \sqrt{nP}$,

$$\begin{array}{rccc} \mathrm{Enc}: & \{0,1\}^{nR} & \to & \mathbb{R}^n \\ & m & \mapsto & \underline{x}(m) \end{array},$$

where R is the *rate* of the system. Alice uses her encoder to encode the set of messages $\{0,1\}^{nR}$ and get a *codebook* $\{\underline{x}(m)\}_{m=1}^{2^{nR}}$ which is simply the collection of codewords.

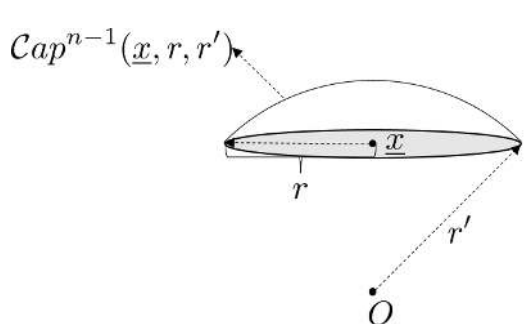
A *deterministic decoder* is a deterministic function which maps Bob's observation to a reconstruction of the message,

$$\begin{array}{rccc} \mathrm{Dec}: & \mathbb{R}^n & \to & \{0,1\}^{nR} \\ & \underline{y} & \mapsto & \widehat{m} \end{array}.$$

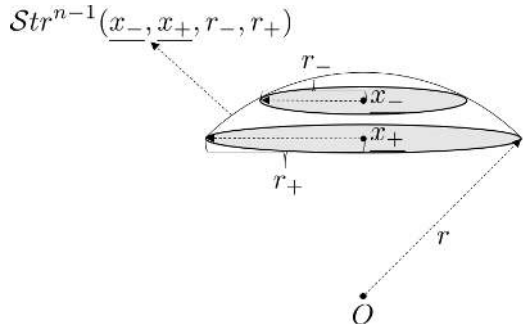
TABLE II

SUMMARY OF RESULTS FOR THE ADVERSARIAL CHANNEL WITH QUADRATIC CONSTRAINTS AND WIRETAP SECRECY. HERE, $\mathbb{1}_{\{P\geq N\}}$ IS 1 WHEN $P \geq N$ AND ZERO OTHERWISE. ALSO, $C_J := \frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right)$. PRIOR WORKS MOSTLY CONSIDERED THE CASE WHERE THE EAVESDROPPER IS INDEPENDENT OF THE JAMMER AND OBSERVES A DEGRADED VERSION OF $\underline{\mathbf{x}}$, WHILE THE JAMMER MUST CHOOSE HIS SIGNALS OBLIVIOUSLY OF $\underline{\mathbf{x}}$ AND $\mathbf{m}$

Reference	Rate	Level of myopia	Common randomness
Goldfeld et al. [34]	$C_{\text{obli,sec}} \geqslant \max\left\{\frac{1}{2}\log\left(1+\frac{P}{N}\right)-\frac{1}{2}\log\left(1+\frac{P}{\sigma^2_{\text{eve}}}\right), 0\right\}$ $C_{\text{obli,sec}} = 0$	$P > N$. Eavesdropper and jammer are independent of each other, and $\sigma^2 = \infty$. Here, σ^2_{eve} denotes noise variance to the eavesdropper. $P \leqslant N$ and $\sigma^2 = \infty$	0
This work (Lemma 18)	$C_{\text{myop,sec}} \geqslant R_{\text{LD}} - C_J$ $C_{\text{myop,sec}} \geqslant R_{\text{LD,myop}} - C_J$ $C_{\text{myop,sec}} = 0$ $C_{\text{myop,sec}} \geqslant R_{GV} - C_J$	$\frac{1}{1-N/P} - 1 \leqslant \frac{\sigma^2}{P} \leqslant \frac{1}{N/P} - 1$ $\frac{\sigma^2}{P} \geqslant \max\left\{\frac{1}{N/P}-1, \frac{1}{1-N/P}-1\right\}$ $\frac{\sigma^2}{P} \leqslant \frac{1}{1-N/P} - 2$ or $\frac{N}{P} \geqslant 1$ otherwise	0
This work (Lemma 17)	$C_{\text{myop}} = 0$ $C_{\text{myop,sec}} \geqslant R_{\text{LD,myop}} - C_J$ $C_{\text{myop,sec}} \geqslant R_{\text{LD}} - C_J$	$\frac{\sigma^2}{P} \leqslant \frac{N}{P} - 1$ $\frac{\sigma^2}{P} \geqslant \max\left\{\frac{1}{N/P}-1, 4\frac{N}{P}-1\right\}$ otherwise	$\Theta(\log n)$
This work (Lemma 16)	$C_{\text{myop}} = 0$ $C_{\text{myop,sec}} \geqslant \min\{R_{\text{LD,myop}}, R_{\text{LD,myop}} - C_J + R_{\text{key}}\}$ $C_{\text{myop,sec}} \geqslant \min\{R_{\text{LD}}, R_{\text{LD}} - C_J + R_{\text{key}}\}$	$\frac{\sigma^2}{P} \leqslant \frac{N}{P} - 1$ $\frac{\sigma^2}{P} \geqslant \max\left\{\frac{1}{N/P}-1, \frac{N}{P}-1\right\}$, and $R_{\text{key}} > \frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right) - R_{\text{LD,myop}}$ and $R_{\text{key}} < \frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right) - R_{\text{LD,myop}}$ otherwise	nR_{key}



(a) An $(n-1)$-dimensional cap $\mathcal{C}ap^{n-1}(\underline{x}, r, r')$ centered at $\underline{x}$ of radius r living on an $(n-1)$-dimensional sphere of radius r'.



(b) An $(n-1)$-dimensional strip $\mathcal{S}tr^{n-1}(\underline{x_-}, \underline{x_+}, r_-, r_+)$ centered at $\underline{x_-}$ and $\underline{x_+}$ of radii r_- and r_+.

Fig. 4. The geometry of caps and strips.

An (n, R, P, N) *deterministic code* $\mathcal{C}$ is a deterministic encoder-decoder pair (Enc, Dec). Sometimes we also slightly abuse the notation and call the set of codewords $\{\underline{x}(m)\}_{m=1}^{2^{nR}}$ a code.

We distinguish between three types of codes:

- *Deterministic codes*: The encoder is a deterministic map from $\{0,1\}^{nR}$ to $\mathbb{R}^n$, and the decoder is a deterministic map from $\mathbb{R}^n$ to $\{0,1\}^{nR}$.
- *Stochastic codes*: The encoder and decoder are allowed to use private randomness. If Alice and Bob have n_A and n_B bits of private randomness respectively, then a stochastic encoder is a map

$$\begin{array}{rccc} \text{Enc}: & \{0,1\}^{nR}\times\{0,1\}^{n_A} & \to & \mathbb{R}^n \\ & (m, k_A) & \mapsto & \underline{x}(m,k) \end{array},$$

while the decoder is a map

$$\begin{array}{rccc} \text{Dec}: & \mathbb{R}^n\times\{0,1\}^{n_B} & \to & \{0,1\}^{nR} \\ & (\underline{y}, k_B) & \mapsto & \widehat{m} \end{array}.$$

Here, k_A is known only to Alice while k_B is known only to Bob.

- *Randomized codes*: Alice and Bob share n_{key} bits of common randomness, which is kept secret from James. They may additionally have n_A and n_B bits of private randomness respectively. The encoder is a map

$$\begin{array}{rccc} \text{Enc}: & \{0,1\}^{nR}\times\{0,1\}^{n_A}\times\{0,1\}^{n_{\text{key}}} & \to & \mathbb{R}^n \\ & (m, k_A, k) & \mapsto & \underline{x}(m,k_A,k) \end{array},$$

while the decoder is a map

$$\begin{array}{rccc} \text{Dec}: & \mathbb{R}^n\times\{0,1\}^{n_B}\times\{0,1\}^{n_{\text{key}}} & \to & \{0,1\}^{nR} \\ & (\underline{y}, k_B, k) & \mapsto & \widehat{m} \end{array}.$$

Here, k is known to Alice and Bob, but not to James. The private randomness k_A is known only to Alice, while k_B is known only to Bob.

In all our code constructions, we will not use any private randomness, i.e., $n_A = n_B = 0$. However, our converse results are true for all values of n_A and n_B.

C. Probability of Error

A decoding error occurs when Bob's reconstruction does not match Alice's message. The *average* (over messages) *probability of error* $P_{\mathrm{e}}^{\mathrm{avg}}$ (also denoted by P_{e} for notational brevity in this paper) is defined as

$$P_{\mathrm{e}} = \sup_{\underline{\mathbf{s}}} \mathbb{P}(\widehat{\mathbf{m}} \neq \mathbf{m}) = \sup_{\underline{\mathbf{s}}} \sum_{m=1}^{2^{nR}} \mathbb{P}(\widehat{\mathbf{m}} \neq \mathbf{m}, \mathbf{m} = m)$$
$$= \sup_{\underline{\mathbf{s}}} \frac{1}{2^{nR}} \sum_{m=1}^{2^{nR}} \mathbb{P}(\widehat{\mathbf{m}} \neq m | \mathbf{m} = m).$$

where the probability is taken over the private randomness in the encoder-decoder pair, the common randomness shared by Alice and Bob, the noise to James, and any additional randomness he may use in choosing $\underline{\mathbf{s}}$. The maximization is taken over all (potentially stochastic) functions $\underline{\mathbf{s}}\colon \mathbb{R}^n \to \mathcal{B}^n(0, \sqrt{nN})$ which map James's observation $\underline{\mathbf{z}}$ to a jamming sequence $\underline{\mathbf{s}}(\underline{\mathbf{z}})$.[6] For deterministic codes,

$$\mathbb{P}(\widehat{\mathbf{m}} \neq m | \mathbf{m} = m)$$
$$= \int_{\mathcal{B}^n(0,\sqrt{nN})} \int_{\mathbb{R}^n} p_{\underline{\mathbf{z}}|\underline{\mathbf{x}}}(\underline{z}|\underline{x}(m)) p_{\underline{\mathbf{s}}|\underline{\mathbf{z}}}(\underline{s}|\underline{z}) \mathbb{1}_{\{\mathrm{Dec}(\underline{x}(m)+\underline{s}) \neq m\}} \, \mathrm{d}\underline{z} \, \mathrm{d}\underline{s}.$$

For stochastic codes,

$$\mathbb{P}(\widehat{\mathbf{m}} \neq m | \mathbf{m} = m)$$
$$= \frac{1}{2^{n_B}} \sum_{k_B=1}^{2^{n_B}} \int_{\mathcal{B}^n(0,\sqrt{nN})} \int_{\mathbb{R}^n} \frac{1}{2^{n_A}} \sum_{k_A=1}^{2^{n_A}} p_{\underline{\mathbf{z}}|\underline{\mathbf{x}}}(\underline{z}|\underline{x}(m,k_A)) p_{\underline{\mathbf{s}}|\underline{\mathbf{z}}}(\underline{s}|\underline{z})$$
$$\mathbb{1}_{\{\mathrm{Dec}(\underline{x}(m,k_A)+\underline{s},k_B) \neq m\}} \, \mathrm{d}\underline{z} \, \mathrm{d}\underline{s}.$$

For randomized codes,

$$\mathbb{P}(\widehat{\mathbf{m}} \neq m | \mathbf{m} = m)$$
$$= \frac{1}{2^{n_B}} \sum_{k_B=1}^{2^{n_B}} \int_{\mathcal{B}^n(0,\sqrt{nN})} \int_{\mathbb{R}^n} \frac{1}{2^{n_A}} \sum_{k_A=1}^{2^{n_A}} \frac{1}{2^{n_{\mathrm{key}}}} \sum_{k=1}^{2^{n_{\mathrm{key}}}}$$
$$p_{\underline{\mathbf{z}}|\underline{\mathbf{x}}}(\underline{z}|\underline{x}(m,k_A,k)) p_{\underline{\mathbf{s}}|\underline{\mathbf{z}}}(\underline{s}|\underline{z}) \mathbb{1}_{\{\mathrm{Dec}(\underline{x}(m,k_A,k)+\underline{s},k_B,k) \neq m\}} \, \mathrm{d}\underline{z} \, \mathrm{d}\underline{s}.$$

D. Rate and Capacity

A rate R is said to be *achievable* if there exists a sequence of (n, R, P, N) codes $\mathcal{C}^{(n)}$ labelled by block-length n such that each code in the sequence has rate $R^{(n)}$ at least R and average probability of error $P_{\mathrm{e}}^{(n)}$ vanishing in n, i.e.,

$$\forall n, \; R^{(n)} \geq R, \quad \text{and} \quad \lim_{n\to\infty} P_{\mathrm{e}}^{(n)} = 0.$$

The *capacity* C of a communication system is the supremum of all achievable rates.

[6]The output of the jammer can also depend on the (potentially stochastic) codebook used by Alice and Bob, and his own private randomness. However, it cannot depend on the common randomness shared only by the encoder-decoder pair. We omit these dependences in the notation for brevity.

E. List-Decoding

Definition 1: Fix $R > 0$ and $n_{\mathrm{key}} \geq 0$. A codebook $\mathcal{C} = \{\underline{x}(m,k) : m \in [2^{nR}], k \in [2^{n_{\mathrm{key}}}]\}$ is said to be (P, N, L)-list-decodable at rate R with n_{key} bits of common randomness if

- $\|\underline{x}(m,k)\|_2 \leq \sqrt{nP}$ for all m, k; and
- for all possible randomized functions $\underline{\mathbf{s}} := \underline{\mathbf{s}}(\mathcal{C}, \underline{\mathbf{x}})$ satisfying $\mathbb{P}(\|\underline{\mathbf{s}}\|_2 \leq \sqrt{nN}) = 1$, we have

$$\mathbb{P}(|\mathcal{B}^n(\underline{\mathbf{x}} + \underline{\mathbf{s}}, \sqrt{nN}) \cap \mathcal{C}^{(\mathbf{k})}| > L) = o(1),$$

where $\mathcal{C}^{(k)} := \{\underline{x}(m,k) : m \in [2^{nR}]\}$ and the shorthand notation $\underline{\mathbf{x}} = \underline{\mathbf{x}}(\mathbf{m}, \mathbf{k})$ is the (potentially stochastic) encoding of $\mathbf{m}$ under common randomness $\mathbf{k}$. In the above equation, the averaging is over the randomness in[7] $\mathbf{m}, \mathbf{k}, \underline{\mathbf{s}}_z$ and $\underline{\mathbf{s}}$.

A rate R is said to be achievable for (P, N, L)-list-decoding with n_{key} bits of common randomness if there exist sequences of codebooks (in increasing n) that are (P, N, L)-list-decodable. The list-decoding capacity is the supremum over all achievable rates.

Remark 1: An error in list-decoding can only occur if the list is too big, i.e., larger than L.

Remark 2: While sticking with maximum power constraint and average probability of error, there also exist other jamming power constraints and error criteria in the literature. They are not always equivalent and we strictly distinguish them in this remark.

1) Maximum vs. average power constraint for James. If we use an average power constraint for James, then no positive rate is achievable under both maximum and average probability of error. This is because James can focus only on a small constant fraction of codewords and allocate large enough power to completely corrupt them to ensure Bob has no hope to decode them. Then Bob's (maximum/average) probability of error is bounded away from zero while James's jamming strategy still satisfies his average power constraint. We will therefore only consider maximum power constraint on James.
2) Maximum vs. average probability of error. As we know, in information theory, an achievability result under maximum probability of error is stronger than that under average one, while it is the other way round for a converse result. In adversarial channel model, if we adopt maximum probability of error, notice that it suffices for James to corrupt the transmission of only one message. Thus we may assume that James knows the transmitted message a priori (but not necessarily the transmitted codeword if Alice uses stochastic encoding where a message potentially corresponds to a bin of codewords).

[7]Note that we are using an average probability of error in our work. This is different from a maximum probability of error, where the list-size is required to be less than or equal to L for every codeword. On the other hand, we are satisfied with this being true for all but a vanishingly small fraction of the codewords.

More formal justification of these criteria and their effect on capacity are given by Hughes and Narayan [8]. The authors defined the notion of λ-capacity for different criteria.

F. Probability Distributions of Interest

Alternatively, the system can be described using probability distributions. We assume the messages are uniformly distributed, i.e., $p_{\mathbf{m}} = \mathrm{Unif}([2^{nR}])$. Given the message to be transmitted, the codewords are distributed according to $p_{\underline{\mathbf{x}}|\mathbf{m}}$. Notice that it is not necessarily a 0-1 distribution, i.e., the codeword may not be entirely determined by the message, as Alice may have access to some private randomness and use *stochastic encoding*. Each message may be associated to many codewords and Alice samples one codeword according to the distribution $p_{\underline{\mathbf{x}}|\mathbf{m}}$ and transmits it. James receives a corrupted version $\underline{\mathbf{z}}$ of $\underline{\mathbf{x}}$ through an AWGN channel specified by

$$p_{\underline{\mathbf{z}}|\underline{\mathbf{x}}}(\underline{z}|\underline{x}) = p^{\otimes n}_{\mathbf{z}|\mathbf{x}}(\underline{z}|\underline{x}) = p^{\otimes n}_{\mathbf{z}|\mathbf{x}}(\underline{x} + (\underline{z} - \underline{x})|\underline{x}) = p^{\otimes n}_{\mathbf{s}_z}(\underline{z} - \underline{x}),$$

where $p^{\otimes n}_{\mathbf{s}_z} = \mathcal{N}(0, \sigma^2 \mathbf{I}_n)$. Based on his observation, James designs an attack vector $\underline{\mathbf{s}}$ according to his jamming strategy specified by $p_{\underline{\mathbf{s}}|\underline{\mathbf{z}}}$. Again, notice that it is not necessarily a 0-1 distribution as James may have access to private randomness and the output of his jamming may not be deterministic. Then Bob receives $\underline{\mathbf{y}}$ which is the sum of $\underline{\mathbf{x}}$ and $\underline{\mathbf{s}}$. In particular, $\underline{\mathbf{y}}$ is a deterministic function of the codeword transmitted in the main channel and the attack vector added to it, i.e.,

$$p_{\underline{\mathbf{y}}|\underline{\mathbf{x}},\underline{\mathbf{s}}}(\underline{y}|\underline{x}, \underline{s}) = \mathbb{1}_{\{\underline{y} = \underline{x} + \underline{s}\}}.$$

Based on his observation, Bob reconstructs $\widehat{\mathbf{m}}$ using a (potentially stochastic) decoder specified by $p_{\widehat{\mathbf{m}}|\underline{\mathbf{y}}}$.

G. Area and Volume

The area of an $(n-1)$-dimensional Euclidean sphere of radius r is given by

Fact 1:

$$\mathrm{Area}(\mathcal{S}^{n-1}(\cdot, r)) = \frac{2\pi^{n/2}}{\Gamma(n/2)} r^{n-1}.$$

The area of an $(n-1)$-dimensional cap centered at $\underline{x}$ of radius r living on an $(n-1)$-dimensional sphere of radius r' can be lower bounded by the volume of an $(n-1)$-dimensional ball centered at $\underline{x}$ of radius r since the intersection of an n-dimensional ball and an $(n-1)$-dimensional hyperplane is an $(n-1)$-dimensional ball, as shown in Figure 5, i.e.,

Fact 2:

$$\mathrm{Area}(\mathcal{C}ap^{n-1}(\underline{x}, r, r')) \geq \mathrm{Vol}(\mathcal{B}^{n-1}(\underline{x}, r)).$$

The area of a cap can also be upper bounded by a sphere of the same radius, i.e.,

Fact 3:

$$\mathrm{Area}(\mathcal{C}ap^{n-1}(\underline{x}, r, r')) \leq \frac{1}{2}\mathrm{Area}(\mathcal{S}^{n-1}(\underline{x}, r)) \leq \mathrm{Area}(\mathcal{S}^{n-1}(\underline{x}, r)).$$

The volume of an n-dimensional Euclidean ball of radius r is given by

Fact 4:

$$\mathrm{Vol}(\mathcal{B}^n(\cdot, r)) = \frac{\pi^{n/2}}{\Gamma(n/2+1)} r^n.$$

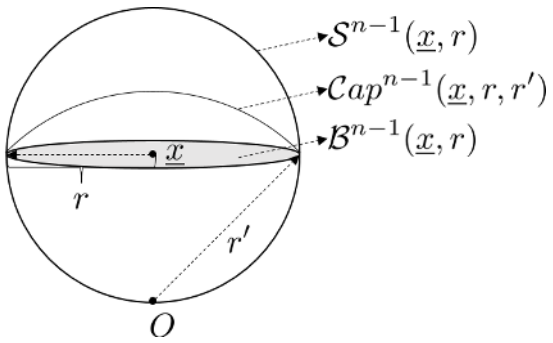


Fig. 5. Approximation of surface area. The surface area of a cap $\mathcal{C}ap^{n-1}(\underline{x}, r, r')$ is upper bounded by that of a sphere $\mathcal{S}^{n-1}(\underline{x}, r)$. It is lower bounded by the volume of a lower dimensional ball $\mathcal{B}^{n-1}(\underline{x}, r)$ since the intersection of an $(n-1)$-dimensional hyperplane parallel to the bottom of the cap passing through $\underline{x}$ and the ball $\mathcal{B}^n(O, r')$ whose surface the cap lives on is an $(n-1)$-dimensional ball $\mathcal{B}^{n-1}(\underline{x}, r)$.

More facts about high-dimensional geometry can be found in the notes by Ball [43].

H. Error Event Decomposition

We will frequently apply the following fact to decompose various decoding error events.

Fact 5: For any two events $\mathcal{A}$ and $\mathcal{B}$, we have $\mathbb{P}(\mathcal{A}) \leq \mathbb{P}(\mathcal{B}) + \mathbb{P}(\mathcal{A}|\mathcal{B}^c)$.

I. Basic Tail Bounds

Standard tail bounds (see, for instance, the monograph by Boucheron et al. [44]) for Gaussians and χ^2-distributions are used at times throughout this paper.

Fact 6: If $\mathbf{g} \sim \mathcal{N}(0, \sigma^2)$, then $\mathbb{P}(|\mathbf{g}| \geq \varepsilon) \leq 2\exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right)$.

Fact 7: If $\underline{\mathbf{g}} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_n)$, then $\|\underline{\mathbf{g}}\|_2^2$ has (scaled) χ^2-distribution and

$$\begin{aligned}
\mathbb{P}(\|\underline{\mathbf{g}}\|_2^2 \geq n\sigma^2(1+\varepsilon)) &\leq \exp\left(-\frac{\varepsilon^2}{4}n\right), \\
\mathbb{P}(\|\underline{\mathbf{g}}\|_2^2 \leq n\sigma^2(1-\varepsilon)) &\leq \exp\left(-\frac{\varepsilon^2}{2}n\right), \\
\mathbb{P}(\|\underline{\mathbf{g}}\|_2^2 \notin n\sigma^2(1\pm\varepsilon)) &\leq 2\exp\left(-\frac{\varepsilon^2}{4}n\right).
\end{aligned}$$

The following lemma is proved in Appendix B.

Lemma 8: Fix $\zeta > 0$ and $\underline{b} \in \mathbb{R}^n$. If $\underline{\mathbf{a}}$ is isotropically distributed on the sphere $\mathcal{S}^{n-1}(0, \|\underline{a}\|_2)$, then

$$\mathbb{P}(|\langle \underline{\mathbf{a}}, \underline{b}\rangle| > n\zeta) \leq 2^{-\frac{(n-1)n^2\zeta^2}{2\|\underline{a}\|_2^2\|\underline{b}\|_2^2}}.$$

We now state a general lemma that will be useful in proving list-decoding results. This essentially says that if codewords are chosen i.i.d. according to a uniform distribution, then the probability that more than $\mathcal{O}(n^2)$ codewords lie within any set of sufficiently small volume is super-exponentially decaying in n. A proof of this lemma for discrete case appeared in Langberg's [45, Lemma 2.1] paper. A proof can be found in Appendix B.

Lemma 9: Suppose $A \subseteq \mathbb{R}^n$ Lebesgue measurable and $\mathcal{C} = \{\underline{x}(m)\}_{m=1}^{2^{nR}} \subseteq A$ contains $2^{nR}, R > 0$ uniform samples from A. If $V \subseteq A$ is Lebesgue measurable and for some small enough constant $\nu > 0$

$$p := \mathbb{P}(\underline{\mathbf{x}} \in V) = \frac{\mu(V)}{\mu(A)} \leq 2^{-(R+\nu)n}, \tag{V.1}$$

where $\underline{\mathbf{x}}$ is sampled uniformly at random from A and $\mu(\cdot)$ denotes the Lebesgue measure of a measurable subset of $\mathbb{R}^n$, then there exists some constant $C = C(c, \nu, R) > 0$, s.t.,

$$\mathbb{P}(|V \cap \mathcal{C}| \geq cn^2) \leq 2^{-Cn^3}.$$

Remark 3: The above lemma can be generalized in a straightforward manner to the case where the measure is, instead of Lebesgue measure, the unique translation-invariant measure on the sphere, i.e., the Haar measure. This variant will be the version that we invoke later in the proof.

Remark 4: The above lemma indicates that the number of codewords falling into V is small (say at most $\mathrm{poly}(n)$) with high probability as long as we can get an exponentially small upper bound on the expected number of such codewords (note that the condition given by Eqn. (V.1) implies $\mathbb{E}(|V \cap \mathcal{C}|) = p2^{nR} \leq 2^{-\nu n}$).

VI. Formal Statements of Main Results

We now formally state the results in this paper with respect to decreasing values of n_{key}. We start with the case when there is an unlimited amount of common randomness available between Alice and Bob.

Lemma 10 ([1]): The capacity of the myopic adversarial channel with an unlimited amount of common randomness is

$$C_{\mathrm{myop,rand}} = \begin{cases} R_{\mathrm{LD}}, & \frac{\sigma^2}{P} \leq \frac{1}{N/P} - 1 \\ R_{\mathrm{LD,myop}}, & \frac{\sigma^2}{P} \geq \max\left\{\frac{1}{N/P} - 1, \frac{N}{P} - 1\right\} \\ 0, & \frac{\sigma^2}{P} \leq \frac{N}{P} - 1, \end{cases} \tag{VI.1}$$

where

$$R_{\mathrm{LD}} := \frac{1}{2}\log\frac{P}{N},$$
$$R_{\mathrm{LD,myop}} := \frac{1}{2}\log\left(\frac{(P+\sigma^2)(P+N) - 2P\sqrt{N(P+\sigma^2)}}{N\sigma^2}\right).$$

These are summarized in Fig. 8.

Proof: The achievability was proved in [1] and the converse is proved in Sec. VII-A. See Sec. VII for details. □

Remark 5: The capacities in Eqn. (VI.1) are continuous at the boundaries of different parameter regimes. Indeed, it is easy to check that

$$R_{\mathrm{LD,myop}} = \begin{cases} R_{\mathrm{LD}}, & \frac{\sigma^2}{P} = \frac{1}{N/P} - 1 \\ 0, & \frac{\sigma^2}{P} = \frac{N}{P} - 1. \end{cases}$$

In fact, in all our results, whenever we have characterizations on both sides of a certain boundary, the capacities are continuous at the boundary.

Remark 6: It is worth noting that $R_{\mathrm{LD}} \leq R_{\mathrm{LD,myop}}$ in all parameter regimes.

We now discuss two possibilities: (1) n_{key} is $\Theta(n)$, and (2) n_{key} is $\Theta(\log n)$. The rate given by Lemma 10 is an upper bound on the capacity in both cases, and we show that this is also achievable in a subregion of the NSRs. The proof will involve a myopic list-decoding argument, which we state below. We will combine this with a known technique [37], [41], [42] which uses $\Theta(\log n)$ bits of common randomness to disambiguate the list and give us a result for unique decoding.

Theorem 11: For $(P, N, \mathcal{O}(n^2))$-list-decoding, the capacity is lower bounded as follows

$$C_{\mathrm{myop,LD}} \geq \begin{cases} R_{\mathrm{LD,myop}}, & \text{if } \frac{\sigma^2}{P} \geq \max\left\{\frac{1}{N/P} - 1, \frac{N}{P} - 1\right\} \\ & \text{and } R_{\mathrm{LD,myop}} + R_{\mathrm{key}} > \frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right) \\ R_{\mathrm{LD}}, & \text{otherwise.} \end{cases}$$

These are summarized in Fig. 7.

Proof: See Sec. VIII-B for a proof sketch and Sec. X for details. □

The rate R_{LD} is achievable even in the presence of an omniscient adversary. A major contribution of this work is in showing that myopia indeed does help, and we can obtain a higher rate of $R_{\mathrm{LD,myop}}$ in a certain regime. It is interesting to note that even when $n_{\mathrm{key}} = 0$, the myopic list-decoding capacity is nonzero for sufficiently large values of σ^2/P. Furthermore, increasing rates of common randomness (R_{key}) help us achieve higher list-decoding rates as seen from Fig. 7.

The above theorem is crucially used in proving the following result for linear amount of common randomness:

Lemma 12: Fix any $R_{\mathrm{key}} > 0$. If Alice and Bob share nR_{key} bits of common randomness, then the capacity is

$$C_{\mathrm{myop}} = \begin{cases} R_{\mathrm{LD,myop}}, & \text{if } \frac{\sigma^2}{P} \geq \max\left\{\frac{1}{N/P} - 1, \frac{N}{P} - 1\right\} \\ & \text{and } R_{\mathrm{key}} > \frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right) - R_{\mathrm{LD,myop}} \\ R_{\mathrm{LD}}, & \frac{\sigma^2}{P} \leq \frac{1}{N/P} - 1 \\ 0, & \frac{\sigma^2}{P} \leq \frac{N}{P} - 1. \end{cases}$$

Furthermore,

$$R_{\mathrm{LD}} \leq C_{\mathrm{myop}} \leq R_{\mathrm{LD,myop}},$$

if

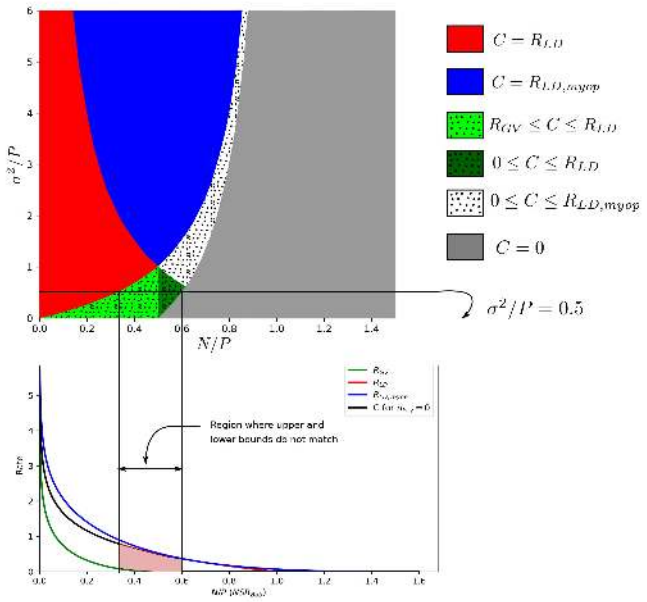
$$\frac{\sigma^2}{P} \geq \max\left\{\frac{1}{N/P} - 1, \frac{N}{P} - 1\right\},$$

and

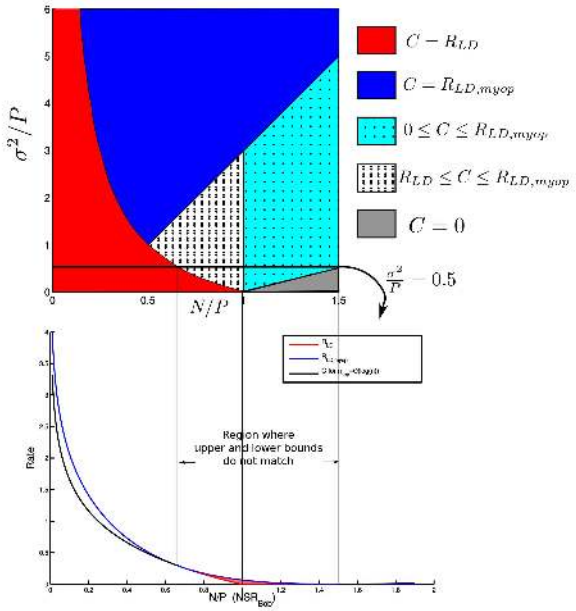
$$R_{\mathrm{key}} < \frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right) - R_{\mathrm{LD,myop}}.$$

Proof: See Sec. VIII-C. □

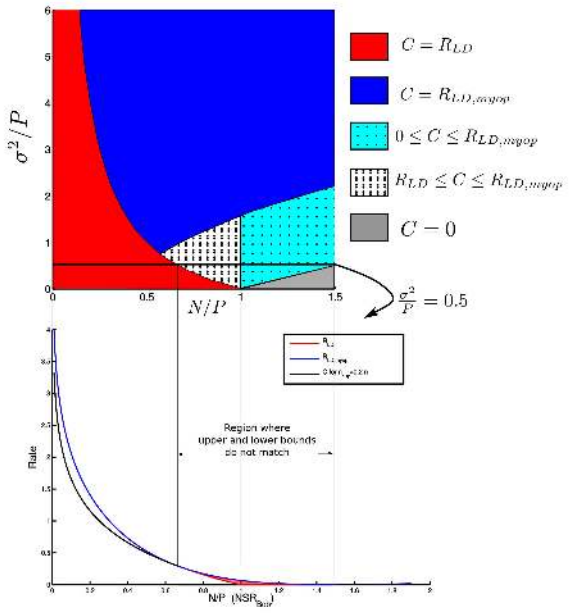
The rates achievable for different values of the NSRs are illustrated in Fig. 6c and Fig. 6d for different values of R_{key}. Our scheme achieves capacity in the red and blue regions of Fig. 6c and 6d. In the white dotted region, R_{LD} is a lower bound on the capacity, as guaranteed by Lemma 12. However, we can achieve $R_{\mathrm{LD,myop}}$ with an infinite amount of common randomness. Hence, there is a small gap between the upper and lower bounds in this region. In the cyan dotted region, $R_{\mathrm{LD}} < 0$ and our lower bound is trivial, while the converse says that the capacity is upper bounded by $R_{\mathrm{LD,myop}}$.



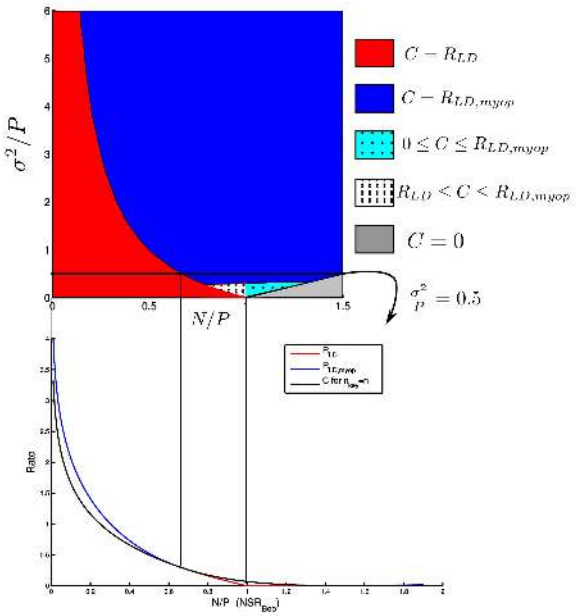
(a) Capacity bounds for the myopic channel with no common randomness. Here, $R_{\mathrm{GV}} := \frac{1}{2}\log\left(\frac{P^2}{4N(P-N)}\right)1_{\{P \gtrsim 2N\}}$.



(b) Capacity bounds with $n_{\mathrm{key}} = \Theta(\log n)$. We have a complete characterization in the solid regions but nonmatching upper and lower bounds in the dotted regions.



(c) An example of our results when $n_{\mathrm{key}} = \Theta(n)$: Capacity bounds with $n_{\mathrm{key}} = 0.2n$.



(d) An example of our results when $n_{\mathrm{key}} = \Theta(n)$: Capacity bounds for $n_{\mathrm{key}} = n$.

Fig. 6. Capacity bounds for different values of n_{key}.

As remarked earlier, we only require $\Theta(\log n)$ bits of common randomness to disambiguate the list (in other words, go from a list-decodable code to a uniquely decodable code). The additional $nR_{\mathrm{key}} - \Theta(\log n)$ bits are used for additional randomization at the encoder to "confuse" James. As is evident from Figures 6c and 6d, and the following lemma, larger values of R_{key} can guarantee that $R_{\mathrm{LD,myop}}$ is achievable in a larger range of NSRs.

Using Theorem 11 for $R_{\mathrm{key}} = 0$, we can show Lemma 13. Note that when $R_{\mathrm{key}} = 0$, the condition $R_{\mathrm{LD,myop}} + R_{\mathrm{key}} > \frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right)$ reduces to $\frac{\sigma^2}{P} \ge 4\frac{N}{P} - 1$.

Lemma 13: When $\log n < n_{\mathrm{key}} = \mathcal{O}(\log n)$, the capacity of the myopic adversarial channel is:

$$C_{\mathrm{myop}} = \begin{cases} R_{\mathrm{LD,myop}}, & \text{if } \frac{\sigma^2}{P} \ge \max\left\{\frac{1}{N/P} - 1, 4\frac{N}{P} - 1\right\} \\ R_{\mathrm{LD}}, & \frac{\sigma^2}{P} \le \frac{1}{N/P} - 1 \\ 0, & \frac{\sigma^2}{P} \le \frac{N}{P} - 1. \end{cases}$$

Furthermore,

$$R_{\mathrm{LD}} \le C_{\mathrm{myop}} \le R_{\mathrm{LD,myop}},$$

if

$$\max\left\{\frac{1}{N/P}-1,\frac{N}{P}-1\right\}\le\frac{\sigma^2}{P}\le 4\frac{N}{P}-1.$$

These results are summarized in Fig. 6b.

Proof: See Sec. VIII-D. □

These results are illustrated in Fig. 6b. We have matching upper and lower bounds in the red, blue, and grey regions. As in the previous subsection, there is a nontrivial gap between the upper and lower bounds in the green and white regions.

When Alice and Bob do not have access to a shared secret key, the achievability proofs in Sec. VIII-C and VIII-D are not valid, and for certain values of the NSRs, tighter converses can be obtained. In this scenario, we will prove the following result.

Theorem 14: The capacity of the myopic adversarial channel with no common randomness is given by

$$C_{\text{myop}}=\begin{cases}R_{\text{LD}}, & \text{if } \frac{1}{1-N/P}-1\le\frac{\sigma^2}{P}\le\frac{1}{N/P}-1\\ R_{\text{LD,myop}}, & \text{if } \frac{\sigma^2}{P}\ge\max\left\{\frac{1}{1-N/P}-1,\frac{1}{N/P}-1\right\}\\ 0, & \text{if } \frac{\sigma^2}{P}\le\frac{1}{1-N/P}-2 \text{ or } \frac{N}{P}\ge 1.\end{cases}$$

In the other regimes, we have

$$R_{\text{GV}}\le C_{\text{myop}}\le \begin{cases}R_{\text{LD}}, & \text{if } \frac{1}{1-N/P}-2\le\frac{\sigma^2}{P}\\ & \le\min\left\{\frac{1}{N/P}-1,\frac{1}{1-N/P}-1\right\}\\ R_{\text{LD,myop}}, & \text{if } \max\left\{\frac{1}{N/P}-1,\frac{1}{1-N/P}-2\right\}\\ & \le\frac{\sigma^2}{P}\le\frac{1}{1-N/P}-1 \text{ and } \frac{N}{P}\le 1.\end{cases}$$

These are summarized in Fig. 6a.

Proof: The achievability follows by combining the myopic list-decoding lemma (Theorem 11 which is proved in Sec. X) and a trick used in [25] for myopic list disambiguation. We adapt the trick to the quadratically constrained case dealt with in this paper and the proof is presented in Sec. XI.

The converse follows by combining the scale-and-babble converse (proved in Sec. VII-A) for Lemma 10 and a symmetrization converse proved in Sec. IX-B.

See Sec. IX-A for an overview of the proof structure. □

A. Wiretap Secrecy

We will show that a simple modification can guarantee secrecy when James also wants to eavesdrop on the message. In addition to achieving a vanishingly small probability of error at the decoder, we wish to ensure that the information leaked to James is vanishingly small, *i.e.*, $I(\underline{\mathbf{x}};\underline{\mathbf{z}})\to 0$ as $n\to\infty$. This can be easily guaranteed by a wiretap code, and we briefly describe the modifications required to ensure this.

When Alice and Bob share infinite common randomness, $nC_{\text{myop,rand}}$ bits of secret key can be used as a one-time pad. In fact, the one-time pad guarantees perfect secrecy: $I(\mathbf{m};\underline{\mathbf{z}})=0$ for all n. In this case, the secrecy capacity can be completely characterized and is equal to $C_{\text{myop,rand}}$.

Lemma 15: The secrecy capacity of the myopic adversarial channel with an unlimited amount of common randomness is

$$C_{\text{myop,rand,sec}}=\begin{cases}R_{\text{LD}}, & \frac{\sigma^2}{P}\le\frac{1}{N/P}-1\\ R_{\text{LD,myop}}, & \frac{\sigma^2}{P}\ge\max\left\{\frac{1}{N/P}-1,\frac{N}{P}-1\right\}\\ 0, & \frac{\sigma^2}{P}\le\frac{N}{P}-1.\end{cases}\tag{VI.2}$$

In the regime where $n_{\text{key}}=\Theta(n)$, we use a standard random binning scheme analogous to [35], [38], [46], [47]. The result of secrecy then essentially follows from [38, Lemma 2]. We give sketch of the proof of the following lemmas in Appendix H.

