Abstract—We study the problem of high-dimensional multiple packing in Euclidean space. Multiple packing is a natural generalization of sphere packing and is defined as follows. Let $N > 0$ and $L \in \mathbb{Z}_{\geq 2}$. A multiple packing is a set $C$ of points in $\mathbb{R}^n$ such that any point in $\mathbb{R}^n$ lies in the intersection of at most $L - 1$ balls of radius $\sqrt{nN}$ around points in $C$. Given a well-known connection with coding theory, multiple packings can be viewed as the Euclidean analog of list-decodable codes, which are well-studied for finite fields. In this paper, we derive the best known lower bounds on the optimal density of list-decodable infinite constellations for constant $L$ under a stronger notion called average-radius multiple packing. To this end, we apply tools from high-dimensional geometry and large deviation theory.

Index Terms—Channel coding, communication channels, channel capacity, combinatorial mathematics, error correction codes.

I. INTRODUCTION

We study the problem of multiple packing in Euclidean space, a natural generalization of the sphere packing problem [2]. Let $N > 0$ and $L \in \mathbb{Z}_{\geq 2}$. We say that a point set $C$ in $\mathbb{R}^n$ forms a $(N, L - 1)$-multiple packing$^1$ if any point in $\mathbb{R}^n$ lies in the intersection of at most $L - 1$ balls of radius $\sqrt{nN}$ around points in $C$. Equivalently, the radius of the smallest ball containing any size-$L$ subset of $C$ is larger than $\sqrt{nN}$. This radius is known as the Chebyshev radius of the $L$-sized subset. If $L = 2$, then $C$ forms a sphere packing, i.e., a point set such that balls of radius $\sqrt{nN}$ around points in $C$ are disjoint, or equivalently, the pairwise distance of points in $C$ is larger than $2\sqrt{nN}$. The density of $C$ is measured by rate (a.k.a. the normalized logarithmic density (NLD)) defined as

$$R(C) := \liminf_{K \to \infty} \frac{1}{n} \ln \frac{|C \cap [-K, K]|}{|\{-K, K\}|},$$

i.e., the (normalized) number of points per volume. Denote by $C_{L-1}(N)$ the largest rate of a $(N, L - 1)$-multiple packing as $n \to \infty$. The goal of this paper is to advance the understanding of $C_{L-1}(N)$.

The problem of multiple packing is closely related to the list-decoding problem [3], [4] in coding theory. Indeed, a multiple packing can be seen exactly as the Euclidean analog of a list-decodable code. We will interchangeably use the terms “packing” and “code” to refer to the point set of interest. To see the connection, note that if any point in a multiple packing is transmitted through an adversarial channel that can inflict an arbitrary additive noise of length at most $\sqrt{nN}$, then the distorted transmission, one can decode to a list of the nearest $L - 1$ points which is guaranteed to contain the transmitted one. The quantity $C_{L-1}(N)$ can therefore be interpreted as the capacity of this channel (in the sense of Poltyrev [5]). Moreover, list-decodable codes can be turned into unique-decodable codes with the aid of side information such as common randomness shared between the transmitter and receiver [6], [7], [8]. List-decoding also serves as a proof technique towards unique-decoding in various communication scenarios; see, e.g., [9], [10].

The $L = 2$ case corresponds to the sphere packing problem, which has a long history since at least the Kepler conjecture [11] in 1611. The best known lower bound is due to Minkowski [12] using a straightforward volume packing argument. The best known upper bound is obtained by reducing it to the bounded case (i.e., packing points in a ball rather than in $\mathbb{R}^n$) for which we have the Kabatiansky–Levenshtein linear programming-type bound [13]. For $L > 2$, Blinovsky [14] claimed a lower bound by analyzing an (expurgated) Poisson Point Process (PPP). However, there were some gaps in the proof that we were unable to resolve. See Section X-F for a discussion. In this work we use a different approach to construct an unbounded packing which achieves the same lower bound as claimed in [14]. The paper [14] also presented an Elias–Bassalygo-type bound without a proof. A complete proof of it can be found in [15].

For the multiple packing problem with $L > 2$, many existing lower bounds are obtained under a stronger notion known as the average-radius multiple packing (see Definition 4 for the exact definition). A set $C$ of $\mathbb{R}^n$-valued points is called an

$^1$Logarithms to the base $e$ are denoted by $\ln(.)$.
average-radius multiple packing if for every \((L - 1)\)-subset of \(C\), the maximum distance from any point in the subset to the centroid of the subset is less than \(\sqrt{nN}\). Here the centroid of a subset is defined as the average of the points in the subset. Denote by \(\overline{C}_{L-1}(N)\) the largest density of average-radius multiple packings. In fact, we study this stronger notion of multiple packing in the present paper. For any finite \(L \in \mathbb{Z}_{\geq 2}\), it is unknown whether the largest multiple packing density under the regular notion is the same as that under the average-radius variant.

For \(L \to \infty\), Zhang and Vatedka [16] determined the limiting value of \(C_{L-1}(N)\). It follows from results in this paper that \(\overline{C}_{L-1}(N)\) converges to the same value as \(L \to \infty\).

Very little is known about structured packings. Grigorescu and Peikert [17] initiated the study of list-decodability of lattices. See also the recent work [18] by Mook and Peikert. Zhang and Vatedka [16] derived lower bounds on the list decoding rates achievable using random lattices and showed that for large \(L\), lattices can achieve the list decoding capacity.

A. Relation to Conference Version

This work was presented in part at the 2022 IEEE International Symposium on Information Theory [1]. Reference [1] only contains the proof of Equation (2) using PPPs. In the current paper, the same result is obtained via infinite constellations whose analysis is simpler and more transparent. Furthermore, results on fundamental properties of different notions of packing density and radius are presented.

II. RELATED WORKS

The sphere packing problem \((L = 2)\) has a long history and has been extensively studied, especially for small dimensions. The largest packing density is open for almost every dimension, except for \(n = 1\) (trivial), \(2\) ( [19], [20]), \(3\) (the Kepler conjecture, [21], [22]), \(8\) ( [23]) and \(24\) ( [24]). For \(n \to \infty\), the best lower and upper bounds remain the trivial sphere packing bound [12] and Kabatiansky–Levenshtein’s linear programming bound [13]. This paper is only concerned with (multiple) packings in high dimensions and we measure the density in the normalized way as mentioned in Section I.

There is a parallel line of research in combinatorial coding theory. Specifically, a uniquely-decodable code (resp. list-decodable code) is nothing but a sphere packing (resp. multiple packing) which has been extensively studied for \(\mathbb{F}_q^n\) equipped with the Hamming metric. Empirically, it seems that the problem is harder for smaller field sizes \(q\).

We first list the best known results for sphere packing (i.e., \(L = 2\)) in Hamming spaces. For \(q = 2\), the best lower and upper bounds are the Gilbert–Varshamov bound [25], [26] proved using a trivial volume packing argument and the second MRRW bound [27] proved using the seminal Delsarte’s linear programming framework [28], respectively. Surprisingly, the Gilbert–Varshamov bound can be improved using algebraic geometry codes [29], [30] for \(q \geq 49\). Note that such a phenomenon is absent in \(\mathbb{R}^n\); as far as we know, no algebraic constructions of Euclidean sphere packings are known to beat the greedy/random constructions. For \(q \geq n\), the largest packing density is known to exactly equal the Singleton bound [31], [32], [33] which is met by, for instance, the Reed–Solomon code [34].