Let $C_J:=\frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right)$. The following rates are achievable:

Lemma 16: If Alice and Bob share nR_{key} bits of common randomness, then the secrecy capacity is

$$C_{\text{myop,sec}}\begin{cases}\ge\min\{R_{\text{LD,myop}},R_{\text{LD,myop}}-C_J+R_{\text{key}}\},\\ \quad\text{if } \frac{\sigma^2}{P}\ge\max\left\{\frac{1}{N/P}-1,\frac{N}{P}-1\right\}\\ \quad\text{and } R_{\text{key}}>\frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right)-R_{\text{LD,myop}}\\ =0, \quad \text{if } \frac{\sigma^2}{P}\le\frac{N}{P}-1\\ \ge\min\{R_{\text{LD}},R_{\text{LD}}-C_J+R_{\text{key}}\}, \quad \text{otherwise.}\end{cases}$$

The results of Lemma 16 are pictorially summarized in Figs. 17c and 17d. Positive rates are guaranteed in the red and blue regions, as in Figs. 6c and 6d.

The analysis above easily extends to the case where $n_{\text{key}}=\Theta(\log n)$ and $n_{\text{key}}=0$. We can obtain the following results.

Lemma 17: When $n_{\text{key}}=\Theta(\log n)$, the secrecy capacity of the myopic adversarial channel is:

$$C_{\text{myop,sec}}\begin{cases}\ge R_{\text{LD,myop}}-C_J, & \text{if } \frac{\sigma^2}{P}\ge\max\left\{\frac{1}{N/P}-1,4\frac{N}{P}-1\right\}\\ =0, & \frac{\sigma^2}{P}\le\frac{N}{P}-1\\ \ge R_{\text{LD}}-C_J, & \text{otherwise.}\end{cases}$$

These results are summarized in Fig. 17a.

Lemma 18: The secrecy capacity of the myopic adversarial channel with no common randomness is given by

$$C_{\text{myop,sec}}\begin{cases}\ge R_{\text{LD}}-C_J, & \text{if } \frac{1}{1-N/P}-1\le\frac{\sigma^2}{P}\le\frac{1}{N/P}-1\\ \ge R_{\text{LD,myop}}-C_J, & \text{if } \frac{\sigma^2}{P}\ge\max\left\{\frac{1}{1-N/P}-1,\frac{1}{N/P}-1\right\}\\ =0, & \text{if } \frac{\sigma^2}{P}\le\frac{1}{1-N/P}-2 \text{ or } \frac{N}{P}\ge 1\\ \ge R_{\text{GV}}-C_J, & \text{otherwise.}\end{cases}$$

The above results are pictorially depicted in Fig. 17b.

We now discuss the proof techniques used to derive our main results.

VII. Lemma 10: Capacity of the Myopic Adversarial Channel When $R_{\text{key}}=\infty$

In this section, we prove Lemma 10. The achievability part uses a random coding argument and can be found in [1]. The

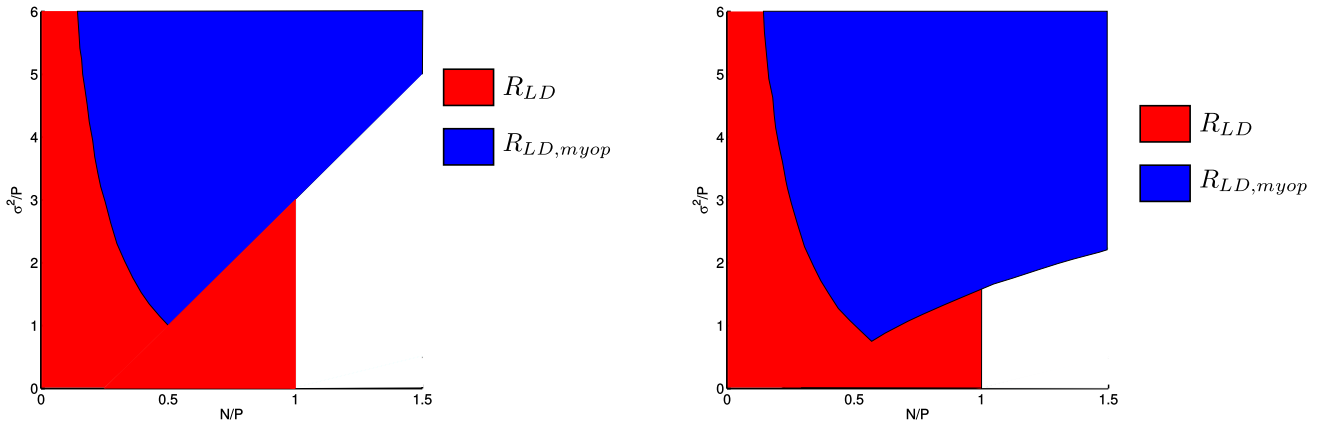


Fig. 7. Achievable rates for myopic list-decoding for 7a $n_{\text{key}} = 0$ and 7b $n_{\text{key}} = 0.2n$ bits of common randomness.

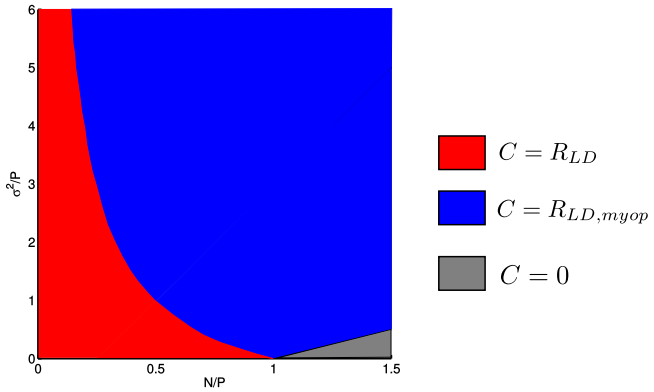


Fig. 8. Capacity of the myopic adversarial channel for $n_{\text{key}} = \infty$ [1]. The x-axis denotes the NSR from Alice to Bob (with noise from James) and the y-axis denotes NSR from Alice to James (with AWGN added).

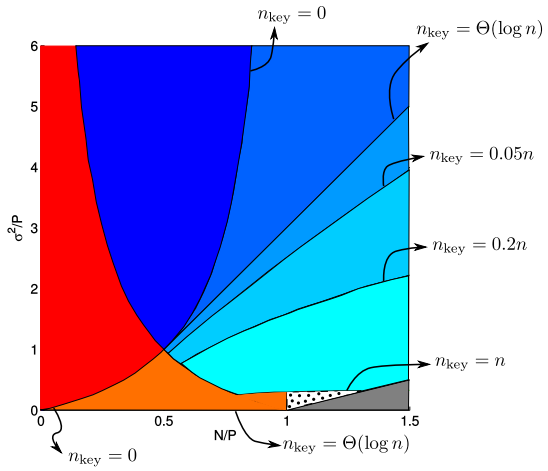


Fig. 9. Expansion of the achievable rate regions for the myopic adversarial channel with different amounts of common randomness. A rate of R_{LD} is achievable in the red and orange regions, whereas $R_{\text{LD,myop}}$ is achievable in the regions with different shades of blue. With $\Omega(n)$ bits of common randomness, a rate of at least R_{LD} is achievable whenever $P > N$.

converse can be proved by specifying an attack for James that instantiates an AWGN channel from Alice to Bob having the capacity given by (VI.1). We now give a proof of the converse (which was omitted in [1]). The bounds that we

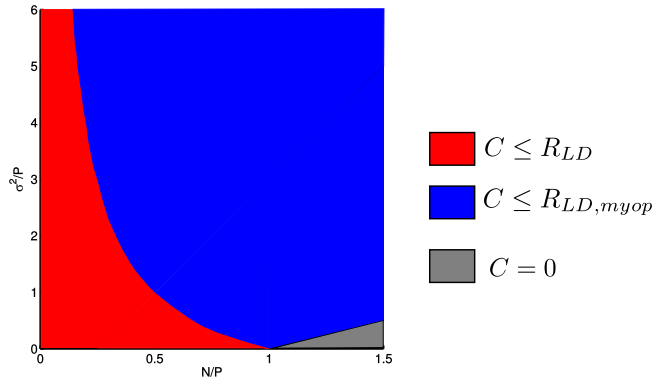


Fig. 10. Upper bounds on capacity as obtained from the scale-and-babble attack. These outer bounds are valid for all values of n_{key}.

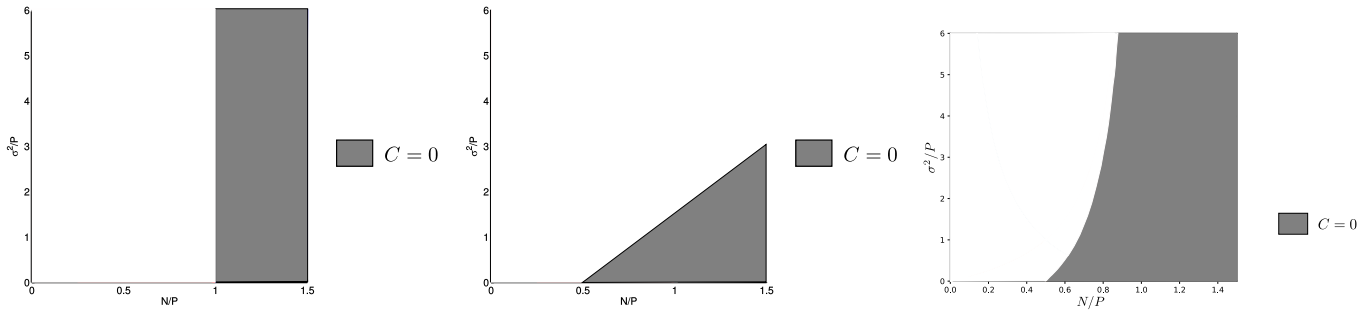
obtain in the following subsection are tight. Coupled with the achievability in [1], we have a complete characterization of the capacity. This also gives us some insight as to what optimal attack strategies for James might be, and hence give a detailed argument.

A. Proof of Converse: "Scale-and-Babble" Attack

In this subsection, we prove the converse part of Lemma 10. The converse involves what we call a "scale-and-babble" attack strategy for James. This attack essentially converts the myopic channel into an equivalent AWGN channel. Since the capacity of the AWGN channel cannot be increased using common randomness, this gives us an upper bound on the capacity for all values of n_{key}. We make no claim about the originality of this proof, as the strategy is well-known (and maybe dates back to Blachman [4], as suggested in [1]. This was also implicitly used in [14]). But we nevertheless provide the details to keep the paper self-contained.

The strategy for James is the following: He uses a certain fraction of his power to subtract a negatively scaled version of his observation; the remainder of his power is used to add AWGN. Specifically,

$$\underline{\mathbf{s}} = \boldsymbol{\beta}(-\alpha\underline{\mathbf{z}} + \underline{\mathbf{g}}) = \boldsymbol{\beta}(-\alpha(\underline{\mathbf{x}} + \underline{\mathbf{s}}_z) + \underline{\mathbf{g}}),$$



(a) (b) (c)

Fig. 11. Regions where the 11a $\underline{\mathbf{z}}$-agnostic symmetrization argument, 11b $\underline{\mathbf{z}}$-aware symmetrization argument and 11c improved $\underline{\mathbf{z}}$-aware symmetrization argument gives zero capacity. Note that these are valid only for $n_{\text{key}} = 0$. However, they continue to hold even if the encoder has access to private randomness.

where[8]

$$\boldsymbol{\beta} = \begin{cases} 1, & \| -\alpha\underline{\mathbf{z}} + \underline{\mathbf{g}}\|_2 \le \sqrt{nN} \\ \frac{\sqrt{nN}}{\| -\alpha\underline{\mathbf{z}} + \underline{\mathbf{g}}\|_2} =: \boldsymbol{\beta}', & \text{otherwise,} \end{cases}$$

$\alpha > 0$ is a constant to be optimized subject to $N - \alpha^2(P + \sigma^2) \ge 0$, i.e., $\alpha \le \sqrt{N/(P+\sigma^2)}$. Also, $\underline{\mathbf{g}} \sim \mathcal{N}(0, \gamma^2 \mathbf{I}_n)$ with $\alpha^2(P+\sigma^2) + \frac{\gamma^2}{1-\varepsilon} = N$, for a small $\varepsilon > 0$. Therefore, $\gamma^2 = (N - \alpha^2(P+\sigma^2))(1-\varepsilon)$, and $\mathbb{P}(\boldsymbol{\beta} \ne 1) = \mathbb{P}(\| -\alpha\underline{\mathbf{z}} + \underline{\mathbf{g}}\|_2 > \sqrt{nN}) = 2^{-\Omega(n)}$. Then

$$\begin{aligned} \underline{\mathbf{y}} &= \underline{\mathbf{x}} + \underline{\mathbf{s}} \\ &= \begin{cases} \underline{\mathbf{x}} - \alpha(\underline{\mathbf{x}} + \underline{\mathbf{s}}_z) + \underline{\mathbf{g}}, & \| -\alpha\underline{\mathbf{z}} + \underline{\mathbf{g}}\|_2 \le \sqrt{nN} \\ \underline{\mathbf{x}} + \boldsymbol{\beta}'(-\alpha(\underline{\mathbf{x}} + \underline{\mathbf{s}}_z) + \underline{\mathbf{g}}), & \text{otherwise} \end{cases} \\ &= \begin{cases} (1-\alpha)\underline{\mathbf{x}} - \alpha\underline{\mathbf{s}}_z + \underline{\mathbf{g}}, & \| -\alpha\underline{\mathbf{z}} + \underline{\mathbf{g}}\|_2 \le \sqrt{nN} \\ (1 - \boldsymbol{\beta}'\alpha)\,\underline{\mathbf{x}} - \boldsymbol{\beta}'\alpha\underline{\mathbf{s}}_z + \boldsymbol{\beta}'\underline{\mathbf{g}}, & \text{otherwise.} \end{cases} \end{aligned} \tag{VII.1}$$

The scaling factor $\boldsymbol{\beta}$ is introduced to make the attack vector satisfy the power constraint. Note that the channel above is not exactly an AWGN channel. However, the probability that $\boldsymbol{\beta} \ne 1$ is exponentially small in n. We have the following claim, formally proved in Appendix E.

Claim 19: Fix $0 < \varepsilon < 1$, and let α be nonnegative. If $\gamma^2 = (N - \alpha^2(P+\sigma^2))(1-\varepsilon)$, then the capacity C_{AWGN} of the channel (VII.1) from $\underline{\mathbf{x}}$ to $\underline{\mathbf{y}}$ is upper bounded as follows

$$C_{\text{AWGN}} < \frac{1}{2}\log\left(1 + \frac{(1-\alpha)^2 P}{\alpha^2\sigma^2 + \gamma^2}\right). \tag{VII.2}$$

Using Claim 19, we have

$$\begin{aligned} C_{\text{myop,rand}} &< \frac{1}{2}\log\left(1 + \frac{(1-\alpha)^2 P}{\alpha^2\sigma^2 + \gamma^2}\right) \\ &= \frac{1}{2}\log\left(1 + \frac{(1-\alpha)^2 P}{N - \alpha^2 P}\right) + g(\varepsilon, P, N, \sigma^2, \alpha), \end{aligned}$$

where $g(\varepsilon, P, N, \sigma^2, \alpha) \to 0$ as $\varepsilon \to 0$. Since this holds for every $\varepsilon > 0$, and every $0 < \alpha \le \sqrt{N/(P+\sigma^2)}$, we can say that

$$C_{\text{myop,rand}} < \min_{0<\alpha\le\sqrt{N/(P+\sigma^2)}} \frac{1}{2}\log\left(1 + \frac{(1-\alpha)^2 P}{N - \alpha^2 P}\right). \tag{VII.3}$$

[8]Here, $\boldsymbol{\beta}$ is a factor introduced to handle atypicality in the noise. As we show subsequently, the value of $\boldsymbol{\beta}$ is 1 with high probability.

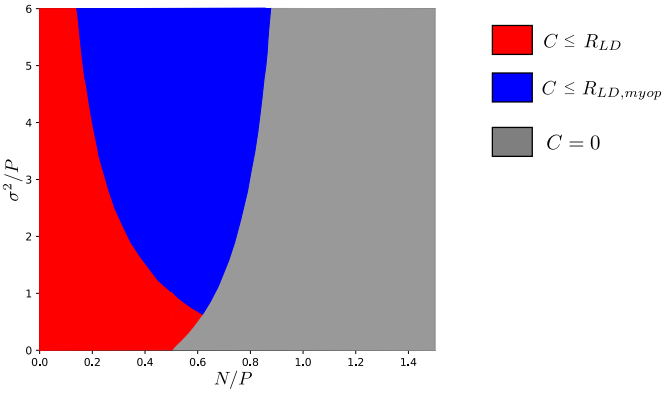


Fig. 12. Upper bounds on capacity for $n_{\text{key}} = 0$. This is obtained by combining the bounds obtained using the scale-and-babble attack in Fig. 10 with the symmetrization arguments in Fig. 11.

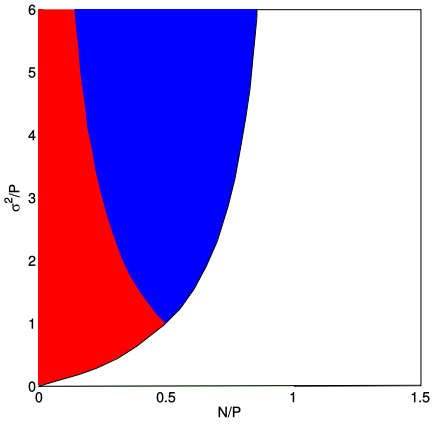


Fig. 13. Currently, our reverse list-decoding arguments, an integral part of our proof techniques for $n_{\text{key}} = 0$, are only valid in the red and blue regions, which present bottlenecks in our analysis of the scenario when no common randomness is available. As ongoing work, we are trying to expand the region where reverse list-decoding is possible.

Denote $f(\alpha) := \frac{(1-\alpha)^2 P}{N - \alpha^2 P}$ and $R(\alpha) := \frac{1}{2}\log(1 + f(\alpha))$. The minimum points of $f(\alpha)$ are $\alpha = N/P$ and $\alpha = 1$.

1) $N/P \ge 1$ (See Figure 18(a) with $N/P = 1.1$).

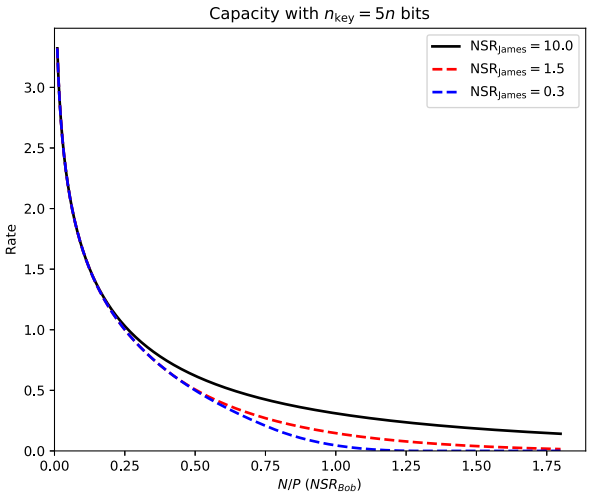


Fig. 14. Capacity of the myopic adversarial channel with $\Omega(n)$ bits of common randomness. Here $\mathrm{NSR}_{\mathrm{James}} := \sigma^2/P$.

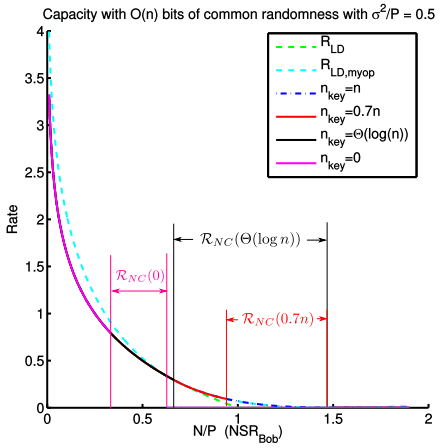


Fig. 15. Capacity of the myopic adversarial channel with different amounts of common randomness. Here, $\mathcal{R}_{NC}(n_{\mathrm{key}})$ denotes the region where our upper and lower bounds do not meet.

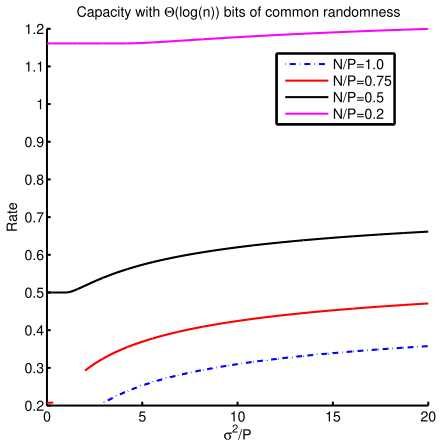


Fig. 16. Capacity of the myopic adversarial channel as a function of σ^2/P.

a) If $N/P < \sqrt{N/(P+\sigma^2)}$, i.e., $\frac{\sigma^2}{P} < \frac{1}{N/P} - 1$, then

$$\min_{0<\alpha\le\sqrt{N/(P+\sigma^2)}} R(\alpha) = \frac{1}{2}\log(1+f(N/P)) = \frac{1}{2}\log\frac{P}{N}.$$

b) If $N/P \ge \sqrt{N/(P+\sigma^2)}$, i.e., $\frac{\sigma^2}{P} \ge \frac{1}{N/P} - 1$, then

$$\min_{0<\alpha\le\sqrt{N/(P+\sigma^2)}} R(\alpha) = \frac{1}{2}\log\left(1+f\left(\sqrt{\frac{N}{P+\sigma^2}}\right)\right)$$
$$= \frac{1}{2}\log\left(\frac{(P+\sigma^2)(P+N)-2P\sqrt{N(P+\sigma^2)}}{N\sigma^2}\right)$$
$$=: R_{\mathrm{LD,myop}}.$$

Notice that if $\sigma \to \infty$, the channel becomes oblivious. It can be directly verified that

$$\lim_{\sigma\to\infty}\frac{1}{2}\log\left(\frac{(P+\sigma^2)(P+N)-2P\sqrt{N(P+\sigma^2)}}{N\sigma^2}\right)$$
$$= \frac{1}{2}\log\left(1+\frac{P}{N}\right),$$

consistent with the oblivious capacity of quadratically constrained channels.

2) $N/P < 1$ (See Figure 18b with $N/P = 0.9$).

a) If $\sqrt{N/(P+\sigma^2)} \ge 1$, i.e., $\frac{\sigma^2}{P} \le \frac{N}{P} - 1$, then

$$\min_{0<\alpha\le\sqrt{N/(P+\sigma^2)}} R(\alpha) = \frac{1}{2}\log(1+f(1)) = 0.$$

b) If $\sqrt{N/(P+\sigma^2)} < 1$, i.e., $\frac{\sigma^2}{P} > \frac{N}{P} - 1$, then

$$\min_{0<\alpha\le\sqrt{N/(P+\sigma^2)}} R(\alpha)$$
$$= \frac{1}{2}\log\left(1+f\left(\sqrt{\frac{N}{P+\sigma^2}}\right)\right) = R_{\mathrm{LD,myop}}.$$

The upper bound on the capacity given by the scale-and-babble attack is shown in Figure 19.

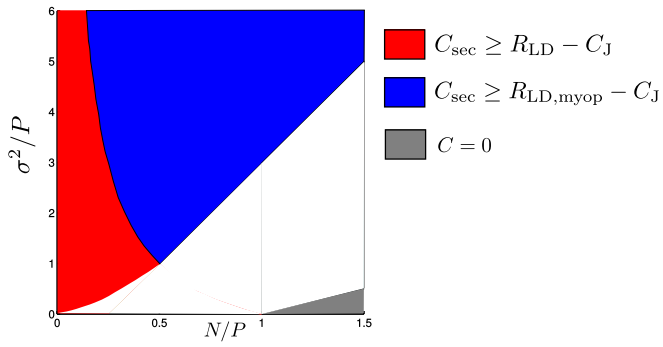
The scale-and-babble attack converts the adversarial channel into an equivalent AWGN channel. The capacity of the point-to-point AWGN channel cannot be increased using private/common randomness. Therefore, the upper bounds obtained using this technique hold regardless of whether deterministic/stochastic/randomized codes are used, and regardless of the amount of common randomness shared by the encoder and decoder.[9]

VIII. Lemmas 12 and 13: Linear and Sublinear Amounts of Common Randomness

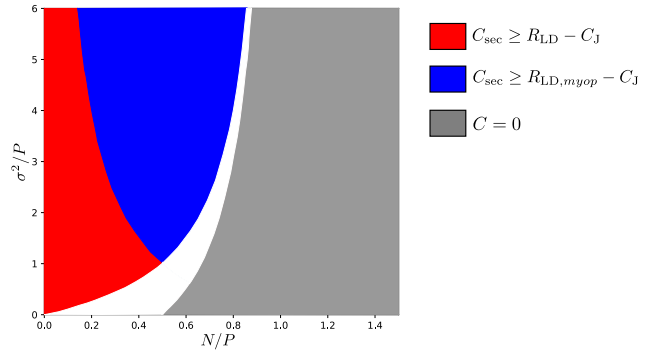
Our approach in these two regimes will involve a myopic list-decoding argument, which we will prove next. We will combine this with a known technique [37], [41], [42] which uses $\Theta(\log n)$ bits of common randomness to disambiguate the list and give us a result for unique decoding.[10] Before we state the main results, we take a brief detour to discuss classical and myopic list-decoding.

[9]This is due to the fact that common randomness/stochastic encoding does not increase the capacity of a memoryless channel. This in turn can be derived from Fano's inequality with common randomness (see Appendix E for more details).

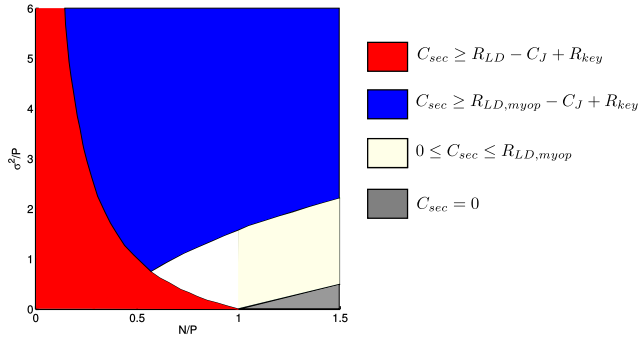
[10]For the case $\sigma^2 = 0$, [42] also showed that $\log n$ bits are necessary to achieve a rate equal to list-decoding capacity with unique decoding.



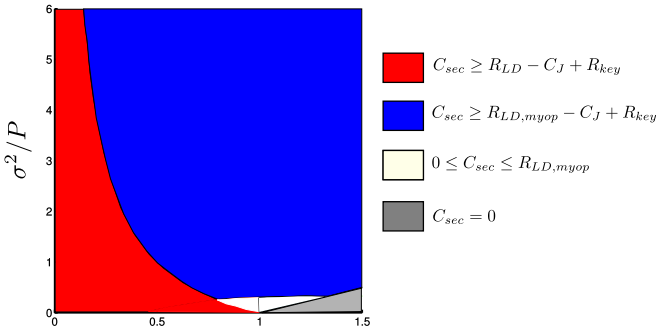
(a) Achievable rates with secrecy for different noise-to-signal ratios when Alice and Bob share $\mathcal{O}(\log n)$ bits of secret key. Positive rates can be achieved in the red and blue regions.



(b) Achievable rates with secrecy for different noise-to-signal ratios when Alice and Bob do not share a secret key. Positive rates can be achieved in the red and blue regions.

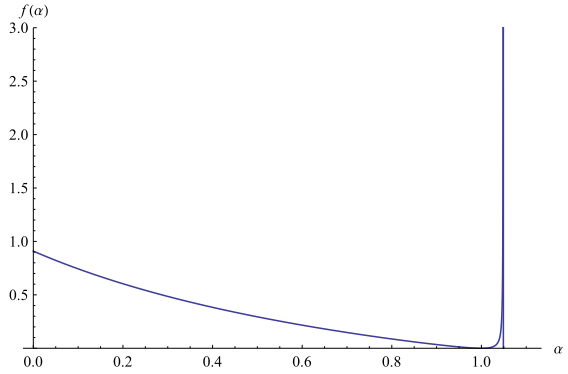


(c) Achievable rates with secrecy for different noise-to-signal ratios when Alice and Bob share $0.2n$ bits of secret key. Positive rates can be achieved in the red and blue regions.

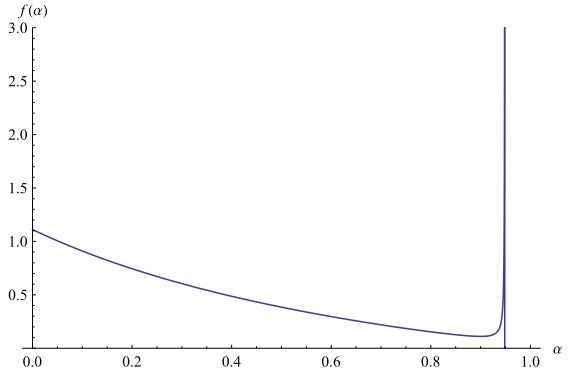


(d) Achievable rates with secrecy for different noise-to-signal ratios when Alice and Bob share n bits of secret key. Positive rates can be achieved in the red and blue regions.

Fig. 17. Achievable rates with secrecy for different values of n_{key}.



(a) $f(\alpha)$ with $N/P = 1.1$.



(b) $f(\alpha)$ with $N/P = 0.9$.

Fig. 18. Optimization of $f(\alpha)$.

A. List-Decoding

Consider the quadratically constrained adversarial channel model where James observes $\underline{\mathbf{z}}$, which is a noisy copy of $\underline{\mathbf{x}}$. In the list-decoding problem, the decoder is not required to recover the transmitted message exactly but can instead output a (small) list of messages with the guarantee that the true message is in the list. We are typically interested in list-sizes that are constant or grow as a low-degree polynomial function of the blocklength.

As a warm-up, let us consider the omniscient adversary (i.e., when $\sigma^2 = 0$). We have the following folk theorem.[11]

Lemma 20: Let $\sigma = 0$ and $n_{\text{key}} = 0$. If $\varepsilon := \frac{1}{2}\log\frac{P}{N} - R > 0$, then R is achievable for $\left(P, N, \Omega\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)\right)$-list-decoding. If $R > \frac{1}{2}\log\frac{P}{N}$, then no sequence of codebooks of rate R is $(P, N, n^{\mathcal{O}(1)})$-list-decodable. Therefore, for polynomial

[11]As will be evident from the proof, the omniscient list-decoding capacity is shown to be $\frac{1}{2}\log\frac{P}{N}$ under a (stronger) maximum probability of error criterion.

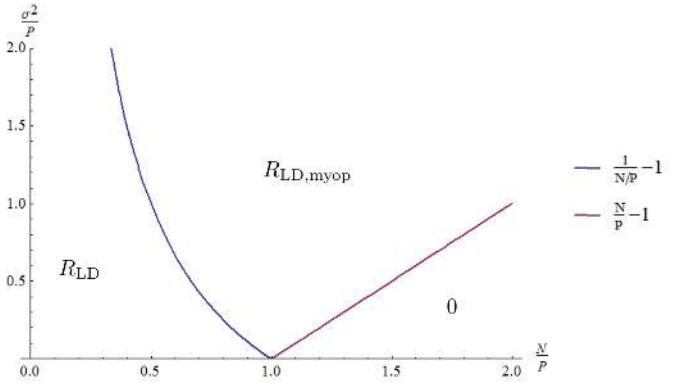


Fig. 19. The upper bound of capacity given by scale-and-babble attack.

list-sizes, the omniscient list-decoding capacity is equal to $\frac{1}{2}\log\frac{P}{N}$.

Although the above result is well-known, we are not aware of a reference with a formal proof of this statement. For completeness, and to keep the paper self-contained, we give a proof in Appendix D.

B. Theorem 11: List-Decoding With a Myopic Adversary

In this subsection, we provide a proof outline of Theorem 11 and forward the detailed proof to Sec. X.

In the case where $\sigma > 0$, we can achieve a higher list-decoding rate for certain values of N/P and σ^2/P. We show that when noise level to James is large enough (he is "sufficiently myopic"), he is unable to exactly determine $\underline{\mathbf{x}}$ from $\underline{\mathbf{z}}$. Conditioned on $\underline{\mathbf{z}}$, the transmitted codeword lies in a thin strip which is roughly $\sqrt{n\sigma^2}$ away from $\underline{\mathbf{z}}$. If R is large enough, then the strip will contain exponentially many codewords, and James cannot distinguish the true codeword from the others. Since we use a random code, these codewords are roughly uniformly distributed over the strip. An effective value of $\underline{\mathbf{s}}$ for one codeword on the strip (in the sense of ensuring maximum confusion for Bob) may be ineffective for most of the remaining codewords. As a result, there is no single direction where James can align $\underline{\mathbf{s}}$ in order to guarantee the level of confusion that Bob could have if he were omniscient. This is what will let us achieve a higher rate.

We will consider the case $n_{\mathrm{key}} = nR_{\mathrm{key}}$, for some $R_{\mathrm{key}} \geq 0$. Even the case when $R_{\mathrm{key}} = 0$ is non-trivial – see Figure 7a for an illustration of the achievable rate.

Theorem 11 (Restatement of Theorem 11): For $(P, N, \mathcal{O}(n^2))$-list-decoding, the capacity is lower bounded as follows

$$C_{\mathrm{myop,LD}} \geq \begin{cases} R_{\mathrm{LD,myop}}, & \text{if } \frac{\sigma^2}{P} \geq \max\left\{\frac{1}{N/P}-1, \frac{N}{P}-1\right\} \\ & \text{and } R_{\mathrm{LD,myop}} + R_{\mathrm{key}} > \frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right) \\ R_{\mathrm{LD}}, & \text{otherwise.} \end{cases}$$

These are summarized in Fig. 7.

We now provide a sketch of the proof, but relegate the details to Sec. X.

Proof Sketch: First, observe that R_{LD} is achievable as long as $N < P$. This is true since R_{LD} is achievable even with an omniscient adversary (Lemma 20). The nontrivial step is in showing that a higher rate of $R_{\mathrm{LD,myop}}$ is achievable in a certain regime of the NSRs. We will prove the achievability using random spherical codes. The $2^{n(R+R_{\mathrm{key}})}$ codewords are sampled independently and uniformly at random from $\mathcal{S}^{n-1}(0, \sqrt{nP})$. This is partitioned randomly into $2^{nR_{\mathrm{key}}}$ codebooks, each containing 2^{nR} codewords. The value of the shared key determines which of the $2^{nR_{\mathrm{key}}}$ codebooks is used for transmission. Since Bob has access to the key, he has to decode one of 2^{nR} codewords. We analyze the probability of error taking James's point of view. To James, one of $2^{n(R+R_{\mathrm{key}})}$ codewords is chosen at random, and the code is $(P, N, \mathcal{O}(n^2))$-list-decodable with high probability if no attack vector $\underline{s}$ can force a list-size of $\Omega(n^2)$ for a nonvanishing fraction of the codewords.

Conditioned on $\underline{\mathbf{z}}$, the true codeword $\underline{\mathbf{x}}$ lies in a thin strip at distance approximately $\sqrt{n\sigma^2}$ to $\underline{\mathbf{z}}$. We show Lemma 24 that as long as the codebook rate $R + R_{\mathrm{key}}$ is greater than $\frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right)$, this strip (with high probability) contains exponentially many codewords, thereby causing sufficient confusion for James. This condition effectively limits the values of the NSRs where $R_{\mathrm{LD,myop}}$ is achievable.

For ease of analysis, we assume that James has access to an oracle, which reveals a particular subset of $2^{\varepsilon n}$ codewords from the strip. This set is guaranteed to contain the true codeword. Clearly, the oracle only makes James more powerful, and any result that holds in this setting also continues to be valid when there is no oracle. The codewords in the oracle-given set (OGS) are all independent (over the randomness in the codebook) and are approximately uniformly (which we call *quasi-uniformly*) distributed over the strip. See Lemma 23 for a formal statement. The codewords outside the oracle-given set are independent of these, and are uniformly distributed over the sphere. From the point of view of James, the true codeword is quasi-uniformly distributed over the OGS.

We fix an attack vector $\underline{s}$, and bound the probability that this forces a list-size greater than L for a significant fraction of the codewords in the oracle-given set. To do so, we find the typical area of the decoding region $\mathcal{B}^n(\underline{\mathbf{x}}+\underline{s}, \sqrt{nN}) \cap \mathcal{S}^{n-1}(0, \sqrt{nP})$ by computing the typical norm of $\underline{\mathbf{y}}$. This decoding region is a cap, whose area is maximized when the radius of the cap (see Sec. IV) is $\sqrt{nN}$. This would be the result of James's attack if he were omniscient. However, due to the randomness in $\underline{\mathbf{s}_z}$ and his uncertainty about $\underline{\mathbf{x}}$, the typical radius is considerably less than $\sqrt{nN}$. It is this reduction in the typical radius that helps us achieve rates above R_{LD}. The value of the typical radius is the solution of an optimization problem, which is identical (under a change of variables) to the one we obtain when analyzing the scale-and-babble attack in Sec. VII-A. See Fig. 20 for an illustration. This is proved in Sec. X-H, with some of the calculations appearing in subsequent subsections.

With an upper bound on the typical decoding volume, we can bound the probability that there are more than L codewords in the decoding region. We separately handle the codewords within and outside the oracle-given set. The probability that a fixed $\underline{s}$ causes list-decoding failure for a

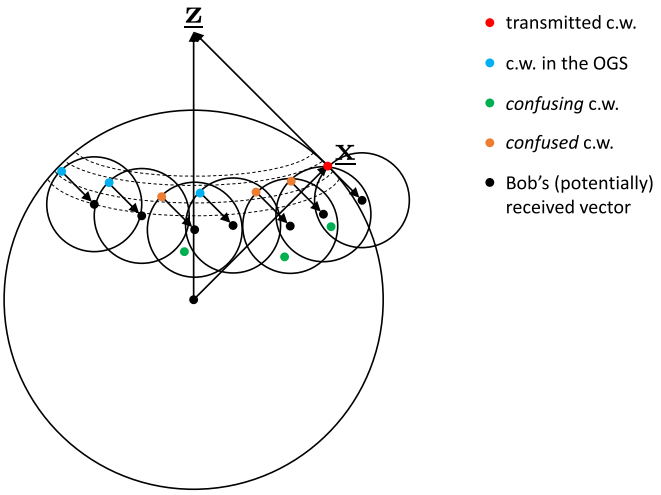


Fig. 20. Intuition behind myopic list-decoding. From the point of view of James, the true codeword is quasi-uniformly distributed over the OGS. The best-case scenario for James would be to "push" the true codeword towards the origin. However, for any fixed attack vector $\underline{\mathbf{s}}$, only a small fraction of the codewords in the OGS lead to a large overlap with $\mathcal{S}^{n-1}(0,\sqrt{nP})$ (which could lead to a large list-size). For a randomly chosen codeword from the OGS, the overlap of the decoding ball with $\mathcal{S}^{n-1}(0,\sqrt{nP})$ is small with high probability.

significant fraction of codewords in the oracle-given set is found to decay super-exponentially in n. We complete the proof using a covering argument for $\underline{s}$ and taking a union bound over all representative attack vectors $\underline{s}_Q$.

C. Achievable Rates Using $\Theta(n)$ Bits of CR

In this subsection, we prove Lemma 12.

The following lemma by Sarwate [37] (originally proved by Langberg [41] for the bit-flip channel, later generalized and improved for more general AVCs and constant list-sizes in [42]) says that a list-decodable code can be converted to a uniquely decodable code with an additional $\mathcal{O}(\log n)$ bits of common randomness. Although the lemma was stated in the context of discrete AVCs with deterministic encoding, the proof goes through even without these restrictions. We can use our list-decodable code as a black box in the following lemma to go from a list-decodable code to a uniquely decodable code.

For an arbitrary AVC $\mathcal{W}$, we define an (n,R,L,ε)-list-decodable code as one which has blocklength n, message rate R, and achieves a list-size of L with probability $1-\varepsilon$.

Lemma 21 (Lemma 13, [37]): Suppose we have a deterministic (n,R,L,ε)-list-decodable code for an AVC $\mathcal{W}$. If the encoder-decoder pair shares n_{key} bits of common randomness, then there exists a blocklength-n code of rate $R-\frac{n_{\text{key}}}{2n}$ such that the decoder can recover the transmitted message with probability $1-\varepsilon-\varepsilon'$, where

$$\varepsilon' := \frac{2nLR}{n_{\text{key}}2^{n_{\text{key}}/2}}.$$

The above lemma says that an additional $n_{\text{key}} = 2\log(nL)$ bits of common randomness is sufficient to disambiguate the list. If $L = n^{\mathcal{O}(1)}$, then the n_{key} required is only logarithmic, and the penalty in the rate $\frac{n_{\text{key}}}{2n}$ is vanishing in n.

We can therefore use this with our myopic list-decoding result in Theorem 11 to obtain achievable rates. Combining this with the converse in Lemma 10, we obtain Lemma 12

D. Achievable Rates With $\Theta(\log n)$ Bits of CR

In this subsection, we prove Lemma 13.

Lemma 21 says that $2\log n$ bits of common randomness is sufficient to disambiguate the list. Using this with Theorem 11 for $R_{\text{key}} = 0$, we have Lemma 13. Note that when $R_{\text{key}} = 0$, the condition $R_{\text{LD,myop}} + R_{\text{key}} > \frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right)$ reduces to $\frac{\sigma^2}{P} \geq 4\frac{N}{P} - 1$.

IX. No Common Randomness

We now discuss the basic ideas required to obtain Theorem 14.

A. Proof Sketch

The proof involves two parts:

- The upper bounds are obtained using Lemma 10, and symmetrization arguments described in Sec. IX-B.
- The achievability involves a combination of list-decoding, reverse list-decoding, and the grid argument. We give a high-level description below. For the rigorous proof, see Sec. XI.

The achievability proof uses several ideas from our discussion on myopic list-decoding in Sec. VIII-B. As discussed in Sec. VIII-B, the transmitted codeword lies in a strip. We can only prove our result in the sufficiently myopic case, i.e., when the strip contains exponentially many codewords. This is possible only if $R > \frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right)$. Just as in the proof of myopic list-decoding, we assume that James has access to an oracle-given set. Our goal is to show that there exists no attack strategy for James that would cause decoding failure for a significant fraction of codewords in the oracle-given set.

See Fig. 20. Let us fix an $\underline{s}$. We say that a codeword $\underline{x}'$ confuses $\underline{x}$ if $\mathcal{B}^n(\underline{x}+\underline{s},\sqrt{nN})$ contains $\underline{x}'$. From James's perspective, the codewords in the oracle-given set are (approximately[12]) equally likely to have been transmitted. We say that decoding fails if the attack vector chosen by James causes $\underline{\mathbf{x}}$ to be confused with another codeword. To analyze this, we study the effect of $\underline{s}$ simultaneously over all codewords in the oracle-given set. The set $\bigcup_{m\in\mathcal{O}_{rcl}}\mathcal{B}^n(\underline{\mathbf{x}}(m)+\underline{s},\sqrt{nN})$ forms a "blob". We show that the probability that the blob contains more than n^4 confusing codewords[13] is vanishingly small. This ensures that there are only a polynomial number of codewords that could potentially confuse the exponentially many codewords in the oracle-given set.

We then find the number of codewords $\underline{\mathbf{x}}(m)$ in the OGS that could potentially be confused by a fixed $\underline{\mathbf{x}}'$. We call this reverse list-decoding, and show that each $\underline{\mathbf{x}}'$ can confuse

[12]We add this qualifier since all points in the strip are not equidistant from $\underline{\mathbf{z}}$.

[13]The degree of the polynomial here is not important, but is chosen to be 4 for convenience. What matters is that the list-size grows polynomially in n.

only polynomially many codewords in the OGS. Our reverse list-decoding argument holds if $\frac{\sigma^2}{P} \geq \frac{1}{1-N/P} - 1$.

We combine the blob list-decoding and reverse list-decoding results and give a combinatorial argument to show that the probability of a fixed $\underline{s}$ causing a decoding error for a significant fraction of codewords in the OGS is super-exponentially decaying in n. We categorise the error into two types:

- Type I: The "confusing" codeword does not lie within the OGS. Blob list-decoding and reverse list-decoding in this case are studied in Sec. XI-A.
- Type II: The "confusing" codeword lies within the OGS. This is studied in Sec. XI-B.

We then use a standard covering argument for $\underline{s}$ and show that the average probability of decoding error is also vanishing in n. This will complete the proof of Theorem 14.

B. An Improved Converse Using Symmetrization

In this section, we prove the symmetrization part of the converse of Theorem 14.

When the encoder is a deterministic map from the set of messages to $\mathbb{R}^n$, we can give a better upper bound for certain values of the NSRs. This attack is based on the scaled babble-and-push attack designed by Li et al. [40] for the quadratically constrained channel with a causal adversary. The basic idea in a symmetrization argument is to make sure that Bob is equally confused between the actual transmitted codeword and a random codeword independently chosen by James. Bob will then be unable to distinguish between the two codewords and therefore makes an error with nonvanishing probability.

Lemma 22: If $n_{\text{key}} = 0$, then $C_{\text{myop}} = 0$ when $\frac{\sigma^2}{P} \leq \frac{1}{1-N/P} - 2$ or $\frac{N}{P} \geq 1$.

Proof: We first present two *suboptimal* jamming strategies referred to as $\underline{\mathbf{z}}$-*agnostic symmetrization* and $\underline{\mathbf{z}}$-*aware symmetrization*. They are simple and natural strategies and give respectively the following bounds *inferior* to the one claimed in Lemma 22.

1) $\underline{\mathbf{z}}$-agnostic symmetrization: If $n_{\text{key}} = 0$, then $C_{\text{myop}} = 0$ when $N \geq P$.
2) $\underline{\mathbf{z}}$-aware symmetrization: If $n_{\text{key}} = 0$, then $C_{\text{myop}} = 0$ when $\frac{\sigma^2}{P} < 4\frac{N}{P} - 2$.

We then slightly modify $\underline{\mathbf{z}}$-aware symmetrization and present an optimal symmetrization-type attack. The analysis follows verbatim that of $\underline{\mathbf{z}}$-aware symmetrization by changing some coefficients. The bound given by such an improved symmetrization subsumes and extends those given by $\underline{\mathbf{z}}$-agnostic/-aware symmetrization.

a) $\underline{\mathbf{z}}$-*agnostic symmetrization:* The first part (Item 1) is considerably simpler, and involves a $\underline{\mathbf{z}}$-*agnostic* symmetrization argument. If $N \geq P$, then James can mimic Alice. A simple attack strategy is the following: He generates a message $\mathbf{m}'$ uniformly at random, and independently of everything else. Using the same encoding strategy that Alice uses; $\mathbf{m}'$ is mapped to a codeword $\underline{\mathbf{s}} = \underline{\mathbf{x}}'$ which is transmitted. Bob receives $\underline{\mathbf{x}} + \underline{\mathbf{x}}'$, and unless $\mathbf{m}' = \mathbf{m}$, he will be unable to determine whether Alice sent $\mathbf{m}$ or $\mathbf{m}'$. Therefore, with probability $1 - 2^{-nR}$ he is unable to decode the correct message, and this is true for all $R > 0$. Therefore, the capacity is zero when $N \geq P$.

b) $\underline{\mathbf{z}}$-*aware symmetrization:* To prove the second part (Item 2), when $\frac{\sigma^2}{P} < 4\frac{N}{P} - 2$, we give a $\underline{\mathbf{z}}$-*aware* symmetrization attack. Here, James picks a random codeword $\underline{\mathbf{x}}'$ uniformly from the codebook and "pushes" $\underline{\mathbf{z}}$ to the midpoint of $\underline{\mathbf{z}}$ and $\underline{\mathbf{x}}'$. Bob is then unable to distinguish between $\underline{\mathbf{x}}$ and $\underline{\mathbf{x}}'$, and will therefore make an error with nonvanishing probability. Specifically, James samples $\mathbf{m}' \sim p_{\mathbf{m}}, \underline{\mathbf{x}}' \sim p_{\underline{\mathbf{x}}|\mathbf{m}}$, both independently of Alice, and sets

$$\underline{\mathbf{s}} = \frac{1}{2}\boldsymbol{\beta}(\underline{\mathbf{x}}' - \underline{\mathbf{z}}) = \frac{1}{2}\boldsymbol{\beta}(\underline{\mathbf{x}}' - \underline{\mathbf{x}} - \underline{\mathbf{s}}_z), \tag{IX.1}$$

where

$$\boldsymbol{\beta} = \begin{cases} 1, & \left\|\frac{1}{2}(\underline{\mathbf{x}}' - \underline{\mathbf{z}})\right\|_2 \leq \sqrt{nN} \\ \frac{\sqrt{nN}}{\left\|\frac{1}{2}(\underline{\mathbf{x}}' - \underline{\mathbf{z}})\right\|_2} =: \boldsymbol{\beta}', & \text{otherwise,} \end{cases} \tag{IX.2}$$

such that

$$\begin{aligned} \underline{\mathbf{y}} &= \underline{\mathbf{x}} + \underline{\mathbf{s}} \\ &= \begin{cases} \underline{\mathbf{x}} + \frac{1}{2}(\underline{\mathbf{x}}' - \underline{\mathbf{x}} - \underline{\mathbf{s}}_z), & \left\|\frac{1}{2}(\underline{\mathbf{x}}' - \underline{\mathbf{z}})\right\|_2 \leq \sqrt{nN} \\ \underline{\mathbf{x}} + \frac{1}{2}\boldsymbol{\beta}'(\underline{\mathbf{x}}' - \underline{\mathbf{x}} - \underline{\mathbf{s}}_z), & \text{otherwise} \end{cases} \\ &= \begin{cases} = \frac{1}{2}(\underline{\mathbf{x}}' + \underline{\mathbf{x}}) - \frac{1}{2}\underline{\mathbf{s}}_z, & \left\|\frac{1}{2}(\underline{\mathbf{x}}' - \underline{\mathbf{z}})\right\|_2 \leq \sqrt{nN} \\ = \left(1 - \frac{1}{2}\boldsymbol{\beta}'\right)\underline{\mathbf{x}} + \frac{1}{2}\boldsymbol{\beta}'\underline{\mathbf{x}}' - \frac{1}{2}\boldsymbol{\beta}'\underline{\mathbf{s}}_z, & \text{otherwise.} \end{cases} \end{aligned}$$

We introduce $\boldsymbol{\beta}$ merely to ensure that the attack vector always satisfies James's power constraint. We don't care much about the second case in Equation (IX.2), since the probability of the second case goes to zero. Hence, even if Bob may be able to decode the message in the second case, we show that in the first case his probability of error is going to be bounded away from zero. Assume we operate at a rate R. Define $\underline{\mathbf{z}}' := \underline{\mathbf{x}}' + \underline{\mathbf{s}}_z$, $\underline{\mathbf{s}}' := \frac{1}{2}(\underline{\mathbf{x}} - \underline{\mathbf{z}}')$ and $\underline{\mathbf{y}}' := \underline{\mathbf{x}}' + \underline{\mathbf{s}}' = \underline{\mathbf{x}}' + \frac{1}{2}(\underline{\mathbf{x}} - \underline{\mathbf{z}}') = \underline{\mathbf{x}}' + \frac{1}{2}(\underline{\mathbf{x}} - \underline{\mathbf{x}}' - \underline{\mathbf{s}}_z) = \frac{1}{2}(\underline{\mathbf{x}}' + \underline{\mathbf{x}}) - \frac{1}{2}\underline{\mathbf{s}}_z = \underline{\mathbf{y}}$. The probability of error can be lower bounded by

$$\begin{aligned} P_e &= \mathbb{P}(\widehat{\mathbf{m}} \neq \mathbf{m}) \\ &\geq \mathbb{P}(\widehat{\mathbf{m}} \neq \mathbf{m}, \underline{\mathbf{x}}' \neq \underline{\mathbf{x}}, \|\underline{\mathbf{s}}\|_2 \leq \sqrt{nN}, \|\underline{\mathbf{s}}'\|_2 \leq \sqrt{nN}) \\ &= \mathbb{P}(\underline{\mathbf{x}}' \neq \underline{\mathbf{x}}, \|\underline{\mathbf{s}}\|_2 \leq \sqrt{nN}, \|\underline{\mathbf{s}}'\|_2 \leq \sqrt{nN}) \\ &\quad \mathbb{P}(\widehat{\mathbf{m}} \neq \mathbf{m} | \underline{\mathbf{x}}' \neq \underline{\mathbf{x}}, \|\underline{\mathbf{s}}\|_2 \leq \sqrt{nN}, \|\underline{\mathbf{s}}'\|_2 \leq \sqrt{nN}) \\ &\geq \frac{1}{2}\,\mathbb{P}(\underline{\mathbf{x}}' \neq \underline{\mathbf{x}}, \|\underline{\mathbf{s}}\|_2 \leq \sqrt{nN}, \|\underline{\mathbf{s}}'\|_2 \leq \sqrt{nN}), \end{aligned}$$

where the last inequality comes from the following argument. Suppose that James has enough power to push the channel output $\underline{\mathbf{y}}$ to $(\underline{\mathbf{x}} + \underline{\mathbf{x}}')/2$ even when $\underline{\mathbf{s}}_z = 0$, and that Bob knew that his observation $\underline{\mathbf{y}}$ is the average of $\underline{\mathbf{x}}$ and $\underline{\mathbf{x}}'$.[14] In this case, Bob cannot distinguish whether $\underline{\mathbf{x}}$ or $\underline{\mathbf{x}}'$ was transmitted and his probability of decoding error is no less than $1/2$. Note that, in the myopic case we are considering, Bob's observation $\underline{\mathbf{y}} = (\underline{\mathbf{x}} + \underline{\mathbf{x}}')/2 - \underline{\mathbf{s}}_z/2$ also contains a (scaled) random noise component other than the average of two codewords. The noise is completely random and independent of everything

[14]Notice that there may exist other pairs of codewords with the same average.

else, hence it does not provide Bob with any information of $\mathbf{m}$ and a decoding error will occur still with probability at least $1/2$. However, there is one more caveat. The output $\underline{\mathbf{y}}$ when $\underline{\mathbf{x}}$ was transmitted by Alice and $\underline{\mathbf{x}}'$ was sampled by James coincide with the output $\underline{\mathbf{y}}'$ when $\underline{\mathbf{x}}'$ was transmitted by Alice and $\underline{\mathbf{x}}$ was sampled by James. If $\underline{\mathbf{s}}'$ violates James's power constraint, then Bob immediately knows that the output is not $\underline{\mathbf{y}}'$, $\underline{\mathbf{x}}$ is the genuine codeword and $\underline{\mathbf{x}}'$ is a spoofing codeword. Hence, to ensure that Bob is fooled by $\underline{\mathbf{x}}$ and $\underline{\mathbf{x}}'$, it had better be the case that $\underline{\mathbf{s}}'$ satisfies his power constraint as well.

The probability $\mathbb{P}(\underline{\mathbf{x}}' \neq \underline{\mathbf{x}}, \|\underline{\mathbf{s}}\|_2 \leq \sqrt{nN}, \|\underline{\mathbf{s}}'\|_2 \leq \sqrt{nN})$ can be bounded as follows.

$$\begin{aligned}
&\mathbb{P}(\underline{\mathbf{x}}' \neq \underline{\mathbf{x}}, \|\underline{\mathbf{s}}\|_2 \leq \sqrt{nN}, \|\underline{\mathbf{s}}'\|_2 \leq \sqrt{nN}) \\
=&\mathbb{P}(\|\underline{\mathbf{s}}\|_2 \leq \sqrt{nN}, \|\underline{\mathbf{s}}'\|_2 \leq \sqrt{nN}) \\
&- \mathbb{P}(\|\underline{\mathbf{s}}\|_2 \leq \sqrt{nN}, \|\underline{\mathbf{s}}'\|_2 \leq \sqrt{nN}, \underline{\mathbf{x}}' = \underline{\mathbf{x}}) \\
\geq&\mathbb{P}(\|\underline{\mathbf{s}}\|_2 \leq \sqrt{nN}, \|\underline{\mathbf{s}}'\|_2 \leq \sqrt{nN}) - \mathbb{P}(\underline{\mathbf{x}}' = \underline{\mathbf{x}}).
\end{aligned}$$

Apparently, $\mathbb{P}(\underline{\mathbf{x}}' = \underline{\mathbf{x}}) = 2^{-nR} \to 0$. It now remains to lower bound the first term.

$$\begin{aligned}
&\mathbb{P}(\|\underline{\mathbf{s}}\|_2 \leq \sqrt{nN}, \|\underline{\mathbf{s}}'\|_2 \leq \sqrt{nN}) \\
=&\mathbb{P}\left(\left\|\frac{1}{2}(\underline{\mathbf{x}}' - \underline{\mathbf{z}})\right\|_2 \leq \sqrt{nN}, \left\|\frac{1}{2}(\underline{\mathbf{x}} - \underline{\mathbf{z}}')\right\|_2 \leq \sqrt{nN}\right) \\
=&\mathbb{P}(\|\underline{\mathbf{x}}' - \underline{\mathbf{x}} - \underline{\mathbf{s}_z}\|_2 \leq 2\sqrt{nN}, \|\underline{\mathbf{x}} - \underline{\mathbf{x}}' - \underline{\mathbf{s}_z}\|_2 \leq 2\sqrt{nN}) \\
=&\mathbb{P}(\|\underline{\mathbf{x}} - \underline{\mathbf{x}}'\|_2^2 + \|\underline{\mathbf{s}_z}\|_2^2 + 2\langle \underline{\mathbf{x}} - \underline{\mathbf{x}}', \underline{\mathbf{s}_z}\rangle \leq 4nN, \\
&\quad \|\underline{\mathbf{x}} - \underline{\mathbf{x}}'\|_2^2 + \|\underline{\mathbf{s}_z}\|_2^2 - 2\langle \underline{\mathbf{x}} - \underline{\mathbf{x}}', \underline{\mathbf{s}_z}\rangle \leq 4nN) \\
\geq&\mathbb{P}(\|\underline{\mathbf{x}} - \underline{\mathbf{x}}'\|_2^2 \leq 2nP(1+\delta_1), \|\underline{\mathbf{s}_z}\|_2^2 \leq n\sigma^2(1+\delta_2), \\
&\quad |\langle \underline{\mathbf{x}} - \underline{\mathbf{x}}', \underline{\mathbf{s}_z}\rangle| \leq n\delta_3) \qquad \text{(IX.3)} \\
\geq&1 - \mathbb{P}(\|\underline{\mathbf{x}} - \underline{\mathbf{x}}'\|_2^2 > 2nP(1+\delta_1)) \\
&- \mathbb{P}(\|\underline{\mathbf{s}_z}\|_2^2 > n\sigma^2(1+\delta_2)) - \mathbb{P}(|\langle \underline{\mathbf{x}} - \underline{\mathbf{x}}', \underline{\mathbf{s}_z}\rangle| > n\delta_3), \qquad \text{(IX.4)}
\end{aligned}$$

where in Eqn. (IX.3) we assume $2P + \sigma^2 = 4N - \varepsilon < 4N$ for some constant $\varepsilon > 0$ and we set $\delta_1 := \frac{\varepsilon}{6P}, \delta_2 := \frac{\varepsilon}{3\sigma^2}, \delta_3 := \varepsilon/6$. The first term in Eqn. (IX.4) can be bounded using Markov's inequality. Specifically,

$$\begin{aligned}
&\mathbb{E}(\|\underline{\mathbf{x}} - \underline{\mathbf{x}}'\|_2^2) \\
=&\mathbb{E}(\|\underline{\mathbf{x}}\|_2^2) + \mathbb{E}(\|\underline{\mathbf{x}}'\|_2^2) - 2\,\mathbb{E}(\langle \underline{\mathbf{x}}, \underline{\mathbf{x}}'\rangle) \\
=&\mathbb{E}(\|\underline{\mathbf{x}}\|_2^2) + \mathbb{E}(\|\underline{\mathbf{x}}'\|_2^2) - 2\sum_{i=1}^{n} \mathbb{E}(\underline{\mathbf{x}}_i \underline{\mathbf{x}}_i') \\
=&\mathbb{E}(\|\underline{\mathbf{x}}\|_2^2) + \mathbb{E}(\|\underline{\mathbf{x}}'\|_2^2) - 2\sum_{i=1}^{n} \mathbb{E}(\underline{\mathbf{x}}_i)\,\mathbb{E}(\underline{\mathbf{x}}_i') \qquad \text{(IX.5)} \\
=&\mathbb{E}(\|\underline{\mathbf{x}}\|_2^2) + \mathbb{E}(\|\underline{\mathbf{x}}'\|_2^2) - 2\sum_{i=1}^{n} (\mathbb{E}(\underline{\mathbf{x}}_i))^2 \qquad \text{(IX.6)} \\
\leq&2nP. \qquad \text{(IX.7)}
\end{aligned}$$

where Equation (IX.5) and Equation (IX.6) follow since $\underline{\mathbf{x}}'$ and $\underline{\mathbf{x}}$ are i.i.d. By Markov's inequality, we have

$$\mathbb{P}(\|\underline{\mathbf{x}} - \underline{\mathbf{x}}'\|_2^2 > 2nP(1+\delta_1)) \leq \frac{2nP}{2nP(1+\delta_1)} = \frac{1}{1+\delta_1}.$$

The second term of Eqn. (IX.4) follows from χ^2 tail bound (Fact 7).

$$\mathbb{P}(\|\underline{\mathbf{s}_z}\|_2^2 > n\sigma^2(1+\delta_2)) \leq \exp(-\delta_2^2 n/4).$$

Since $\langle \underline{\mathbf{x}} - \underline{\mathbf{x}}', \underline{\mathbf{s}_z}\rangle \sim \mathcal{N}(0, \|\underline{\mathbf{x}} - \underline{\mathbf{x}}'\|_2^2 \sigma^2 \mathbf{I}_n)$, by Gaussian tal bound (Fact 6),

$$\begin{aligned}
\mathbb{P}(|\langle \underline{\mathbf{x}} - \underline{\mathbf{x}}', \underline{\mathbf{s}_z}\rangle| > n\delta_3) \leq&2\exp\left(-\frac{(n\delta_3)^2}{2\|\underline{\mathbf{x}} - \underline{\mathbf{x}}'\|_2^2 \sigma^2}\right) \\
\leq&2\exp\left(-\frac{n^2\delta_3^2}{4nP\sigma^2}\right) \\
=&2\exp\left(-\frac{n\delta_3^2}{4P\sigma^2}\right).
\end{aligned}$$

Finally, we have

$$\begin{aligned}
P_e \geq&\frac{1}{2}(1 - \mathbb{P}(\|\underline{\mathbf{x}} - \underline{\mathbf{x}}'\|_2^2 > 2nP(1+\delta_1)) - \\
&\mathbb{P}(\|\underline{\mathbf{s}_z}\|_2^2 > n\sigma^2(1+\delta_2)) - \mathbb{P}(|\langle \underline{\mathbf{x}} - \underline{\mathbf{x}}', \underline{\mathbf{s}_z}\rangle| > n\delta_3) \\
&- \mathbb{P}(\underline{\mathbf{x}}' = \underline{\mathbf{x}})) \\
\geq&\frac{1}{2}\left(1 - \frac{1}{1+\delta_1} - \exp(-\delta_2^2 n/4)\right. \\
&\left.-2\exp\left(-\frac{n\delta_3^2}{4P\sigma^2}\right) - 2^{-nR}\right) \\
=&\frac{1}{2}\left(\frac{\varepsilon/6P}{1+\varepsilon/6P} - \exp\left(-\frac{n\varepsilon^2}{36\sigma^4}\right)\right. \\
&\left.-2\exp\left(-\frac{n\varepsilon^2}{144P\sigma^2}\right) - 2^{-nR}\right) \\
\to&\frac{\varepsilon/6P}{2(1+\varepsilon/6P)},
\end{aligned}$$

which is bounded away from zero. Thus no positive rate is achievable when $\frac{\sigma^2}{P} < 4\frac{N}{P} - 2$.

c) Improved $\underline{\mathbf{z}}$-aware symmetrization: Finally, we modify the previous $\underline{\mathbf{z}}$-aware symmetrization by optimizing the coefficients in front of $\underline{\mathbf{x}}'$ and $\underline{\mathbf{z}}$ in the design of $\underline{\mathbf{s}}$ (Eqn. (IX.1)).

Let

$$\underline{\mathbf{s}} = \alpha\underline{\mathbf{z}} + \beta\underline{\mathbf{x}}' + \underline{\mathbf{g}}, \qquad \text{(IX.8)}$$

where $\alpha < 0, \beta > 0$ are to be determined momentarily, $\underline{\mathbf{g}} \sim \mathcal{N}(0, \gamma^2\mathbf{I}_n)$ for some $\gamma > 0$ to be determined later, and $\underline{\mathbf{x}}'$ is a random codeword sampled uniformly from Alice's codebook.[15]

Under the choice of $\underline{\mathbf{s}}$ defined in Eqn. (IX.8), Bob receives

$$\underline{\mathbf{y}} = \underline{\mathbf{x}} + \underline{\mathbf{s}} = \underline{\mathbf{x}} + \alpha\underline{\mathbf{z}} + \beta\underline{\mathbf{x}}' + \underline{\mathbf{g}} = (1+\alpha)\underline{\mathbf{x}} + \beta\underline{\mathbf{x}}' + \alpha\underline{\mathbf{s}_z} + \underline{\mathbf{g}}. \qquad \text{(IX.9)}$$

[15]Strictly speaking, as in Eqn. (IX.1), we should also multiply $\underline{\mathbf{s}}$ by a normalization factor $\boldsymbol{\beta}$. It ensures that $\underline{\mathbf{s}}$ satisfies James's power constraint with probability *one*. The way to handle it is precisely the same as in the previous part IX-B.0.b and we omit the technical details in this part.

We observe the following two points from Eqn. (IX.9). Firstly, to "symmetrize" the channel from Alice to Bob, James had better set $1+\alpha=\beta$. This ensures that Bob has no idea whether $\underline{\mathbf{x}}$ or $\underline{\mathbf{x}}'$ was transmitted even if he somehow magically knew the value of $\alpha\underline{\mathbf{s}}_z+\underline{\mathbf{g}}$. Secondly, to save his power, James had better set $\gamma=0$, that is, not add additional Gaussian noise in $\underline{\mathbf{s}}$. Therefore we set $\underline{\mathbf{s}}=\alpha\underline{\mathbf{z}}+(1+\alpha)\underline{\mathbf{x}}'$ where $\alpha<0$ and $1+\alpha>0$, i.e., $\alpha>-1$.

We now evaluate $\frac{1}{n}\mathbb{E}(\|\underline{\mathbf{s}}\|_2^2)$ and contrast it with James's power constraint N.

$$\begin{aligned}\mathbb{E}(\|\underline{\mathbf{s}}\|_2^2)=&\mathbb{E}(\|\alpha\underline{\mathbf{z}}+(1+\alpha)\underline{\mathbf{x}}'\|_2^2)\\ =&\mathbb{E}(\|\alpha\underline{\mathbf{x}}+(1+\alpha)\underline{\mathbf{x}}'+\alpha\underline{\mathbf{s}}_z\|_2^2)\\ =&\alpha^2\,\mathbb{E}(\|\underline{\mathbf{x}}\|_2^2)+(1+\alpha)^2\,\mathbb{E}(\|\underline{\mathbf{x}}'\|_2^2)+\alpha^2\,\mathbb{E}(\|\underline{\mathbf{s}}_z\|_2^2)\\ &+2\alpha(1+\alpha)\,\mathbb{E}(\langle\underline{\mathbf{x}},\underline{\mathbf{x}}'\rangle)+\alpha^2\,\mathbb{E}(\langle\underline{\mathbf{x}},\underline{\mathbf{s}}_z\rangle)\\ &+\alpha(1+\alpha)\,\mathbb{E}(\langle\underline{\mathbf{x}}',\underline{\mathbf{s}}_z\rangle)\\ \le&\alpha^2\cdot nP+(1+\alpha)^2\cdot nP+\alpha^2\cdot n\sigma^2+0+0 \qquad \text{(IX.10)}\\ =&n(2P+\sigma^2)\alpha^2+2nP\alpha+nP. \qquad \text{(IX.11)}\end{aligned}$$

Eqn. (IX.10) follows from the same calculation as in Eqn. (IX.7). Minimizing Eqn. (IX.11) over $\alpha\in(-1,0)$ (so as to minimize the amount of power James spent), we obtain the minimizer

$$\alpha_*=-\frac{P}{2P+\sigma^2},\ \beta_*=1+\alpha_*=\frac{P+\sigma^2}{2P+\sigma^2}.$$

The above calculation can be directly substituted into the previous part (IX-B.0.b). This implies that $C_{\text{myop}}=0$ as long as the power James spent in transmitting $\underline{\mathbf{s}}$ defined in Eqn. (IX.1) is at most $\sqrt{nN}$. That is, the RHS of Eqn. (IX.11) evaluated at $\alpha=\alpha_*$ is at most N: $(2P+\sigma)^2\alpha_*^2+2P\alpha_*+P\le N$. This reduces to the condition $\frac{\sigma^2}{P}\le\frac{1}{1-N/P}-2$, as promised in Lemma 22. □

Remark 7: The argument above generalizes the Plotkin bound (via the Cauchy–Schwarz inequality – see, e.g., Li et al. [40]) to scenarios with additional randomness in $\underline{\mathbf{s}}_z$.

X. Myopic List-Decoding

We now describe our coding scheme and prove that it achieves the rate in Theorem 11 (restated in Theorem 11). In Sec. X-A, we formally describe the scheme. The proof of Theorem 11 proceeds by analyzing various error events which are formally defined in Sec. X-C. The proof is outlined in Sec. X-D, and the probabilities of the various error events are analyzed in the following subsections.

A. Coding Scheme

1) Codebook Construction: We use random spherical codes. Before the communication, Alice samples $2^{n(R+R_{\text{key}})}$ codewords $\{\underline{x}(m,k):m\in[2^{nR}],k\in[2^{nR_{\text{key}}}]\}$ independently and uniformly at random from the sphere $\mathcal{S}^{n-1}\left(0,\sqrt{nP}\right)$. Once sampled, the codebook is fixed and revealed to every party: Alice, Bob and James. Notice that all codewords satisfy Alice's power constraint $\|\underline{x}\|_2\le\sqrt{nP}$. We define $\mathcal{C}^{(k)}:=\{\underline{x}(m,k):m\in[2^{nR}]\}$ to be the kth codebook. We also distinguish the message rate R from the codebook rate $R_{\text{code}}:=R+R_{\text{key}}$.

2) Encoder: Let $k\in[2^{nR_{\text{key}}}]$ be the realization of the secret key shared by Alice and Bob. Alice sends $\underline{x}(m,k)$ if she wants to transmit message m to Bob.

3) Decoder: Bob uses a minimum distance decoder. Having received $\underline{y}$, he outputs $\widehat{m}$ such that the corresponding codeword $\underline{x}(\widehat{m},k)$ is the nearest (in Euclidean distance) one in $\mathcal{C}^{(k)}$ to his observation, i.e.,

$$\widehat{m}=\operatorname*{argmin}_{m'\in\{0,1\}^{nR}}\|\underline{x}(m',k)-\underline{y}\|_2.$$

Remark 8: We emphasize that the above coding scheme is designed for the original *unique* decoding problem. To approach the proof of unique decodability, we have to go through a novel notion of list-decoding referred to as myopic list-decoding[16] as the title of this section suggests. However, myopic list-decoding appears only as a proof technique and neither Bob nor James really performs a step of myopic list-decoding. Note that the above decoder for unique decoding will not be used until Section XI.

For the rest of this section, we fix two quantities: ε is a small positive constant independent of n, and δ is a parameter that decays as $\Theta((\log n)/n)$. The latter parameter δ is used in Eqn. (X.2) to parameterize the thickness of each strip $\mathcal{S}tr^{n-1}(\underline{z}_Q,i)$.

B. The Strips and the Oracle-Given Set (OGS)

Let $L=3n^2$. To simplify the proof, we prove the achievability part under a more powerful adversary who has access to an oracle in addition to $\underline{\mathbf{z}}$. The oracle reveals a random subset of $2^{\varepsilon n}$ codewords that contains the transmitted codeword and others that are all at approximately the same distance to $\underline{\mathbf{z}}$. We call it an oracle-given set, denoted $\mathcal{O}rcl(\underline{\mathbf{z}},\underline{\mathbf{x}})$. Conditioned on James's knowledge, the transmitted codeword is independent of all codewords outside the oracle-given set. We now describe the rule that assigns a pair $(\underline{\mathbf{x}},\underline{\mathbf{z}})$ to an OGS.

Choose any optimal covering $\mathcal{Z}$ of $\mathcal{S}h^n(0,\sqrt{n(P+\sigma^2)(1\pm\varepsilon)})$ such that $\min_{\underline{z}'\in\mathcal{Z}}\|\underline{z}-\underline{z}'\|_2\le\sqrt{n\delta_{\mathcal{Z}}}$ for all $\underline{z}$ in the shell. The size of such a covering can be bounded as follows.

$$\begin{aligned}|\mathcal{Z}|&\le\left(\frac{\text{Vol}(\mathcal{B}^n(0,\sqrt{n(P+\sigma^2)(1+\varepsilon)}+\sqrt{n\delta_{\mathcal{Z}}}))}{\text{Vol}(\mathcal{B}^n(0,\sqrt{n\delta_{\mathcal{Z}}}))}\right)^{1+o(1)}\\ &=\left(\frac{\sqrt{(P+\sigma^2)(1+\varepsilon)}+\sqrt{\delta_{\mathcal{Z}}}}{\sqrt{\delta_{\mathcal{Z}}}}\right)^{n(1+o(1))}=:c^n_{\varepsilon,\delta_{\mathcal{Z}}}.\end{aligned} \qquad \text{(X.1)}$$

Given $\underline{z}$, let $\underline{z}_Q:=\arg\min_{\underline{z}'\in\mathcal{Z}}\|\underline{z}-\underline{z}'\|_2$ denote the closest point to $\underline{z}$ in $\mathcal{Z}$ (a.k.a. the *quantization* of $\underline{z}$). For each $\underline{z}_Q\in\mathcal{Z}$, and $i\in\{-\varepsilon/\delta+1,\dots,\varepsilon/\delta\}$, define the i-th strip

$$\begin{aligned}&\mathcal{S}tr^{n-1}(\underline{z}_Q,i):=\mathcal{S}^{n-1}(0,\sqrt{nP})\\ &\cap\mathcal{S}h^n(\underline{z}_Q,\sqrt{n\sigma^2(1+(i-1)\delta)},\sqrt{n\sigma^2(1+i\delta)})\end{aligned} \qquad \text{(X.2)}$$

[16]See Section X-D below for what it means for a code to be non-myopic-list-decodable.

to be the set of all points on the coding sphere at a distance of at least $\sqrt{n\sigma^2(1+(i-1)\delta)}$ but at most $\sqrt{n\sigma^2(1+i\delta)}$ away from $\underline{z_Q}$. It is not hard to see that the union of the strips is the whole power sphere: $\bigcup_{\underline{z_Q}}\bigcup_i \mathcal{S}tr^{n-1}(\underline{z_Q},i) = \mathcal{S}^{n-1}(0,\sqrt{nP})$.[17] Let $\mathcal{C} := \{\underline{x}(m,k) : m \in [2^{nR}], k \in [2^{nR_{\rm key}}]\}$ denote the codebook. Define

$$\mathcal{M}_{\rm str}(\underline{z_Q},i) := \{(m,k) : \underline{x}(m,k) \in \mathcal{S}tr^{n-1}(\underline{z_Q},i)\} \quad \text{(X.3)}$$

to be the set of indices of the codewords that lie in $\mathcal{S}tr^{n-1}(\underline{z_Q},i)$. We partition this set of indices into blocks of size $2^{n\varepsilon}$ each (except perhaps the last block) in the lexicographic order of (m,k). Let $\{\mathcal{O}rcl^{(j)}(\underline{z_Q},i)\}_{j=1}^{\ell}$ denote the partition, where $\ell := \lceil |\mathcal{M}_{\rm str}(\underline{z_Q},i)|/2^{n\varepsilon}\rceil$. Each of these blocks constitutes an oracle-given set. If (m,k) corresponding to the transmitted codeword $\underline{x}(m,k)$ lies in the μth block $\mathcal{O}rcl^{(\mu)}(\underline{z_Q},\lambda)$ of the partition of the λth strip $\mathcal{M}_{\rm str}(\underline{z_Q},\lambda)$ for some $\lambda \in \{-\varepsilon/\delta+1,\ldots,\varepsilon/\delta\}$ and $\mu \in [\ell]$, then the oracle reveals $\mathcal{O}rcl(\underline{z_Q},\underline{x}) := \mathcal{O}rcl^{(\mu)}(\underline{z_Q},\lambda)$ to James.

Remark 9: It is important to note that all sets defined above (strips, OGSs, etc.) are designed a priori, before communication takes place.

C. Error Events

Define

$$\begin{aligned}&\mathcal{L}^{(k)}(\underline{x}(m),\underline{s})\\ :=&\{w \in [2^{nR}] : \underline{x}(w,k) \in \mathcal{B}^n(\underline{x}(m,k)+\underline{s},\sqrt{nN}) \cap \mathcal{C}^{(k)}\}\\ =&\{w \in [2^{nR}] : \|\underline{x}(w,k)-\underline{x}(m,k)-\underline{s}\|_2 \le \sqrt{nN}\}\end{aligned}$$

for $m \in [2^{nR}], k \in [2^{nR_{\rm key}}], \underline{s} \in \mathcal{B}^n(0,\sqrt{nN})$. Recall that by the proof sketch in Section VIII-B, to prove the existence of a myopic list-decodable code, we want to show that with high probability over the randomness in codebook construction, the uniform selection of the messages, the key shared by encoder-decoder and the channel from Alice to James, only a vanishing fraction of codewords in the OGS (say, $2^{n\varepsilon/4}$ out of $2^{n\varepsilon}$ codewords in the OGS) have list-size larger than L under *some* attack vector by James. Formally, we want to show that a random spherical code ensemble as constructed in Section X-A is *myopically list-decodable* in the following sense.

Definition 2 (Myopic List-Decodability): A code ensemble $\{\mathcal{C}^{(\mathbf{k})}\}$ with common randomness $\mathbf{k}$ shared by Alice and Bob is said to be myopic list-decodable if

$$\mathbb{P}\Big(\exists \underline{s} \in \mathcal{B}^n(0,\sqrt{nN}),\ |\{(m,\mathbf{k}) \in \mathcal{O}rcl(\underline{\mathbf{z}},\underline{\mathbf{x}}) : |\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s})| > L\}| > 2^{n(\varepsilon-h(\varepsilon,\tau,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})/2)}\Big) = o(1),$$

where $0 < h(\varepsilon,\tau,\delta_{\mathcal{S}},\delta_{\mathcal{Z}}) < 2\varepsilon$ is a vanishing function in each of its variables. In particular, we can take $h(\varepsilon) = \frac{3}{2}\varepsilon$ by setting τ, $\delta_{\mathcal{S}}$ and $\delta_{\mathcal{Z}}$ to be suitable functions of ε.

Here and throughout the rest of this paper, we often write $\underline{\mathbf{x}}$ as a shorthand for $\underline{\mathbf{x}}(\mathbf{m},\mathbf{k})$ where $\mathbf{m}$ is a uniform message, $\mathbf{k}$ is a random shared key and for any given pair $(\mathbf{m},\mathbf{k})$ the corresponding codeword $\underline{\mathbf{x}}(\mathbf{m},\mathbf{k})$ is chosen uniformly from the power sphere $\mathcal{S}^{n-1}(0,\sqrt{nP})$. To analyze the above probability, we define a number of error events.

Choose any optimal covering $\mathcal{S}$ of $\mathcal{B}^n(0,\sqrt{nN})$ such that $\min_{\underline{s}' \in \mathcal{S}} \|\underline{s}-\underline{s}'\|_2 \le \sqrt{n\delta_{\mathcal{S}}}$ for all $\underline{s}$ in the ball. Given $\underline{s}$, let $s_Q := \arg\min_{\underline{s}' \in \mathcal{S}} \|\underline{s}-\underline{s}'\|_2$ denote the closest point to $\underline{s}$ in $\mathcal{S}$ (a.k.a. the quantization of $\underline{s}$). By similar calculation to Equation (X.1), we have

$$|\mathcal{S}| \le \left(\frac{\sqrt{N}+\sqrt{\delta_{\mathcal{S}}}}{\sqrt{\delta_{\mathcal{S}}}}\right)^{n(1+o(1))} =: c_{\delta_{\mathcal{S}}}^n. \quad \text{(X.4)}$$

We say that list-decoding fails if any of the following events occurs. We will ultimately show that the probability of failure is negligible. The error events that we analyze are listed below:

$\mathcal{E}_{\rm atyp}$: James's observation $\underline{\mathbf{z}}$ behaves atypically, or equivalently, the noise $\underline{\mathbf{s}_z}$ to James behaves atypically.

$$\mathcal{E}_{\rm atyp_1} := \{\|\underline{\mathbf{s}_z}\|_2 \notin \sqrt{n\sigma^2(1\pm\varepsilon)}\}. \quad \text{(X.5)}$$

$$\mathcal{E}_{\rm atyp_2} := \{|\cos(\angle_{\underline{\mathbf{x}},\underline{\mathbf{s}_z}})| \ge \varepsilon\}. \quad \text{(X.6)}$$

$$\mathcal{E}_{\rm atyp_3} := \{\|\underline{\mathbf{z}}\|_2 \notin \sqrt{n(P+\sigma^2)(1\pm\varepsilon)}\}. \quad \text{(X.7)}$$

Hence the error event $\mathcal{E}_{\rm atyp}$ is the union of the above three events.

$$\mathcal{E}_{\rm atyp} := \mathcal{E}_{\rm atyp_1} \cup \mathcal{E}_{\rm atyp_2} \cup \mathcal{E}_{\rm atyp_3}. \quad \text{(X.8)}$$

$\mathcal{E}_{\rm str}$: One of the strips $\{\mathcal{S}tr^{n-1}(\underline{z_Q},i)\}_i$ contains fewer than $2^{3\varepsilon n}$ codewords.

$$\mathcal{E}_{\rm str}(i) := \{|\mathcal{M}_{\rm str}(\underline{z_Q},i)| < 2^{3\varepsilon n}\}. \quad \text{(X.9)}$$

$$\mathcal{E}_{\rm str} := \bigcup_i \mathcal{E}_{\rm str}(i). \quad \text{(X.10)}$$

Note that for a fixed $\underline{z_Q}$, the randomness in error events defined in Eqn. (X.9) and (X.10) comes from codebook construction, message selection and common randomness. That is, the number of message-key pairs in each $\mathcal{M}_{\rm str}(\underline{z_Q},i)$ is a random variable.