Less is known for multiple packing in Hamming spaces. We first discuss the binary case (i.e., \(q = 2\)). For every \(L \in \mathbb{Z}_{\geq 2}\), the best lower bound appears to be Blinovskys’s bound [35, Theorem 2, Chapter 2] proved under the stronger notion of average-radius list-decoding. The best upper bound for \(L = 3\) is due to Ashikhmin et al. [36] who combined the MRRW bound [27] and Litsyn’s bound [37] on distance distribution. For any \(L \geq 4\), the best upper bound is essentially due to Blinovskys again [38], [35, Theorem 3, Chapter 2], though there are some partial improvements. In particular, the idea in [36] was recently generalized to larger \(L\) by Polyanskiy [39] who improved Blinovskys’s upper bound for \(L = 2\) (i.e., odd \(L - 1\)) and sufficiently large \(L\). Similar to [36], the proof also makes use of a bound on distance distribution due to Kalai and Linial [40] which in turn relies on Delsarte’s linear programming bound. For larger \(q\), Blinovskys’s lower and upper bounds\(^3\) [41], [42], [43, Chapter III, Lecture 9, §1 and 2] remain the best known.

As \(L \to \infty\), the limiting value of the largest multiple packing density is a folklore in the literature known as the “list-decoding capacity” theorem.\(^4\) Moreover, the limiting value remains the same under the average-radius notion.

The problem of list-decoding was also studied for settings beyond the Hamming errors, e.g., list-decoding against erasures [45], [46], insertions/deletions [47], asymmetric errors [48], etc. Zhang et al. considered list-decoding over general adversarial channels [49]. List-decoding against other types of adversaries with limited knowledge such as oblivious or myopic adversaries were also considered in the literature [9], [50], [51], [52], [53]. The current paper can be viewed as a collection of results for list-decodable codes for adversarial channels over \(\mathbb{R}\) with \(\ell_2\) constraints.

III. OUR RESULTS

We derive the best known lower bound on the largest multiple packing density. Let \(C_{L-1}(N)\) and \(\overline{C}_{L-1}(N)\) denote the largest density of multiple packings under the standard and the average-radius notions, respectively.

We juxtapose our bound with various existing bounds for the \((N, L - 1)\)-multiple packing problem. In Theorem 8, we prove the following lower bound on the optimal density for \((N, L - 1)\)-average-radius list-decoding (which is stronger than \((N, L - 1)\)-list-decoding):

\[
\overline{C}_{L-1}(N) \geq \frac{1}{2} \ln \frac{L - 1}{2\pi e N L} - \frac{\ln L}{2(L - 1)}.
\]

This bound turns out to be the largest known lower bound on both \(C_{L-1}(N)\) and \(\overline{C}_{L-1}(N)\) for all \(N \geq 0\) and \(L \in \mathbb{Z}_{\geq 2}\). In [14], Blinovskys considered PPPs and arrived at the same

\(^3\)Some gaps in the proof of the upper bound in [41], [42] are recently observed. These gaps are closed in [44] and the results therein are extended to the list-recovery setting which is a generalization of \(q\)-ary list-decoding.

\(^4\)It is an abuse of terminology to use “list-decoding capacity” here to refer to the large \(L\) limit of the \((L - 1)\)-list-decoding capacity.
bound. See Section X-F for a discussion. It is worthwhile pointing out that the above bound can also be obtained under \((N, L-1)\)-list-decoding (which is weaker than \((N, L)\)-average-radius list-decoding) via a connection with error exponents [54]. The techniques for \textit{bounded} packings (in which all points lie either in \(B^n(0, \sqrt{n}P)\) or on \(S^{n-1}(0, \sqrt{n}P)\) for some \(P > 0\)) in [55] can be adapted to the unbounded setting (where points can lie anywhere in \(\mathbb{R}^n\)) considered in this paper and be strengthened to work for the stronger notion of \textit{average-radius} multiple packing. They yield the following lower bound on \(C_{L-1}(N)\):

\[
\overline{C}_{L-1}(N) \geq \frac{1}{2} \ln \frac{L-1}{4\pi eNL}. \tag{3}
\]

As for upper bounds, the techniques in [14, 55, 56] can be adapted to the unbounded setting as well which yield the following upper bound on \(C_{L-1}(N)\):

\[
C_{L-1}(N) \leq \frac{1}{2} \ln \frac{L-1}{2\pi eNL}. \tag{4}
\]

Finally, it is known (see, e.g., [16]) that as \(L \to \infty\), \(C_{L-1}(N)\) converges to the following expression:

\[
C_{LD}(N) = \frac{1}{2} \ln \frac{1}{2\pi eN}. \tag{5}
\]

Note that, by the lower and upper bounds (Equations (2) and (4)) on \(\overline{C}_{L-1}(N)\) for finite \(L\), the limiting value of \(\overline{C}_{L-1}(N)\) as \(L \to \infty\) is also the above expression.

All the above bounds for \((N, L-1)\)-multiple packing are plotted in Figure 1 with \(L = 5\). The horizontal axis is \(N\) and the vertical axis is the value of various bounds. The largest lower bound turns out to be Equation (2) (for all \(N \geq 0\) and \(L \in \mathbb{Z}_{\geq 2}\)). This bound together with the Elias–Bassalygo-type upper bound in Equation (4) are plotted in Figure 2 for \(L = 3, 4, 5\). They both converge from below to Equation (5) as \(L\) increases.

\(^3\)Here we use \(B^n(z, r)\) and \(S^{n-1}(z, r)\) to denote the \(n\)-dimensional Euclidean ball and \((n-1)\)-dimensional Euclidean sphere of radius \(r\) centered at \(z\), respectively.

IV. LIST-DECODING CAPACITY FOR LARGE \(L\)

All bounds in this paper hold for any \textit{fixed} \(L\). In this section, we discuss the impact of our finite-\(L\) bounds on the understanding of the limiting values of the largest multiple packing density as \(L \to \infty\). Some of these results were known previously and others follow from the bounds in the current paper.

Characterizing \(C_{L-1}(N)\) or \(\overline{C}_{L-1}(N)\) is a difficult task that is out of reach given the current techniques. However, if the list-size \(L\) is allowed to grow, we can actually characterize \(C_{LD}(N) := \lim_{L \to \infty} C_{L-1}(N)\), \(\overline{C}_{LD}(N) := \lim_{L \to \infty} \overline{C}_{L-1}(N)\), where the subscript \(LD\) denotes List-Decoding.

The value of \(C_{LD}(N)\) is characterized in [16] which equals \(\frac{1}{2} \ln \frac{1}{2\pi eN}\).