$\mathcal{E}_{\rm orcl}$: Since the number of messages need not be an integer multiple of $2^{n\varepsilon}$, the last OGS may be substantially smaller than the others, and hence could have a higher probability of error. But the probability that the transmitted codeword happens to fall into the last set is small. Call $\mathcal{E}_{\rm orcl}$ the event that the message corresponding to the transmitted codeword belongs to the last block $\mathcal{O}rcl^{(\ell)}(\underline{z_Q},\lambda)$ of the partition of $\mathcal{M}_{\rm str}(\underline{z_Q},\lambda)$.

$$\mathcal{E}_{\rm orcl} := \{\boldsymbol{\mu} = \ell\}. \quad \text{(X.11)}$$

In the above definition (Eqn. (X.11)), the randomness in the random variable $\boldsymbol{\mu}$ results from the uniform selection of message-key pair $(\mathbf{m},\mathbf{k})$.

$\mathcal{E}_{\rm LD-rad}$: Assume that none of $\mathcal{E}_{\rm atyp},\mathcal{E}_{\rm str},\mathcal{E}_{\rm orcl}$ occurs. For any $(m,k) \in \mathcal{O}rcl^{(j)}(\underline{z_Q},i)$ and $\underline{s_Q} \in \mathcal{S}$, let $\sqrt{n\cdot r(m,\underline{s_Q})}$ be the radius of the cap $\mathcal{B}^n(\underline{x}(m,k)+\underline{s_Q},\sqrt{nN}+\sqrt{n\delta_{\mathcal{S}}}) \cap \mathcal{S}^{n-1}(0,\sqrt{nP})$

[17] Indeed, to see this, note that the quantization $\underline{z_Q} \in \mathcal{Z}$ essentially ranges over all directions. By making the quantization level $\delta_{\mathcal{Z}}$ sufficiently fine compared to the thickness parameters ε and δ of the strips, one is able to cover the whole sphere using strips.

(which is the list-decoding region) of $\underline{x}(m,k)$ under $\underline{s_Q}$. We will show that $\mathbf{r} := \mathbf{r}(m, \underline{s_Q})$ concentrates around a certain typical value $r_{\rm opt}(\underline{s_Q})$:

$$r_{\rm opt}(\underline{s_Q}) = \mathbb{E}(\mathbf{r}). \tag{X.12}$$

In the worst case, the maximum (over the choice of $\underline{s_Q}$) of this typical value is given by $r_{\rm opt}$ which is substantially smaller than $\sqrt{nN}$. The exact value of $r_{\rm opt}$ is the minimizer of a certain optimization problem (X.49). We will refer to $r_{\rm opt}$ as the *typical radius*. Let $\mathcal{E}_{\rm LD-rad}(m, \underline{s_Q})$ denote the event that $\mathbf{r}(m, \underline{s_Q})$ is atypical, *i.e.*, $\mathbf{r}$ is significantly larger than its expectation $r_{\rm opt}$.

$$\mathcal{E}_{\rm LD-rad}(m, \underline{s_Q}) := \{\mathbf{r}(m, \underline{s_Q}) > r_{\rm opt}(\underline{s_Q})(1 + f_{11}(\varepsilon, \delta_{\mathcal{S}}))\}, \tag{X.13}$$

for some small function $f_{11}(\varepsilon, \delta_{\mathcal{S}})$ to be determined later.

$\mathcal{E}_{\underline{s_Q}}$: Assume that none of $\mathcal{E}_{\rm atyp}, \mathcal{E}_{\rm str}, \mathcal{E}_{\rm orcl}$ occurs. Define

$$\begin{aligned}
&\psi(\underline{z_Q}, i, j, \underline{s_Q}) \\
:=&|\{(m,k) \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i) : \\
&\quad \mathbf{r}(m, \underline{s_Q}) > r_{\rm opt}(\underline{s_Q})(1 + f_{11}(\varepsilon, \delta_{\mathcal{S}}))\}| \\
=& \sum_{(m,k) \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \mathbb{1}_{\{\mathbf{r}(m, \underline{s_Q}) > r_{\rm opt}(\underline{s_Q})(1 + f_{11}(\varepsilon, \delta_{\mathcal{S}}))\}} \\
=& \sum_{(m,k) \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \mathbb{1}_{\mathcal{E}_{\rm LD-rad}(m, \underline{s_Q})}
\end{aligned} \tag{X.14}$$

to be the number of messages in the OGS whose encodings have atypical list-decoding radii. Let $\mathcal{E}_{\underline{s_Q}}(\underline{z_Q}, i, j)$ be the event that there are more than n^2 such messages in the OGS.

$$\mathcal{E}_{\underline{s_Q}}(\underline{z_Q}, i, j) := \{\psi(\underline{z_Q}, i, j, \underline{s_Q}) > n^2\}. \tag{X.15}$$

$\mathcal{E}_{\rm LD}$: Given that none of $\mathcal{E}_{\rm atyp}, \mathcal{E}_{\rm str}, \mathcal{E}_{\rm orcl}$ occurs, there exists an attack vector $\underline{s_Q} \in \mathcal{S}$ that results in a list-size greater than L for at least one codeword in the oracle-given set. For each $k, \underline{z_Q}, i, j, \underline{s_Q}$, define

$$\begin{aligned}
&\chi(\underline{z_Q}, i, j, \underline{s_Q}) \\
=:&|\{(m,k) \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i) : |\mathcal{L}^{(k)}(\underline{x}(m), \underline{s_Q})| > L\}| \\
=& \sum_{(m,k) \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \mathbb{1}_{\{|\mathcal{L}^{(k)}(\underline{x}(m), \underline{s_Q})| > L\}}
\end{aligned} \tag{X.16}$$

to be the number of codewords in an oracle-given set that result in a large list-size when perturbed by $\underline{s_Q}$. Then, we define

$$\begin{aligned}
&\mathcal{E}_{\rm LD}(\underline{z_Q}, i, j, \underline{s_Q}) \\
:=& \Big\{ \Big| \Big\{ (m,k) \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i) : \\
&\quad |\mathcal{L}^{(k)}(\underline{x}(m), \underline{s_Q})| > L \Big\} \Big| > n^2 + 1 \Big\} \\
=& \{\chi(\underline{z_Q}, i, j, \underline{s_Q}) > n^2 + 1\}
\end{aligned} \tag{X.17}$$

to be the event that there exists codewords that can be perturbed by $\underline{s_Q}$ to give a large list-size. The error event $\mathcal{E}_{\rm LD}$ is defined as

$$\mathcal{E}_{\rm LD} := \bigcup_{\underline{z_Q}} \bigcup_{i} \bigcup_{j} \bigcup_{\underline{s_Q}} \mathcal{E}_{\rm LD}(\underline{z_Q}, i, j, \underline{s_Q}). \tag{X.18}$$

D. The Probability of Myopic List-Decoding Error – Proof of Theorem 11

We now prove Theorem 11 by upper bounding the probability that an error occurs in myopic list-decoding. Let

$$\begin{aligned}
&\mathcal{L}^{(k)}(\underline{x}(m), \underline{s_Q}) := \{w \in [2^{nR}] : \\
&\quad \underline{x}(w,k) \in \mathcal{B}^n(\underline{x}(m,k) + \underline{s_Q}, \sqrt{nN} + \sqrt{n\delta_{\mathcal{S}}}) \cap \mathcal{C}^{(k)}\}.
\end{aligned}$$

We can decompose the failure probability in the following manner using Fact 5.

$$\begin{aligned}
&\mathbb{P}\left(\exists \underline{s} \in \mathcal{B}^n(0, \sqrt{nN}),\ |\{(m, \mathbf{k}) \in \mathcal{O}rcl(\underline{\mathbf{z}}, \underline{\mathbf{x}}) : \right. \\
&\qquad \left. |\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m), \underline{s})| > L\}| > 2^{n(\varepsilon - h(\varepsilon,\tau,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})/2)}\right) \\
\le& \mathbb{P}\left(\exists \underline{s_Q} \in \mathcal{S},\ |\{(m, \mathbf{k}) \in \mathcal{O}rcl(\underline{\mathbf{z}_Q}, \underline{\mathbf{x}}) : \right. \\
&\qquad \left. |\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m), \underline{s_Q})| > L\}| > 2^{n(\varepsilon - h(\varepsilon,\tau,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})/2)}\right) \\
\le& \mathbb{P}\left(\mathcal{E}_{\rm atyp} \cup \mathcal{E}_{\rm str} \cup \mathcal{E}_{\rm orcl} \cup \mathcal{E}_{\rm LD}\right) \\
=& \mathbb{P}\left(\mathcal{E}_{\rm atyp} \cup \bigcup_i \mathcal{E}_{\rm str}(i) \cup \mathcal{E}_{\rm orcl} \cup \bigcup_{\underline{z_Q}} \bigcup_i \bigcup_j \bigcup_{\underline{s_Q}} \mathcal{E}_{\rm LD}(\underline{z_Q}, i, j, \underline{s_Q})\right) \\
\le& 1 - (1 - \mathbb{P}(\mathcal{E}_{\rm atyp})) \\
&\cdot \left(1 - \sum_i \mathbb{P}(\mathcal{E}_{\rm str}(i) | \mathcal{E}_{\rm atyp}^c)\right) \\
&\cdot (1 - \mathbb{P}(\mathcal{E}_{\rm orcl} | \mathcal{E}_{\rm atyp}^c \cap \mathcal{E}_{\rm str}^c)) \\
&\cdot \left(1 - \sum_{\underline{z_Q}} \sum_i \sum_j \sum_{\underline{s_Q}} \mathbb{P}(\mathcal{E}_{\rm LD}(\underline{z_Q}, i, j, \underline{s_Q}) | \mathcal{E}_{\rm atyp}^c \cap \mathcal{E}_{\rm str}^c \cap \mathcal{E}_{\rm orcl}^c)\right).
\end{aligned}$$

It is therefore sufficient to show that each of the error terms is vanishing in n.

The analysis of $\mathcal{E}_{\rm atyp}$, $\mathcal{E}_{\rm str}$ and $\mathcal{E}_{\rm orcl}$ follows from somewhat standard concentration inequalities which are formally justified in Section X-E, Section X-F and Section X-G, respectively. Notice that in the analysis of $\mathcal{E}_{\rm str}$, James is said to be *sufficiently myopic* if, given his observation $\underline{\mathbf{z}}$, his uncertainty set, i.e., any strip $\mathcal{S}tr^{n-1}(\underline{\mathbf{z}}, i)$, contains at least exponentially many codewords. This is guaranteed if $R + R_{\rm key} > \frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right)$.

Much of the complication of our work is devoted to the analysis of $\mathcal{E}_{\rm LD}(\underline{z_Q}, i, j, \underline{s_Q})$. We further factorize it into sub-events and treat them separately. Define $\mathcal{E} := \mathcal{E}_{\rm atyp} \cup \mathcal{E}_{\rm str} \cup \mathcal{E}_{\rm orcl}$. Fix $\underline{z_Q}$, i, j and $\underline{s_Q}$. We are able to show that

$$\mathbb{P}(\mathcal{E}_{\rm LD}(\underline{z_Q}, i, j, \underline{s_Q}) | \mathcal{E}^c) \le 2^{-\Omega(n^3)},$$

which allows us to take a union bound over exponentially many objects. To this end, we need to further decompose the error event $\mathcal{E}_{\rm LD}(\underline{z_Q}, i, j, \underline{s_Q})$ in a careful manner.

Since the codewords all lie on a sphere, the effective decoding region is equal to $\mathcal{B}(\underline{\mathbf{x}}(m,\mathbf{k})+\underline{s_Q},\sqrt{nN}+\sqrt{n\delta_{\mathcal{S}}})\cap \mathcal{S}^{n-1}(0,\sqrt{nP})$. We will first show in Lemma 25 that for most codewords in the oracle-given set, the area of the effective decoding region under any fixed $\underline{s_Q}$ is not too large. The list-sizes for the remaining codewords can be controlled using two-step list-decoding argument and grid argument in Lemma 27. Specifically,

$$\mathbb{P}(\mathcal{E}_{\mathrm{LD}}(\underline{z_Q},i,j,\underline{s_Q})|\mathcal{E}^c)\le \mathbb{P}(\mathcal{E}_{\underline{s_Q}}|\mathcal{E}^c) + \mathbb{P}(\mathcal{E}_{\mathrm{LD}}(\underline{z_Q},i,j,\underline{s_Q})|\mathcal{E}^c\cap\mathcal{E}^c_{\underline{s_Q}}).$$

We first compute the explicit value of the typical volume of list-decoding region. Using Chernoff's bound, we show in Lemma 26 that $\mathbb{P}(\mathcal{E}_{\underline{s_Q}}|\mathcal{E}^c)\le 2^{-\Omega(n^3)}$. The second term can be shown to also be $2^{-\Omega(n^3)}$. This implies that for any given attack vector $\underline{s_Q}$, the probability that $\underline{s_Q}$ can force a large list-size for any codewords is super-exponentially small. To complete the proof, we take a union bound over $\underline{z_Q}$, the strips, the OGSs and $\underline{s_Q}$.

The whole bounding procedure (including myopic list-decoding in this section and unique decoding in Section XI) is depicted in Figure 21.

E. Event $\mathcal{E}_{\mathrm{atyp}}$: Analysis of Atypical Behaviour of James's Observation

Though, as already mentioned in Section X-C, we still use the shorthand notation $\underline{\mathbf{x}}$ to denote $\underline{\mathbf{x}}(\mathbf{m},\mathbf{k})$, results in this subsection in fact hold regardless of the distribution of $\underline{\mathbf{x}}$ and in particular the readers, if they want, can take $\underline{\mathbf{x}}$ to be any fixed vector $\underline{x}$ on $\mathcal{S}^{n-1}(0,\sqrt{nP})$.

From Fact 7, we know the probability that $\underline{\mathbf{s}}_z,\underline{\mathbf{x}}$ are not jointly typical vanishes as $n\to\infty$. Specifically, the AWGN to James is independent of everything else, hence has norm concentrating around $\sqrt{n\sigma^2}$ and is appproximately orthogonal to $\underline{\mathbf{x}}$.

$$\mathbb{P}(\mathcal{E}_{\mathrm{atyp}_1})=\mathbb{P}(\|\underline{\mathbf{s}}_z\|_2\notin\sqrt{n\sigma^2(1\pm\varepsilon)}) \le 2\exp(-\varepsilon^2 n/4)=:2^{-f_1(\varepsilon)n}.$$

Since $\underline{\mathbf{s}}_z$ is AWGN independent of $\underline{\mathbf{x}}$, the average dot product between the two vectors is zero. We can further bound the probability that the cosine of the angle between the two exceeds ε.

$$\begin{aligned}
&\mathbb{P}(\mathcal{E}_{\mathrm{atyp}_2})\\
&=\mathbb{P}(|\cos(\angle_{\underline{\mathbf{x}},\underline{\mathbf{s}}_z})|\ge\varepsilon)\\
&=\mathbb{P}\left(\left|\frac{\langle\underline{\mathbf{x}},\underline{\mathbf{s}}_z\rangle}{\|\underline{\mathbf{x}}\|_2\|\underline{\mathbf{s}}_z\|_2}\right|\ge\varepsilon\right)\\
&=\mathbb{P}\left(\left|\frac{\langle\underline{e}_1,\underline{\mathbf{s}}_z\rangle}{\|\underline{\mathbf{s}}_z\|_2}\right|\ge\varepsilon\right)\\
&=\mathbb{P}(|\underline{\mathbf{s}}_{z_1}|\ge\varepsilon\|\underline{\mathbf{s}}_z\|_2)\\
&\le\mathbb{P}(|\underline{\mathbf{s}}_{z_1}|\ge\varepsilon\|\underline{\mathbf{s}}_z\|_2,\ \|\underline{\mathbf{s}}_z\|_2\in\sqrt{n\sigma^2(1\pm\varepsilon)})\\
&\quad+\mathbb{P}(\|\underline{\mathbf{s}}_z\|_2\notin\sqrt{n\sigma^2(1\pm\varepsilon)})\\
&\le\mathbb{P}(|\underline{\mathbf{s}}_{z_1}|\ge\varepsilon\sqrt{n\sigma^2(1-\varepsilon)})+\mathbb{P}(\|\underline{\mathbf{s}}_z\|_2\notin\sqrt{n\sigma^2(1\pm\varepsilon)})\\
&\le 2\exp(-\varepsilon^2(1-\varepsilon)n/2)+2^{-f_1(\varepsilon)n}\\
&=:2^{-f_2(\varepsilon)n},
\end{aligned}\tag{X.19}$$

where in Equation (X.19), without loss of generality, we assume $\underline{\mathbf{x}}/\|\underline{\mathbf{x}}\|_2=\underline{e_1}$, where $\underline{e_1}=(1,0,\ldots,0)^T$ is the unit vector along the first dimension. Notice that

$$\|\underline{\mathbf{z}}\|_2^2=\|\underline{\mathbf{x}}+\underline{\mathbf{s}}_z\|_2^2=\|\underline{\mathbf{x}}\|_2^2+\|\underline{\mathbf{s}}_z\|_2^2+2\langle\underline{\mathbf{x}},\underline{\mathbf{s}}_z\rangle.$$

Choose the smallest $\varepsilon_1:=\varepsilon_1(\varepsilon)$ that satisfies $n(P+\sigma^2)(1\pm\varepsilon)\subset nP+n\sigma^2(1\pm\varepsilon_1)\pm 2\sqrt{nP}\sqrt{n\sigma^2(1+\varepsilon_1)}\varepsilon_1$. We can also concentrate James's observation. In the following, the probability is computed with respect to the product distribution of $(\underline{\mathbf{x}},\underline{\mathbf{s}}_z)$ since they are independent.

$$\begin{aligned}
&\mathbb{P}(\mathcal{E}_{\mathrm{atyp}_3})=\mathbb{P}(\|\underline{\mathbf{z}}\|_2^2\notin n(P+\sigma^2)(1\pm\varepsilon))\\
&\le\mathbb{P}(\|\underline{\mathbf{z}}\|_2^2\notin nP+n\sigma^2(1\pm\varepsilon_1)\pm\sqrt{nP}\sqrt{n\sigma^2(1+\varepsilon_1)}\varepsilon_1)\\
&\le\mathbb{P}(\|\underline{\mathbf{s}}_z\|_2^2\notin n\sigma^2(1\pm\varepsilon_1))\\
&\quad+\mathbb{P}(|\langle\underline{\mathbf{x}},\underline{\mathbf{s}}_z\rangle|\ge\sqrt{nP}\sqrt{n\sigma^2(1+\varepsilon_1)}\varepsilon_1)\\
&=\mathbb{P}(\|\underline{\mathbf{s}}_z\|_2^2\notin n\sigma^2(1\pm\varepsilon_1))\\
&\quad+\mathbb{P}(|\langle\underline{\mathbf{x}},\underline{\mathbf{s}}_z\rangle|\ge\sqrt{nP}\sqrt{n\sigma^2(1+\varepsilon_1)}\varepsilon_1,\|\underline{\mathbf{s}}_z\|_2^2\in n\sigma^2(1\pm\varepsilon_1))\\
&\quad+\mathbb{P}(|\langle\underline{\mathbf{x}},\underline{\mathbf{s}}_z\rangle|\ge\sqrt{nP}\sqrt{n\sigma^2(1+\varepsilon_1)}\varepsilon_1,\|\underline{\mathbf{s}}_z\|_2^2\notin n\sigma^2(1\pm\varepsilon_1))\\
&\le 2\,\mathbb{P}(\|\underline{\mathbf{s}}_z\|_2^2\notin n\sigma^2(1\pm\varepsilon_1))+\mathbb{P}(|\langle\underline{\mathbf{x}},\underline{\mathbf{s}}_z\rangle|\ge\|\underline{\mathbf{x}}\|_2\|\underline{\mathbf{s}}_z\|_2\varepsilon_1)\\
&\le 2\cdot 2^{-f_1(\varepsilon_1)n}+2^{-f_2(\varepsilon_1)n}\\
&=:2^{-f_3(\varepsilon)n}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{P}(\mathcal{E}_{\mathrm{atyp}})&\le\mathbb{P}(\mathcal{E}_{\mathrm{atyp}_1})+\mathbb{P}(\mathcal{E}_{\mathrm{atyp}_2})+\mathbb{P}(\mathcal{E}_{\mathrm{atyp}_3})\\
&\le 2^{-f_1(\varepsilon)n}+2^{-f_2(\varepsilon)n}+2^{-f_3(\varepsilon)n}=:2^{-nf_{\mathrm{atyp}}(\varepsilon)},
\end{aligned}$$

where $f_{\mathrm{atyp}}(\varepsilon)$ is positive as long as $\varepsilon>0$ and $\lim_{\varepsilon\downarrow 0}f_{\mathrm{atyp}}(\varepsilon)=0$.

F. Event $\mathcal{E}_{\mathrm{str}}$: Number and Distribution of Codewords in Strips: Exponential Behaviour and Quasi-Uniformity

The intersection of $\mathcal{S}h^n(\underline{\mathbf{z}_Q},\sqrt{n\sigma^2(1\pm\varepsilon)})$ with $\mathcal{S}^{n-1}(0,\sqrt{nP})$ forms a thick strip $\mathcal{S}tr^{n-1}(\underline{\mathbf{z}_Q}):=\bigcup_i\mathcal{S}tr^{n-1}(\underline{\mathbf{z}_Q},i)$, i.e., the union of all thin strips, which we study next. For ease of incoming calculations, let us first translate the slacks in the distances from the strips to $\underline{\mathbf{z}_Q}$, i.e., ε and δ as afore-defined, to slacks in the radii $\sqrt{nr_{\mathrm{str}}}$ of strips, i.e, ρ and τ, respectively. Recall that the set of strips is defined as follows

$$\begin{aligned}
\mathcal{S}tr^{n-1}(\underline{\mathbf{z}_Q},i)&=\mathcal{S}^{n-1}(0,\sqrt{nP})\\
&\cap\mathcal{S}h^n(\underline{\mathbf{z}_Q},\sqrt{n\sigma^2(1+(i-1)\delta)},\sqrt{n\sigma^2(1+i\delta)}),\\
&\qquad\forall i\in\{-\varepsilon/\delta+1,\ldots,\varepsilon/\delta\}.
\end{aligned}\tag{X.20}$$

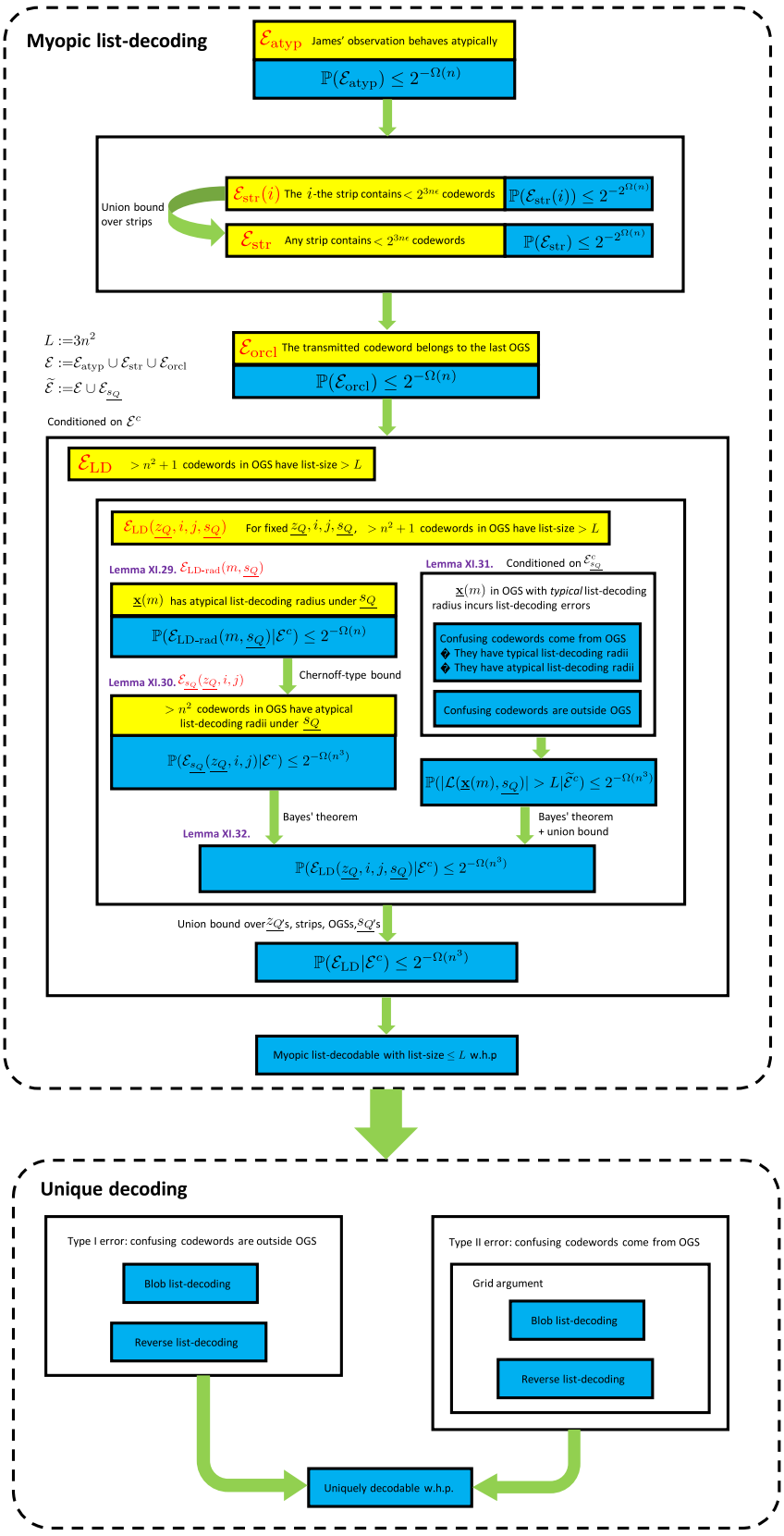


Fig. 21. A flowchart describing the procedure of bounding error probability. Some notation is simplified for ease of drawing.

Now we write it in a slightly different form

$$\mathcal{S}tr^{n-1}(\underline{\mathbf{z}_Q}, i) = \mathcal{C}ap^{n-1}(\cdot, \sqrt{nr_{\text{str}}(1+i\tau)}, \sqrt{nP}) \setminus \mathcal{C}ap^{n-1}(\cdot, \sqrt{nr_{\text{str}}(1+(i-1)\tau)}, \sqrt{nP}), \quad \forall i \in \{-\rho/\tau+1, \ldots, \rho/\tau\}. \tag{X.21}$$

Note that $\rho/\tau = \varepsilon/\delta$. Define $d_i := \sqrt{n\sigma^2(1+i\delta)}$ and $r_{\text{str},i} := \sqrt{nr_{\text{str}}(1+i\tau)}$, for any $i \in \{-\varepsilon/\delta+1, \ldots, \varepsilon/\delta\}$.

Then by Heron's formula,

$$\frac{1}{2}\|\underline{z_Q}\|_2 r_{\mathrm{str},i} = \sqrt{s(s-d_i)(s-\|\underline{z_Q}\|_2)(s-\sqrt{nP})},$$

where

$$s = \frac{1}{2}(d_i + \|\underline{z_Q}\|_2 + \sqrt{nP}).$$

Solving the equation, we have

$$r_{\mathrm{str},i} = \frac{2}{\|\underline{z_Q}\|_2}\sqrt{s(s-d_i)(s-\|\underline{z_Q}\|_2)(s-\sqrt{nP})}.$$

It follows that ρ and τ only differ by a constant factor from ε and δ, respectively.

As mentioned, codewords are *almost* uniformly distributed in the strip from James's point of view. Given $\underline{\mathbf{z}}$, we now characterize the *quasi-uniformity* in terms of τ. Define quasi-uniformity factor

$$\Delta(\tau) := \sup_{\underline{z}\in\mathcal{S}h^n(0,\sqrt{n(P+\sigma^2)(1\pm\varepsilon)})} \max_i \sup_{\underline{x}^{(1)},\underline{x}^{(2)}\in\mathcal{S}tr^{n-1}(\underline{z_Q},i)} \frac{p_{\underline{\mathbf{x}}|\underline{\mathbf{z}}}(\underline{x}^{(1)}|\underline{z_Q})}{p_{\underline{\mathbf{x}}|\underline{\mathbf{z}}}(\underline{x}^{(2)}|\underline{z_Q})}, \tag{X.22}$$

where the conditional density is determined by the joint law $p_{\underline{\mathbf{x}},\underline{\mathbf{z}}}$ where $\underline{\mathbf{x}} \sim \mathrm{Unif}(\mathcal{S}^{n-1}(0,\sqrt{nP}))$, $\underline{\mathbf{z}} = \underline{\mathbf{x}} + \underline{\mathbf{s}}_z$ and $\underline{\mathbf{s}}_z \sim \mathcal{N}(0,\sigma^2\mathbf{I}_n)$.

We will show that under appropriate choice of parameters, the above ratio is small maximized over $\underline{x}^{(1)},\underline{x}^{(2)}$ in $\mathcal{S}tr^{n-1}(\underline{z_Q},i)$. By our random code construction, each codeword indeed follows the distribution $\underline{\mathbf{x}}(m,k) \sim \mathrm{Unif}(\mathcal{S}^{n-1}(0,\sqrt{nP}))$ for any (m,k). Hence the ratio remains small if $\underline{x}^{(1)},\underline{x}^{(2)}$ are respectively replaced by $\underline{x}(m_1,k_2),\underline{x}(m_2,k_2)$ for some $(m_1,k_1),(m_2,k_2)$ in $\mathcal{M}_{\mathrm{str}}(\underline{z_Q},i)$.

Lemma 23 (Quasi-Uniformity): For appropriate choices of the small constant ρ and $\tau = \Theta((\log n)/n)$, conditioned on $\mathcal{E}^c_{\mathrm{atyp}}$, we have $\Delta(\tau) = \mathcal{O}(\mathrm{poly}(n))$.

Proof: As shown in Figure 22, obviously, for fixed $\underline{z}$ and i, the inner supremum is achieved by a point $\underline{x}^-$ on the upper boundary (closer to $\underline{z}$) of the strip and a point $\underline{x}^+$ on the lower boundary (further from $\underline{z}$) of the strip. Let r_{str} be such that $|\overline{xO'}| = \sqrt{nr_{\mathrm{str}}}$. Calculations in Appendix F show that

$$\sup_{\underline{x}^{(1)},\underline{x}^{(2)}\in\mathcal{S}tr^{n-1}(\underline{z},i)} \frac{p_{\underline{\mathbf{x}}|\underline{\mathbf{z}}}(\underline{x}^{(1)}|\underline{z_Q})}{p_{\underline{\mathbf{x}}|\underline{\mathbf{z}}}(\underline{x}^{(2)}|\underline{z_Q})} = \exp\left(\frac{\|\underline{z}\|_2+\sqrt{n\delta_{\mathcal{Z}}}}{\sigma^2}\frac{2nr_{\mathrm{str}}\tau}{\sqrt{n(P-r_-)}+\sqrt{n(P-r_+)}}\right),$$

where $r_- := r_{\mathrm{str}}(1-\tau)$ and $r_+ := r_{\mathrm{str}}(1+\tau)$ as defined in Eqn. (F.1). Then, conditioned on $\mathcal{E}^c_{\mathrm{atyp}}$, the quasi-uniformity factor is upper bounded by

$$\begin{aligned}\Delta(\tau) \le& \sup_{\underline{z}\in\mathcal{S}h^n(0,\sqrt{n(P+\sigma^2)(1\pm\varepsilon)})} \max_i \exp\left(\frac{\|\underline{z}\|_2+\sqrt{n\delta_{\mathcal{Z}}}}{\sigma^2}\frac{2nr_{\mathrm{str}}\tau}{\sqrt{n(P-r_-)}+\sqrt{n(P-r_+)}}\right)\\ \le& \exp\left(\frac{\sqrt{n(P+\sigma^2)(1+\varepsilon)}+\sqrt{n\delta_{\mathcal{Z}}}}{\sigma^2}\frac{2nr_{\mathrm{str}}\tau}{\sqrt{n(P-r_{\mathrm{str}}(1-\tau))}+\sqrt{n(P-r_{\mathrm{str}}(1+\tau))}}\right)\\ \le& \exp\left(\frac{\sqrt{(P+\sigma^2)(1+\varepsilon)}+\sqrt{\delta_{\mathcal{Z}}}}{\sigma^2}\frac{2n\tau\frac{P\sigma^2(1+\varepsilon)}{(P+\sigma^2)(1-\varepsilon)}}{\sqrt{P-\frac{P\sigma^2(1+\varepsilon)(1-\tau)}{(P+\sigma^2)(1-\varepsilon)}}+\sqrt{P-\frac{P\sigma^2(1+\varepsilon)(1+\tau)}{(P+\sigma^2)(1-\varepsilon)}}}\right).\end{aligned} \tag{X.23}$$

Eqn. (X.23) follows since the bound is increasing in r_{str}. Bounds on r_{str} can be obtained as follows. In the triangle ΔxzO, we have

$$\frac{1}{2}\|\underline{z}\|_2\sqrt{n\mathbf{r}_{\mathrm{str}}} = \frac{1}{2}\|\underline{\mathbf{x}}\|_2\|\underline{\mathbf{s}}_z\|\sin(\angle_{\underline{\mathbf{x}},\underline{\mathbf{s}}_z}),$$

which implies

$$\mathbf{r}_{\mathrm{str}} = \frac{P\|\underline{\mathbf{s}}_z\|_2(1-\cos(\angle_{\underline{\mathbf{x}},\underline{\mathbf{s}}_z})^2)}{\|\underline{z}\|_2}.$$

Conditioned on $\mathcal{E}^c_{\mathrm{atyp}}$,

$$\frac{P\sigma^2(1-\varepsilon)}{(P+\sigma^2)(1+\varepsilon)} \le r_{\mathrm{str}} \le \frac{P\sigma^2(1+\varepsilon)}{P\sigma^2(1-\varepsilon)}. \tag{X.24}$$

We have $\Delta(\tau) = \mathcal{O}(\mathrm{poly}(n))$ by taking $\tau = \Theta((\log n)/n)$. □

Next, we show that if the codebook rate is large enough, then with high probability (over the randomness in the codebook generation) every strip will contain exponentially many codewords.

Lemma 24 (Exponentially Many Codewords in Strips): Let $R_{\mathrm{code}} > \frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right)$. Then, from James's perspective, he is confused with exponentially many codewords in the strip with probability doubly exponentially close to one.

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{\mathrm{str}}|\mathcal{E}^c_{\mathrm{atyp}}) =& \mathbb{P}\left(\bigcup_i \mathcal{E}_{\mathrm{str}}(i)\middle|\mathcal{E}^c_{\mathrm{atyp}}\right)\\ =& \mathbb{P}\left(\exists i,\ |\mathcal{M}_{\mathrm{str}}(\underline{z_Q},i)| \le 2^{3\varepsilon n}\middle|\mathcal{E}^c_{\mathrm{atyp}}\right)\\ =& \mathbb{P}\left(\exists i,\ |\mathcal{S}tr^{n-1}(\underline{z_Q},i)\cap\mathcal{C}| \le 2^{3\varepsilon n}\middle|\mathcal{E}^c_{\mathrm{atyp}}\right)\\ \le& 2^{-2^{\Omega(n)}}.\end{aligned}$$

Remark 10: Note that James's uncertainty set contains codewords from the whole codebook $\mathcal{C}$, not only from $\mathcal{C}^{(\mathbf{k})}$ for some particular $\mathbf{k}$, since the shared key is assumed to be kept secret from James.

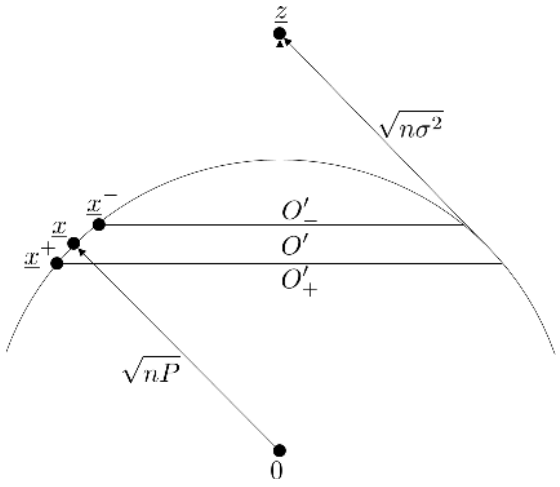


Fig. 22. For any James's observation $\underline{z}$, a thin strip containing the transmitted codeword $\underline{x}$ is constructed on the coding sphere $\mathcal{S}^{n-1}(0,\sqrt{nP})$. Typically, the geometry is shown in above figure. From James's perspective, given $\underline{z}$, codewords in the strip are approximately equally likely to be transmitted by Alice. The quasi-uniformity is defined as the maximum deviation of probability of codewords in the strip. Notice that any codeword at the same latitude has exactly the same probability. Codewords on the upper (respectively lower) boundary of the strip, say $\underline{x}^-$ (respectively $\underline{x}^+$), are most (respectively least) likely in the strip to be transmitted. For small enough ($\mathcal{O}((\log n)/n)$) thickness of the strip, the quasi-uniformity factor is a polynomial in n.

Proof: First, in Appendix G, we show that, for any typical $\underline{z_Q}$ and i,

$$\mathbb{E}\left(|\mathcal{S}tr^{n-1}(\underline{z_Q},i)\cap\mathcal{C}|\Big|\,\mathcal{E}^c_{\rm atyp}\right)\geq 2^{4\varepsilon n}.$$

Note that the random variable $|\mathcal{S}tr^{n-1}(\underline{z_Q},i)\cap\mathcal{C}|$ can be written as a sum of a bunch of independent indicator variables

$$|\mathcal{S}tr^{n-1}(\underline{z_Q},i)\cap\mathcal{C}|=\sum_{(m,k)\in[2^{n(R+R_{\rm key})}]}\mathbb{1}_{\{\underline{\mathbf{x}}(m,k)\in\mathcal{S}tr^{n-1}(\underline{z_Q},i)\}},$$

or in slightly different notation

$$|\mathcal{M}_{\rm str}(\underline{z_Q},i)|=\sum_{(m,k)\in[2^{n(R+R_{\rm key})}]}\mathbb{1}_{\{(m,k)\in\mathcal{M}_{\rm str}(\underline{z_Q},i)\}}.$$

Thus by Chernoff bound,

$$\mathbb{P}(\mathcal{E}_{\rm str}(i)|\mathcal{E}^c_{\rm atyp})=\mathbb{P}(|\mathcal{S}tr^{n-1}(\underline{z_Q},i)\cap\mathcal{C}|\leq 2^{3\varepsilon n}|\mathcal{E}^c_{\rm atyp})\leq 2^{-2^{\Omega(n)}}.$$

Note that there are at most $2\rho/\tau=\mathcal{O}(n/\log n)$ many i's. Lemma 24 is then obtained by taking a union bound over all i's. □

G. Event $\mathcal{E}_{\rm orcl}$: Transmitted Codeword Falls Into the Last Block

Conditioned on $\underline{\mathbf{z}}$ and the OGS, the transmitted codeword is quasi-uniformly distributed over the strip that contains the OGS. Given $\mathcal{E}^c_{\rm atyp}$ and $\mathcal{E}^c_{\rm str}$, there are at least $2^{3\varepsilon n}$ many codewords in each strip. Also, notice that each OGS (except perhaps the last one) is of size $2^{n\varepsilon}$. Therefore, the probability over $\mathbf{m},\mathbf{k}$ that $\mathcal{E}_{\rm orcl}$ occurs can be bounded as follows.

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{\rm orcl}|\mathcal{E}^c_{\rm atyp}\cap\mathcal{E}^c_{\rm str})&\leq\frac{1}{\ell}\Delta(\tau)\\&=\frac{1}{\lceil|\mathcal{M}_{\rm str}(\underline{z_Q},i)|/2^{n\varepsilon}\rceil}\Delta(\tau)\\&\leq\frac{1}{2^{3\varepsilon n}/2^{n\varepsilon}}\Delta(\tau)=2^{-2\varepsilon n}\Delta(\tau).\end{aligned}$$

H. Event $\mathcal{E}_{\rm LD}$: Existence of a "Bad" Attack Vector

At first, we fix $\underline{z_Q}\in\mathcal{Z},i\in\{-\varepsilon/\delta+1,\ldots,\varepsilon/\delta\}$, $j\in[\ell],\underline{s_Q}\in\mathcal{S}$. We will show that the probability that the list-size is greater than L is super-exponentially small in n. We will finally use a quantization argument and take a union bound over $\underline{z_Q},i,j,\underline{s_Q}$ to show that $\mathbb{P}(\mathcal{E}_{\rm LD}|\mathcal{E}^c)=o(1)$.

For any $(m,\mathbf{k})\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i)$, to prove that $\mathbb{P}(|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})|>L|\mathcal{E}^c)$ is super-exponentially decaying, we will find the typical value of $\sqrt{n\mathbf{r}}$, where $\mathbf{r}:=\mathbf{r}(m,\underline{s_Q})$ is the normalized radius of the list-decoding region $\mathcal{B}^n(\underline{\mathbf{x}}(m,\mathbf{k})+\underline{s_Q},\sqrt{nN}+\sqrt{n\delta_{\mathcal{S}}})\cap\mathcal{S}^{n-1}(0,\sqrt{nP})$ (which is nothing but a cap), and use this to obtain an upper bound on the probability that the list-size is large.

Recall that the typical radius is defined as $r_{\rm opt}(\underline{s_Q}):=\mathbb{E}(\mathbf{r}(m,\underline{s_Q}))$ (Eqn. (X.12)). This will be obtained as the solution to an optimization problem (X.49) and actually corresponds to the worst-case $\underline{s}$ that James can choose for the given OGS. Recall that $\mathcal{E}_{\rm LD-rad}(m,\underline{s_Q})$ denotes the event that the radius of the list-decoding region $\mathcal{B}^n(\underline{\mathbf{x}}(m,\mathbf{k})+\underline{s_Q},\sqrt{nN}+\sqrt{n\delta_{\mathcal{S}}})\cap\mathcal{S}^{n-1}(0,\sqrt{nP})$ is not typical, *i.e.* that $\mathbf{r}$ is much larger than $r_{\rm opt}$. Let $\mathcal{J}$ denote the event that $(\mathbf{z}_Q,\mathbf{s}_Q)=(\underline{z_Q},\underline{s_Q})$ and the transmitted codeword lies in $\mathcal{O}rcl^{(j)}(\underline{z_Q},i)$. In Section X-I, we will show the following:

Lemma 25: Fix $\underline{z_Q}$, i, j and $\underline{s_Q}$. There exists $f_{11}(\varepsilon,\delta_{\mathcal{S}})$ satisfying $f_{11}(\varepsilon,\delta_{\mathcal{S}})\to 0$ as $\varepsilon\to 0$ such that for every (m,k) in the OGS,

$$\begin{aligned}&\mathbb{P}(\mathcal{E}_{\rm LD-rad}(m,\underline{s_Q})|\mathcal{E}^c\cap\mathcal{J})\\&\quad=\mathbb{P}\left(\mathbf{r}(m,\underline{s_Q})>r_{\rm opt}(\underline{s_Q})(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))\Big|\,\mathcal{E}^c\cap\mathcal{J}\right)\\&\qquad\leq 2^{-f_9(\varepsilon,\eta,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})n},\end{aligned}$$

where $\mathcal{E}=\mathcal{E}_{\rm atyp}\cup\mathcal{E}_{\rm str}\cup\mathcal{E}_{\rm orcl}$ and $f_9(\varepsilon,\eta,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})$ can be taken as $\frac{3}{2}\varepsilon$ by choosing proper η, $\delta_{\mathcal{S}}$ and $\delta_{\mathcal{Z}}$.

Recall that $\mathcal{E}_{\underline{s_Q}}$ is the event that the radius of the list-decoding region $\mathbf{r}$ is greater than $r_{\rm opt}(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))$ for more than n^2 codewords in the OGS. Since codewords in OGS are quasi-uniformly distributed and independent, using Chernoff-type bound similar to Lemma 9 and the above lemma, we get that

Lemma 26: Fix $\underline{z_Q}$, i, j and $\underline{s_Q}$. Then

$$\begin{aligned}&\mathbb{P}(\mathcal{E}_{\underline{s_Q}}(\underline{z_Q},i,j)|\mathcal{E}^c\cap\mathcal{J})\\&=\mathbb{P}\Big(|\{(m,\mathbf{k})\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i):\\&\qquad\mathbf{r}(m,\underline{s_Q})>r_{\rm opt}(\underline{s_Q})(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))\}|>n^2\Big|\,\mathcal{E}^c\cap\mathcal{J}\Big)\end{aligned}$$

$$\leq \mathbb{P}\left(\sum_{(m,k)\in \mathcal{O}rcl^{(j)}(\underline{z_Q},i)} \mathbb{1}_{\mathcal{E}_{\mathrm{LD-rad}}(m,\underline{s_Q})} > n^2 \middle| \mathcal{E}^c \cap \mathcal{J} \right)$$
$$\leq 2^{-\Omega(n^3)}.$$

In Section X-J, we will show the following:

Lemma 27: Fix $\underline{z_Q}$, i, j and $\underline{s_Q}$. For any $(m,\mathbf{k}) \in \mathcal{O}rcl^{(j)}(\underline{z_Q},i)$ for which $\mathbf{r}$ is typical, we have

$$\mathbb{P}(|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})| > L | \widetilde{\mathcal{E}}^c \cap \mathcal{J}) \leq 2^{-\Omega(n^3)},$$

where L is set to be $3n^2$ and $\widetilde{\mathcal{E}}$ denotes $\mathcal{E} \cup \mathcal{E}_{\underline{s_Q}}(\underline{z_Q},i,j)$.

Lemmas 25 and 27 allow us to conclude that the probability that there exist n^2+1 codewords with list-sizes greater than L is super-exponentially small. Recall that $\mathcal{J}$ denotes the realization of James's knowledge, i.e., the event that $(\underline{\mathbf{z}}_Q, \underline{\mathbf{s}}_Q) = (\underline{z_Q}, \underline{s_Q})$ and $(\mathbf{m},\mathbf{k}) \in \mathcal{O}rcl^{(j)}(\underline{z_Q},i)$.

Lemma 28: Fix $\underline{z_Q}$, i, j and $\underline{s_Q}$. Then

$$\mathbb{P}(\mathcal{E}_{\mathrm{LD}}(\underline{z_Q},i,j,\underline{s_Q}) | \mathcal{E}^c \cap \mathcal{J}) \leq 2^{-\Omega(n^3)}.$$

Proof: Using Fact 5 we obtain:

$$\mathbb{P}(\mathcal{E}_{\mathrm{LD}}(\underline{z_Q},i,j,\underline{s_Q}) | \mathcal{E}^c \cap \mathcal{J}) \leq \mathbb{P}(\mathcal{E}_{\mathrm{LD}}(\underline{z_Q},i,j,\underline{s_Q}) | \widetilde{\mathcal{E}}^c \cap \mathcal{J}) + \mathbb{P}(\mathcal{E}_{\underline{s_Q}}(\underline{z_Q},i,j) | \mathcal{E}^c \cap \mathcal{J}).$$

The second term is bounded by Lemma 26.

The first term is super-exponentially small by Lemma 27 and union bound. The probabilities are computed with respect to the quasi-uniform (due to the conditioning) distribution of the codewords in the OGS over the strip, and the uniform distribution of the remaining codewords over the sphere (conditioned on the OGS, they are independent of James's observations).

$$\mathbb{P}\left(\mathcal{E}_{\mathrm{LD}}(\underline{z_Q},i,j,\underline{s_Q}) \middle| \widetilde{\mathcal{E}}^c \cap \mathcal{J} \right)$$
$$= \mathbb{P}\left(\chi(\underline{z_Q},i,j,\underline{s_Q}) > n^2+1 \middle| \widetilde{\mathcal{E}}^c \cap \mathcal{J} \right) \tag{X.25}$$
$$= \mathbb{P}\left(\sum_{(m,\mathbf{k})\in \mathcal{O}rcl^{(j)}(\underline{z_Q},i)} \mathbb{1}_{\{|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})|>L\}} > n^2+1 \middle| \widetilde{\mathcal{E}}^c \cap \mathcal{J} \right) \tag{X.26}$$
$$= \mathbb{P}\left(\sum_{(m,\mathbf{k})\in \mathcal{O}rcl^{(j)}(\underline{z_Q},i)} \mathbb{1}_{\{|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})|>L\}} \left(\mathbb{1}_{\mathcal{E}_{\mathrm{LD-rad}}(m,\underline{s_Q})} + \mathbb{1}_{\mathcal{E}_{\mathrm{LD-rad}}(m,\underline{s_Q})^c} \right) > n^2+1 \middle| \widetilde{\mathcal{E}}^c \cap \mathcal{J} \right)$$
$$\leq \mathbb{P}\left(\sum_{(m,\mathbf{k})\in \mathcal{O}rcl^{(j)}(\underline{z_Q},i)} \mathbb{1}_{\{|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})|>L\}} \mathbb{1}_{\mathcal{E}_{\mathrm{LD-rad}}(m,\underline{s_Q})} > n^2 \middle| \widetilde{\mathcal{E}}^c \cap \mathcal{J} \right)$$
$$+ \mathbb{P}\left(\sum_{(m,\mathbf{k})\in \mathcal{O}rcl^{(j)}(\underline{z_Q},i)} \mathbb{1}_{\{|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})|>L\}} \mathbb{1}_{\mathcal{E}_{\mathrm{LD-rad}}(m,\underline{s_Q})^c} > 1 \middle| \widetilde{\mathcal{E}}^c \cap \mathcal{J} \right) \tag{X.27}$$
$$\leq \mathbb{P}\left(\sum_{(m,\mathbf{k})\in \mathcal{O}rcl^{(j)}(\underline{z_Q},i)} \mathbb{1}_{\mathcal{E}_{\mathrm{LD-rad}}(m,\underline{s_Q})} > n^2 \middle| \widetilde{\mathcal{E}}^c \cap \mathcal{J} \right) \tag{X.28}$$
$$+ \mathbb{P}\left(\sum_{(m,\mathbf{k})\in \mathcal{O}rcl^{(j)}(\underline{z_Q},i)} \mathbb{1}_{\{|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})|>L\}} \mathbb{1}_{\mathcal{E}_{\mathrm{LD-rad}}(m,\underline{s_Q})^c} \geq 1 \middle| \widetilde{\mathcal{E}}^c \cap \mathcal{J} \right)$$
$$= \mathbb{P}\left(\psi(\underline{z_Q},i,j,\underline{s_Q}) > n^2 \middle| \widetilde{\mathcal{E}}^c \cap \mathcal{J} \right) \tag{X.29}$$
$$+ \mathbb{P}\Big(\exists (m,\mathbf{k}) \in \mathcal{O}rcl^{(j)}(\underline{z_Q},i), \mathbf{r}(m,\underline{s_Q}) \leq r_{\mathrm{opt}}(\underline{s_Q})(1+f_{11}(\varepsilon,\delta_{\mathcal{S}})), |\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})| > L \Big| \widetilde{\mathcal{E}}^c \cap \mathcal{J} \Big) \tag{X.30}$$
$$\leq \mathbb{P}\left(\mathcal{E}_{\underline{s_Q}}(\underline{z_Q},i,j) \middle| \widetilde{\mathcal{E}}^c \cap \mathcal{J} \right) + 2^{n\varepsilon}\, \mathbb{P}\left(|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})| > L \middle| \widetilde{\mathcal{E}}^c \cap \mathcal{J} \right) \tag{X.31}$$
$$\leq 0 + 2^{n\varepsilon} 2^{-\Omega(n^3)} \tag{X.32}$$
$$= 2^{-\Omega(n^3)}$$

Eqn. (X.25) is by the definition of $\mathcal{E}_{\mathrm{LD}}(\underline{z_Q},i,j,\underline{s_Q})$ (Eqn. (X.17)). Eqn. (X.26) is by the definition of $\chi(\underline{z_Q},i,j,\underline{s_Q})$ (Eqn. (X.16)). Eqn. (X.27) is by the union bound. In Eqn. (X.28), we upper bound $\mathbb{1}_{\mathcal{E}_{\mathrm{LD-rad}}(m,\underline{s_Q})}$ by 1. Eqn. (X.29) follows from the definition of $\psi(\underline{z_Q},i,j,\underline{s_Q})$ (Eqn. (X.14)). Eqn. (X.30) follows from the definition of $\mathcal{E}_{\mathrm{LD-rad}}(m,\underline{s_Q})$ (Eqn. (X.13)). In Eqn. (X.31), the first term follows from the definition of $\mathcal{E}_{\underline{s_Q}}(\underline{z_Q},i,j)$ (Eqn. (X.15)). To get the second term in Eqn. (X.31), we drop the first event $\{\mathbf{r}(m,\underline{s_Q}) \leq r_{\mathrm{opt}}(\underline{s_Q})(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))\}$ and take a union bound. Note that conditioned on $\mathcal{E}^c_{\mathrm{orcl}} \subset \mathcal{E}^c \subset \widetilde{\mathcal{E}}^c$, we have $|\mathcal{O}rcl^{(j)}(\underline{z_Q},i)| = 2^{n\varepsilon}$. In Eqn. (X.32), the first term follows since we conditioned on $\widetilde{\mathcal{E}}^c$ where $\widetilde{\mathcal{E}} = \mathcal{E} \cup \mathcal{E}_{\underline{s_Q}}(\underline{z_Q},i,j)$. The second term of Eqn. (X.32) follows from Lemma 27.

This finishes the proof of Lemma 28. □

There are at most $2^{\mathcal{O}(n)}$ many $\underline{z_Q}$'s in the covering $\mathcal{Z}$ of $\underline{z}$'s. For each $\underline{z_Q}$, there are at most $2\varepsilon/\delta = \Theta(n/\log n)$ strips. For fixed $\underline{z_Q}$ and i, a loose upper bound on the number of oracle-given sets in the strip $\mathcal{S}tr^{n-1}(\underline{z_Q},i)$ is $2^{n(R_{\mathrm{code}}-\varepsilon)}$. We are required to quantize $\underline{s}$ using a finite covering of $\mathcal{B}^n(0,\sqrt{nN})$. The steps mimic the proof of Lemma 32, and we omit the details. The argument for $\mathbb{P}(\mathcal{E}_{\mathrm{LD}}|\mathcal{E}^c) = o(1)$ follows by taking a union bound over

$$2^{\mathcal{O}(n)} \cdot \Theta(n/\log n) \cdot 2^{n(R_{\mathrm{code}}-\varepsilon)} \cdot 2^{\mathcal{O}(n)} = 2^{\mathcal{O}(n)}$$

many configurations of the four-tuple $\underline{z_Q}$, i, j and $\underline{s_Q}$. This completes the basic ingredients needed to obtain Theorem 11.

All that remains is to prove Lemma 25 and Lemma 27.

I. Proof of Lemma 25

In this section, we will upper bound the average area of the list-decoding region over James's uncertainty in the OGS, and

show that with high probability it will not exceed the typical value largely.

Let us begin by taking a closer look at the geometry. Let $\underline{\mathbf{x}}$ be quasi-uniformly distributed over the strip $\mathcal{S}tr^{n-1}(\underline{z_Q}, i)$. For reasons that will be clear in the subsequent calculations, we decompose $\underline{\mathbf{x}}$ and $\underline{s_Q}$ into sums of vectors parallel and orthogonal to $\underline{z_Q}$,

$$\underline{\mathbf{x}} = \underline{\mathbf{x}_Q}^{\parallel} + \underline{\mathbf{x}_Q}^{\perp}, \quad \underline{s_Q} = \underline{s_Q}^{\parallel} + \underline{s_Q}^{\perp},$$

where $\underline{e}^{\parallel}$ denotes the unit vector along $\underline{z_Q}$ and

$$\begin{aligned} \underline{\mathbf{x}_Q}^{\parallel} &:= \sqrt{n\boldsymbol{\alpha}_x}\underline{e}^{\parallel}, & \underline{\mathbf{x}_Q}^{\perp} &:= \sqrt{n\boldsymbol{\beta}_x}\underline{\mathbf{e}}^{\perp}, \\ \underline{s_Q}^{\parallel} &:= -\sqrt{n\alpha_s}\underline{e}^{\parallel}, & \underline{s_Q}^{\perp} &:= \sqrt{n\beta_s}\underline{e_s}^{\perp}, \\ \sqrt{n\boldsymbol{\alpha}_x} &= \frac{\langle \underline{\mathbf{x}}, \underline{z_Q}\rangle}{\|\underline{z_Q}\|_2}, & -\sqrt{n\alpha_s} &= \frac{\langle \underline{s_Q}, \underline{z_Q}\rangle}{\|\underline{z_Q}\|_2}. \end{aligned}$$

Remark 11: In this remark, we clarify the notation in the above decomposition. Since $\underline{z_Q}$ is given, the unit vector $\underline{e}^{\parallel}$ along $\underline{z_Q}$ is fixed (hence in non-boldface). Since $\underline{\mathbf{x}}$ is quasi-uniformly distributed over $\mathcal{S}tr^{n-1}(\underline{z_Q}, i)$, the perpendicular component of $\underline{\mathbf{x}}$ is isotropically distributed in an annulus centered at 0, perpendicular to $\underline{z_Q}$ and hence its direction $\underline{\mathbf{e}}^{\perp}$ is also isotropically distributed (therefore in boldface). Furthermore, since $\mathcal{S}tr^{n-1}(\underline{z_Q}, i)$ has certain thickness, the angle between $\underline{\mathbf{x}}$ and $\underline{z_Q}$ gets a slight variation, so $\boldsymbol{\alpha}_x, \boldsymbol{\beta}_x$ are in fact random variables (hence in boldface) as well.

On the other hand, since $\underline{z_Q}, \underline{s_Q}$ are both fixed, both components of $\underline{s_Q}$ are fixed and in particular the direction $\underline{e_s}^{\perp}$ of its perpendicular component is in non-boldface.

Note that $\underline{\mathbf{x}}$ is on $\mathcal{S}^{n-1}(0, \sqrt{nP})$ and thus $\boldsymbol{\alpha}_x + \boldsymbol{\beta}_x = P$. Note also that $\underline{s}$, and thus $\underline{s_Q}$, should satisfy James's power constraint, i.e, $\alpha_s + \beta_s \leq N$.

This decomposition will bring us analytic ease to compute the list-size, which is equivalent to computing the radius $\sqrt{n\mathbf{r}}$ of the myopic list-decoding region. It is noted that, for any $(m, k) \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)$, the list-decoding region of $\underline{x}(m, k)$ under $\underline{s_Q}$ is nothing but a cap

$$\begin{aligned} &\mathcal{B}^n(\underline{\mathbf{y}_Q}, \sqrt{nN} + \sqrt{n\delta_{\mathcal{S}}}) \cap \mathcal{S}^{n-1}(0, \sqrt{nP}) \\ =&\mathcal{B}^n(\underline{\mathbf{x}}(m, \mathbf{k}) + \underline{s_Q}, \sqrt{nN} + \sqrt{n\delta_{\mathcal{S}}}) \cap \mathcal{S}^{n-1}(0, \sqrt{nP}) \\ =&\mathcal{C}ap^{n-1}(\cdot, \sqrt{n\mathbf{r}}, \sqrt{nP}), \end{aligned}$$

where $\underline{\mathbf{y}_Q} := \underline{\mathbf{x}} + \underline{s_Q}$.

Notice that, averaged over the codebook generation, $\underline{\mathbf{e}}^{\perp}$ is uniformly distributed[18] over the unit sphere $\mathcal{S}^{n-2}(0, 1)$ orthogonal to $\underline{z_Q}$. We will use Lemma 8 to bound the tail of inner product of $\underline{\mathbf{x}_Q}^{\perp}$ and $\underline{z_Q}$.

Heuristically, in expectation, without quantization, we can compute the scale of each component of $\underline{\mathbf{x}}$ and $\underline{s_Q}$. A glimpse

[18]Technically speaking, this is indeed the case only when we decompose $\underline{\mathbf{x}}$ with respect to $\underline{z}$, but not its quantization $\underline{z_Q}$. It is not exactly, yet still approximately true when we take the quantization $\underline{z_Q}$ of $\underline{z}$. This quantization error will be taken into account via Lemma 29.

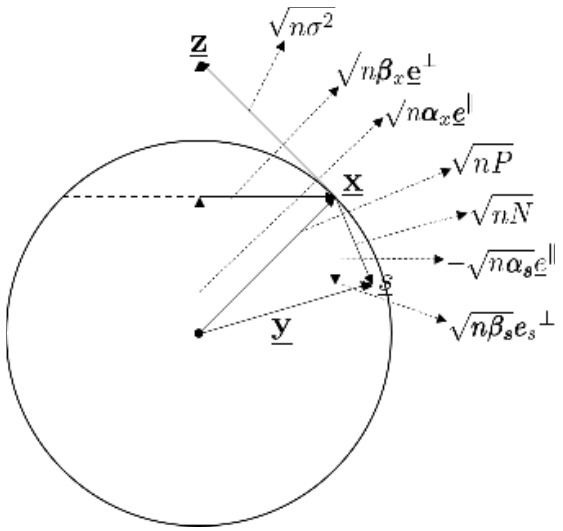


Fig. 23. In expectation, the noise to James $\underline{\mathbf{s}_z}$ is orthogonal to the codeword $\underline{\mathbf{x}}$ and of length $\sqrt{n\sigma^2}$. Fix a legitimate attack vector $\underline{s}$ and decompose $\underline{\mathbf{x}}$ and $\underline{s}$ into sums of components orthogonal and parallel to James's observation $\underline{\mathbf{z}}$. Each component in above figure can be concentrated and is robust to quantization errors.

at the geometry (shown in Figure 23) immediately gives us the following relations:

$$\frac{\sqrt{n\boldsymbol{\alpha}_x}}{\sqrt{nP}} = \frac{\sqrt{nP}}{\sqrt{n(P+\sigma^2)}} \implies \boldsymbol{\alpha}_x = \frac{P^2}{P+\sigma^2}, \tag{X.33}$$

$$\frac{\sqrt{n\boldsymbol{\beta}_x}}{\sqrt{n\sigma^2}} = \frac{\sqrt{nP}}{\sqrt{n(P+\sigma^2)}} \implies \boldsymbol{\beta}_x = \frac{P\sigma^2}{P+\sigma^2}, \tag{X.34}$$

both of which follow from similarity of triangles. In fact, the lengths of both components are well concentrated. Recall that conditioned on $\mathcal{E}^c \cap \mathcal{J}$, the transmitted codeword is quasi-uniformly distributed over the strip. For any $0 < \eta < 1$, we have

$$\begin{aligned} &\mathbb{P}\left(\boldsymbol{\alpha}_x \notin \frac{P^2}{P+\sigma^2}(1\pm\eta)\middle| \mathcal{E}^c \cap \mathcal{J}\right) \\ =&\mathbb{P}\left(\frac{\langle \underline{\mathbf{x}}, \underline{z_Q}\rangle}{\|\underline{z_Q}\|_2} \notin \sqrt{n\frac{P^2}{P+\sigma^2}(1\pm\eta)}\middle| \mathcal{E}^c \cap \mathcal{J}\right) \\ =&\mathbb{P}\left(\langle \underline{\mathbf{x}}, \underline{z_Q}\rangle \notin \sqrt{n\frac{P^2}{P+\sigma^2}(1\pm\eta)}\|\underline{z_Q}\|_2\middle| \mathcal{E}^c \cap \mathcal{J}\right) \\ \leq&\mathbb{P}\left(\langle \underline{\mathbf{x}}, \underline{z_Q}\rangle \notin \sqrt{n\frac{P^2}{P+\sigma^2}(1\pm\eta)}(\|\underline{z}\|_2 \mp \sqrt{n\delta_{\mathcal{Z}}})\middle| \mathcal{E}^c \cap \mathcal{J}\right) \\ \leq&\mathbb{P}\left(\langle \underline{\mathbf{x}}, \underline{z_Q}\rangle \notin \sqrt{n\frac{P^2}{P+\sigma^2}(1\pm\eta)}\right. \\ &\left.(\sqrt{n(P+\sigma^2)(1\mp\varepsilon)} \mp \sqrt{n\delta_{\mathcal{Z}}})\middle| \mathcal{E}^c \cap \mathcal{J}\right) \qquad \text{(X.35)} \\ \leq&\mathbb{P}(\langle \underline{\mathbf{x}}, \underline{z_Q}\rangle \notin nP(1 \pm f_4(\varepsilon, \eta, \delta_{\mathcal{Z}}))|\mathcal{E}^c \cap \mathcal{J}) \\ =:&2^{-f_5(\varepsilon,\eta,\delta_{\mathcal{Z}})n}, \qquad \text{(X.36)} \end{aligned}$$

where Inequality (X.35) follows from that $\|\underline{z}\|_2 \in \sqrt{n(P+\sigma^2)(1\pm\varepsilon)}$ since we condition on $\mathcal{E}^c \cap \mathcal{J}$, and

Inequality (X.36) follows from the following calculations.

$$\begin{aligned}
&\mathbb{P}(\langle \underline{\mathbf{x}}, \underline{z_Q}\rangle \notin nP(1\pm f_4(\varepsilon,\eta,\delta_{\mathcal{Z}}))|\mathcal{E}^c\cap\mathcal{J})\\
=&\mathbb{P}(\langle \underline{\mathbf{x}}, \underline{z}+\underline{z_e}\rangle \notin nP(1\pm f_4(\varepsilon,\eta,\delta_{\mathcal{Z}}))|\mathcal{E}^c\cap\mathcal{J}) \quad \text{(X.37)}\\
=&\mathbb{P}(\langle \underline{\mathbf{x}}, \underline{z}\rangle + \langle \underline{\mathbf{x}}, \underline{z_e}\rangle \notin nP(1\pm f_4(\varepsilon,\eta,\delta_{\mathcal{Z}}))|\mathcal{E}^c\cap\mathcal{J})\\
\le&\mathbb{P}(\langle \underline{\mathbf{x}}, \underline{z}\rangle \notin nP(1\pm f_4(\varepsilon,\eta,\delta_{\mathcal{Z}}))\mp\|\underline{\mathbf{x}}\|_2\|\underline{z_e}\|_2|\mathcal{E}^c\cap\mathcal{J}) \quad \text{(X.38)}\\
\le&\mathbb{P}(\langle \underline{\mathbf{x}}, \underline{z}\rangle \notin nP(1\pm f_4(\varepsilon,\eta,\delta_{\mathcal{Z}}))\mp n\sqrt{P\delta_{\mathcal{Z}}}|\mathcal{E}^c\cap\mathcal{J})\\
=&\mathbb{P}(\langle \underline{\mathbf{x}}, \underline{\mathbf{x}}+\underline{\mathbf{s}_z}\rangle \notin nP(1\pm f_4(\varepsilon,\eta,\delta_{\mathcal{Z}})\mp\sqrt{P\delta_{\mathcal{Z}}})|\mathcal{E}^c\cap\mathcal{J})\\
\le&\mathbb{P}(\langle \underline{\mathbf{x}}, \underline{\mathbf{s}_z}\rangle + \|\underline{\mathbf{x}}\|_2^2 \notin nP(1\pm f_4'(\varepsilon,\eta,\delta_{\mathcal{Z}}))|\mathcal{E}^c\cap\mathcal{J})\\
=&\mathbb{P}(|\langle \underline{\mathbf{x}}, \underline{\mathbf{s}_z}\rangle| > nPf_4'(\varepsilon,\eta,\delta_{\mathcal{Z}})|\mathcal{E}^c\cap\mathcal{J})\\
=&\mathbb{P}\left(|\langle \underline{e_1}, \underline{\mathbf{e}_{s_z}}\rangle| > \frac{nPf_4'(\varepsilon,\eta,\delta_{\mathcal{Z}})}{\sqrt{nP}\|\underline{\mathbf{s}_z}\|_2}\middle|\mathcal{E}^c\cap\mathcal{J}\right)\\
\le&\mathbb{P}\left(|\langle \underline{e_1}, \underline{\mathbf{e}_{s_z}}\rangle| > \frac{nPf_4'(\varepsilon,\eta,\delta_{\mathcal{Z}})}{\sqrt{nP}\sqrt{n\sigma^2(1+\varepsilon)}}\middle|\mathcal{E}^c\cap\mathcal{J}\right) \quad \text{(X.39)}\\
\le&2^{-\frac{Pf_4'(\varepsilon,\eta,\delta_{\mathcal{Z}})^2}{2\sigma^2(1+\varepsilon)}(n-1)} \quad \text{(X.40)}\\
=:&2^{-nf_5(\varepsilon,\eta,\delta_{\mathcal{Z}})}, \quad \text{(X.41)}
\end{aligned}$$

where in the above chain of (in)equalities, we use the following facts.

1) In (X.37), we write $\underline{z_Q}=\underline{z}+\underline{z_e}$, where $\underline{z_e}$ denotes the decomposition error which has norm at most $\sqrt{n\delta_{\mathcal{Z}}}$ by the choice of the covering $\mathcal{Z}$.
2) Inequality (X.38) follows from Cauchy–Schwarz inequality $-\|\underline{\mathbf{x}}\|_2\|\underline{z_e}\|_2\le\langle\underline{\mathbf{x}},\underline{z_e}\rangle\le\|\underline{\mathbf{x}}\|_2\|\underline{z_e}\|_2$.
3) In Inequality (X.39), $\|\underline{\mathbf{s}_z}\|_2\in\sqrt{n\sigma^2(1\pm\varepsilon)}$ since we condition on $\mathcal{E}^c\cap\mathcal{J}$.
4) Inequality (X.40) is a straightforward application of Lemma 8.

It follows that with probability at least $1-2^{-f_5(\varepsilon,\eta,\delta_{\mathcal{Z}})n}$ the scale of the perpendicular component is also concentrated around its expected value,

$$\begin{aligned}
\boldsymbol{\beta}_x\in&P-\frac{P^2}{P+\sigma^2}(1\pm\eta)\\
=&\frac{P\sigma^2}{P+\sigma^2}(1\mp\eta/\sigma^2)=:\frac{P\sigma^2}{P+\sigma^2}(1\mp\eta_1),
\end{aligned}$$

All the above concentration is over the randomness in the channel between Alice and James and the codebook generation.

We are now ready to compute the expected value of the radius $\sqrt{n\mathbf{r}}$ of the list-decoding region and concentrate it. To this end, let us first do some rough calculations to see what we should aim for. Loosely speaking, we expect the following quantity

$$\begin{aligned}
\langle \underline{\mathbf{x}}, -\underline{s_Q}\rangle =&\langle\sqrt{n\boldsymbol{\alpha}_x}\underline{e}^{\|}+\sqrt{n\boldsymbol{\beta}_x}\underline{\mathbf{e}}^{\perp},\sqrt{n\alpha_s}\underline{e}^{\|}-\sqrt{n\beta_s}\underline{e_s}^{\perp}\rangle\\
=&\langle\sqrt{n\boldsymbol{\alpha}_x}\underline{e}^{\|},\sqrt{n\alpha_s}\underline{e}^{\|}\rangle-\langle\sqrt{n\boldsymbol{\beta}_x}\underline{\mathbf{e}}^{\perp},\sqrt{n\beta_s}\underline{e_s}^{\perp}\rangle \quad \text{(X.42)}\\
=&n\sqrt{\boldsymbol{\alpha}_x\alpha_s}-n\sqrt{\boldsymbol{\beta}_x\beta_s}\langle\underline{\mathbf{e}}^{\perp},\underline{e_s}^{\perp}\rangle
\end{aligned}$$

to satisfy

$$\mathbb{E}(\langle\underline{\mathbf{x}},-\underline{s_Q}\rangle)\approx nP\sqrt{\frac{\alpha_s}{P+\sigma^2}}. \quad \text{(X.43)}$$

This is because:

1) In Equation (X.42), by decomposition, crossing terms vanish.
2) When there is no quantization error in $\underline{z}$, it holds that $\mathbb{E}(\langle\underline{\mathbf{e}}^{\perp},\underline{e_s}^{\perp}\rangle)=0$ as $\underline{\mathbf{e}}^{\perp}$ is isotropically distributed on a ring $\mathcal{S}^{n-2}(0,1)$. By previous heuristic calculations (see Equation (X.33)), we expect $\boldsymbol{\alpha}_x$ to be roughly $\frac{P^2}{P+\sigma^2}$.

These indicate that Equation (X.43) should be reasonably correct modulo some small error terms, bounded below in Lemma 29.

Remark 12: Notice that, for a fixed β_x, we actually know exactly the distribution of $\langle\sqrt{n\beta_x}\underline{\mathbf{e}}^{\perp},\sqrt{n\beta_s}\underline{e_s}^{\perp}\rangle$. Indeed, in $\mathbb{R}^n$, given a fixed unit vector $\underline{e}$ and a random unit vector $\underline{\mathbf{e}}$ isotropically distributed on the unit sphere $\mathcal{S}^{n-1}(0,1)$, $\frac{|\langle\underline{\mathbf{e}},\underline{e}\rangle|^2+1}{2}$ follows Beta distribution $\mathrm{Beta}\left(\frac{n-1}{2},\frac{n-1}{2}\right)$. Plugging this into our calculation, indeed, we get that $\langle\sqrt{n\beta_x}\underline{\mathbf{e}}^{\perp},\sqrt{n\beta_s}\underline{e_s}^{\perp}\rangle$ has mean 0. However, this result does not bring us any analytic advantage. Rather, in the analysis, when caring about concentration, we use Lemma 8 to approximate the tail of this distribution.

Although given $\underline{z}$, $\underline{\mathbf{e}}^{\perp}$ is perfectly isotropic on the unit ring $\mathcal{S}^{n-2}(0,1)$, it is not exactly the case when we are working with $\underline{z_Q}$. It is, however, probably approximately correct (PAC). Specifically, this issue can be fixed by the following lemma. As shown in Figure 24b, recall that we denote by $\underline{\mathbf{x}}=\underline{\mathbf{x}_Q}^{\|}+\underline{\mathbf{x}_Q}^{\perp}$ the decomposition with respect to $\underline{z_Q}$. Also, denote by $\underline{\mathbf{x}}=\underline{\mathbf{x}}^{\|}+\underline{\mathbf{x}}^{\perp}$ the decomposition with respect to $\underline{z}$. Define the error vectors as $\underline{\mathbf{x}_e}:=\underline{\mathbf{x}_Q}^{\perp}-\underline{\mathbf{x}}^{\perp}$ and $\underline{s_e}:=\underline{s_Q}^{\perp}-\underline{\mathbf{s}}^{\perp}$.

Lemma 29: Fix $\zeta>0$. Also fix $\underline{z}$ and $\underline{s}$, and thereby $\underline{z_Q}$ and $\underline{s_Q}$. Let $\zeta':=\zeta-\sqrt{P\delta_{\mathcal{S}}}-\sqrt{\frac{NP\delta_{\mathcal{Z}}}{(P+\sigma^2)(1-\varepsilon)}}-\sqrt{\frac{P\delta_{\mathcal{S}}\delta_{\mathcal{Z}}}{(P+\sigma^2)(1-\varepsilon)}}$. Then

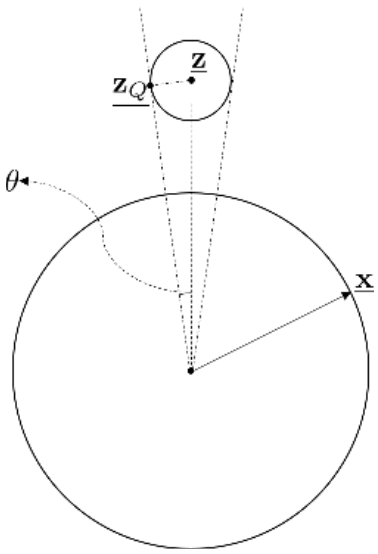
$$\mathbb{P}(|\langle\underline{\mathbf{x}_Q}^{\perp},\underline{s_Q}^{\perp}\rangle|\ge n\zeta|\mathcal{E}^c\cap\mathcal{J})\le 2^{-\frac{(n-2)\zeta'^2}{2NP}}.$$

Proof: We write

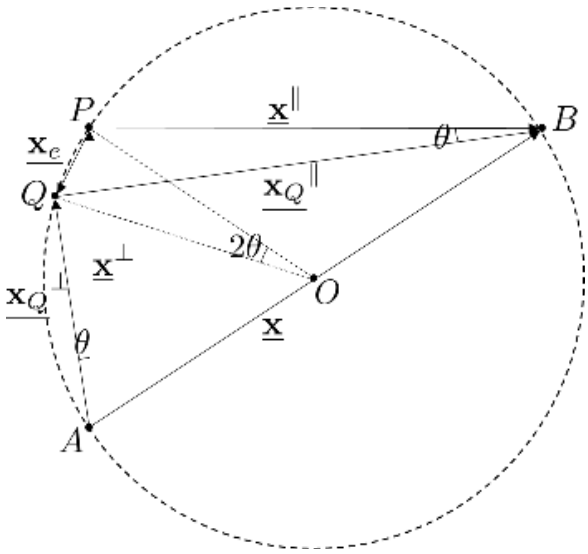
$$\begin{aligned}
&\mathbb{P}(|\langle\underline{\mathbf{x}_Q}^{\perp},\underline{s_Q}^{\perp}\rangle|\ge n\zeta|\mathcal{E}^c\cap\mathcal{J})\\
=&\mathbb{P}(|\langle\underline{\mathbf{x}}^{\perp}+\underline{\mathbf{x}_e},\underline{\mathbf{s}}^{\perp}+\underline{s_e}\rangle|\ge n\zeta|\mathcal{E}^c\cap\mathcal{J})\\
\le&\mathbb{P}(|\langle\underline{\mathbf{x}}^{\perp},\underline{\mathbf{s}}^{\perp}\rangle|+|\langle\underline{\mathbf{x}}^{\perp},\underline{s_e}\rangle|+|\langle\underline{\mathbf{x}_e},\underline{\mathbf{s}}^{\perp}\rangle|+|\langle\underline{\mathbf{x}_e},\underline{s_e}\rangle|\ge n\zeta|\mathcal{E}^c\cap\mathcal{J})\\
\le&\mathbb{P}(|\langle\underline{\mathbf{x}}^{\perp},\underline{\mathbf{s}}^{\perp}\rangle|+\|\underline{\mathbf{x}}^{\perp}\|_2\|\underline{s_e}\|_2+\|\underline{\mathbf{x}_e}\|_2\|\underline{\mathbf{s}}^{\perp}\|_2+\|\underline{\mathbf{x}_e}\|_2\|\underline{s_e}\|_2\ge n\zeta|\mathcal{E}^c\cap\mathcal{J})\\
\le&\mathbb{P}(|\langle\underline{\mathbf{x}}^{\perp},\underline{\mathbf{s}}^{\perp}\rangle|+n\sqrt{P\delta_{\mathcal{S}}}+\sqrt{nN}\|\underline{\mathbf{x}_e}\|_2+\sqrt{n\delta_{\mathcal{S}}}\|\underline{\mathbf{x}_e}\|_2\ge n\zeta|\mathcal{E}^c\cap\mathcal{J}).
\end{aligned}$$

It then suffices to upper bound $\|\underline{\mathbf{x}_e}\|_2$. Write $\underline{z_Q}=\underline{z}+\underline{z_e}$ where $\underline{z_e}$ denotes the quantization error for $\underline{z}$ with respect to the covering $\mathcal{Z}$. As $\underline{z_e}$ is at most $\sqrt{n\delta_{\mathcal{Z}}}$, for maximum θ shown in Figure 24a, we have $\sin(\theta)=\frac{\|\underline{z_e}\|_2}{\|\underline{z}\|_2}\le\frac{\sqrt{n\delta_{\mathcal{Z}}}}{\|\underline{z}\|_2}$. Notice that $\angle QAP=\angle QBP=\angle QOP/2=\theta$. Consider the triangle ΔQOP. We have $\frac{|\overline{QP}|/2}{|\overline{QO}|}=\frac{\|\underline{\mathbf{x}_e}\|_2/2}{\|\underline{\mathbf{x}}\|_2/2}=\sin(\theta)\le\frac{\sqrt{n\delta_{\mathcal{Z}}}}{\|\underline{z}\|_2}$, i.e., $\|\underline{\mathbf{x}_e}\|_2\le\frac{\sqrt{nP}}{\sqrt{n(P+\sigma^2)(1-\varepsilon)}}\sqrt{n\delta_{\mathcal{Z}}}=\sqrt{\frac{nP\delta_{\mathcal{Z}}}{(P+\sigma^2)(1-\varepsilon)}}$. Now Lemma 8 can be applied.

$$\mathbb{P}(|\langle\underline{\mathbf{x}_Q}^{\perp},\underline{s_Q}^{\perp}\rangle|\ge n\zeta)$$



(a) For the ease of union bounds over James's observation, any given $\underline{\mathbf{z}}$ is quantized to $\underline{\mathbf{z}}_Q$ using the $\sqrt{n\delta_{\mathcal{Z}}}$-net $\mathcal{Z}$. The quantization error introduced by covering will have an impact on the rest of the analysis. We measure this using the angular distance between $\underline{\mathbf{z}}$ and $\underline{\mathbf{z}}_Q$. The maximum angle θ is given by the geometry shown in the above figure. θ is small given a typical $\underline{\mathbf{z}}$ due to the covering property of $\mathcal{Z}$. We will keep track how errors θ are propagated.



(b) Ideally we would like to decompose everything into directions along and perpendicular to $\underline{\mathbf{z}}$. However, what we really work with is the quantized version $\underline{\mathbf{z}}_Q$ of $\underline{\mathbf{z}}$. This introduces approximation errors that we must keep track of when computing the parallel and perpendicular components of the vectors we are interested in. A comparison of the geometry of decomposition with respect to $\underline{\mathbf{z}}_Q$ and $\underline{\mathbf{z}}$ is shown in the above figure. In particular, the error in the perpendicular component $\underline{\mathbf{x}}^{\perp}$ of $\underline{\mathbf{x}}$ stemming from the quantization of $\underline{\mathbf{z}}$ can be bounded in terms of the angular quantization error θ.