\textit{Theorem 1} [16]: Let \(N > 0\). Then for any \(\varepsilon > 0\),

1) There exist \((N, L-1)\)-multiple packings of rate \(\frac{1}{2} \ln \frac{1}{2\pi eN} - \varepsilon\) for some \(L = O\left(\frac{1}{\ln \frac{1}{2\pi eN}}\right)\);

2) Any \((N, L-1)\)-multiple packing of rate \(\frac{1}{2} \ln \frac{1}{2\pi eN} + \varepsilon\) must satisfy \(L = e^{\Theta(n\varepsilon)}\).

Therefore, \(C_{LD}(N) = \frac{1}{2} \ln \frac{1}{2\pi eN}\).

Moreover, we claim \(\overline{C}_{LD}(N) = \frac{1}{2} \ln \frac{1}{2\pi eN}\). For an upper bound, recall that average-radius list-decodability implies (regular) list-decodability. Therefore, any upper bound on \(C_{L-1}(N)\) is also an upper bound on \(\overline{C}_{L-1}(N)\). We already saw an upper bound on \(C_{L-1}(N)\) in Equation (4) that approaches \(\frac{1}{2} \ln \frac{1}{2\pi eN}\) as \(L \to \infty\). Indeed, according to Theorem 8, for sufficiently large \(L\), our construction achieves \(\frac{1}{2} \ln \frac{1}{2\pi eN}\) under average-radius multiple packing.

\textit{Theorem 2}: For any \(N > 0\), \(\overline{C}_{LD}(N) = \frac{1}{2} \ln \frac{1}{2\pi eN}\).

V. OUR TECHNIQUES

We summarize our techniques below.

To obtain lower bounds on the largest multiple packing density, our basic strategy is random coding with expurgation, a standard tool from information theory. To show the existence of a list-decodable code of rate \(R\), we can sample \(e^{nR}\) points independently at random drawn according to a certain
The list-decoding radius is the Chebyshev radius of a size-$L$ list, i.e., the maximum probability of the error event, i.e., the probability of the error event, after the removal process. We then get a list-decodable code of rate $R$ by noting that the remaining code contains no bad lists.

In the above framework, the key ingredient is a good bound on the probability of the error event, i.e., the probability that the list-decoding radius of a size-$L$ list is smaller than $\sqrt{nN}$. Under the standard notion of multiple packing, the list-decoding radius is the Chebyshev radius of the list, i.e., the radius of the smallest ball containing the list. Under the average-radius notion of multiple packing, the (squared) list-decoding radius is the average squared radius of the list, i.e., the average squared distance from each point in the list to the centroid of the list.

Using the above idea, we first construct a finite codebook with minimum average squared radius $nN$ and supported over the hypercube $[-K,K]^n$ for a suitably chosen $K$. This is obtained by expurgating a random codebook obtained by choosing points independently and uniformly from $[-K,K]^n$. The finite codebook is then tiled across $\mathbb{R}^n$ to obtain an infinite constellation with the aforementioned density and minimum average squared radius $nN$. This construction is loosely inspired by the infinite constellations [5] which was originally studied in the context of coding for the additive white Gaussian noise channel. A similar construction was used by [16] to derive lower bounds on $C_{\text{LD}}(N)$ for large $L$.

The exact exponents of the probability of the error event are obtained using Cramér’s large deviation principle and the Laplace’s method.

As a technical contribution, we discover several new representations of the average radius and the Chebyshev radius. They play crucial roles in facilitating the analyses and the applications of some of these representations go beyond the scope of this paper. To name a few, the average squared radius of a list can be written as a quadratic form associated with the list. This representation is used to analyze Gaussian codes and spherical codes in [15] and infinite constellations in this paper. The average squared radius can also be written as the difference between the average (squared) norm of points in the list and the (squared) norm of the centroid of the list. This representation is used to analyze spherical codes and ball codes in [15]. The average squared radius can be further written as the average pairwise distance of the list. This allows us to give a one-line proof of the Blachman–Few reduction and its strengthened version [15]. Yet another way of writing the average squared radius using the average norm and the average pairwise correlation turns out to be useful for the proof of the Plotkin-type bound [15].

VI. ORGANIZATION OF THE PAPER

This paper derives a lower bound on the largest multiple packing density which turns out to be the best known so far. The rest of the paper is organized as follows. Notational conventions and preliminary definitions/facts are listed in Sections VII and VIII, respectively. After that, we present in Section IX the formal definitions of multiple packing and pertaining notions. We also discuss different notions of density of codes used in the literature. Furthermore, we obtain several novel representations of the Chebyshev radius and the average squared radius which are crucial for estimating their tail probabilities. In Section X, we formally introduce our construction and prove the main result. We end the paper with several open questions in Section XI.

VII. NOTATION

A. Conventions

Sets are denoted by capital letters in calligraphic typeface, e.g., $\mathcal{C}, \mathcal{B}$, etc. Random variables are denoted by lower case letters in boldface or capital letters in plain typeface, e.g., $x, S$, etc. Their realizations are denoted by corresponding lower case letters in plain typeface, e.g., $x, s$, etc. Vectors (random
or fixed) of length \( n \), where \( n \) is the blocklength without further specification, are denoted by lower case letters with underlines, e.g., \( \underline{x}, \underline{g}, \underline{z}, \underline{y} \). Vectors of length different from \( n \) are denoted by an arrow on top and the length will be specified whenever used, e.g., \( \vec{t}, \vec{\alpha} \). The \( i \)-th entry of a vector \( \underline{x} \in \mathbb{A}^n \) is denoted by \( \underline{x}(i) \) since we can alternatively think of \( \underline{x} \) as a function from \([n]\) to \( \mathbb{A} \). Same for a random vector \( \underline{x} \). Matrices are denoted by capital letters, e.g., \( A, \Sigma \), etc. Similarly, the \((i, j)\)-th entry of a matrix \( G \in \mathbb{F}^{m \times n} \) is denoted by \( G(i, j) \). We sometimes write \( G_{n \times m} \) to explicitly specify its dimension. For square matrices, we write \( G_n \) for short. Letter \( I \) is reserved for identity matrix.

### B. Functions

We use the standard Bachmann–Landau (Big-Oh) notation for asymptotics of real-valued functions in positive integers.

For two real-valued functions \( f(n), g(n) \) of positive integers, we say that \( f(n) \) asymptotically equals \( g(n) \), denoted \( f(n) \asymp g(n) \), if

\[
\lim_{n \to \infty} \frac{f(n)}{g(n)} = 1.
\]

For instance, \( 2^{n \log n} \asymp 2^n n^{2^n} \). We write \( f(n) \preceq g(n) \) (read \( f(n) \) is much smaller than \( g(n) \)) if the coefficients of the dominant terms in the exponents of \( f(n) \) and \( g(n) \) match,

\[
\lim_{n \to \infty} \frac{\log f(n)}{\log g(n)} = 1.
\]

For instance, \( 2^{3n} \preceq 2^{5n/2} \). Note that if \( f(n) \asymp g(n) \) implies \( f(n) \preceq g(n) \), but the converse is not true.

For any \( q \in \mathbb{R}_{>0} \), we write \( \log_q(x) \) for the logarithm to the base \( q \). In particular, let \( \log(x) \) and \( \ln(x) \) denote logarithms to the base \( 2 \) and \( e \), respectively.

For any \( \mathcal{A} \subseteq \Omega \), the indicator function of \( \mathcal{A} \) is defined as, for any \( x \in \Omega \),

\[
1_{\mathcal{A}}(x) := \begin{cases} 1 & x \in \mathcal{A} \\ 0 & x \notin \mathcal{A} \end{cases}.
\]

At times, we will slightly abuse notation by saying that \( 1_{\mathcal{A}} \) is 1 when event \( \mathcal{A} \) happens and 0 otherwise. Note that \( 1_{\emptyset}(\cdot) \) yields the following formula

\[
\int_{\mathcal{U}} f(\underline{u}) d\underline{u} = \int_{\mathcal{U}} f(\varphi(\underline{u})) |\det(\nabla \varphi(\underline{u}))| d\underline{u},
\]

where \( \nabla \varphi \in \mathbb{R}^{n \times n} \) denotes the Jacobian matrix of \( \varphi \).

#### 3 (Change of Variable): Let \( \mathcal{U} \subset \mathbb{R}^n \) be an open set and \( \varphi: \mathcal{U} \to \mathbb{R}^n \) be an injective differentiable function with continuous partial derivatives, the Jacobian of which is nonzero for every \( \underline{x} \in \mathcal{U} \). Then for any compactly supported, continuous function \( f: \varphi(\mathcal{U}) \to \mathbb{R} \), the substitution \( \underline{x} = \varphi(\underline{u}) \) yields the following formula

\[
\int_{\mathcal{U}} f(\underline{u}) d\underline{u} = \int_{\mathcal{U}} f(\varphi(\underline{u})) |\det(\nabla \varphi(\underline{u}))| d\underline{u},
\]

where \( \nabla \varphi \in \mathbb{R}^{n \times n} \) denotes the Jacobian matrix of \( \varphi \).

#### 4 (Laplace’s Method): Suppose \( f: \mathbb{R}^d \to \mathbb{R} \) is a twice continuously differentiable function on \( \mathcal{A} \subset \mathbb{R}^d \), and there exists a unique point \( \bar{t} \) in \( \text{int}(\mathcal{A}) \) (where \( \text{int}(\cdot) \) denotes the interior of a set) such that

\[
f(\bar{t}) = \min_{\underline{t} \in \mathcal{A}} f(\underline{t}), \quad (\text{Hess } f)(\bar{t}) > 0,
\]

\[\text{Recall that we use an arrow on top to denote a vector of length different from the blocklength } n. \text{ Here } \bar{t} \text{ has length } d.\]
where $Hess f \in \mathbb{R}^{d \times d}$ denotes the Hessian matrix of $f$. Suppose $g(\hat{t})$ is positive. Then
\[
\int_{\mathcal{A}} g(\hat{t}) e^{-M f(\hat{t})} d\hat{t} \sim \frac{2\pi}{M} \frac{d/2}{\sqrt{\det((Hess f)(\hat{t}^0))}}.
\]

**Theorem 5 (Cramér):** Let \( \{x_i\}_{i=1}^n \) be a sequence of i.i.d. real-valued random variables. Let \( s_n := \frac{1}{n} \sum_{i=1}^n x_i \). Then for any closed \( \mathcal{F} \subset \mathbb{R} \),
\[
\limsup_{n \to \infty} \frac{1}{n} \ln \Pr[s_n \in \mathcal{F}] \leq - \inf_{x \in \mathcal{F}} \{ \lambda x - \ln \mathbb{E}[e^{\lambda x}] \};
\]
and for any open \( \mathcal{G} \subset \mathbb{R} \),
\[
\liminf_{n \to \infty} \frac{1}{n} \ln \Pr[s_n \in \mathcal{G}] \geq - \inf_{x \in \mathcal{G}} \{ \lambda x - \ln \mathbb{E}[e^{\lambda x}] \}.
\]
Furthermore, when \( \mathcal{F} \) or \( \mathcal{G} \) corresponds to the upper (resp. lower) tail of \( s_n \), the maximizer \( \lambda \geq 0 \) (resp. \( \lambda \leq 0 \)).

**IX. Basic Definitions and Facts**

Given the intimate connection between packing and error-correcting codes, we will interchangeably use the terms “multiple packing” and “list-decodable code”. The parameter \( L \in \mathbb{Z}_{\geq 2} \) is called the multiplicity of overlap or the list-size. The parameter \( N \) is called the noise power constraints. Elements of a packing are called either points or codewords. We will call a size-\( L \) subset of a packing an \( L \)-list. This paper is only concerned with the fundamental limits of multiple packing for asymptotically large dimension \( n \to \infty \). When we say “\( a \) code \( \mathcal{C} \), we always mean an infinite sequence of codes \( \{ \mathcal{C}_i \}_{i \geq 1} \) where \( \mathcal{C}_i \subset \mathbb{R}^{n_i} \) and \( \{ n_i \}_{i \geq 1} \) is an increasing sequence of positive integers.

In the rest of this section, we list a sequence of formal definitions and some facts associated with these definitions.

**Definition 1 (Multiple Packing):** Let \( N > 0 \) and \( L \in \mathbb{Z}_{\geq 2} \). A subset \( \mathcal{C} \subset \mathbb{R}^N \) is called a \((N, L-1)\)-list-decodable code (a.k.a. an \((N, L-1)\)-multiple packing) if for every \( y \in \mathbb{R}^N \),
\[
|\mathcal{C} \cap B^N(y, \sqrt{nN})| \leq L - 1.
\]
The rate (a.k.a. density) of \( \mathcal{C} \) is defined as
\[
R(\mathcal{C}) := \liminf_{K \to \infty} \frac{1}{n} \frac{|\mathcal{C} \cap (KB)|}{|KB|},
\]
where \( B \) is an arbitrary centrally symmetric connected compact set in \( \mathbb{R}^n \) with nonempty interior.

**Remark 1:** The limit may not exist therefore \( \liminf \) is used. It is known [57] that with the choice \( KB = B(y, K) \) (where \( y \in \mathbb{R}^n \)), if the limit exists for one \( y \), it exists for all \( y \) and is equal for all \( y \). In that case, \( \mathcal{C} \) is said to have uniform density, and for such \( \mathcal{C} \) [57] showed that for every compact set \( B \) that is the closure of its interior and every \( y \), the density is equal to
\[
\lim_{K \to \infty} \frac{1}{n} \frac{|\mathcal{C} \cap (KB + y)|}{|KB|}.
\]

Though the limit in general may not exist, the upper density is always well-defined:
\[
\overline{R}(\mathcal{C}) := \lim_{K \to \infty} \sup_{y \in \mathbb{R}^n} \frac{1}{n} \ln \frac{|\mathcal{C} \cap (KB + y)|}{|KB|}.
\]

**Remark 2:** Common choices of \( B \) include the unit ball \( B^n \), the unit cube \([-1, 1]^n \), the fundamental Voronoi region \( V_{\Lambda} \) of a (full-rank) lattice \( \Lambda \subset \mathbb{R}^n \), etc. Since the limit is the same for different choices of \( B \) when \( \mathcal{C} \) has uniform density, we may choose the \( B \) based on convenience in proofs. Therefore, we do not fix the choice of \( B \) in Definition 1.