Fig. 24. The geometry of the propagation of the quantization error of $\underline{\mathbf{z}}$.

$$\begin{aligned}
&\leq \mathbb{P}\left(|\langle \underline{\mathbf{x}}^{\perp}, \underline{s}^{\perp}\rangle| \geq n\zeta - n\sqrt{P\delta_{\mathcal{S}}} - n\sqrt{\frac{NP\delta_{\mathcal{Z}}}{(P+\sigma^2)(1-\varepsilon)}}\right.\\
&\qquad \left. -n\sqrt{\frac{P\delta_{\mathcal{S}}\delta_{\mathcal{Z}}}{(P+\sigma^2)(1-\varepsilon)}}\right)\\
&\leq 2^{-\frac{(n-2)n^2\zeta'^2}{2\|\underline{\mathbf{x}}^{\perp}\|_2^2\|\underline{s}^{\perp}\|_2^2}}\\
&\leq 2^{-\frac{(n-2)\zeta'^2}{2NP}}.
\end{aligned}$$

This completes the proof for Lemma 29. □

Now we give a concentrated version of Equation (X.43).

$$\begin{aligned}
&\mathbb{P}\left(\langle -\underline{\mathbf{x}}, \underline{s_Q}\rangle \notin nP\sqrt{\frac{\alpha_s}{P+\sigma^2}(1\pm\varepsilon)}\,\middle|\,\mathcal{E}^c\cap\mathcal{J}\right) \qquad \text{(X.44)}\\
=&\mathbb{P}\left(\langle\sqrt{n\boldsymbol{\alpha}_x}\underline{\mathbf{e}}^{\|}, \sqrt{n\alpha_s}\underline{\mathbf{e}}^{\|}\rangle - \langle\sqrt{n\boldsymbol{\beta}_x}\underline{\mathbf{e}}^{\perp}, \sqrt{n\beta_s}\underline{e_s}^{\perp}\rangle\right.\\
&\qquad \left.\notin nP\sqrt{\frac{\alpha_s}{P+\sigma^2}(1\pm\varepsilon)}\,\middle|\,\mathcal{E}^c\cap\mathcal{J}\right)\\
=&\mathbb{P}\left(\langle\sqrt{n\boldsymbol{\alpha}_x}\underline{\mathbf{e}}^{\|}, \sqrt{n\alpha_s}\underline{\mathbf{e}}^{\|}\rangle - \langle\sqrt{n\boldsymbol{\beta}_x}\underline{\mathbf{e}}^{\perp}, \sqrt{n\beta_s}\underline{e_s}^{\perp}\rangle\right.\\
&\qquad \left.\notin nP\sqrt{\frac{\alpha_s}{P+\sigma^2}(1\pm\varepsilon)},\ \boldsymbol{\alpha}_x \in \frac{P^2}{P+\sigma^2}(1\pm\eta)\,\middle|\,\mathcal{E}^c\cap\mathcal{J}\right)\\
&+\mathbb{P}\left(\langle\sqrt{n\boldsymbol{\alpha}_x}\underline{\mathbf{e}}^{\|}, \sqrt{n\alpha_s}\underline{\mathbf{e}}^{\|}\rangle - \langle\sqrt{n\boldsymbol{\beta}_x}\underline{\mathbf{e}}^{\perp}, \sqrt{n\beta_s}\underline{e_s}^{\perp}\rangle\right.\\
&\qquad \left.\notin nP\sqrt{\frac{\alpha_s}{P+\sigma^2}(1\pm\varepsilon)},\ \boldsymbol{\alpha}_x \notin \frac{P^2}{P+\sigma^2}(1\pm\eta)\,\middle|\,\mathcal{E}^c\cap\mathcal{J}\right)\\
\leq&\mathbb{P}\left(\langle\sqrt{n\boldsymbol{\beta}_x}\underline{\mathbf{e}}^{\perp}, \sqrt{n\beta_s}\underline{e_s}^{\perp}\rangle \notin nP\sqrt{\frac{\alpha_s}{P+\sigma^2}(1\pm\eta)}\right.\\
&\qquad \left. -nP\sqrt{\frac{\alpha_s}{P+\sigma^2}(1\pm\varepsilon)}\,\middle|\,\mathcal{E}^c\cap\mathcal{J}\right) \qquad \text{(X.45)}\\
&+\mathbb{P}\left(\boldsymbol{\alpha}_x \notin \frac{P^2}{P+\sigma^2}(1\pm\eta)\,\middle|\,\mathcal{E}^c\cap\mathcal{J}\right). \qquad \text{(X.46)}
\end{aligned}$$

By (X.36), the second term (X.46) is at most $2^{-f_5(\varepsilon,\eta,\delta_{\mathcal{Z}})n}$. The first term (X.45) can be bounded as follows.

$$\begin{aligned}
&\mathbb{P}\left(|\langle\sqrt{n\boldsymbol{\beta}_x}\underline{\mathbf{e}}^{\perp}, \sqrt{n\beta_s}\underline{e_s}^{\perp}\rangle| > nP\sqrt{\frac{\alpha_s}{P+\sigma^2}}\right.\\
&\qquad \left.(\sqrt{1+\eta}+\sqrt{1+\varepsilon})\,\middle|\,\mathcal{E}^c\cap\mathcal{J}\right)\\
\leq&\mathbb{P}(|\langle\sqrt{n\boldsymbol{\beta}_x}\underline{\mathbf{e}}^{\perp}, \sqrt{n\beta_s}\underline{e_s}^{\perp}\rangle| > nf_8(\varepsilon,\tau)|\mathcal{E}^c\cap\mathcal{J})\\
=&\mathbb{P}(|\langle\underline{\mathbf{x}_Q}^{\perp}, \underline{s_Q}^{\perp}\rangle| \geq nf_8(\varepsilon,\eta)|\mathcal{E}^c\cap\mathcal{J})\\
\leq&2^{-\frac{(n-2)f_8'(\varepsilon,\eta,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})^2}{2NP}}, \qquad \text{(X.47)}
\end{aligned}$$

where inequality (X.47) follows from Lemma 29 and $f_8'(\varepsilon,\eta,\delta_{\mathcal{S}},\delta_{\mathcal{Z}}) := f_8(\varepsilon,\eta) - \sqrt{P\delta_{\mathcal{S}}} - \sqrt{\frac{NP\delta_{\mathcal{Z}}}{(P+\sigma^2)(1-\varepsilon)}} - \sqrt{\frac{P\delta_{\mathcal{S}}\delta_{\mathcal{Z}}}{(P+\sigma^2)(1-\varepsilon)}}$. Thus combining bounds (X.36) and (X.47) on terms (X.46) and (X.45) respectively, the probability (X.44) is well bounded by

$$2^{-\frac{(n-2)f_8'(\varepsilon,\eta,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})}{2NP}} + 2^{-f_5(\varepsilon,\eta,\delta_{\mathcal{Z}})n} =: 2^{-f_9(\varepsilon,\eta,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})n}.$$

The above results allow us to compute the typical length of $\underline{\mathbf{y}}_Q$ under the translation of the prescribed $\underline{s}_Q$. We can do so by writing the "angular correlation" between $\underline{\mathbf{x}}$ and $\underline{s}_Q$ in analytic and geometric ways separately. Specifically, it follows

from the above concentration result that with probability at least $1-2^{-f_9(\varepsilon,\eta,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})n}$,

$$\cos(\angle_{-\underline{\mathbf{x}},\underline{s_Q}}) = \frac{\langle -\underline{\mathbf{x}}, \underline{s_Q}\rangle}{\|\underline{\mathbf{x}}\|_2\|\underline{s_Q}\|_2} \in \frac{nP\sqrt{\frac{\alpha_s}{P+\sigma^2}(1\pm\varepsilon)}}{\sqrt{nP}\sqrt{nN}} = \sqrt{\frac{P\alpha_s}{N(P+\sigma^2)}(1\pm\varepsilon)}.$$

On the other hand, by law of cosines, we have (with probability one)

$$\cos(\angle_{-\underline{\mathbf{x}},\underline{s_Q}}) = \frac{\|\underline{\mathbf{x}}\|_2^2+\|\underline{s_Q}\|_2^2-\|\underline{\mathbf{y}_Q}\|_2^2}{2\|\underline{\mathbf{x}}\|_2\|\underline{s_Q}\|_2} = \frac{nP+nN-\|\underline{\mathbf{y}_Q}\|_2^2}{2\sqrt{nP}\sqrt{nN}}.$$

It immediately follows that

$$\|\underline{\mathbf{y}_Q}\|_2^2 = n(P+N) - 2nP\sqrt{\frac{\alpha_s}{P+\sigma^2}(1\pm\varepsilon)},$$

with probability at least $1-2^{-f_9(\varepsilon,\eta,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})n}$.

Denote by $\sqrt{n\mathbf{r}}$ the radius of the intersection $\mathcal{B}^n(\underline{\mathbf{y}_Q}, \sqrt{nN}+\sqrt{n\delta_{\mathcal{S}}}) \cap \mathcal{S}^{n-1}(0,\sqrt{nP})$. Staring at the triangle $\Delta OAO'$ shown in Figure 25, we know that, with probability at least $1-2^{-f_9(\varepsilon,\eta,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})n}$,

$$\left(\frac{\sqrt{n\mathbf{r}}}{\sqrt{nP}}\right)^2 = (\sin(\angle AOO'))^2 = 1-(\cos(\angle AOO'))^2,$$

i.e.,

$$\frac{\mathbf{r}}{P} = 1-\left(\frac{nP+\|\underline{\mathbf{y}_Q}\|_2^2-(\sqrt{nN}+\sqrt{n\delta_{\mathcal{S}}})^2}{2\sqrt{nP}\|\underline{\mathbf{y}_Q}\|_2}\right)^2$$

$$\Longrightarrow \mathbf{r} = \frac{N-P\frac{\alpha_s}{P+\sigma^2}(1\pm\varepsilon)+f_{10}(\varepsilon,\delta_{\mathcal{S}})}{P+N-2P\sqrt{\frac{\alpha_s}{P+\sigma^2}(1\pm\varepsilon)}}P,$$

where $f_{10}(\varepsilon,\delta_{\mathcal{S}}) := -\frac{(\delta_{\mathcal{S}}+2\sqrt{N\delta_{\mathcal{S}}})^2}{4P} + \left(1-\sqrt{\frac{\alpha_s}{P+\sigma^2}(1-\varepsilon)}\right)(\delta_{\mathcal{S}}+2\sqrt{N\delta_{\mathcal{S}}})$. That is to say, we have

$$\mathbb{P}\left(\mathbf{r}\notin \frac{N-P\frac{\alpha_s}{P+\sigma^2}}{P+N-2P\sqrt{\frac{\alpha_s}{P+\sigma^2}}}P(1\pm f_{11}(\varepsilon,\delta_{\mathcal{S}}))\middle|\mathcal{E}^c\cap\mathcal{J}\right) \le 2^{-f_9(\varepsilon,\eta,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})n}. \quad \text{(X.48)}$$

From James's perspective, he aims to maximize the above quantity to confuse Bob to the largest extent. He will take the jamming strategy corresponding to the optimal solution of the following optimization problem.[19] The average (over the randomness in $\underline{\mathbf{z}}$) worst-case (over α_s) value of $\mathbf{r}$ obtained in this manner is what we call r_{opt}.

$$\begin{array}{ll}\min_{\alpha_s} & \frac{P+N-2P\sqrt{\frac{\alpha_s}{P+\sigma^2}}}{N-P\frac{\alpha_s}{P+\sigma^2}} \\ \text{subject to} & 0\le\alpha_s\le N \\ & \frac{\sigma^2}{P}\ge\frac{1}{1-N/P}-1 \\ & P,N,\sigma^2\ge 0\end{array} \quad \text{(X.49)}$$

[19]Although James could possibly choose other strategies, the proof still goes through by taking a union bound over all possible type classes of $\underline{\mathbf{s}}$ (corresponding to all possible attack strategies). Since there are only polynomially many of them, and none of them can be as bad as the worst-case strategy, the arguments still hold.

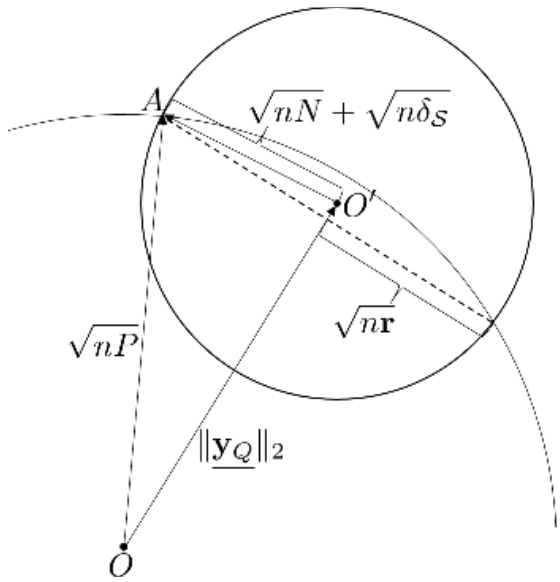


Fig. 25. Fix a legitimate attack vector $\underline{s_Q}$. We compute the expected radius of the list-decoding region, which is the cap $\mathcal{C}ap^{n-1}(\underline{\mathbf{y}_Q}, \sqrt{n\mathbf{r}}, \sqrt{nP}) = \mathcal{C}ap^{n-1}(\underline{\mathbf{x}}+\underline{s_Q}, \sqrt{n\mathbf{r}}, \sqrt{nP})$ shown in above figure, over codewords in the strip.

Remark 13: Notice that the objective function of the above optimization problem (X.49) is essentially the same as that of the optimization (VII.3) in the scale-and-babble converse argument under the map $\alpha\mapsto\sqrt{\frac{\alpha_s}{P+\sigma^2}}$.

Solving this optimization problem[20] and combining it with Lemma 27 which will be proved in Section X-J, we get that, if the code operates at a rate

$$R\le\begin{cases}R_{\text{LD}}, & \frac{1}{1-N/P}-1\le\frac{\sigma^2}{P}\le\frac{1}{N/P}-1\\ R_{\text{LD,myop}}, & \frac{\sigma^2}{P}\ge\max\left\{\frac{1}{1-N/P}-1,\frac{1}{N/P}-1\right\},\end{cases} \quad \text{(X.50)}$$

where

$$R_{\text{LD}} := \frac{1}{2}\log\frac{P}{N},$$

$$R_{\text{LD,myop}} := \frac{1}{2}\log\left(\frac{(P+\sigma^2)(P+N)-2P\sqrt{N(P+\sigma^2)}}{N\sigma^2}\right),$$

[20]The solution to the optimization problem (X.49) is obtained by elementary algebraic manipulation which is omitted. We attach the following Mathematica codes for verification (where `\[Alpha]` denotes α_S and `M` denotes N since the symbol `N` is reserved by Mathematica).

```
Minimize[{(P+M-2 P Sqrt[\[Alpha]/(P+\[Sigma]^2)])/
(M-(P \[Alpha])/(P+\[Sigma]^2)),0<=\[Alpha]<=M,
P/M>=1+P/\[Sigma]^2,P>=0,M>=0,\[Sigma]>=0},\[Alpha]]
```

The above codes are displayed as

$$\text{Minimize}\left[\left\{\frac{P+M-2P\sqrt{\frac{\alpha}{P+\sigma^2}}}{M-\frac{P\alpha}{P+\sigma^2}}, 0\le\alpha\le M, \frac{P}{M}\ge 1+\frac{P}{\sigma^2}, P\ge 0, M\ge 0, \sigma\ge 0\right\}, \alpha\right]$$

in the graphical user interface of Mathematica.

then no matter towards which direction James is going to push the transmitted codeword, a vast majority (an exponentially close to one fraction) of codewords in the strip have small list-sizes (at most a low-degree polynomial in n).

In conclusion,

$$\mathbb{P}\left(\mathbf{r} > r_{\text{opt}}(1+f_{11}(\varepsilon,\delta_{\mathcal{S}})|\mathcal{E}^c\cap\mathcal{J}\right) \le 2^{-f_9(\varepsilon,\eta,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})n}.$$

where r_{opt} is the average list-decoding radius obtained by James choosing the worst-case attack vector minimizing the rate corresponding to the optimization problem (X.49). One can choose $\eta,\delta_{\mathcal{S}}$ and $\delta_{\mathcal{Z}}$ so that $f_9(\varepsilon,\eta,\delta_{\mathcal{S}},\delta_{\mathcal{Z}})=\frac{3}{2}\varepsilon$. This completes the proof of Lemma 25.

J. Proof of Lemma 27

In this section, we will show that myopic list-decoding succeeds with high probability conditioned on everything behaves typically (which is true as we have analyzed in previous sections).

Let r_{opt} denote the optimal solution of the above optimization. In what follows, we will prove that the probability that there are too many (more than $L=3n^2$) codewords in the list-decoding region is super-exponentially small. According to where the codewords in the list-decoding region come from, the list-decoding error can be divided into two types. If the confusing codewords come from the OGS, then they are quasi-uniformly distributed on the strip that the OGS belongs to, given James observation and extra information revealed to him. Otherwise, if the confusing codewords are outside the OGS, then they are uniformly distributed by the codebook generation.

1) Confusing codewords in the list-decoding region come from OGS. We further subdivide these confusing codewords into two types: those which have a typical $\mathbf{r}$ and those which do not.
   - Confusing codeword has atypical $\mathbf{r}$: Conditioned on $\mathcal{E}^c_{\underline{s_Q}}$, there are at most n^2 codewords with atypical $\mathbf{r}$. The distribution of these codewords is hard to obtain, and we will pessimistically assume that all these codewords are included in the list.
   - Confusing codeword has typical $\mathbf{r}$: The codewords with a typical $\mathbf{r}$ are all independent but no longer uniformly distributed over the strip given $\mathcal{E}^c_{\underline{s_Q}}$. However, the distribution is almost uniform. For any set $\mathcal{A}\subset\mathbb{R}^n$, we have

$$\begin{aligned}&\mathbb{P}(\underline{\mathbf{x}}(m,\mathbf{k})\in\mathcal{A}|\mathcal{E}^c_{\underline{s_Q}}\cap\mathcal{E}^c\cap\mathcal{J})\\ \le&\frac{\mathbb{P}(\underline{\mathbf{x}}(m,\mathbf{k})\in\mathcal{A}|\mathcal{E}^c\cap\mathcal{J})}{\mathbb{P}(\mathcal{E}^c_{\underline{s_Q}}|\mathcal{E}^c\cap\mathcal{J})}\\ =&\mathbb{P}(\underline{\mathbf{x}}(m,\mathbf{k})\in\mathcal{A}|\mathcal{E}^c\cap\mathcal{J})(1+o(1)).\end{aligned}$$

     Therefore, this conditioning does not significantly affect our calculations. Conditioned on bad events aforementioned not happening the average probability that a codeword falls into the list-decoding region is

$$\begin{aligned}&\mathbb{P}(m'\in\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})\cap\mathcal{O}rcl^{(j)}(\underline{z_Q},i)|\widetilde{\mathcal{E}}^c\cap\mathcal{J})\\ \le&\frac{\text{Area}(\mathcal{C}ap^{n-1}(\cdot,\sqrt{nr_{\text{opt}}(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))},\sqrt{nP}))}{\text{Area}(\mathcal{S}tr^{n-1}(\underline{z_Q},i))}\\ &\cdot\Delta(\tau)(1+o(1))\\ =&\sqrt{\frac{r_{\text{opt}}(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))}{r_{\text{str}}}}\\ &\cdot 2^{n\left(-\frac{1}{2}\log\left(\frac{P}{r_{\text{opt}}}\right)+\frac{1}{2}\log\left(\frac{P}{r_{\text{str}}}\right)+f_{11}(\varepsilon,\delta_{\mathcal{S}})\right)}\\ &\cdot\Delta(\tau)(1+o(1))\\ \le&\sqrt{\frac{r_{\text{opt}}(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))}{\frac{P\sigma^2(1-\varepsilon)}{(P+\sigma^2)(1+\varepsilon)}}}\\ &\cdot 2^{n\left(-\frac{1}{2}\log\left(\frac{P}{r_{\text{opt}}}\right)+\frac{1}{2}\log\left(1+\frac{P}{\sigma^2}\right)+2\varepsilon+f_{11}(\varepsilon,\delta_{\mathcal{S}})\right)}\\ &\cdot\Delta(\tau)(1+o(1)),\end{aligned}$$

     which is exponentially small. Then by similar calculations to Lemma 9, we have

$$\begin{aligned}&\mathbb{P}(|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})\cap\mathcal{O}rcl^{(j)}(\underline{z_Q},i)|>2n^2|\widetilde{\mathcal{E}}^c\cap\mathcal{J})\\ &\qquad\le 2^{-\Omega(n^3)}.\end{aligned}$$

2) Confusing codewords in the list-decoding region do not belong to the OGS.

$$\begin{aligned}&\mathbb{E}\left(|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})\backslash\mathcal{O}rcl^{(j)}(\underline{z_Q},i)|\,\middle|\,\widetilde{\mathcal{E}}^c\cap\mathcal{J}\right)\\ \le&\frac{\text{Area}(\mathcal{C}ap^{n-1}(\cdot,\sqrt{nr_{\text{opt}}(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))},\sqrt{nP}))}{\text{Area}(\mathcal{S}^{n-1}(0,\sqrt{nP}))}2^{nR_{\text{code}}}\\ =&\sqrt{\frac{P}{r_{\text{opt}}(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))}}2^{n\left(R_{\text{code}}-\frac{1}{2}\log\left(\frac{P}{r_{\text{opt}}}\right)+f_{11}(\varepsilon,\delta_{\mathcal{S}})\right)},\end{aligned}$$

   which is exponentially small if R_{code} is below the threshold. Then immediately by Lemma 9 (and Remark 4 following it),

$$\mathbb{P}(|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})\backslash\mathcal{O}rcl^{(j)}(\underline{z_Q},i)|>n^2|\widetilde{\mathcal{E}}^c\cap\mathcal{J})\le 2^{-\Omega(n^3)}. \tag{X.51}$$

Taking into account two types of error in Equation (X.51) and Equation (X.51), respectively, we get

$$\begin{aligned}&\mathbb{P}(|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})|>L|\widetilde{\mathcal{E}}^c\cap\mathcal{J})\\ \le&\mathbb{P}(|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})\cap\mathcal{O}rcl^{(j)}(\underline{z_Q},i)|>2n^2|\widetilde{\mathcal{E}}^c\cap\mathcal{J})\\ &+\mathbb{P}(|\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})\backslash\mathcal{O}rcl^{(j)}(\underline{z_Q},i)|>n^2|\widetilde{\mathcal{E}}^c\cap\mathcal{J})\\ \le&2^{-\Omega(n^3)}.\end{aligned} \tag{X.52}$$

This completes the proof of Lemma 25.

In the following section, we will argue that Bob will enjoy a vanishing probability of error below the threshold given by optimization (X.49), which matches the converse in Section VII-A in the corresponding region.

XI. Achievability in the Sufficiently Myopic Regime – From Myopic List-Decoding to Unique Decoding

In this section, given the myopic list-decodability results proved in Sec. X, we provide the proof of the second half

(unique decodability) of the achievability part of Theorem 14 and hence finish the achievability proof.

We first sketch the roadmap to proving that Bob can uniquely decode m with high probability. From James's point of view, $\underline{\mathbf{x}}$ is quasi-uniformly distributed over the strip. Loosely speaking, we will say that a message m_1 confuses a message m_2 if Bob declares his estimate to be m_1 when the actual message is m_2. The probability of error is small if

- the total number of messages (call them *confusing codewords*) that can confuse *any* message in the OGS is small (say $\mathrm{poly}(n)$) with probability super-exponentially close to one; and
- any message can only confuse a small number (say $\mathrm{poly}(n)$) of messages in the OGS (call them *confused codewords*) with probability super-exponentially close to one.

The first statement follows from a *blob list-decoding* argument and second follows from a *reverse list-decoding* argument. Technically, as we have seen in the analysis of myopic list-decoding error, we have to analyze the decoding error for two cases – the case where the confusing codewords come from the OGS and the case where they are outside the OGS. Note that these two cases are distinguished. Details are elaborated in Section XI-A and Section XI-B.

A. Type I Error

For type I error, confusing codewords come from $[2^{nR}]\backslash\mathcal{O}rcl^{(j)}(\underline{z_Q},i)$. We will prove that there are at most polynomially many (out of exponentially many) codewords in the OGS which can be erroneously decoded due to the confusion with some message outside the OGS.

Recall the definition of $\mathcal{E}:=\mathcal{E}_{\mathrm{atyp}}\cup\mathcal{E}_{\mathrm{str}}\cup\mathcal{E}_{\mathrm{orcl}}$, which is the union of several error events. Conditioned on $\mathcal{E}$, we have that simultaneously $\underline{\mathbf{z}}$ behaves typically, the strip contains a large number of codewords, and the transmitted codewords does not fall into the last oracle-given set which can potentially have too small size.

Lemma 30:

$$\mathbb{P}\Big(\exists\underline{z_Q},\exists i,\exists j,\exists\underline{s_Q},\ |\{m\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i):\ \exists m'\in[2^{nR}]\backslash\mathcal{O}rcl^{(j)}(\underline{z_Q},i),\ \underline{\mathbf{x}}(m')\in\mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m),\underline{s_Q})\}|\geq n^4\Big|\ \mathcal{E}^c\Big)\leq 2^{-\Omega(n^3)}.$$

Proof: We prove the lemma using a two-step list-decoding argument. The idea behind this type of argument is along the lines of [25]. Fix $\underline{z_Q}$, i, j and $\underline{s_Q}$. Notice that $\{\underline{\mathbf{x}}(m)\in\mathcal{C}^{(\mathbf{k})}: m\in[2^{nR}]\backslash\mathcal{O}rcl^{(j)}(\underline{z_Q},i)\}$ are independently and uniformly distributed.

It is a folklore (Appendix D) in the literature that a spherical code $\mathcal{C}$ of rate $\frac{1}{2}\log\frac{P}{N}-\varepsilon$ are $(P,N,\widetilde{\mathcal{O}}(1/\varepsilon))$-list-decodable (with exponential concentration), thus are also $(P,N,\mathcal{O}(n^2))$-list-decodable (with super-exponential concentration by Lemma 9).

1) Myopic Blob List-Decoding: Let $\mathcal{X}(\underline{s_Q}):=\{\underline{\mathbf{x}}(m'): m'\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i),\ \text{and}\ \mathbf{r}(m',\underline{s_Q})<r_{\mathrm{opt}}(1+f_{11}(\varepsilon))\}$ be the set of all codewords in the OGS having typical list-decoding region. Since as shown in Lemma 25, there is only an exponentially small fraction of codewords in the OGS that do not fall into $\mathcal{X}(\underline{s_Q})$, it suffices to prove that a $1-o(1)$ fraction of the codewords in $\mathcal{X}(\underline{s_Q})$ can with high probability be decoded uniquely for every $\underline{s_Q}$. In fact, we will show that, out of exponentially many codewords in $\mathcal{X}(\underline{s_Q})$, there are only polynomially many codewords that will incur decoding errors. Define

$$\mathcal{B}lob=\bigcup_{\underline{\mathbf{x}}\in\mathcal{X}(\underline{s_Q})}\mathcal{B}^n(\underline{\mathbf{x}}+\underline{s_Q},\sqrt{nN}+\sqrt{n\delta_{\mathcal{S}}}).$$

Let us fix a realization $\mathcal{J}$ of $(\underline{z_Q},i,j,\underline{s_Q})$. If $\widetilde{\mathcal{E}}=\mathcal{E}\cup\mathcal{E}_{\underline{s_Q}}(\underline{z_Q},i,j)$, then the number of codewords in the blob is expected to be

$$\begin{aligned}
&\mathbb{E}\left(|\mathcal{B}lob\cap(\mathcal{C}^{(\mathbf{k})}\backslash\mathcal{O}rcl^{(j)}(\underline{z_Q},i))|\Big|\ \widetilde{\mathcal{E}}^c\cap\mathcal{J}\right)\\
\leq&\frac{\mathrm{Area}(\mathcal{C}ap^{n-1}(\cdot,\sqrt{nr_{\mathrm{opt}}(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))},\sqrt{nP}))2^{n\varepsilon}}{\mathrm{Area}(\mathcal{S}^{n-1}(0,\sqrt{nP}))}2^{nR}\\
\leq&\frac{\mathrm{Area}(\mathcal{S}^{n-1}(\cdot,\sqrt{nr_{\mathrm{opt}}(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))}))2^{n\varepsilon}}{\mathrm{Area}(\mathcal{S}^{n-1}(0,\sqrt{nP}))}2^{nR}\\
=&\frac{\sqrt{nr_{\mathrm{opt}}(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))}^{n-1}}{\sqrt{nP}^{n-1}}2^{n\varepsilon}2^{nR}\\
=&\sqrt{\frac{P}{r_{\mathrm{opt}}(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))}}2^{n\left(R-\frac{1}{2}\log\left(\frac{P}{r_{\mathrm{opt}}(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))}\right)+\varepsilon\right)}\\
\leq&\sqrt{\frac{P}{r_{\mathrm{opt}}(1+f_{11}(\varepsilon,\delta_{\mathcal{S}}))}}2^{n\left(R-\frac{1}{2}\log\left(\frac{P}{r_{\mathrm{opt}}}\right)+f_{11}(\varepsilon,\delta_{\mathcal{S}})+\varepsilon\right)},
\end{aligned}\tag{XI.1}$$

which is exponentially small. The last inequality (XI.1) follows since $\log(1+x)\leq x$ for $x\leq 1$. Then by Lemma9, the actual number of codewords exceeds n^2 with probability at most $2^{-\Omega(n^3)}$.

2) Reverse List-Decoding: Conditioned on $\mathcal{E}^c_{\underline{s_Q}}$ and $\mathcal{J}$, the codewords in $\mathcal{X}(\underline{s_Q})$ are independent but not uniformly distributed over the strip. However, the distribution is almost uniform and does not affect our calculations except for adding a $(1+o(1))$ term. More precisely, for any set $\mathcal{A}\subset\mathbb{R}^n$, we have

$$\begin{aligned}
\mathbb{P}(\underline{\mathbf{x}}(m,\mathbf{k})\in\mathcal{A}|\mathcal{E}^c_{\underline{s_Q}}\cap\mathcal{J})&\leq\frac{\mathbb{P}(\underline{\mathbf{x}}(m,\mathbf{k})\in\mathcal{A})}{\mathbb{P}(\mathcal{E}_{\underline{s_Q}})}\\
&=\mathbb{P}(\underline{\mathbf{x}}(m,\mathbf{k})\in\mathcal{A})(1+o(1)).
\end{aligned}$$

The expected number of codewords corresponding to messages in OGS translated by $\underline{s_Q}$ lying in the ball $\mathcal{B}^n\left(\underline{\mathbf{x}}(m'),\sqrt{nN}+\sqrt{n\delta_{\mathcal{S}}}\right)$ for any $m'\in[2^{nR}]\backslash\mathcal{O}rcl^{(j)}(\underline{z_Q},i)$ in (XI.2), as shown at the bottom of the next page, where in Eqn. (XI.2) we use $\log(1-x)\geq-2x$ for small enough $x>0$. The above quantity is exponentially small according to the sufficient myopia assumption. Thus the actual number is at most n^2 with probability at least $1-2^{-\Omega(n^3)}$.

3) Union Bound: By Section XI-A.1 and Section XI-A.2, for any $\underline{z_Q}$, i, j and $\underline{s_Q}$, there are at most n^4 messages from $\mathcal{O}rcl^{(j)}(\underline{z_Q},i)$ satisfying the condition in the lemma with probability at least $1-2^{-\Omega(n^3)}$. To see this, let $\mathbf{A}$

be a $2^{n\varepsilon} \times (2^{nR} - 2^{n\varepsilon})$ matrix defined as $\mathbf{A}(m, m') = \mathbb{1}_{\{\underline{\mathbf{x}}(m') \in \mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m), \underline{s_Q})\}}$ for any $m \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)$ and $m' \in [2^{nR}] \setminus \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)$. The quantity we would like to concentrate $|\{m \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i) : \exists m' \in [2^{nR}] \setminus \mathcal{O}rcl^{(j)}(\underline{z_Q}, i),\ \underline{\mathbf{x}}(m') \in \mathcal{L}^{(\mathbf{k})}(\underline{\mathbf{x}}(m), \underline{s_Q})\}|$ can be written as $\sum_{m \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \mathbb{1}_{\{\exists m' \in [2^{nR}] \setminus \mathcal{O}rcl^{(j)}(\underline{z_Q}, i),\ \mathbf{A}_{m,m'} = 1\}}$, i.e., the number of nonzero rows[21] of $\mathbf{A}$. We can bound it above as follows.

$$
\begin{aligned}
&\sum_{m \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \mathbb{1}_{\{\exists m' \in [2^{nR}] \setminus \mathcal{O}rcl^{(j)}(\underline{z_Q}, i),\ \mathbf{A}(m,m') = 1\}} \\
&\le \sum_{m \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \sum_{m' \in [2^{nR}] \setminus \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \mathbb{1}_{\{\mathbf{A}(m,m') = 1\}} \qquad \text{(XI.3)} \\
&= \sum_{m \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \sum_{m' \in [2^{nR}] \setminus \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \mathbb{1}_{\{\mathbf{A}(m,m') = 1\}} \\
&\qquad \mathbb{1}_{\{\exists m_0 \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i),\ \mathbf{A}(m_0,m') = 1\}} \qquad \text{(XI.4)} \\
&= \sum_{m' \in [2^{nR}] \setminus \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \Bigg[\mathbb{1}_{\{\exists m_0 \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i),\ \mathbf{A}(m_0,m') = 1\}} \\
&\qquad \left(\sum_{m \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \mathbb{1}_{\{\mathbf{A}(m,m') = 1\}} \right) \Bigg] \qquad \text{(XI.5)} \\
&= \left(\sum_{m' \in [2^{nR}] \setminus \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \mathbb{1}_{\{\exists m_0 \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i),\ \mathbf{A}(m_0,m') = 1\}} \right) \\
&\qquad \times \left(\max_{m' \in [2^{nR}] \setminus \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \sum_{m \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} \mathbb{1}_{\{\mathbf{A}(m,m') = 1\}} \right). \qquad \text{(XI.6)}
\end{aligned}
$$

Eqn. (XI.3) is by union bound. Eqn. (XI.4) holds since the event in the first indicator implies that in the second one. In Eqn. (XI.5), we rearrange the summations. In Eqn. (XI.6), we bound the inner summation (which depends on m') by the largest one among all $m' \in [2^{nR}] \setminus \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)$. In the above chain of (in)equalities, we are essentially bounding the number of nonzero rows of a matrix by the number of nonzero entries which is further bounded by the number of nonzero columns times the largest weight of a column. Note that the number of nonzero columns (the first term in Eqn. (XI.6)) is the number of *confusing* codewords (outside the OGS) that can confuse some codewords in the OGS, that is, the number of codewords in the blob. As shown in Section XI-A.1, this number is at most n^2 with probability $1 - 2^{-\Omega(n^3)}$. The maximum weight of a column is the largest number of *confused* codewords in the OGS that a codeword outside the OGS can confuse. This number, as shown in Section XI-A.2, is at most n^2 with probability $1 - 2^{-\Omega(n^3)}$. Therefore, by Eqn. (XI.6), the number of codewords in the OGS that might be confused by some codewords outside the OGS is at most n^4 with probability $1 - 2^{-\Omega(n^3)}$. Finally, a union bound over all assumptions we have made completes the proof. □

[21] We say that a row (resp. column) is nonzero if not all of its entries are zero.

B. Type II Error

For type II error, confusing codewords come from $\mathcal{O}rcl^{(j)}(\underline{z_Q}, i)$. We will prove that there are at most polynomially many codewords which are erroneously decoded due to confusion with another message in the same OGS. Once

$$
\begin{aligned}
&\mathbb{E}\left(|\mathcal{B}^n(\underline{\mathbf{x}}(m'), \sqrt{nN} + \sqrt{n\delta_{\mathcal{S}}}) \cap (\{\underline{\mathbf{x}}(m)\}_{m \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)} + \underline{s_Q})| \,\middle|\, \widetilde{\mathcal{E}}^c \cap \mathcal{J} \right) \\
&= \mathbb{E}\left(|\mathcal{B}^n(\underline{\mathbf{x}}(m') - \underline{s_Q}, \sqrt{nN} + \sqrt{n\delta_{\mathcal{S}}}) \cap \{\underline{\mathbf{x}}(m)\}_{m \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)}| \,\middle|\, \widetilde{\mathcal{E}}^c \cap \mathcal{J} \right) \\
&\le \frac{\mathrm{Area}(\mathcal{S}^{n-1}(\cdot, \sqrt{nN} + \sqrt{n\delta_{\mathcal{S}}}))}{\mathrm{Area}(\mathcal{S}tr^{n-1}(O'_-, O'_+, \sqrt{nr_-}, \sqrt{nr_+}))} 2^{n\varepsilon}(1 + o(1)) \\
&= \frac{\mathrm{Area}(\mathcal{S}^{n-1}(\cdot, \sqrt{nN} + \sqrt{n\delta_{\mathcal{S}}}))}{\mathrm{Area}(\mathcal{C}ap^{n-1}(O'_+, \sqrt{nr_+}, \sqrt{nP})) - \mathrm{Area}(\mathcal{C}ap^{n-1}(O'_-, \sqrt{nr_-}, \sqrt{nP}))} 2^{n\varepsilon}(1 + o(1)) \\
&\le \frac{\mathrm{Area}(\mathcal{S}^{n-1}(\cdot, \sqrt{nN} + \sqrt{n\delta_{\mathcal{S}}}))}{\mathrm{Vol}(\mathcal{B}^{n-1}(O'_+, \sqrt{nr_+})) - \mathrm{Area}(\mathcal{S}^{n-1}(O'_-, \sqrt{nr_-}))} 2^{n\varepsilon}(1 + o(1)) \\
&\le \sqrt{\frac{N}{r_{\mathrm{str}}}} 2^{n\left(-\frac{1}{2}\log\left(\frac{P}{N + \delta_{\mathcal{S}} + 2\sqrt{N\delta_{\mathcal{S}}}} \right) + \frac{1}{2}\log\left(\frac{P}{r_{\mathrm{str}}} \right) + \varepsilon \right)} \frac{(1 + o(1))}{2^{(n-1)\frac{1}{2}\log(1+\tau) - \Theta(\log n)} - 2^{(n-1)\frac{1}{2}\log(1-\tau)}} \\
&= \sqrt{\frac{N}{r_{\mathrm{str}}}} 2^{n\left(-\frac{1}{2}\log\frac{P}{N} - \frac{1}{2}\log\left(1 - \frac{\delta_{\mathcal{S}} + 2\sqrt{N\delta_{\mathcal{S}}}}{N + \delta_{\mathcal{S}} + 2\sqrt{N\delta_{\mathcal{S}}}} \right) + \frac{1}{2}\log\left(\frac{P}{r_{\mathrm{str}}} \right) + \varepsilon \right)} \frac{(1 + o(1))}{2^{(n-1)\frac{1}{2}\log(1+\tau) - \Theta(\log n)} - 2^{(n-1)\frac{1}{2}\log(1-\tau)}} \\
&\le \sqrt{\frac{N}{r_{\mathrm{str}}}} 2^{n\left(-\frac{1}{2}\log\frac{P}{N} + \frac{1}{2}\log\left(\frac{P}{r_{\mathrm{str}}} \right) + 2\frac{\delta_{\mathcal{S}} + 2\sqrt{N\delta_{\mathcal{S}}}}{N + \delta_{\mathcal{S}} + 2\sqrt{N\delta_{\mathcal{S}}}} + \varepsilon \right)} \frac{(1 + o(1))}{2^{(n-1)\frac{1}{2}\log(1+\tau) - \Theta(\log n)} - 2^{(n-1)\frac{1}{2}\log(1-\tau)}} \qquad \text{(XI.2)} \\
&\le \sqrt{\frac{N(P + \sigma^2)(1 + \varepsilon)}{P\sigma^2(1 - \varepsilon)}} 2^{n\left(-\frac{1}{2}\log\frac{P}{N} + \frac{1}{2}\log\left(1 + \frac{P}{\sigma^2} \right) + 2\frac{\delta_{\mathcal{S}} + 2\sqrt{N\delta_{\mathcal{S}}}}{N + \delta_{\mathcal{S}} + 2\sqrt{N\delta_{\mathcal{S}}}} + 3\varepsilon \right)} \frac{(1 + o(1))}{2^{(n-1)\frac{1}{2}\log(1+\tau) - \Theta(\log n)} - 2^{(n-1)\frac{1}{2}\log(1-\tau)}},
\end{aligned}
$$

again we will only analyze the probability of error only for codewords having typical list-decoding volume.

Lemma 31:

$$\mathbb{P}\Bigg(\exists \underline{z_Q}, \exists i, \exists j, \exists \underline{s_Q},\ |\{m \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i) : \exists m' \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i) \setminus \{m\}, \underline{\mathbf{x}}(m') \in \mathcal{L}^{(\mathbf{k})}(\underline{x}(m), \underline{s_Q})\}| \geq 2 \cdot 2^{n\varepsilon/2} \cdot n^4 \Big| \mathcal{E}^c \cap \mathcal{J}\Bigg) \leq 2^{-\Omega(n^3)}.$$

Proof: Notice that for different $m \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)$, the events $\{\exists m' \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i) \setminus \{m\},\ \underline{\mathbf{x}}(m') \in \mathcal{B}^n(\underline{\mathbf{x}}(m) + \underline{s_Q}, \sqrt{nN} + \sqrt{n\delta_{\mathcal{S}}})\}$ are not independent. This issue is resolved by arranging messages in OGS into a $2^{n\varepsilon/2} \times 2^{n\varepsilon/2}$ square matrix $\mathbf{M}$ lexicographically and applying blob list-decoding and reverse list-decoding to any row or column, denoted $\mathcal{R}$, of $\mathbf{M}$. Again, using Lemma 9, it suffices to bound the expected blob list-size and reverse list-size from above by some exponentially small quantity.