**Remark 3:** The rate of a packing (as per Equation (7)) is also called the normalized logarithmic density in the literature. It measures the normalized number of points per unit volume.

Note that the condition given by Equation (6) is equivalent to that for any \( (x_1, \ldots, x_L) \in \left( \frac{1}{L} \right)^L \),
\[
\bigcap_{i=1}^L B^n(x_i, \sqrt{nN}) = \emptyset.
\]

**Definition 2 (Chebyshev Radius and Average Squared Radius of a List):** Let \( x_1, \ldots, x_L \) be \( L \) points in \( \mathbb{R}^n \). Then the squared Chebyshev radius \( \overline{\text{rad}}^2(x_1, \ldots, x_L) \) of \( x_1, \ldots, x_L \) is defined as the (squared) radius of the smallest ball containing \( x_1, \ldots, x_L \), i.e.,
\[
\overline{\text{rad}}^2(x_1, \ldots, x_L) := \min_{y \in \mathbb{R}^n} \max_{i \in [L]} \| x_i - y \|^2.
\]
The average squared radius \( \overline{\text{rad}}^2(x_1, \ldots, x_L) \) of \( x_1, \ldots, x_L \) is defined as the average squared distance to the centroid, i.e.,
\[
\overline{\text{rad}}^2(x_1, \ldots, x_L) := \frac{1}{L} \sum_{i=1}^L \| x_i - \bar{x} \|^2,
\]
where \( \bar{x} := \frac{1}{L} \sum_{i=1}^L x_i \) is the centroid of \( x_1, \ldots, x_L \). We refer to the square root of the average squared radius as the average radius of the list.

**Remark 4:** One should note that for an \( L \)-list \( \mathcal{L} \) of points, the smallest ball containing \( \mathcal{L} \) is not necessarily the same as the circumscribed ball, i.e., the ball such that all points in \( \mathcal{L} \) live on the boundary of the ball. The circumscribed ball of the polytope \( \text{conv}(\mathcal{L}) \) spanned by the points in \( \mathcal{L} \) may not exist. If it does exist, it is not necessarily the smallest one containing \( \mathcal{L} \). However, whenever it exists, the smallest ball containing \( \mathcal{L} \) must be the circumscribed ball of a certain subset of \( \mathcal{L} \). (See, e.g., the Wikipedia page for circumscribed sphere [60].)

**Remark 5:** We remark that the motivation behind the definition of average squared radius (Equation (10)) is to replace...
the maximization in Equation (9) with average.

\[
\min_{y \in \mathbb{R}^n} \mathbb{E}_{i \sim [L]} \left[ ||x_i - y||_2^2 \right] \\
= \min_{y \in \mathbb{R}^n} \frac{1}{L} \sum_{i=1}^L \sum_{j=1}^n (x_i(j) - y(j))^2 \\
= \min_{(y_1, \ldots, y_n) \in \mathbb{R}^n} \frac{1}{L} \sum_{i=1}^L \sum_{j=1}^n (x_i(j) - y_j)^2 \\
= \sum_{j=1}^n \min_{y_j \in \mathbb{R}} \frac{1}{L} \sum_{i=1}^L (x_i(j) - y_j)^2 \\
= \frac{1}{L} \sum_{i=1}^L \sum_{j=1}^n (x_i(j) - y_j)^2 \\
= \frac{1}{L} \sum_{i=1}^L \|x_i - y\|^2.
\]

Equation (12) holds since the inner summation \(\frac{1}{L} \sum_{i=1}^L (x_i(j) - y(j))^2\) in Equation (11) only depends on \(y_j\) among all \(y_1, \ldots, y_n\). Equation (13) follows since for each \(j\), the minimizer of the minimization in Equation (12) is \(y_j^* = \frac{1}{L} \sum_{i=1}^L x_i(j)\). In Equation (14), the minimizer of \(y_j^* = \frac{1}{L} \sum_{i=1}^L x_i(j)\) equals \(\bar{x}_j = \frac{1}{L} \sum_{i=1}^L x_i(j)\).

**Definition 3 (Chebyshev Radius and Average Squared Radius of a Code):** Given a code \(C \subset \mathbb{R}^n\) of rate \(R\), the squared \((L-1)\)-list-decoding radius of \(C\) is defined as

\[
\text{rad}_x^2(C) := \min_{\mathcal{L}} \text{rad}^2(L).
\]

The \((L-1)\)-average squared radius of \(C\) is defined as

\[
\text{rad}^2_x(C) := \min_{\mathcal{L}} \text{rad}^2(L).
\]

**Definition 4 (Average-Radius Multiple Packing):** A subset \(C \subset \mathbb{R}^n\) is called an \((N, L-1)\)-average-radius list-decodable code (a.k.a. an \((N, L-1)\)-average-radius multiple packing) if \(\text{rad}_x^2(C) > nN\). The rate (a.k.a. density) \(R(C)\) of \(C\) is given by Equation (7). The \((N, L-1)\)-average-radius list-decoding capacity (a.k.a. \((N, L-1)\)-average-radius multiple packing density) is defined as

\[
\mathcal{C}_{L-1}(N) := \lim_{n \to \infty} \inf_{C \subset \mathbb{R}^n : \text{rad}^2_x(C) \geq nN} R(C).
\]

The squared \((L-1)\)-average-radius list-decoding radius at rate \(R\) (without input constraint) is defined as

\[
\text{rad}_x^2(R) := \lim_{n \to \infty} \sup_{C \subset \mathbb{R}^n : \text{rad}^2_x(C) \geq R} R(C).
\]

Note that \((L-1)\)-list-decodability defined by Equation (8) is equivalent to \(\text{rad}_x^2(C) > nN\). We also define the \((N, L-1)\)-list-decoding capacity (a.k.a. \((N, L-1)\)-multiple packing density) \(\mathcal{C}_{L-1}(N)\) and the \((L-1)\)-list-decoding radius \(\text{rad}_x^2(R)\) at rate \(R\):

\[
\mathcal{C}_{L-1}(N) := \lim_{n \to \infty} \sup_{C \subset \mathbb{R}^n : \text{rad}^2_x(C) \geq nN} R(C),
\]

\[
\text{rad}_x^2(R) := \lim_{n \to \infty} \sup_{C \subset \mathbb{R}^n : \text{rad}^2_x(C) \geq R} R(C).
\]

Since the average radius is at most the Chebyshev radius, average-radius list-decodability is stronger than regular list-decodability. Any lower (resp. upper) bound on \(\mathcal{C}_{L-1}(N)\) (resp. \(\mathcal{C}_{L-1}(N)\)) is automatically a lower (resp. upper) bound on \(\mathcal{C}_{L-1}(N)\) (resp. \(\mathcal{C}_{L-1}(N)\)). Proving upper/lower bounds on \(\mathcal{C}_{L-1}(N)\) (resp. \(\mathcal{C}_{L-1}(N)\)) is equivalent to proving upper/lower bounds on \(\text{rad}_x^2(R)\) (resp. \(\text{rad}_x^2(R)\)) for a given density.

**A. Different Notions of Density of Packings**

We measure the density of a packing using Equation (7). In the literature, there exists another commonly used notion of density for multiple packings. It measures the fraction of space occupied by the union of the balls of radius \(\sqrt{nN}\) centered around points in the packing. Specifically, for an \((N, L-1)\)-packing \(C \subset \mathbb{R}^n\),

\[
\Delta(C) := \lim_{P \to \infty} \inf \frac{1}{n} \ln \frac{1}{n} \left| \bigcup_{x \in \mathcal{C}} B^n(x, \sqrt{nN}) \cap B^n(\sqrt{nP}) \right|.
\]

We prove the following statement.

**Theorem 6:** Let \(N > 0\) and \(L \in \mathbb{Z}_{\geq 2}\). Let \(C \subset \mathbb{R}^n\) be an \((N, L-1)\)-multiple packing. Then \(\Delta(C) \leq R(C) + \frac{1}{2} \ln(2\pi eN)\).

**Proof:** Define an auxiliary density \(\tilde{\Delta}(C)\) of \(C \subset \mathbb{R}^n\) as

\[
\tilde{\Delta}(C) := \lim_{P \to \infty} \inf \frac{1}{n} \ln \frac{1}{n} \left| \bigcup_{x \in \mathcal{C}} B^n(x, \sqrt{nN}) \right|.
\]

The proof consists of two steps. We first show the desired relation for \(\Delta(C)\) vs. \(R(C)\). The we argue that \(\Delta(C) = \Delta(C)\).

We now proceed with the first step. Obviously,

\[
\left| \bigcup_{x \in \mathcal{C}} B^n(x, \sqrt{nN}) \right| \leq \left| \mathcal{C} \cap B^n(\sqrt{nP}) \right| \cdot \left| B^n(\sqrt{nP}) \right|,
\]

where the RHS assumes that all balls are disjoint. This implies

\[
\tilde{\Delta}(C) \leq \lim_{P \to \infty} \inf \frac{1}{n} \ln \frac{1}{n} \left| \mathcal{C} \cap B^n(\sqrt{nP}) \right| \cdot \left| B^n(\sqrt{nP}) \right|,
\]

\[
= R(C) + \frac{1}{n} \ln \left| B^n(\sqrt{nP}) \right| \geq R(C) + \frac{1}{2} \ln(2\pi eN).
\]

On the other hand, since each point is covered by at most \(L-1\) balls, we have

\[
\left| \bigcup_{x \in \mathcal{C}} B^n(x, \sqrt{nN}) \right| \geq \frac{1}{L-1} \left| \mathcal{C} \cap B^n(\sqrt{nP}) \right| \cdot \left| B^n(\sqrt{nP}) \right|,
\]

\[
= R(C) + \frac{1}{n} \ln \left| B^n(\sqrt{nP}) \right| \geq R(C) + \frac{1}{2} \ln(2\pi eN).
\]
which, by similar calculations as above, implies
\[ \Delta(C) \geq R(C) + \frac{1}{2} \ln(2\pi eN) - o(1). \]  
(19)

Combining Equations (18), and (19), we have that for any 
\((N, L - 1)\)-packing \(C\),
\[ \Delta(C) \leq R(C) + \frac{1}{2} \ln(2\pi eN). \]

In the second step, we show \( \Delta(C) = \Delta(C) \) which suffices to 
conclude the theorem. Observe
\[
\begin{align*}
\Delta(C) & = \lim_{n \to \infty} \frac{1}{n} \ln \left| \bigcup_{\hat{z} \in C \cap B_n(\sqrt{np})} B_n(\hat{z}, \sqrt{nN}) \cap B_n(\sqrt{nP} - \sqrt{nN}) \right| \\
& \leq \lim_{n \to \infty} \frac{1}{n} \ln \left| \bigcup_{\hat{z} \in C \cap B_n(\sqrt{np})} B_n(\hat{z}, \sqrt{nN}) \right|. 
\end{align*}
\]

Normalizing both sides by \( \lim \inf_{n \to \infty} \frac{1}{n} \ln \left| B_n(\sqrt{np}) \right| \) and using 
the definitions of \( \Delta(C) \) and \( \Delta(C) \), we have the LHS
\[
\begin{align*}
\lim_{n \to \infty} \frac{1}{n} \ln \left| \bigcup_{\hat{z} \in C \cap B_n(\sqrt{np})} B_n(\hat{z}, \sqrt{nN}) \cap B_n(\sqrt{nP} - \sqrt{nN}) \right| \\
= \lim_{n \to \infty} \left( \frac{1}{n} \ln \left| B_n(\sqrt{nP} - \sqrt{nN}) \right| \\
+ \frac{1}{n} \left| \bigcup_{\hat{z} \in C} B_n(\hat{z}, \sqrt{nN}) \cap B_n(\sqrt{nP} - \sqrt{nN}) \right| \right) \\
= \lim_{n \to \infty} \frac{1}{n} \ln \left( \frac{\sqrt{nP} - \sqrt{nN}}{\sqrt{nN}} \right)^n \\
+ \lim_{n \to \infty} \frac{1}{n} \left| \bigcup_{\hat{z} \in C} B_n(\hat{z}, \sqrt{nN}) \cap B_n(\sqrt{nP} - \sqrt{nN}) \right| \\
= \Delta(C),
\end{align*}
\]

and the RHS \( \Delta(C) \). Hence \( \Delta(C) \leq \Delta(C) \). The other direction 
follows similarly:
\[
\begin{align*}
\Delta(C) & = \lim_{n \to \infty} \frac{1}{n} \ln \left| \bigcup_{\hat{z} \in C \cap B_n(\sqrt{np})} B_n(\hat{z}, \sqrt{nN}) \right| \\
& \leq \lim_{n \to \infty} \frac{1}{n} \ln \left( \bigcup_{\hat{z} \in C \cap B_n(\sqrt{np})} B_n(\hat{z}, \sqrt{nN}) \cap B_n(\sqrt{nP} + \sqrt{nN}) \right) \\
& \leq \Delta(C).
\end{align*}
\]

which implies
\[
\Delta(C) \leq \lim_{n \to \infty} \frac{1}{n} \ln \left| B_n(\sqrt{np}) \right| \\
+ \lim_{n \to \infty} \frac{1}{n} \ln \left| \bigcup_{\hat{z} \in C} B_n(\hat{z}, \sqrt{nN}) \cap B_n(\sqrt{nP} + \sqrt{nN}) \right| \\
= \Delta(C).
\]

This finishes the proof of the theorem. \( \square \)

B. Chebyshev Radius and Average Radius

In this section, we present several different representations of the 
Chebyshev radius and average squared radius. Some of 
them will be crucially used in the subsequent sections of this 
paper. These representations are summarized in the following 
theorern which will be proved in the subsequent subsections.