1) Blob List-Decoding: The expected number of codewords in the intersection of the blob and the codewords corresponding to the OGS is

$$\begin{aligned}
&\mathbb{E}\left(|\mathcal{B}lob \cap (\mathcal{O}rcl^{(j)}(\underline{z_Q}, i) \setminus \mathcal{R})| \Big| \widetilde{\mathcal{E}}^c \cap \mathcal{J}\right) \\
\leq &\frac{\mathrm{Area}(\mathcal{C}ap^{n-1}(\cdot, \sqrt{nr_{\mathrm{opt}}(1 + f_{11}(\varepsilon, \delta_{\mathcal{S}}))}, \sqrt{nP}))2^{n\varepsilon}}{\mathrm{Area}(\mathcal{S}tr^{n-1}(O'_-, O'_+, \sqrt{nr_-}, \sqrt{nr_+}))} \\
&\cdot 2^{n\varepsilon/2}\Delta(\tau)(1 + o(1)) \\
\leq &\sqrt{\frac{r_{\mathrm{opt}}(1 + f_{11}(\varepsilon, \delta_{\mathcal{S}}))}{r_{\mathrm{str}}}} \\
&\cdot 2^{n\left(-\frac{1}{2}\log\left(\frac{P}{r_{\mathrm{opt}}(1 + f_{11}(\varepsilon, \delta_{\mathcal{S}}))}\right) + \frac{1}{2}\log\left(\frac{P}{r_{\mathrm{str}}}\right) + 3\varepsilon/2\right)} \\
&\cdot \frac{\Delta(\tau)(1 + o(1))}{2^{(n-1)\frac{1}{2}\log(1+\tau) - \Theta(\log n)} - 2^{(n-1)\frac{1}{2}\log(1-\tau)}} \\
\leq &\sqrt{\frac{r_{\mathrm{opt}}(1 + f_{11}(\varepsilon, \delta_{\mathcal{S}}))}{r_{\mathrm{str}}}} \\
&\cdot 2^{n\left(-\frac{1}{2}\log\left(\frac{P}{r_{\mathrm{opt}}}\right) + \frac{1}{2}\log\left(\frac{P}{r_{\mathrm{str}}}\right) + f_{11}(\varepsilon, \delta_{\mathcal{S}}) + 3\varepsilon/2\right)} \\
&\cdot \frac{\Delta(\tau)(1 + o(1))}{2^{(n-1)\frac{1}{2}\log(1+\tau) - \Theta(\log n)} - 2^{(n-1)\frac{1}{2}\log(1-\tau)}} \\
\leq &\sqrt{\frac{r_{\mathrm{opt}}(1 + f_{11}(\varepsilon, \delta_{\mathcal{S}}))}{\frac{P\sigma^2(1-\varepsilon)}{(P+\sigma^2)(1+\varepsilon)}}} \\
&\cdot 2^{n\left(-\frac{1}{2}\log\left(\frac{P}{r_{\mathrm{opt}}}\right) + \frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right) + f_{11}(\varepsilon, \delta_{\mathcal{S}}) + 7\varepsilon/2\right)} \\
&\cdot \frac{\Delta(\tau)(1 + o(1))}{2^{(n-1)\frac{1}{2}\log(1+\tau) - \Theta(\log n)} - 2^{(n-1)\frac{1}{2}\log(1-\tau)}}.
\end{aligned}$$

2) Reverse List-Decoding: The expected number of messages in the row (column) which can be confused by a single message is

$$\begin{aligned}
&\mathbb{E}\left(|\mathcal{B}^n(\underline{\mathbf{x}}(m'), \sqrt{nN} + \sqrt{n\delta_{\mathcal{S}}}) \cap (\{\underline{\mathbf{x}}(m)\}_{m \in \mathcal{R}} + \underline{s_Q})| \Big| \widetilde{\mathcal{E}}^c \cap \mathcal{J}\right) \\
= &\mathbb{E}\left(|\mathcal{B}^n(\underline{\mathbf{x}}(m') - \underline{s_Q}, \sqrt{nN} + \sqrt{n\delta_{\mathcal{S}}}) \cap \{\underline{\mathbf{x}}(m)\}_{m \in \mathcal{R}}|| \widetilde{\mathcal{E}}^c \cap \mathcal{J}\right) \\
\leq &\frac{\mathrm{Area}(\mathcal{S}^{n-1}(\cdot, \sqrt{nN} + \sqrt{n\delta_{\mathcal{S}}}))}{\mathrm{Area}(\mathcal{S}tr^{n-1}(O'_-, O'_+, \sqrt{nr_-}, \sqrt{nr_+}))} 2^{n\varepsilon/2}\Delta(\tau)(1 + o(1)) \\
\leq &\sqrt{\frac{N}{r_{\mathrm{str}}}} 2^{n\left(-\frac{1}{2}\log\frac{P}{N} + \frac{1}{2}\log\left(\frac{P}{r_{\mathrm{str}}}\right) + 2\frac{\delta_{\mathcal{S}} + 2\sqrt{N\delta_{\mathcal{S}}}}{N + \delta_{\mathcal{S}} + 2\sqrt{N\delta_{\mathcal{S}}}} + \varepsilon/2\right)} \\
&\cdot \frac{\Delta(\tau)(1 + o(1))}{2^{(n-1)\frac{1}{2}\log(1+\tau) - \Theta(\log n)} - 2^{(n-1)\frac{1}{2}\log(1-\tau)}} \\
\leq &\sqrt{\frac{N(P + \sigma^2)(1 + \varepsilon)}{P\sigma^2(1 - \varepsilon)}} \\
&\cdot 2^{n\left(-\frac{1}{2}\log\frac{P}{N} + \frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right) + 2\frac{\delta_{\mathcal{S}} + 2\sqrt{N\delta_{\mathcal{S}}}}{N + \delta_{\mathcal{S}} + 2\sqrt{N\delta_{\mathcal{S}}}} + 5\varepsilon/2\right)} \\
&\cdot \frac{\Delta(\tau)(1 + o(1))}{2^{(n-1)\frac{1}{2}\log(1+\tau) - \Theta(\log n)} - 2^{(n-1)\frac{1}{2}\log(1-\tau)}}.
\end{aligned}$$

3) Grid Argument: By Section XI-B.1 and Section XI-B.2, for any $\underline{z_Q}$, i, j, $\underline{s_Q}$ and $\mathcal{R}$, there are at most $n^2 \cdot n^2 = n^4$ messages satisfying the condition in the lemma. Thus there are at most $2 \cdot 2^{n\varepsilon/2} \cdot n^4$ such "bad" messages in $\mathbf{M}$, i.e., the OGS. This can be formally proved as follows. Let $\mathcal{E}_{\mathrm{conf}}(m, m')$ be the event that m and m' are confusable, i.e.,

$$\mathcal{E}_{\mathrm{conf}}(m, m') := \{\underline{\mathbf{x}}(m') \in \mathcal{L}^{(\mathbf{k})}(\underline{x}(m), \underline{s_Q})\}.$$

Let $\mathbf{M}(k, \cdot)$ and $\mathbf{M}(\cdot, \ell)$ denote the k-th row and ℓ-th column of $\mathbf{M}$, respectively. We now bound the number of codewords in the OGS that can be confused by some other codewords in the OGS as follows. Eqn. (XI.7), as shown at the bottom of the next page, follows since $\mathcal{O}rcl^{(j)}(\underline{z_Q}, i) \setminus \{m\} = (\mathcal{O}rcl^{(j)}(\underline{z_Q}, i) \setminus \mathbf{M}(m, \cdot)) \cup (\mathcal{O}rcl^{(j)}(\underline{z_Q}, i) \setminus \mathbf{M}(\cdot, m))$. Eqn. (XI.8), as shown at the bottom of the next page, is by the union bound. Note that the inner summation of first (resp. second) term in Eqn. (XI.9), as shown at the bottom of the next page, is the number of codewords in the k-th row (resp. ℓ-th column) that might be confused by codewords from other rows (resp. columns). Both inner summations are at most n^4 with probability $1 - 2^{-\Omega(n^3)}$. Since there are $2^{n\varepsilon/2}$ rows/columns, the total number of codewords in the OGS that might be confused by some other codewords in the OGS is at most $2 \cdot 2^{n\varepsilon/2} \cdot n^4$ with probability $1 - 2^{-\Omega(n^3)}$. A union bound over $\underline{z_Q}$, i, j and $\underline{s_Q}$ completes the proof. □

XII. Concluding Remarks/Future Directions

In this work, we studied the capacity of a myopic adversarial channel with quadratic constraints. We did so for different amounts of common randomness, and were able to find a complete characterization for certain regimes of the noise-to-signal ratios of Bob and James.

1) For different regimes of the NSRs (Figs. 6a, 6b, 6c, and 6d), we were able to characterize the capacity in the red, blue and grey regions. We only have nonmatching upper and lower bounds on the capacity in the green and white regions.
2) We also derived a myopic list-decoding result in the general case when Alice and Bob share a linear amount of common randomness. We believe that this is a useful technique that is worth exploring for general channels.

3) When Alice uses a deterministic encoder, we believe that an improved converse using linear programming-type bounds might be obtained in the green and white regions.
4) The $\underline{\mathbf{z}}$-aware symmetrization argument could also be extended to obtain Plotkin-type upper bounds on the rate in the green and white regions.
5) We also believe that superposition codes could be used to obtain improved achievability results in the green and white regions for the case when there is no common randomness. In particular, we feel that rates exceeding R_{GV} should be achievable using superposition codes in the green and white regions.
6) A natural problem is to find the minimum amount of common randomness required to achieve the capacity in Fig. 8. We know for certain values of the NSRs, this is achievable with no common randomness (blue and red regions in Fig. 6a). Even $\Theta(\log n)$ bits is sufficient to achieve R_{LD} in the entire red region in Fig. 8, while the blue region can be expanded with increasing amounts ($\Omega(n)$ bits) of shared secret key. A lower bound on n_{key} needed to achieve capacity along the lines of [41] would be of interest.
7) In this article, we studied the impact of an adversary who has noncausal access to a noisy version of the transmitted signal. However, in reality, James can only choose his attack vector based on a *causal* observation of the transmission. Li et al. [40] have some recent results for the quadratically constrained adversarial channel where the jammer can choose the ith symbol of his transmission based on the first i symbols of the transmitted codeword. An interesting direction is to look at the impact of myopia in this setup.
8) Our work was inspired by the study of the discrete myopic adversarial channel [25]. A part of their work involved studying the capacity of a binary channel with a bit-flipping adversary who can flip at most np bits of the transmitted codeword (for some $0 < p < 1/2$). The adversary can choose his attack vector based on a noncausal observation of the output of a binary symmetric channel with crossover probability q. Dey et al. [25] observed that if $q > p$ (sufficiently myopic), then the adversary is essentially "blind," i.e., the capacity is equal to $1 - H(p)$. This is what one would obtain when James were oblivious to the transmitted codeword. In other words, as long as the channel from Alice to Bob has capacity greater than that of the channel seen by James, damage that James can do is minimal. What we observe in the quadratically constrained case is slightly different. A sufficient condition for our results to go through is that the *list-decoding capacity* for Bob be greater than the *Shannon capacity* for the channel seen by James. Even then, we can never hope to achieve the oblivious capacity $\frac{1}{2}\log(1+\frac{P}{N})$ for any finite σ. What we can achieve is the *myopic list-decoding capacity*. In the bit-flipping adversarial case, the list-decoding capacity is equal to the capacity of the channel with an oblivious adversary. No amount of myopia can let us obtain a higher list-decoding capacity. However, the two capacities are different in the quadratically constrained scenario.
9) The difference between oblivious and list-decoding capacities might explain the gap between the upper and lower bounds for general discrete myopic adversarial channels [25]. Our technique of using myopic list-decoding could potentially be used to close this gap in certain regimes.
10) While the use of list-decoding as a technique for obtaining capacity of general AVCs is not new [37], we believe that myopic list-decoding and reverse list-decoding can be generalized to arbitrary AVCs to obtain results even in the case where the encoder-decoder pair do not share common randomness.
11) The code construction in our achievability proof is the random spherical code ensemble which cannot be decoded efficiently. It is of interest to construct efficiently decodable codes for myopic channels. Note that lattice codes do not work due to its linearity. Specifically, the nearest codeword of any codeword in a lattice code is along the same direction. As a result, if an adversarial attack $\underline{s}$ confuses one codeword, the same attack can confuse any other codeword in the code. Therefore, lattice codes are bad under *average* probability of error.

APPENDIX A
TABLE OF NOTATION

See Table III.

$$\sum_{m\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i)} \mathbb{1}_{\{\exists m'\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i)\setminus\{m\},\, \mathcal{E}_{\mathrm{conf}}(m,m')\}}$$

$$= \sum_{m\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i)} \mathbb{1}_{\{\exists m'\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i)\setminus\mathbf{M}(m,\cdot),\, \mathcal{E}_{\mathrm{conf}}(m,m') \text{ or } \exists m'\in\mathcal{O}rcl^{(j)}\underline{(z_Q},i)\setminus\mathbf{M}(\cdot,m),\, \mathcal{E}_{\mathrm{conf}}(m,m')\}} \tag{XI.7}$$

$$\leq \sum_{m\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i)} \mathbb{1}_{\{\exists m'\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i)\setminus\mathbf{M}(m,\cdot),\, \mathcal{E}_{\mathrm{conf}}(m,m')\}} + \sum_{m\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i)} \mathbb{1}_{\{\exists m'\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i)\setminus\mathbf{M}(\cdot,m),\, \mathcal{E}_{\mathrm{conf}}(m,m')\}}, \tag{XI.8}$$

$$\leq \sum_{k\in[2^{n\varepsilon/2}]} \sum_{m\in\mathbf{M}(k,\cdot)} \mathbb{1}_{\{\exists m'\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i)\setminus\mathbf{M}(m,\cdot),\, \mathcal{E}_{\mathrm{conf}}(m,m')\}} + \sum_{\ell\in[2^{n\varepsilon/2}]} \sum_{m\in\mathbf{M}(\cdot,\ell)} \mathbb{1}_{\{\exists m'\in\mathcal{O}rcl^{(j)}(\underline{z_Q},i)\setminus\mathbf{M}(\cdot,m),\, \mathcal{E}_{\mathrm{conf}}(m,m')\}}. \tag{XI.9}$$

TABLE III
TABLE OF NOTATION

Symbol	Description	Value/Range
C	Capacity	$\limsup_{n\to\infty} R^{(n)}$
$\mathcal{C}$	Codebook	$\{\underline{x}(m,k) : m \in [2^{nR}], k \in [2^{n_{\text{key}}}]\} \subseteq \mathcal{B}^n(0,\sqrt{nP})$
C_{AWGN}	Capacity of AWGN channels	$\frac{1}{2}\log\left(1+\frac{P}{N}\right)$
$\mathcal{C}^{(\mathbf{k})}$	Codebook shared by Alice and Bob specified by common randomness $\mathbf{k}$	$\mathcal{C}^{(\mathbf{k})} = \{\underline{x}(m,\mathbf{k})\}_{m=1}^{2^{nR}}$
C_{myop}	Capacity of myopic adversarial channels with bounded common randomness	See Lemma 12, Lemma 13 and Theorem 14
$C_{\text{myop,sec}}$	Secrecy capacity of myopic adversarial channels with bounded common randomness	See Lemma 16, Lemma 17 and Lemma 18
$C_{\text{myop,LD}}$	Myopic list-decoding capacity	See Theorem 11
$C_{\text{myop,rand}}$	Capacity of myopic adversarial channels with unbounded common randomness	See Lemma 10
$C_{\text{myop,rand,sec}}$	Secrecy capacity of myopic adversarial channels with unbounded common randomness	See Lemma 15
$\mathcal{E}$	Shorthand notation for the union of several error events	$\mathcal{E} = \mathcal{E}_{\text{atyp}} \cup \mathcal{E}_{\text{str}} \cup \mathcal{E}_{\text{orcl}}$
$\widetilde{\mathcal{E}}$	Shorthand notation for the union of several error events	$\widetilde{\mathcal{E}} = \mathcal{E} \cup \mathcal{E}_{\underline{s_Q}}(\underline{z_Q}, i, j)$
$\mathcal{E}_{\text{atyp}}$	The error event that James's observation behaves atypically	See Equation (X.8)
$\mathcal{E}_{\text{LD}}$	The error event that more than n^2+1 codewords have large list-sizes	See Equation (X.18)
$\mathcal{E}_{\text{LD}-\text{rad}}(m, \underline{s_Q})$	The error event that $\underline{\mathbf{x}}(m)$ has an atypical list-decoding radius under $\underline{s_Q}$	See Equation (X.13)
$\mathcal{E}_{\text{orcl}}$	The error event that the transmitted codewords falls into the last OGS in a strip	See Item X-C
$\mathcal{E}_{\text{str}}$	The error event that there are not enough codewords in a strip	See Equation (X.10)
$\mathcal{E}_{\underline{s_Q}}(\underline{z_Q}, i, j)$	The error event that more than n^2 codewords have atypical list-decoding radii under $\underline{s_Q}$	See Equation (X.15)
$\underline{\mathbf{g}}$	Gaussian part of scale-and-babble attack	$\underline{\mathbf{g}} \sim \mathcal{N}(0, \gamma^2 \mathbf{I}_n)$
$\mathcal{J}$	The event that $\underline{z_Q}$, the strip, the OGS and $\underline{s_Q}$ are instantiated	$(\underline{\mathbf{z}_Q}, \underline{\mathbf{s}_Q}) = (\underline{z_Q}, \underline{s_Q}),\ m \in \mathcal{O}rcl^{(j)}(\underline{z_Q}, i)$
$\mathbf{k}$	Common randomness shared by Alice and Bob	$\mathbf{k} \in \{0,1\}^{n_{\text{key}}}$
ℓ	Number of OGSs in a strip	$\ell = \lceil \lvert\mathcal{M}_{\text{str}}(\underline{z_Q}, i)\rvert / 2^{n\varepsilon} \rceil$
$\mathcal{L}^{(\mathbf{k})}(\underline{x}(m), \underline{s})$	List of $\underline{x}(m) + \underline{s}$	$\mathcal{B}^n(\underline{x}(m)+\underline{s}, \sqrt{nN}) \cap \mathcal{C}^{(\mathbf{k})}$
$\mathcal{L}^{(\mathbf{k})}(\underline{x}(m), \underline{s_Q})$	List of $\underline{x}(m) + \underline{s_Q}$	$\mathcal{B}^n(\underline{x}(m)+\underline{s_Q}, \sqrt{nN}+\sqrt{n\delta_{\mathcal{S}}}) \cap \mathcal{C}^{(\mathbf{k})}$
$\mathbf{m}$	Message held by Alice	$\mathbf{m} \sim \text{Unif}([2^{nR}])$
$\widehat{\mathbf{m}}$	Bob's reconstruction	$\widehat{\mathbf{m}} \in \{0,1\}^{nR}$
$\mathcal{M}_{\text{str}}(\underline{z_Q}, i)$	Set of codewords in $\mathcal{S}tr^{n-1}(\underline{z_Q}, i)$	$\underline{z_Q} \in \mathcal{Z}, i \in \{-\varepsilon/\delta+1, \cdots, \varepsilon/\delta\}$
N	James's power constraint, i.e., $\lVert\underline{s}\rVert_2 \leqslant \sqrt{nN}$	$N \in \mathbb{R}_{>0}$
n	Blocklength/number of channel uses	$n \in \mathbb{Z}_{>0}$
n_{key}	Amount of common randomness	$n_{\text{key}} = nR_{\text{key}}$
$\mathcal{O}rcl^{(j)}(\underline{z_Q}, i)$	Oracle-given set	$\underline{z_Q} \in \mathcal{Z}, i \in \{-\varepsilon/\delta+1, \cdots, \varepsilon/\delta\}, j \in [\ell]$
$\mathcal{O}rcl(\underline{z_Q}, \underline{\mathbf{x}})$	The oracle-given set containing the transmitted $\underline{\mathbf{x}}$	$\mathcal{O}rcl(\underline{z_Q}, \underline{\mathbf{x}}) = \mathcal{O}rcl^{(\iota)}(\underline{z_Q}, i)$
P	Alice's power constraint, i.e., $\lVert\underline{x}\rVert_2 \leqslant \sqrt{nP}$	$P \in \mathbb{R}_{>0}$
R	Rate	$\frac{\log\lvert\mathcal{C}^{(\mathbf{k})}\rvert}{n} \in \mathbb{R}_{\geqslant 0}$
R_{code}	Codebook rate	$R_{\text{code}} = R + R_{\text{key}} = \frac{\log\lvert\mathcal{C}\rvert}{n}$

TABLE III
(Continued.) TABLE OF NOTATION

R_{GV}	Gilbert–Varshamov bound, a lower bound on capacity of quadratically constrained omniscient adversarial channels	$\frac{1}{2}\log\left(\frac{P^2}{4N(P-N)}\right)\mathbb{1}_{\{P\geqslant 2N\}}$
R_{key}	Key rate	$R_{\mathrm{key}}=n_{\mathrm{key}}/n$
R_{LD}	List-decoding capacity of quadratically constrained omniscient adversarial channels	$\frac{1}{2}\log\frac{P}{N}$
R_{LP}	Linear programming bound, an upper bound on capacity of quadratically constrained omniscient adversarial channels	$(\alpha\log\alpha-\beta\log\beta)\mathbb{1}_{\{P\geqslant 2N\}}$, where $\alpha=\frac{P+2\sqrt{N(P-N)}}{4\sqrt{N(P-N)}}, \beta=\frac{P-2\sqrt{N(P-N)}}{4\sqrt{N(P-N)}}$
$R_{\mathrm{LD,myop}}$	N/A	$\frac{1}{2}\log\left(\frac{(P+\sigma^2)(P+N)-2P\sqrt{N(P+\sigma^2)}}{N\sigma^2}\right)$
R_{Rankin}	Rankin bound, an upper bound on capacity of quadratically constrained omniscient adversarial channels	$\frac{1}{2}\log\left(\frac{P}{2N}\right)\mathbb{1}_{\{P\geqslant 2N\}}$
$\mathbf{r}(m,\underline{s_Q})$	Radius of list-decoding region $\mathcal{C}ap^{n-1}(\underline{\mathbf{x}}(m)+\underline{s_Q},\sqrt{n\mathbf{r}})=\mathcal{B}^n(\underline{\mathbf{x}}(m)+\underline{s_Q},\sqrt{nN}+\sqrt{n\delta_{\mathcal{S}}})$ for $m\in\mathcal{O}rcl(z_Q,\underline{\mathbf{x}})$	See Equation (X.48)
$r_{\mathrm{opt}}(\underline{s_Q})$	Optimal solution of optimization (X.49)	See Equation (X.50)
$\mathbf{r}_{\mathrm{str}}$	Radius $\sqrt{n\mathbf{r}_{\mathrm{str}}}$ of a strip	$\mathbf{r}_{\mathrm{str}}\in\frac{P\sigma^2}{P+\sigma^2}(1\pm\varepsilon)$ w.h.p.
$\mathcal{S}$	An optimal covering of $\mathcal{B}^n(0,\sqrt{nP})$ with quantization error at most $\sqrt{n\delta_{\mathcal{S}}}$	$\mathcal{S}=\{\underline{s_Q}^{(i)}\}_{i=1}^{\lvert\mathcal{S}\rvert}$
$\mathcal{S}tr^{n-1}(\underline{z_Q},i)$	Strip	$\underline{z_Q}\in\mathcal{Z}, i\in\{-\varepsilon/\delta+1,\cdots,\varepsilon/\delta\}$
$\underline{\mathbf{s}}$	James's attack vector	$\underline{\mathbf{s}}\in\mathcal{B}^n(0,\sqrt{nN})$
$\underline{s_Q}$	Quantization of $\underline{s}$	$\underline{s_Q}\in\mathcal{S}$
$\underline{\mathbf{s}_z}$	AWGN to James	$\underline{\mathbf{s}_z}\sim\mathcal{N}(0,\sigma^2\mathbf{I}_n)$
$\underline{\mathbf{x}}$	Alice's transmitted codeword	$\underline{\mathbf{x}}\in\mathcal{C}$
$\underline{\mathbf{y}}$	Bob's observation	$\underline{\mathbf{y}}=\underline{\mathbf{x}}+\underline{\mathbf{s}}\in\mathcal{B}^n(0,\sqrt{nP}+\sqrt{nN})$
$\mathcal{Z}$	An optimal covering of $\mathcal{S}h^n(0,\sqrt{n(P+\sigma^2)(1\pm\varepsilon)})$ with quantization error at most $\sqrt{n\delta_{\mathcal{Z}}}$	$\mathcal{Z}=\{\underline{z_Q}^{(i)}\}_{i=1}^{\lvert\mathcal{Z}\rvert}$
$\underline{\mathbf{z}}$	James's noisy observation of $\underline{\mathbf{x}}$	$\underline{\mathbf{z}}=\underline{\mathbf{x}}+\underline{\mathbf{s}_z}\in\mathbb{R}^n$
$\underline{z_Q}$	Quantization of $\underline{z}$	$\underline{z_Q}\in\mathcal{Z}$
$\Delta(\tau)$	Quasi-uniformity factor	$\max_{\underline{z}}\max_i\frac{\max_{\underline{x}\in\mathcal{S}tr^{n-1}(\underline{z_Q},i)}p_{\underline{\mathbf{x}}\mid\underline{\mathbf{z}_Q}}(\underline{x}\mid\underline{z_Q})}{\min_{\underline{x}\in\mathcal{S}tr^{n-1}(\underline{z_Q},i)}p_{\underline{\mathbf{x}}\mid\underline{\mathbf{z}_Q}}(\underline{x}\mid\underline{z_Q})}$
δ	Thickness of a strip (See Equation (X.20))	$\mathcal{O}\left(\frac{\log n}{n}\right)$
$\delta_{\mathcal{S}}$	Quantization error parameter for $\underline{s}$, i.e., for any $\underline{s}\in\mathcal{B}^n(0,\sqrt{nN})$, there exists $\underline{s}'\in\mathcal{S}$, such that $\lVert\underline{s}-\underline{s}'\rVert_2\leqslant\sqrt{n\delta_{\mathcal{S}}}$	$\mathcal{O}(1)$
$\delta_{\mathcal{Z}}$	Quantization error parameter for $\underline{s}$, i.e., for any $\underline{z}\in\mathcal{S}h^n(0,\sqrt{n(P+\sigma^2)(1\pm\varepsilon)})$, there exists $\underline{z}'\in\mathcal{Z}$, such that $\lVert\underline{z}-\underline{z}'\rVert_2\leqslant\sqrt{n\delta_{\mathcal{Z}}}$	$\mathcal{O}(1)$
ε	N/A	$\mathcal{O}(1)$
ρ	N/A	$\mathcal{O}(1)$
σ	Standard deviation of channel noise to James	$\sqrt{\mathrm{Var}(\mathbf{s}_z)}$
τ	Thickness of a strip (See Equation (X.21))	$\mathcal{O}\left(\frac{\log n}{n}\right)$
$\chi(\underline{z_Q},i,j,\underline{s_Q})$	Number of codewords in an OGS with large list-sizes	See Equation (X.16)
$\psi(\underline{z_Q},i,j,\underline{s_Q})$	Number of codewords in an OGS with atypical list-decoding radii	See Equation (X.14)

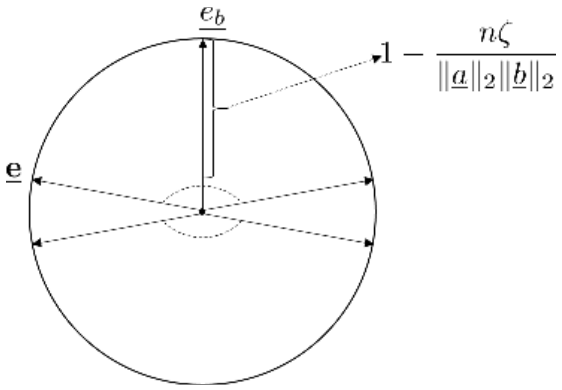


Fig. 26. The geometry corresponding to the tail bound of $|\langle \underline{\mathbf{a}}, \underline{b}\rangle|$ in Lemma 8.

APPENDIX B
PROOFS OF BASIC LEMMAS

A. Proof of Lemma 8

Let $\underline{e_b}$ denote the unit vector along $\underline{b}$, i.e., $\underline{e_b} = \underline{b}/\|\underline{b}\|_2$. Let $\underline{\mathbf{e}}$ denote the random unit vector along $\underline{\mathbf{a}}$ which is isotropically distributed on the unit sphere $\mathcal{S}^{n-1}(0,1)$, i.e., $\underline{\mathbf{e}} = \underline{\mathbf{a}}/\|\underline{a}\|_2$. Notice that $|\langle \underline{\mathbf{a}}, \underline{b}\rangle| > n\zeta$ if and only if $|\langle \underline{\mathbf{e}}, \underline{e_b}\rangle| > \frac{n\zeta}{\|\underline{a}\|_2\|\underline{b}\|_2}$, i.e., if and only if $\underline{\mathbf{e}}$ lies on one of two caps (shown in Figure 26) of height $1-\frac{n\zeta}{\|\underline{a}\|_2\|\underline{b}\|_2}$. Thus we have

$$\begin{aligned}
&\mathbb{P}(|\langle \underline{\mathbf{a}}, \underline{b}\rangle| > n\zeta)\\
=&\mathbb{P}\left(|\langle \underline{\mathbf{e}}, \underline{e_b}\rangle| > \frac{n\zeta}{\|\underline{a}\|_2\|\underline{b}\|_2}\right)\\
=&\frac{2\,\mathrm{Area}\left(\mathcal{C}ap^{n-1}\left(\frac{n\zeta}{\|\underline{a}\|_2\|\underline{b}\|_2}\frac{\underline{b}}{\|\underline{b}\|_2}, \sqrt{1-\frac{n^2\zeta^2}{\|\underline{a}\|_2^2\|\underline{b}\|_2^2}}, 1\right)\right)}{\mathrm{Area}(\mathcal{S}^{n-1}(0,1))}\\
\le&\frac{\mathrm{Area}\left(\mathcal{S}^{n-1}\left(0, \sqrt{1-\frac{n^2\zeta^2}{\|\underline{a}\|_2^2\|\underline{b}\|_2^2}}\right)\right)}{\mathrm{Area}(\mathcal{S}^{n-1}(0,1))}\\
=&2^{-\frac{n-1}{2}\log\left(\frac{1}{1-\frac{n^2\zeta^2}{\|\underline{a}\|_2^2\|\underline{b}\|_2^2}}\right)}\\
\le&2^{-\frac{(n-1)n^2\zeta^2}{2\|\underline{a}\|_2^2\|\underline{b}\|_2^2}},
\end{aligned}$$

where the last step follows from the inequality $\log\left(\frac{1}{1-x}\right) \ge \log(1+x) \ge x$ for small enough positive x. □

B. Proof of Lemma 9

Since p is the probability that a point chosen uniformly at random from A lies in V, we have

$$\begin{aligned}
\mathbb{P}(|V\cap\mathcal{C}| \ge cn^2) &= \sum_{i=cn^2}^{2^{nR}} \binom{2^{nR}}{i} p^i (1-p)^{2^{nR}-i}\\
&\le \sum_{i=cn^2}^{2^{nR}} \binom{2^{nR}}{i} p^i\\
&\le \sum_{i=cn^2}^{2^{nR}} \binom{2^{nR}}{i} 2^{-in(R+\nu)}
\end{aligned}$$

where the last step follows from the assumption that $p \le 2^{-n(R+\nu)}$. Using bounds on the binomial coefficient, the probability can be upper bounded as follows for large enough n:

$$\begin{aligned}
\mathbb{P}(|V\cap\mathcal{C}| \ge cn^2) &\le \sum_{i=cn^2}^{2^{nR}} \left(\frac{2^{nR}e}{i}\right)^i 2^{-in(R+\nu)}\\
&\le 2^{nR}\left(\frac{2^{nR}e}{cn^2}\right)^{cn^2} 2^{-cn^2\cdot n(R+\nu)}\\
&= 2^{nR+cRn^3+c(\log e)n^2-cn^2\log(cn^2)-c(R+\nu)n^3}\\
&= 2^{-c\nu n^3-2cn^2\log n+(c\log e-c\log c)n^2+nR}\\
&\le 2^{-Cn^3}.
\end{aligned}$$

This completes the proof. □

APPENDIX C
STOCHASTIC VS. DETERMINISTIC ENCODING AGAINST AN OMNISCIENT ADVERSARY

Suppose we are given a sequence of $(n, R_{\mathrm{stoch}}^{(n)}, P, N)$ stochastic codes $\mathcal{C}_{\mathrm{stoch}}^{(n)} = \{\underline{x}(m,k) : m \in [2^{nR_{\mathrm{stoch}}^{(n)}}], k \in [2^{nR_{\mathrm{key}}^{(n)}}]\}$ of blocklength n, message rate $R_{\mathrm{stoch}}^{(n)}$, bounded private key rate $R_{\mathrm{key}}^{(n)}$, subject to maximum power constraint P for Alice and maximum power constraint N for James, with a deterministic decoder and average probability of error $P_{\mathrm{e,stoch}}^{(n)} \overset{n\to\infty}{\to} 0$. Fix any n, we will turn $\mathcal{C}_{\mathrm{stoch}}^{(n)}$ into a $(n, R_{\mathrm{det}}^{(n)}, P, N)$ deterministic code $\mathcal{C}_{\mathrm{det}}^{(n)}$. The deterministic decoder associated with $\mathcal{C}_{\mathrm{stoch}}^{(n)}$ partitions $\mathbb{R}^n$ (the space that James's observation $\underline{y}$ lives in) into $2^{nR_{\mathrm{stoch}}^{(n)}}$ cells $\{\mathcal{Y}^{(n)}(m) \subset \mathbb{R}^n : m \in [2^{nR_{\mathrm{stoch}}^{(n)}}]\}$, where $\mathcal{Y}^{(n)}(m) := \{\underline{y} \in \mathcal{B}^n(0, \sqrt{nP} + \sqrt{nN}) : \mathrm{Dec}(\underline{y}) = m\}$. Collect all "good" messages into $\mathcal{M}^{(n)} = \{m \in [2^{nR_{\mathrm{stoch}}^{(n)}}] : \mathbb{P}(\widehat{\mathbf{m}} \ne \mathbf{m} | \mathbf{m} = m) < 1\}$. Assume that James's jamming strategy is deterministic. For any good message $m \in \mathcal{M}^{(n)}$, there must exist at least one "good" codeword $\underline{x}(m,k)$ such that

$$p_{\underline{\mathbf{x}}|\mathbf{m}}(\underline{x}(m,k)|m) > 0,$$

and

$$\forall \underline{s} \in \mathcal{B}^n(0, \sqrt{nN}),\ \underline{x}(m,k) + \underline{s} \in \mathcal{Y}^{(n)}(m).$$

The second condition is equivalent to $\mathcal{B}^n(\underline{x}(m,k), \sqrt{nN}) \subseteq \mathcal{Y}^{(n)}(m)$, i.e., James does not have enough power to push $\underline{x}(m,k)$ outside $\mathcal{Y}^{(n)}(m)$. For any good message, take any one of good codewords and we get a deterministic codebook $\mathcal{C}_{\mathrm{det}}^{(n)} = \{\underline{x}(m,\cdot) \in \mathcal{C}_{\mathrm{stoch}}^{(n)} : \underline{x}(m,\cdot) \text{ is good},\ m \in [2^{nR_{\mathrm{stoch}}^{(n)}}]\}$. By construction, this deterministic code with the same decoding region partition restricted to the messages in $\mathcal{M}^{(n)}$ enjoys zero probability of error. We then argue that it has asymptotically the same rate $R_{\mathrm{det}} = \lim_{n\to n} R_{\mathrm{stoch}}^{(n)}$ as $\mathcal{C}_{\mathrm{stoch}}^{(n)}$.

$$\begin{aligned}
P_{\mathrm{e,stoch}}^{(n)} &= \frac{1}{2^{nR_{\mathrm{stoch}}^{(n)}}} \sum_{m=1}^{2^{nR_{\mathrm{stoch}}^{(n)}}} \mathbb{P}(\widehat{\mathbf{m}} \ne \mathbf{m} | \mathbf{m} = m)\\
&\ge \frac{1}{2^{nR_{\mathrm{stoch}}^{(n)}}} |\{m \in [2^{nR_{\mathrm{stoch}}^{(n)}}] : \mathbb{P}(\widehat{\mathbf{m}} \ne \mathbf{m} | \mathbf{m} = m) = 1\}|
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2^{nR^{(n)}_{\text{stoch}}}} |(\mathcal{M}^{(n)})^c| \\
&= \mathbb{P}(\mathbf{m} \notin \mathcal{M}^{(n)}) \to 0.
\end{aligned}
$$

Define $\mathbf{e} = \mathbb{1}_{\{\mathbf{m} \in \mathcal{M}^{(n)}\}}$. We have

$$
\begin{aligned}
nR^{(n)}_{\text{stoch}} &= H(\mathbf{m}) \\
&= H(\mathbf{e}) + H(\mathbf{m}|\mathbf{e}) \\
&= H(\mathbf{e}) + \mathbb{P}(\mathbf{e}=1) H(\mathbf{m}|\mathbf{e}=1) + \mathbb{P}(\mathbf{e}=0) H(\mathbf{m}|\mathbf{e}=0) \\
&= H(\mathbf{e}) + \mathbb{P}(\mathbf{m} \in \mathcal{M}^{(n)}) H(\mathbf{m}|\mathbf{m} \in \mathcal{M}^{(n)}) \\
&\quad + \mathbb{P}(\mathbf{m} \notin \mathcal{M}^{(n)}) H(\mathbf{m}|\mathbf{m} \notin \mathcal{M}^{(n)}).
\end{aligned}
$$

It follows that

$$
\begin{aligned}
R^{(n)}_{\text{det}} &= \frac{1}{n} H(\mathbf{m}|\mathbf{m} \in \mathcal{M}^{(n)}) \\
&= \frac{1}{n\,\mathbb{P}(\mathbf{m} \in \mathcal{M}^{(n)})} (nR^{(n)}_{\text{stoch}} - H(\mathbf{e}) \\
&\quad - \mathbb{P}(\mathbf{m} \notin \mathcal{M}^{(n)}) H(\mathbf{m}|\mathbf{m} \notin \mathcal{M}^{(n)})) \\
&\geq \frac{1}{n\,\mathbb{P}(\mathbf{m} \in \mathcal{M}^{(n)})} (nR^{(n)}_{\text{stoch}} - 1 - P^{(n)}_{\text{e,stoch}} nR^{(n)}_{\text{stoch}}) \\
&= \frac{1}{1 - \mathbb{P}(\mathbf{m} \notin \mathcal{M}^{(n)})} \left((1 - P^{(n)}_{\text{e,stoch}}) R^{(n)}_{\text{stoch}} - \frac{1}{n} \right) \\
&\to \lim_{n \to \infty} R^{(n)}_{\text{stoch}}.
\end{aligned}
$$

APPENDIX D
QUADRATICALLY CONSTRAINED LIST-DECODING CAPACITY WITH AN OMNISCIENT ADVERSARY

A. Achievability

We use a random spherical code, i.e., the 2^{nR}, $R = \frac{1}{2}\log\frac{P}{N} - \varepsilon$ codewords $\mathcal{C} = \{\underline{x}(m)\}_{m=1}^{2^{nR}}$ are chosen independently and uniformly at random from the Euclidean sphere centered at the origin of radius $\sqrt{nP}$. Since James has a power constraint of $\sqrt{nN}$, the received vector $\underline{\mathbf{y}} = \underline{\mathbf{x}} + \underline{\mathbf{s}}$ is guaranteed to lie within the shell $\mathcal{S}h^n(0, \sqrt{nP} \pm \sqrt{nN})$. We will prove the following result:

Lemma 32: There exists a constant $c > 0$ independent of n and ε, but possibly on P, N, and R, such that

$$
\mathbb{P}\left(\forall \underline{y} \in \mathcal{S}h^n(0, \sqrt{nP} \pm \sqrt{nN}), \right.
\left. |\mathcal{B}^n(\underline{y}, \sqrt{nN}) \cap \mathcal{C}| < c\frac{1}{\varepsilon}\log\frac{1}{\varepsilon} \right) \geq 1 - 2^{-\Omega(n)}.
$$

Proof: Let $L := c\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}$ be the desired list-size, for some absolute constant c to be determined later. Define $\delta := N\varepsilon^2/8$. At first, we increase the list-decoding radius by a small amount $\sqrt{n\delta}$. As we will see later, this will be helpful when we take a union bound over possible $\underline{y}$'s. We first show that for any fixed $\underline{y}$, the probability (over the codebook) that there are more than L codewords within a distance $\sqrt{nN} + \sqrt{n\delta}$ to $\underline{y}$ is sufficiently small.

Observe that $|\mathcal{B}^n(\underline{y}, \sqrt{nN} + \sqrt{n\delta}) \cap \mathcal{C}| = |\mathcal{C}ap^{n-1}(\underline{y}, \sqrt{nN} + \sqrt{n\delta}, \sqrt{nP}) \cap \mathcal{C}|$. We claim that for any fixed $\underline{y}$,

$$
\mathbb{P}\left(|\mathcal{B}^n(\underline{y}, \sqrt{nN} + \sqrt{n\delta}) \cap \mathcal{C}| > L \right) \leq c_2 2^{-n(L+1)\varepsilon/2}. \tag{D.1}
$$

for some constant c_2 independent of n and ε.

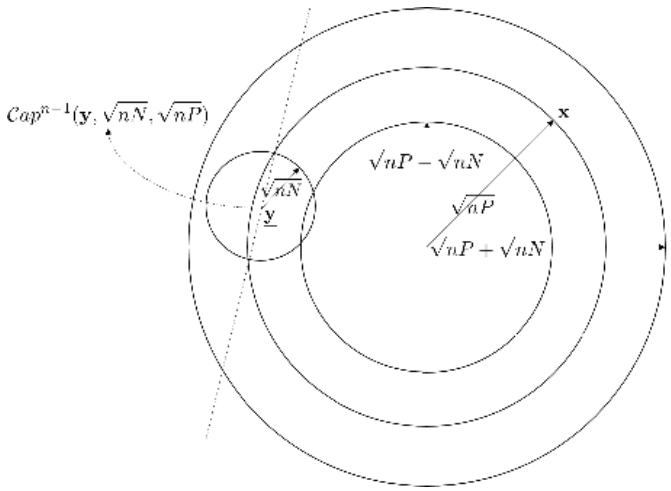


Fig. 27. Maximal intersection of the decoding ball with the sphere $\mathcal{S}^{n-1}(0, \sqrt{nP})$. In the figure, we omit the quantization parameter δ that goes into the actual proof, in particular dilates the radius of the noise ball by an additive factor $\sqrt{n\delta}$.