**Theorem 7:** Let \( L \in \mathbb{Z}_{\geq 2} \) and \( \hat{z}_1, \ldots, \hat{z}_L \in \mathbb{R}^n \). Then the 
squared Chebyshev radius of \( \hat{z}_1, \ldots, \hat{z}_L \) admits the following 
alternative representations:

1) \( \text{rad}^2(\hat{z}_1, \ldots, \hat{z}_L) = \max_{\hat{z} \in B_L} \sum_{i=1}^L \| \hat{z}_i - \hat{z} \|_2^2 \),

where \( \hat{z} := \sum_{i=1}^L \hat{z}_i / L \) and \( B_L \) denotes the 
probability simplex on \( [L] \);

2) \( \text{rad}^2(\hat{z}_1, \ldots, \hat{z}_L) = \lim_{p \to \infty} \text{rad}^{(p)}(\hat{z}_1, \ldots, \hat{z}_L) \),

where \( \text{rad}^{(p)}(\hat{z}_1, \ldots, \hat{z}_L) := \left( \frac{1}{L} \sum_{i=1}^L \| \hat{z}_i - \hat{z} \|_p^2 \right)^{1/p} \);

3) there exists a unique \( 1 < q < \infty \) depending on \( \hat{z}_1, \ldots, \hat{z}_L \) such that

\[
\text{rad}^2(\hat{z}_1, \ldots, \hat{z}_L) = \left( \frac{1}{L} \sum_{i=1}^L \| \hat{z}_i - \hat{z} \|_q^2 \right)^{1/q},
\]

and \( \hat{z} := \frac{1}{L} \sum_{i=1}^L \hat{z}_i \).

The average squared radius of \( \hat{z}_1, \ldots, \hat{z}_L \) admits the following 
alternative representations:

1) \( \text{rad}^2(\hat{z}_1, \ldots, \hat{z}_L) = \frac{1}{L} \sum_{i=1}^L \| \hat{z}_i \|_2^2 - \| \hat{z} \|_2^2 \);

2) \( \text{rad}^2(\hat{z}_1, \ldots, \hat{z}_L) = \frac{L-1}{L^2} \sum_{i=1}^L \| \hat{z}_i \|_2^2 - \frac{1}{L} \sum_{(i,j) \in [L]^2, i \neq j} \langle \hat{z}_i, \hat{z}_j \rangle ;

3) \( \text{rad}^2(\hat{z}_1, \ldots, \hat{z}_L) = \frac{1}{L^2} \sum_{(i,j) \in [L]^2, i \neq j} \| \hat{z}_i - \hat{z}_j \|_2^2 \).

1) Another Representation of the Chebyshev Radius: The 
Chebyshev radius involves a minimax expression which is in 
general tricky to handle. One can use minimax theorem to 
interchange the min and max and then compute the inner min
explicitly,

\[
\text{rad}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) = \min_{y \in \mathbb{R}^n} \max_{\mathbf{z} \in \Delta_L} \|\mathbf{z} - y\|^2_2
\]

\[
= \min_{y \in \mathbb{R}^n} \max_{\mathbf{z} \in \Delta_L} \sum_{i=1}^L \mathbb{E}[(\mathbf{z}(i) - y(i))^2]
\]

The last equality follows since the maximum is always achieved by a singleton \( \mathbb{E} \in \{0, 1\}^L \). Note that the objective function on the RHS is linear (hence concave) in \( \mathbb{E} \) and quadratic (hence convex) in \( y \). Therefore the max and min can be interchanged and we get

\[
\max_{\mathbf{z} \in \Delta_L} \min_{y \in \mathbb{R}^n} \sum_{i=1}^L \mathbb{E}[(\mathbf{z}(i) - y(i))^2]
\]

\[
= \max_{\mathbf{z} \in \Delta_L} \sum_{j=1}^n \sum_{i=1}^L \mathbb{E}[(\mathbf{z}(i) - y(j))^2]
\]

\[
= \max_{\mathbf{z} \in \Delta_L} \sum_{j=1}^n \min_{y \in \mathbb{R}^n} \sum_{i=1}^L \mathbb{E}[(\mathbf{z}(i) - y(j))^2]
\]

The last equality follows since each inner summation \( \sum_{i=1}^L \mathbb{E}[(\mathbf{z}(i) - y(j))^2] \) only depends on \( y(j) \) among all \( y(1), \ldots, y(n) \). For each \( j \), the minimizing \( y_j^* \) equals

\[
y_j^* := \frac{\sum_{i=1}^L \mathbb{E}(x_i(j))}{\sum_{i=1}^L \mathbb{E}(x_i)} = \sum_{i=1}^L \mathbb{E}(x_i(j)),
\]

where the last equality is because \( \mathbb{E} \in \Delta_L \). Therefore

\[
\text{rad}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) = \max_{\mathbf{z} \in \Delta_L} \sum_{i=1}^L \mathbb{E}[(\mathbf{z}(i) - y_j^*)]^2
\]

where \( y_j^* := \sum_{i=1}^L \mathbb{E}(x_i(j)) \).

2) Higher-Order Approximations to the Chebyshev Radius:

As explained in Remark 5, the average squared radius is a linear relaxation of the squared Chebyshev radius:

\[
\text{rad}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) = \min_{y \in \mathbb{R}^n} \mathbb{E} \left[ \|\mathbf{z} - y\|^2_2 \right]
\]

One may obtain better and better approximations to the squared Chebyshev radius by taking higher and higher order relaxations:

\[
\text{rad}^{(p)}(\mathbf{x}_1, \ldots, \mathbf{x}_L) = \min_{y \in \mathbb{R}^n} \left( \mathbb{E} \left[ \|\mathbf{z} - y\|^{2p}_2 \right] \right)^{1/p}
\]

\[
= \left( \min_{y \in \mathbb{R}^n} \mathbb{E} \left[ \|\mathbf{z} - y\|^{2p}_2 \right] \right)^{1/p}, \quad (22)
\]

where \( p \geq 1 \). The second equality in Equation (22) follows since the \( f(\cdot) = (\cdot)^{1/p} \) is monotone increasing. Note that \( \text{rad}^{(1)}(\mathbf{x}_1, \ldots, \mathbf{x}_L) = \text{rad}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) \). Moreover, since \( \mathbb{E}[(\cdot)^{p}/p] \) is increasing in \( p \), we have

\[
\text{rad}^{(p)}(\mathbf{x}_1, \ldots, \mathbf{x}_L) \leq \text{rad}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L).
\]

However, we do not know how to analyze \( \text{rad}^{(p)} \). It seems difficult to get a closed-form solution of the minimization since the minimizer \( y_j^* \) cannot be obtained by minimizing over \( y(j) \) for different \( j \) separately.