The maximal intersection of a ball $\mathcal{B}^n(\underline{y}, \sqrt{nN} + \sqrt{n\delta})$ and the Euclidean sphere $\mathcal{S}^{n-1}(0, \sqrt{nP})$ is shown in Figure 27 (with $\sqrt{n\delta}$ being dropped since it is a proof artifact rather than an essential factor in the geometry of list-decoding). It can be seen that the corresponding $\underline{y}$ has length $\sqrt{n(P-N)}$.

The probability of a codeword falling into the cap can be upper bounded by

$$
\begin{aligned}
p &:= \mathbb{P}\left(\underline{\mathbf{x}} \in \mathcal{C}ap^{n-1}(\underline{y}, \sqrt{nN} + \sqrt{n\delta}, \sqrt{nP}) \right) \\
&= \frac{\text{Area}(\mathcal{C}ap^{n-1}(\underline{y}, \sqrt{nN} + \sqrt{n\delta}, \sqrt{nP}))}{\text{Area}(\mathcal{S}^{n-1}(0, \sqrt{nP}))} \\
&\leq \frac{\text{Area}(\mathcal{S}^{n-1}(\underline{y}, \sqrt{nN} + \sqrt{n\delta}))}{\text{Area}(\mathcal{S}^{n-1}(0, \sqrt{nP}))} \\
&= \left(\frac{N}{P}\right)^{(n-1)/2} \left(1 + \sqrt{\frac{\delta}{N}}\right)^{n-1} \\
&= 2^{-\frac{n-1}{2}\left(\log\left(\frac{P}{N}\right) + 2\log\left(1 + \sqrt{\frac{\delta}{N}}\right)\right)} \\
&\leq c_1 2^{-n\left(\frac{1}{2}\log\frac{P}{N} + 2\sqrt{\frac{\delta}{N}}\right)} \\
&= c_1 2^{-n\left(\frac{1}{2}\log\frac{P}{N} + \frac{\varepsilon}{2}\right)},
\end{aligned}
$$

where $c_1 := \sqrt{\frac{P}{N}}$. In the last step, we have used the fact that $\delta = N\varepsilon^2/8$. Now consider the left-hand side of ((D.1)).

$$
\begin{aligned}
&\mathbb{P}\left(|\mathcal{C}ap^{n-1}(\underline{y}, \sqrt{nN} + \sqrt{n\delta}, \sqrt{nP}) \cap \mathcal{C}| > L \right) \\
&= \sum_{i=L+1}^{2^{nR}} \binom{2^{nR}}{i} p^i (1-p)^{2^{nR} - i}
\end{aligned}
$$

$$\leq 2^{nR}\binom{2^{nR}}{L+1}p^{L+1} \tag{D.2}$$

$$\leq 2^{nR}\left(\frac{2^{nR}e}{L+1}\right)^{L+1}\left(c2^{-n\frac{1}{2}\log\frac{P}{N}+\frac{\varepsilon n}{2}}\right)^{L+1}$$

$$= c_2 2^{nR+n(L+1)\left(R-\frac{1}{2}\log\frac{P}{N}+\frac{\varepsilon}{2}\right)}$$

$$= c_2 2^{-n(L+1)\varepsilon/2+nR}, \tag{D.3}$$

where $c_2 := \left(\frac{ec_1}{L+1}\right)^{L+1}$. Since we are interested in constant list-sizes, c_2 does not depend on n.

Define $\mathcal{Y}$ to be an optimal covering of $\mathcal{S}h^n(0,\sqrt{nP}\pm\sqrt{nN})$ by balls of radius $\sqrt{n\delta}$. In other words, $\mathcal{Y}$ is a finite set of points in $\mathcal{S}h^n(0,\sqrt{nP}\pm\sqrt{nN})$ such that $\min_{\underline{y}'\in\mathcal{Y}}\|\underline{y}-\underline{y}'\| \leq \sqrt{n\delta}$ for all $\underline{y}\in\mathcal{S}h^n(0,\sqrt{nP}\pm\sqrt{nN})$. In addition, $\mathcal{Y}$ is the smallest (in cardinality) over all possible coverings. One can achieve (for e.g., using lattice codes [48, Chapter 2])

$$|\mathcal{Y}| \leq \left(\frac{\mathrm{Vol}(\mathcal{B}^n(0,\sqrt{nP}+\sqrt{nN}+\sqrt{n\delta}))}{\mathrm{Vol}(\mathcal{B}^n(0,\sqrt{n\delta}))}\right)^{1+o(1)}$$

$$= \left(\frac{\sqrt{P}+\sqrt{N}+\sqrt{\delta}}{\sqrt{\delta}}\right)^{n(1+o(1))} =: \left(\frac{c_3}{\varepsilon}\right)^n. \tag{D.4}$$

We now have everything to prove Lemma 32.

$$\mathbb{P}\left(\exists\underline{y}\in\mathcal{S}h^n(0,\sqrt{nP}\pm\sqrt{nN}),\ |\mathcal{C}ap^{n-1}(\underline{y},\sqrt{nN},\sqrt{nP})\cap\mathcal{C}|>L\right)$$

$$\leq \mathbb{P}\left(\exists\underline{y_Q}\in\mathcal{Y},\ |\mathcal{C}ap^{n-1}(\underline{y_Q},\sqrt{nN}+\sqrt{n\delta},\sqrt{nP})\cap\mathcal{C}|>L\right)$$

$$\leq \sum_{\underline{y_Q}\in\mathcal{Y}}\mathbb{P}\left(|\mathcal{C}ap^{n-1}(\underline{y_Q},\sqrt{nN}+\sqrt{n\delta},\sqrt{nP})\cap\mathcal{C}|>L\right)$$

Now using (D.4) and (D.3), we have

$$\mathbb{P}\left(\exists\underline{y}\in\mathcal{S}h^n(0,\sqrt{nP}\pm\sqrt{nN}),\ |\mathcal{C}ap^{n-1}(\underline{y},\sqrt{nN},\sqrt{nP})\cap\mathcal{C}|>L\right)$$

$$\leq c_2 2^{-n(L+1)\varepsilon/2+nR}\left(\frac{c_3}{\varepsilon}\right)^n$$

$$= 2^{-\Omega(n)}$$

as long as

$$(L+1)\varepsilon/2 - R - \log\left(\frac{c_3}{\varepsilon}\right) > 0$$

or equivalently, $L > c\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}$ for a suitable constant c. This completes the proof of Lemma 32. □

B. Converse

Now we turn to the converse part of the list-decoding capacity theorem over quadratically constrained channels.

Lemma 33: If $R > \frac{1}{2}\log\frac{P}{N}$, then no sequence of codebooks of rate R is $(P,N,n^{\mathcal{O}(1)})$-list-decodable.

Proof: We will show that for any code $\mathcal{C}$ of rate $R = \frac{1}{2}\log\frac{P}{N}+\varepsilon$ with 2^{nR} codewords (not necessarily randomly) chosen from $\mathcal{S}^{n-1}(0,\sqrt{nP})$, there must some $\underline{y}$ with list-size exceeding $\mathcal{O}(1/\varepsilon)$. Let's choose $\underline{\mathbf{y}}$ uniformly at random on $\mathcal{S}^{n-1}(0,\sqrt{n(P-N)})$. As we saw, such $\underline{\mathbf{y}}$'s result in the largest list-decoding regions. Define

$$p := \mathbb{P}\left(\underline{x}\in\mathcal{C}ap^{n-1}(\underline{\mathbf{y}},\sqrt{nN},\sqrt{nP})\right)$$

$$= \mathbb{P}\left(\underline{\mathbf{y}}\in\mathcal{C}ap^{n-1}(\underline{x},\sqrt{nN},\sqrt{nP})\right).$$

First notice that p can be lower bounded by

$$p \geq \frac{\mathrm{Vol}(\mathcal{B}^{n-1}(0,\sqrt{nN}))}{\mathrm{Area}(\mathcal{S}^{n-1}(0,\sqrt{nP}))}$$

$$= \frac{1}{2\sqrt{\pi}}\left(\frac{\sqrt{2}}{\sqrt{n}}+\mathcal{O}(n^{-3/2})\right)\left(\frac{N}{P}\right)^{(n-1)/2}$$

$$= c_n\left(\frac{N}{P}\right)^{n/2}$$

$$= c_n 2^{-n\frac{1}{2}\log\frac{P}{N}},$$

where $c_n := \frac{1}{2\sqrt{\pi}}\left(\frac{\sqrt{2}}{\sqrt{n}}+\mathcal{O}(n^{-3/2})\right)\sqrt{\frac{P}{N}}$. The expected number of codewords in the intersection is

$$\mathbb{E}(|\mathcal{C}ap^{n-1}(\underline{\mathbf{y}},\sqrt{nN},\sqrt{nP})\cap\mathcal{C}|) = p2^{nR}$$

$$\geq c_n 2^{n\left(R-\frac{1}{2}\log\frac{P}{N}\right)} = c_n 2^{n\varepsilon}.$$

Hence there must exist some $\underline{y}$ in $\mathcal{S}^{n-1}(0,\sqrt{n(P-N)})$ such that

$$|\mathcal{C}ap^{n-1}(\underline{y},\sqrt{nN},\sqrt{nP})\cap\mathcal{C}|$$

$$\geq \mathop{\mathbb{E}}_{\underline{\mathbf{y}}\sim\mathrm{Unif}(\mathcal{S}^{n-1}(0,\sqrt{n(P-N)}))}(|\mathcal{C}ap^{n-1}(\underline{\mathbf{y}},\sqrt{nN},\sqrt{nP})\cap\mathcal{C}|)$$

$$\geq c_n 2^{n\varepsilon}.$$

By a black-box reduction from ball codes to spherical codes [18], this converse holds for any code satisfying Alice's power constraint. □

APPENDIX E
PROOF OF CLAIM 19

We want to prove that the scale-and-babble attack instantiates a channel whose capacity is equal to that of an equivalent AWGN channel.

We begin with the following observation, which follows from a simple application of the chain rule of mutual information.

Lemma 34: Consider any joint distribution $p_{\underline{\mathbf{x}},\underline{\mathbf{y}}}$ on $(\underline{\mathbf{x}},\underline{\mathbf{y}})$ such that $\underline{\mathbf{x}}$ has differential entropy which grows as $2^{o(n)}$. Let $\boldsymbol{\xi}$ be a $\{0,1\}$-valued random variable (possibly depending on $\underline{\mathbf{x}},\underline{\mathbf{y}}$) that takes value 0 with probability 2^{-cn} for some constant $c>0$. Then, $I(\underline{\mathbf{x}};\underline{\mathbf{y}}) \in I(\underline{\mathbf{x}};\underline{\mathbf{y}}|\boldsymbol{\xi}=1)(1-2^{-cn}) + 2^{-cn(1-o(1))}$.

Proof: The claim follows from straightforward computation.

$$I(\underline{\mathbf{x}};\underline{\mathbf{y}}) = \mathbb{P}(\boldsymbol{\xi}=1)I(\underline{\mathbf{x}};\underline{\mathbf{y}}|\boldsymbol{\xi}=1) + \mathbb{P}(\boldsymbol{\xi}=0)I(\underline{\mathbf{x}};\underline{\mathbf{y}}|\boldsymbol{\xi}=0)$$

$$= (1-2^{-cn})I(\underline{\mathbf{x}};\underline{\mathbf{y}}|\boldsymbol{\xi}=1) + 2^{-cn}(H(\underline{\mathbf{x}}|\boldsymbol{\xi}=0) - H(\underline{\mathbf{x}}|\underline{\mathbf{y}},\boldsymbol{\xi}=0))$$

$$= (1-2^{-cn})I(\underline{\mathbf{x}};\underline{\mathbf{y}}|\boldsymbol{\xi}=1) + 2^{-cn}\mathcal{O}(H(\underline{\mathbf{x}}))$$

$$= (1-2^{-cn})I(\underline{\mathbf{x}};\underline{\mathbf{y}}|\boldsymbol{\xi}=1) + 2^{-cn(1-o(1))}.$$ □

Let $\underline{\widetilde{\mathbf{y}}} := (1-\alpha)\underline{\mathbf{x}} + \underline{\widetilde{\mathbf{g}}}$, $\underline{\widetilde{\mathbf{g}}} := \underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z}$. The channel from $\underline{\mathbf{x}}$ to $\underline{\widetilde{\mathbf{y}}}$ is a standard AWGN channel with capacity (VII.2). Let

$$\boldsymbol{\xi} = \begin{cases} \boldsymbol{\beta}, & \text{if } \boldsymbol{\beta} = 1 \\ 0, & \text{if } 0 < \boldsymbol{\beta} < 1 \end{cases} = \begin{cases} 1, & \text{if } \| - \alpha\underline{\mathbf{z}} + \underline{\mathbf{g}}\|_2 \le \sqrt{nN} \\ 0, & \text{otherwise,} \end{cases}$$

i.e., the indicator random variable if James's power constraint is satisfied. Clearly, $I(\underline{\mathbf{x}}; \underline{\widetilde{\mathbf{y}}}|\boldsymbol{\beta} = 1) = I(\underline{\mathbf{x}}; \underline{\mathbf{y}}|\boldsymbol{\beta} = 1)$.

Lemma 35:

$$\mathbb{P}(\boldsymbol{\beta} \neq 1) = 2^{-\Omega(n)}.$$

Proof: Recall that

$$\begin{aligned} \underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z} &\sim \mathcal{N}(0, (\gamma^2 + \alpha^2\sigma^2)\mathbf{I}_n) \\ &= \mathcal{N}(0, (N - \alpha^2 P - (N - \alpha^2(P + \sigma^2))\varepsilon)\mathbf{I}_n) \\ &=: \mathcal{N}(0, (N - \alpha^2 P - \varepsilon')\mathbf{I}_n). \end{aligned}$$

Now we can bound the probability

$$\begin{aligned} &\mathbb{P}(\boldsymbol{\beta} \neq 1) \\ =&\mathbb{P}(\| - \alpha\underline{\mathbf{z}} + \underline{\mathbf{g}}\|_2 > \sqrt{nN}) \\ =&\mathbb{P}(\alpha^2\|\underline{\mathbf{x}}\|_2^2 + \|\underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z}\|_2^2 - 2\langle\alpha\underline{\mathbf{x}}, \underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z}\rangle > nN) \\ \le&\mathbb{P}(\|\underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z}\|_2^2 - 2\langle\alpha\underline{\mathbf{x}}, \underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z}\rangle > n(N - \alpha^2 P)) \\ =&\mathbb{P}(\|\underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z}\|_2^2 - 2\langle\alpha\underline{\mathbf{x}}, \underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z}\rangle > n(N - \alpha^2 P), \\ &\quad 2|\langle\alpha\underline{\mathbf{x}}, \underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z}\rangle| > n\varepsilon'/2) \\ &+ \mathbb{P}(\|\underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z}\|_2^2 - 2\langle\alpha\underline{\mathbf{x}}, \underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z}\rangle > n(N - \alpha^2 P), \\ &\quad 2|\langle\alpha\underline{\mathbf{x}}, \underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z}\rangle| \le n\varepsilon'/2) \\ \le&\mathbb{P}(|\langle\alpha\underline{\mathbf{x}}, \underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z}\rangle| > n\varepsilon'/4) \qquad \text{(E.1)} \\ &+ \mathbb{P}(\|\underline{\mathbf{g}} - \alpha\underline{\mathbf{s}_z}\|_2^2 > n(N - \alpha^2 P - \varepsilon'/2)). \qquad \text{(E.2)} \end{aligned}$$

We bound terms (E.1) and (E.2) separately.

$$\begin{aligned} \text{(E.1)} =&\mathbb{P}(|\mathcal{N}(0, \alpha^2\|\underline{\mathbf{x}}\|_2^2(N - \alpha^2 P - \varepsilon'))| > n\varepsilon'/4) \\ \le&2\exp\left(-\frac{(n\varepsilon'/4)^2}{2\alpha^2\|\underline{\mathbf{x}}\|_2^2(N - \alpha^2 P - \varepsilon')}\right) \\ \le&2\exp\left(-\frac{\varepsilon'^2}{32\alpha^2 P(N - \alpha^2 P - \varepsilon')}n\right) \\ =:&2^{-g_1(\varepsilon')n}, \end{aligned}$$

and

$$\begin{aligned} &\text{(E.2)} \\ =&\mathbb{P}(\|\mathcal{N}(0, (N - \alpha^2 P - \varepsilon')\mathbf{I}_n)\|_2^2 > n(N - \alpha^2 P - \varepsilon'/2)) \\ \le&\exp\left(\left\{-\frac{\varepsilon'/2}{N - \alpha^2 P - \varepsilon'} + \ln\left(1 + \frac{\varepsilon'/2}{N - \alpha^2 P - \varepsilon'}\right)\right\}\frac{n}{2}\right) \\ \le&\exp\left(-\frac{1}{4}\left(\frac{\varepsilon'/2}{N - \alpha^2 P - \varepsilon'}\right)^2\frac{n}{2}\right) \\ =&\exp\left(-\frac{\varepsilon'^2}{32(N - \alpha^2 P - \varepsilon')^2}n\right) \\ =:&2^{-g_2(\varepsilon')n}. \end{aligned}$$

□

Using Lemmas 34 and 35, we have that

$$\begin{aligned} I(\underline{\mathbf{x}}; \underline{\mathbf{y}}) =&I(\underline{\mathbf{x}}; \underline{\mathbf{y}}|\boldsymbol{\beta} = 1)(1 - 2^{-\Omega(n)}) + 2^{-\Omega(n)} \\ =&I(\underline{\mathbf{x}}; \underline{\widetilde{\mathbf{y}}}|\boldsymbol{\beta} = 1)(1 - 2^{-\Omega(n)}) + 2^{-\Omega(n)} \\ =&I(\underline{\mathbf{x}}; \underline{\widetilde{\mathbf{y}}})(1 + o(1)). \end{aligned}$$

Here we have used the fact that $\underline{\mathbf{x}}$ is power constrained, and hence has differential entropy $\Theta(n)$.

The rest of the proof follows along the same lines as the standard converse for the AWGN channel [49, Sec. 9.2]. Let us briefly outline the steps involved.

As a first step, note that Fano's inequality still holds even in the presence of common randomness. Let $\mathbf{k}$ denote the shared secret key. Specifically, if $\mathbf{m}$, $\widehat{\mathbf{m}}$, and P_{e}, respectively, denote the message chosen, Bob's estimate of the message, and the probability of error, then

$$H(\mathbf{m}|\widehat{\mathbf{m}}, \mathbf{k}) \le H(\mathbf{m}|\widehat{\mathbf{m}}) \le H(P_{\mathrm{e}}) + nRP_{\mathrm{e}},$$

where the first step follows because conditioning reduces entropy, and the second follows from (standard) Fano's inequality.

Now, if we demand that the probability of error be vanishingly small in n, then

$$\begin{aligned} nR = H(\mathbf{m}) = H(\mathbf{m}|\mathbf{k}) &= I(\mathbf{m}; \widehat{\mathbf{m}}|\mathbf{k}) + H(\mathbf{m}|\widehat{\mathbf{m}}, \mathbf{k}) \\ &\le I(\mathbf{m}; \widehat{\mathbf{m}}|\mathbf{k}) + o(n) \\ &\le I(\underline{\mathbf{x}}; \underline{\mathbf{y}}|\mathbf{k}) + o(n) \\ &\le I(\underline{\mathbf{x}}; \underline{\widetilde{\mathbf{y}}}|\mathbf{k})(1 + o(1)) + o(n) \\ &\le \frac{1}{2}\log\left(1 + \frac{(1-\alpha)^2 P}{\alpha^2\sigma^2 + \gamma^2}\right)n. \end{aligned} \qquad \text{(E.3)}$$

We have skipped a number of arguments in obtaining the last step, but these follow from the standard converse proof for the AWGN channel. The only property of the codebook used there is that it satisfies an average power constraint (which is indeed satisfied as we have a more restrictive max power constraint). This completes the proof of Claim 19.

Appendix F
Quasi-Uniformity – Proof of Lemma 23

As shown in Fig. 22, define $|\overline{xO'}| := \sqrt{nr_{\mathrm{str}}}$. By the construction of the strip,

$$\begin{aligned} |\overline{x^- O'_-}| &:= \sqrt{nr_-} := \sqrt{nr_{\mathrm{str}}(1 - \tau)}, \\ |\overline{x^+ O'_+}| &:= \sqrt{nr_+} := \sqrt{nr_{\mathrm{str}}(1 + \tau)}. \end{aligned} \qquad \text{(F.1)}$$

Then the quasi-uniformity factor can be computed as follows

$$\begin{aligned} &\frac{\sup_{\underline{x} \in \mathcal{S}tr^{n-1}(O'_-, O'_+, \sqrt{nr_-}, \sqrt{nr_+})} p_{\underline{\mathbf{x}}|\underline{\mathbf{z}}}(\underline{x}|\underline{z})}{\inf_{\underline{x} \in \mathcal{S}tr^{n-1}(O'_-, O'_+, \sqrt{nr_-}, \sqrt{nr_+})} p_{\underline{\mathbf{x}}|\underline{\mathbf{z}}}(\underline{x}|\underline{z})} \\ =&\frac{p_{\underline{\mathbf{x}}|\underline{\mathbf{z}}}(\underline{x}^-|\underline{z})}{p_{\underline{\mathbf{x}}|\underline{\mathbf{z}}}(\underline{x}^+|\underline{z})} \\ =&\frac{p_{\underline{\mathbf{x}},\underline{\mathbf{z}}}(\underline{x}^-, \underline{z})}{p_{\underline{\mathbf{x}},\underline{\mathbf{z}}}(\underline{x}^+, \underline{z})} \\ =&\frac{p_{\underline{\mathbf{z}}|\underline{\mathbf{x}}}(\underline{z}|\underline{x}^-)p_{\underline{\mathbf{x}}}(\underline{x}^-)}{p_{\underline{\mathbf{z}}|\underline{\mathbf{x}}}(\underline{z}|\underline{x}^+)p_{\underline{\mathbf{x}}}(\underline{x}^+)} \end{aligned}$$

$$
\begin{aligned}
&\mathbb{E}\left(\left|\mathcal{S}tr^{n-1}(O'_-, O'_+, \sqrt{nr_-}, \sqrt{nr_+}) \cap \mathcal{C}\right| \middle| \mathcal{E}^c_{\text{atyp}}\right) \\
&\geq \frac{\text{Area}(\mathcal{C}ap^{n-1}(O'_+, \sqrt{nr_+}, \sqrt{nP})) - \text{Area}(\mathcal{C}ap^{n-1}(O'_-, \sqrt{nr_-}, \sqrt{nP}))}{\text{Area}(\mathcal{S}^{n-1}(0, \sqrt{nP}))} 2^{nR_{\text{code}}} \Delta(\tau)^{-1} \\
&\geq \frac{\text{Vol}(\mathcal{B}^{n-1}(O'_+, \sqrt{nr_+})) - \text{Area}(\mathcal{S}^{n-1}(O'_-, \sqrt{nr_-}))}{\text{Area}(\mathcal{S}^{n-1}(0, \sqrt{nP}))} 2^{nR_{\text{code}}} \Delta(\tau)^{-1} \\
&\asymp \left[\frac{1}{\sqrt{n-1}} \left(\frac{r_+}{P}\right)^{(n-1)/2} - \left(\frac{r_-}{P}\right)^{(n-1)/2}\right] 2^{nR_{\text{code}}} \Delta(\tau)^{-1} \qquad \text{(G.1)} \\
&= \left[n^{-1/2} \left(\frac{r_{\text{str}}}{P}(1+\tau)\right)^{(n-1)/2} - \left(\frac{r_{\text{str}}}{P}(1-\tau)\right)^{(n-1)/2}\right] 2^{nR_{\text{code}}} \Delta(\tau)^{-1} \\
&= \left(2^{(n-1)\left(\frac{1}{2}\log\left(\frac{r_{\text{str}}}{P}\right) + \frac{1}{2}\log(1+\tau)\right) - \frac{1}{2}\log n} - 2^{(n-1)\left(\frac{1}{2}\log\left(\frac{r_{\text{str}}}{P}\right) + \frac{1}{2}\log(1-\tau)\right)}\right) 2^{nR_{\text{code}}} \Delta(\tau)^{-1} \\
&= \sqrt{\frac{P}{r_{\text{str}}}} 2^{-n\frac{1}{2}\log\left(\frac{P}{r_{\text{str}}}\right)} \left(2^{(n-1)\frac{1}{2}\log(1+\tau) - \frac{1}{2}\log n} - 2^{(n-1)\frac{1}{2}\log(1-\tau)}\right) 2^{nR_{\text{code}}} \Delta(\tau)^{-1} \\
&\geq \sqrt{\frac{(P+\sigma^2)(1+\varepsilon)}{\sigma^2(1-\varepsilon)}} 2^{-n\frac{1}{2}\log\left(\frac{(P+\sigma^2)(1-\varepsilon)}{\sigma^2(1+\varepsilon)}\right)} \left(2^{(n-1)\frac{1}{2}\log(1+\tau) - \frac{1}{2}\log n} - 2^{(n-1)\frac{1}{2}\log(1-\tau)}\right) 2^{nR_{\text{code}}} \Delta(\tau)^{-1}, \qquad \text{(G.2)}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{p_{\underline{\mathbf{z}}|\underline{\mathbf{x}}}(\underline{z}|\underline{x}^-)}{p_{\underline{\mathbf{z}}|\underline{\mathbf{x}}}(\underline{z}|\underline{x}^+)} \\
&= \exp\left(\frac{\|\underline{z} - \underline{x}^+\|_2^2 - \|\underline{z} - \underline{x}^-\|_2^2}{2\sigma^2}\right) \\
&= \exp\left(\frac{\|\underline{z}\|_2(\sqrt{n(P-r_-)} - \sqrt{n(P-r_+)})}{\sigma^2}\right) \qquad \text{(F.2)} \\
&= \exp\left(\frac{\|\underline{z}\|_2}{\sigma^2} \frac{2nr_{\text{str}}\tau}{\sqrt{n(P-r_-)} + \sqrt{n(P-r_+)}}\right),
\end{aligned}
$$

where Eqn. (F.2) holds since

$$
\begin{aligned}
\|\underline{z} - \underline{x}^{\pm}\|_2^2 &= nr_{\pm} + (\|\underline{z}\|_2 - \sqrt{nP - nr_{\pm}})^2 \\
&= \|\underline{z}\|_2^2 + nP - 2\|\underline{z}\|_2\sqrt{nP - nr_{\pm}}.
\end{aligned}
$$

In the above calculation, recall that as mentioned after the definition in Eqn. (X.22), the joint density $p_{\underline{\mathbf{x}},\underline{\mathbf{z}}}$ is given by $\underline{\mathbf{x}} \sim \text{Unif}(\mathcal{S}^{n-1}(0, \sqrt{nP}))$ and $\underline{\mathbf{z}} = \underline{\mathbf{x}} + \underline{\mathbf{s}_z}$ where $\underline{\mathbf{s}_z} \sim \mathcal{N}(0, \sigma^2\mathbf{I}_n)$.

Appendix G
Exponentially Many Codewords in the Strip – Proof of Lemma 24

The expected number of codewords in a strip can be estimated in (G.1) and (G.2), as shown at the top of this page. where Eqn. (G.1) follows from the fact that

$$
\begin{aligned}
\text{Vol}(\mathcal{B}^n(0,1)) &\asymp \frac{1}{\sqrt{\pi n}} \left(\frac{2\pi e}{n}\right)^{n/2}, \\
\text{Area}(\mathcal{S}^{n-1}(0,1)) &\asymp \sqrt{\frac{n}{\pi}} \left(\frac{2\pi e}{n}\right)^{n/2}.
\end{aligned}
$$

Eqn. (G.2) follows from Eqn. (X.24). The factor in the parentheses is a polynomial in n if we properly set $\tau = \mathcal{O}((\log n)/n)$. If the coding rate is strictly above the threshold $\frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right)$, then, in expectation, there are at least $2^{4\varepsilon n}$ codewords in every strip.

Appendix H
Proof of Lemma 16

The coding scheme is determined by parameters (R, R_{key}, R_e), where R_{key} denotes the rate of the secret key. We generate $2^{n(R+R_{\text{key}}+R_e)}$ codewords uniformly at random from the sphere $\mathcal{S}^{n-1}(0, \sqrt{nP})$. Let us index the codewords using the triple $(i_1, i_2, i_3) \in [2^{nR}] \times [2^{nR_{\text{key}}}] \times [2^{nR_e}]$. The messages are chosen uniformly at random from $[2^{nR}]$. Given a message $\mathbf{m} \in [2^{nR}]$ and key $\mathbf{k} \in [2^{nR_{\text{key}}}]$, the encoder picks $\mathbf{r}$ uniformly at random from $[2^{nR_e}]$, and transmits the $(\mathbf{m}, \mathbf{k}, \mathbf{r})$th codeword $\underline{\mathbf{x}}(\mathbf{m}, \mathbf{k}, \mathbf{r})$. Bob knows $\mathbf{k}$, and has to decode $\mathbf{m}$ from $\underline{\mathbf{y}}$.

We will choose the parameters so as to satisfy:

$$R + R_e < C_{\text{myop}},$$

and

$$R_e = \max\left\{0,\ \frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right) - R_{\text{key}} + \delta\right\}$$

for some small $\delta > 0$.

As long as $R + R_e < C_{\text{myop}}$, the probability of decoding error is $o(1)$ from Lemma 12. Since we are using random spherical code (which is a good resolvability code for the AWGN channel), we can directly invoke [38, Lemma 2] and [38, Remark 3] which show that the mutual information $I(\mathbf{m}, \underline{\mathbf{z}})$ is exponentially vanishing in n as long as $R' = R_e + R_{\text{key}} \geq \text{p-lim sup}_{n\to\infty} \frac{1}{n} I(\underline{\mathbf{x}}; \underline{\mathbf{z}}) + \delta = \frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right) + \delta$. □

References

[1] A. D. Sarwate, "An AVC perspective on correlated jamming," in *Proc. Int. Conf. Signal Process. Commun. (SPCOM)*, 2012, pp. 1–5.

[2] Y. Zhang, S. Vatedka, S. Jaggi, and A. D. Sarwate, "Quadratically constrained myopic adversarial channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 471–475.
[3] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacity of a class of channels under random coding," *Ann. Math. Statist.*, vol. 31, no. 3, pp. 558–567, 1960.
[4] N. Blachman, "On the capacity of a band-limited channel perturbed by statistically dependent interference," *IEEE Trans. Inf. Theory*, vol. IT-8, no. 1, pp. 48–55, Jan. 1962.
[5] R. A. Rankin, "The closest packing of spherical caps in n dimensions," *Proc. Glasgow Math. Assoc.*, vol. 2, no. 3, pp. 139–144, 1955.
[6] G. A. Kabatiansky and V. I. Levenshtein, "On bounds for packings on a sphere and in space," *Problemy Peredachi Informatsii*, vol. 14, no. 1, pp. 3–25, 1978.
[7] H. Cohn and Y. Zhao, "Sphere packing bounds via spherical codes," *Duke Math. J.*, vol. 163, no. 10, pp. 1965–2002, 2014.
[8] B. Hughes and P. Narayan, "Gaussian arbitrarily varying channels," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 2, pp. 267–284, Mar. 1987.
[9] I. Csiszár and P. Narayan, "Capacity of the Gaussian arbitrarily varying channel," *IEEE Trans. Inf. Theory*, vol. 37, no. 1, pp. 18–26, Jan. 1991.
[10] B. Hughes and P. Narayan, "The capacity of a vector Gaussian arbitrarily varying channel," *IEEE Trans. Inf. Theory*, vol. IT-34, no. 5, pp. 995–1003, Sep. 1988.
[11] T. G. Thomas and B. Hughes, "Exponential error bounds for random codes on Gaussian arbitrarily varying channels," *IEEE Trans. Inf. Theory*, vol. 37, no. 3, pp. 643–649, May 1991.
[12] A. Sarwate and M. Gastpar, "Randomization bounds on Gaussian arbitrarily varying channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2006, pp. 2161–2165.
[13] F. Haddadpour, M. J. Siavoshani, M. Bakshi, and S. Jaggi, "On AVCs with quadratic constraints," in *Proc. IEEE Int. Symp. Inf. Theory*, 2013, pp. 271–275.
[14] M. Médard, "Capacity of correlated jamming channels," in *Proc. Annu. Allerton Conf. Commun. Control Comput.*, vol. 35. Champaign, IL, USA: Univ. of Illinois, 1997, pp. 1043–1052.
[15] S. Shafiee and S. Ulukus, "Mutual information games in multiuser channels with correlated jamming," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4598–4607, Oct. 2009.
[16] C. R. Baker and I.-F. Chao, "Information capacity of channels with partially unknown noise. I. Finite-dimensional channels," *SIAM J. Appl. Math.*, vol. 56, no. 3, pp. 946–963, Jun. 1996.
[17] F. Hosseinigoki and O. Kosut, "List-decoding capacity of the Gaussian arbitrarily-varying channel," *Entropy*, vol. 21, no. 6, p. 575, Jun. 2019.
[18] Y. Zhang and S. Vatedka, "List decoding random Euclidean codes and infinite constellations," 2019, *arXiv:1901.03790*.
[19] U. Pereg and Y. Steinberg, "The arbitrarily varying Gaussian relay channel with sender frequency division," 2018, *arXiv:1805.12595*.
[20] A. Beemer, O. Kosut, J. Kliewer, E. Graves, and P. Yu, "Authentication against a myopic adversary," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Jun. 2019, pp. 1–5.
[21] Y. Zhang, S. Vatedka, and S. Jaggi, "Quadratically constrained two-way adversarial channels," 2020, *arXiv:2001.02575*.
[22] A. J. Budkuley, B. K. Dey, S. Jaggi, M. Langberg, A. D. Sarwate, and C. Wang, "Symmetrizability for myopic AVCs," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2020, pp. 2103–2107.
[23] X. Wang, A. J. Budkuley, A. Bogdanov, and S. Jaggi, "When are large codes possible for AVCs?" in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019, pp. 632–636.
[24] B. K. Dey, S. Jaggi, M. Langberg, A. D. Sarwate, and C. Wang, "The interplay of causality and myopia in adversarial channel models," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019, pp. 1002–1006.
[25] B. K. Dey, S. Jaggi, and M. Langberg, "Sufficiently myopic adversaries are blind," *IEEE Trans. Inf. Theory*, vol. 65, no. 9, pp. 5718–5736, Sep. 2019.
[26] Q. Zhang, M. Bakshi, and S. Jaggi, "Covert communication over adversarially jammed channels," *IEEE Trans. Inf. Theory*, vol. 67, no. 9, pp. 6096–6121, Sep. 2021.
[27] J. Song, Q. Zhang, S. Kadhe, M. Bakshi, and S. Jaggi, "Stealthy communication over adversarially jammed multipath networks," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7473–7484, Dec. 2020.
[28] E. MolavianJazi, M. Bloch, and J. N. Laneman, "Arbitrary jamming can preclude secure communication," in *Proc. 47th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2009, pp. 1069–1075.
[29] I. Bjelaković, H. Boche, and J. Sommerfeld, "Capacity results for arbitrarily varying wiretap channels," in *Information Theory, Combinatorics, and Search Theory*. Berlin, Germany: Springer, 2013, pp. 123–144.
[30] X. He, A. Khisti, and A. Yener, "MIMO multiple access channel with an arbitrarily varying eavesdropper: Secrecy degrees of freedom," *IEEE Trans. Inf. Theory*, vol. 59, no. 8, pp. 4733–4745, Aug. 2013.
[31] X. He and A. Yener, "MIMO wiretap channels with unknown and varying eavesdropper channel states," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6844–6869, Nov. 2014.
[32] J. Nötzel, M. Wiese, and H. Boche, "The arbitrarily varying wiretap channel—Secret randomness, stability, and super-activation," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3504–3531, Jun. 2016.
[33] M. Wiese, J. Notzel, and H. Boche, "A channel under simultaneous jamming and eavesdropping attack—Correlated random coding capacities under strong secrecy criteria," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 3844–3862, Jul. 2016.
[34] Z. Goldfeld, P. Cuff, and H. H. Permuter, "Arbitrarily varying wiretap channels with type constrained states," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7216–7244, Dec. 2016.
[35] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Aug. 1975.
[36] B. K. Dey, S. Jaggi, M. Langberg, and A. D. Sarwate, "The benefit of a 1-bit jump-start, and the necessity of stochastic encoding, in jamming channels," 2016, *arXiv:1602.02384*.
[37] A. D. Sarwate, "Robust and adaptive communication under uncertain interference," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Univ. California, Berkeley, Berkeley, CA, USA, Jul. 2008.
[38] M. R. Bloch and J. N. Laneman, "Strong secrecy from channel resolvability," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8077–8098, Dec. 2013.
[39] I. Csiszar and P. Narayan, "The capacity of the arbitrarily varying channel revisited: Positivity, constraints," *IEEE Trans. Inf. Theory*, vol. IT-34, no. 2, pp. 181–193, Mar. 1988.
[40] T. Li, B. Kumar Dey, S. Jaggi, M. Langberg, and A. D. Sarwate, "Quadratically constrained channels with causal adversaries," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2018, pp. 621–625.
[41] M. Langberg, "Private codes or succinct random codes that are (almost) perfect," in *Proc. 45th Annu. IEEE Symp. Found. Comput. Sci.*, Rome, Italy, Oct. 2004, pp. 325–334.
[42] S. Bhattacharya, A. J. Budkuley, and S. Jaggi, "Shared randomness in arbitrarily varying channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2019, pp. 627–631.
[43] K. Ball, "An elementary introduction to modern convex geometry," *Flavors Geometry*, vol. 31, nos. 1–58, pp. 1–58, 1997.
[44] S. Boucheron, G. Lugosi, and P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford, U.K.: Oxford Univ. Press, 2013.
[45] M. Langberg, "Oblivious communication channels and their capacity," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 424–429, Jan. 2008.
[46] S. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 4, pp. 451–456, Jul. 1978.
[47] W. Kang and N. Liu, "Wiretap channel with shared key," in *Proc. IEEE Inf. Theory Workshop*, Aug. 2010, pp. 1–5.
[48] J. H. Conway and N. J. A. Sloane. *Sphere Packings, Lattices And Groups* (Fundamental Principles of Mathematical Sciences), vol. 290, E. Bannai *et al.*, Eds., 3rd ed. New York, NY, USA: Springer-Verlag, 1999.
[49] T. Cover and J. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.

Yihan Zhang received the B.Eng. degree in computer science and technology from Northeastern University, Shenyang, China, in June 2016, and the Ph.D. degree from the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong, in August 2020. He was a Post-Doctoral Researcher at the Henry and Marilyn Taub Faculty of Computer Science, Technion—Israel Institute of Technology, from October 2020 to October 2021. He has been a Post-Doctoral Researcher at the Institute of Science and Technology Austria since October 2021. His research interests include information theory, coding theory, and statistics theory.

Shashank Vatedka (Member, IEEE) received the B.E. degree in electronics and communication from Visvesvaraya Technological University, Belgaum, and the M.Sc. (Engineering) and Ph.D. degrees from the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru, in 2017. From 2016 to 2018, he was a Research Assistant and a Post-Doctoral Researcher at the Institute of Network Coding, The Chinese University of Hong Kong, Hong Kong. From 2018 to 2019, he was a Post-Doctoral Fellow at Telecom Paris, France. Since October 2019, he has been an Assistant Professor with the Department of Electrical Engineering, Indian Institute of Technology, Hyderabad. His primary research interests include information theory and coding, with applications to data compression, security, and statistical inference. He was a recipient of the TCS Fellowship during his Ph.D. studies and the Seshagiri Kaikini Medal for Best Ph.D. Thesis at the Department of Electrical Communication Engineering, IISc Bengaluru. He also received the Best Paper Award-Honorable Mention at SPCOM 2020 and the Best Poster Award at the 2021 Stanford Compression Workshop.

Sidharth (Sid) Jaggi received the B.Tech. degree from IIT Bombay in 2000, and the M.S. and Ph.D. degrees from the CalTech in 2001 and 2006, respectively, all in electrical engineering. He was a Post-Doctoral Associate with LIDS MIT in 2006. He joined the Department of Information Engineering with The Chinese University of Hong Kong in 2007, and the School of Mathematics with the University of Bristol in 2020, where he is currently an Associate Professor. His research interests include the intersection of network information theory, coding theory, and algorithms. His research group thus (somewhat unwillingly) calls itself the CAN-DO-IT Team (Codes, Algorithms, Networks: Design and Optimization for Information Theory). Topics he has dabbled in include sparse recovery/group-testing, covert communication, network coding, and adversarial channels.

Anand D. Sarwate (Senior Member, IEEE) received the B.S. degrees in electrical engineering and computer science and mathematics from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2002, and the M.S. and Ph.D. degrees in electrical engineering from the Department of Electrical Engineering and Computer Sciences (EECS), University of California, Berkeley (UC Berkeley), Berkeley, CA, USA, in 2005 and 2008, respectively. He was a Post-Doctoral Researcher with the University of California, San Diego, CA, from 2008 to 2011. He was a Research Assistant Professor with the Toyota Technological Institute at Chicago from 2011 to 2013. He is a currently an Associate Professor with the Department of Electrical and Computer Engineering, The State University of New Jersey, New Brunswick, NJ, USA, where he has been since January 2014. His research interests include information theory, machine learning, signal processing, optimization, and privacy and security. He is a member of Phi Beta Kappa and Eta Kappa Nu. He received the NSF CAREER award in 2015 and the Rutgers Board of Governors Research Fellowship for Scholarly Excellence in 2020.