3) More Representations of the Average Squared Radius:

Recall that Equation (21), as a lower bound on \( \text{rad}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) \), admits an explicit formula given by Equation (14):

\[
\text{rad}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) \leq \frac{1}{L} \sum_{i=1}^L \|\mathbf{z}_i - \bar{\mathbf{x}}\|^2_2, \quad (23)
\]

where \( \bar{\mathbf{x}} := \frac{1}{L} \sum_{i=1}^L \mathbf{z}_i \) denotes the centroid of \( \mathbf{z}_1, \ldots, \mathbf{z}_L \). On the other hand, we have

\[
\text{rad}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) = \min_{y \in \mathbb{R}^n} \mathbb{E} \left[ \|\mathbf{z} - y\|^2_2 \right]
\]

\[
= \max_{\mathbf{z} \in \mathbb{R}^n} \sum_{i=1}^L \mathbb{E}[(\mathbf{z}(i) - y(i))^2]
\]

\[
\leq \max_{\mathbf{z} \in \mathbb{R}^n} \sum_{i=1}^L \mathbb{E}[(\mathbf{z}(i) - y(j))^2]
\]

\[
=: \text{rad}^2_{\max}(\mathbf{x}_1, \ldots, \mathbf{x}_L). \quad (24)
\]

Contrasting Equations (23) and (24), by monotonicity and continuity of \( \|\cdot\|_p \) in \( p \), we know that there exists \( 1 \leq p \leq \infty \) such that

\[
\text{rad}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) = \left( \frac{1}{L} \sum_{i=1}^L \|\mathbf{z}_i - \bar{\mathbf{x}}\|^{2p}_2 \right)^{1/p}.
\]

However, we do not know how to use the above observation for the following two reasons. Firstly, the above expression seems tricky to handle. Secondly and more importantly, the number \( p \) depends on \( x_1, \ldots, x_L \) and is typically different for different lists.

Finally, we provide several alternative expressions for \( \text{rad}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) \) which will be useful in the proceeding sections of this paper.

\[
\text{rad}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) = \frac{1}{L} \sum_{i=1}^L \|\mathbf{z}_i - \bar{\mathbf{x}}\|^{2p}_2
\]

\[
= \frac{1}{L} \sum_{i=1}^L \|\mathbf{z}_i - \bar{\mathbf{x}} - \bar{\mathbf{x}}_i - \bar{\mathbf{x}}\|
\]

\[
= \frac{1}{L} \sum_{i=1}^L \left( \|\mathbf{z}_i\|^2 - 2\langle \mathbf{z}_i, \bar{\mathbf{x}} \rangle + \|\bar{\mathbf{x}}\|^2 \right)
\]

\[
= \frac{1}{L} \sum_{i=1}^L \|\mathbf{z}_i\|^2 - 2\langle \mathbf{z}_i, \bar{\mathbf{x}} \rangle + \|\bar{\mathbf{x}}\|^2 = \frac{1}{L} \sum_{i=1}^L \|\mathbf{z}_i\|^2 - \|\bar{\mathbf{x}}\|^2.
\]

The above expression can be further written as

\[
\text{rad}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) = \frac{1}{L} \sum_{i=1}^L \|\mathbf{z}_i\|^2_2 - \|\bar{\mathbf{x}}\|^2_2
\]

\[
= \frac{1}{L} \sum_{i=1}^L \|\mathbf{z}_i\|^2_2 - \frac{1}{L^2} \sum_{(i,j) \in [L]^2} \langle \mathbf{z}_i, \mathbf{z}_j \rangle
\]

\[
= \frac{1}{L} \sum_{i=1}^L \|\mathbf{z}_i\|^2_2 - \frac{1}{L^2} \sum_{i=1}^L \|\mathbf{z}_i\|^2_2 - \frac{1}{L^2} \sum_{(i,j) \in [L]^2, i \neq j} \langle \mathbf{z}_i, \mathbf{z}_j \rangle \]
At last, Equation (26) can in turn be rewritten as
\[
\overline{\text{rad}}^2(\mathbf{z}_1, \ldots, \mathbf{z}_L) = \frac{L - 1}{L^2} \sum_{i=1}^{L} \left\| \mathbf{z}_i \right\|^2 - \frac{1}{L^2} \sum_{(i,j) \in [L]^2; i \neq j} \langle \mathbf{z}_i, \mathbf{z}_j \rangle.
\]  
(26)

To analyze average-radius list-decodability of \( C \), we first construct an average-radius list-decodable code \( C_K \) supported within \( A = \mathcal{I}^n \) where \( \mathcal{I} := [-K, K] \) is a sufficiently large interval for some \( K > 0 \). We later tile this codebook over \( \mathbb{R}^n \) to obtain an infinite constellation having the same average squared radius as the finite codebook.

The finite codebook \( C_K \) is obtained by drawing \( M := \lambda_n |A| \) points independently and uniformly at random from \( A \) and expurgating the resulting codebook. Let \( C_K := \{ \mathbf{x}_1, \ldots, \mathbf{x}_M \} \) denote the \( M \) independent points uniformly distributed over \( A \).

\textbf{Lemma 9}: There exists a finite codebook \( C_K \) supported over \( A \), having minimum average squared radius at least \( \sqrt{nN} \) and density
\[
\frac{1}{n} \ln \frac{|C_K|}{|A|} \geq \frac{1}{2} \ln \frac{L - 1}{2\pi e NL} - \frac{1}{2} (L - 1) \ln L + o(1).
\]

The first step is to bound
\[
\Pr \left[ \overline{\text{rad}}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) \leq nN \right].
\]  
(31)

for every subset of \( L \) codewords in \( C_K \). In fact, we will prove the following lemma.

\textbf{Lemma 10}: For any \( \mathbf{x}_1, \ldots, \mathbf{x}_L \) drawn independently and uniformly at random from \( A \), we have
\[
\Pr \left[ \overline{\text{rad}}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) \leq nN \right] = e^{nE(K)+o(n)},
\]

where the \( o(n) \) term is independent of \( K \), and
\[
E(K) := \frac{L - 1}{2} \ln \frac{L - 1}{2\pi e NL} - \frac{1}{2} (L - 1) \ln (L - 1).
\]

First, we note that Equation (31) can be alternatively written as
\[
\Pr \left[ \overline{\text{rad}}^2(\mathbf{x}_1, \ldots, \mathbf{x}_L) \leq nN \right] = \Pr \left[ \frac{1}{L} \sum_{i=1}^{L} \| \mathbf{x}_i \|^2 \leq nN \right].
\]  
(32)

\textbf{Remark 6}: Note that the above bound (Equation (30)) approaches \( \frac{1}{2} \ln \frac{1}{2\pi e} \) as \( L \to \infty \). The latter quantity is known to be the list-decoding capacity for asymptotically large \( L \) (see Section IV). On the extreme, when \( L = 2 \), the above bound becomes \( \frac{1}{L} \ln \frac{1}{2\pi e} \), which recovers the best known bound due to Minkowski \cite{Minkowski}.

To prove the above theorem, let \( R < \frac{1}{2} \ln \frac{L - 1}{2\pi e NL} - \frac{1}{2} (L - 1) \ln L \) and \( \lambda_n \equiv e^{nR} \). The exact choice of \( \lambda_n \) is given by Equation (51).
where $\mathbf{x}_i(j) \stackrel{i.i.d.}{\sim} \text{Unif}(\mathcal{I})$ for each $i \in [L]$ and $j \in [n]$.

We note that the function $g(\vec{t})$ is a quadratic form of $\vec{t} \in \mathbb{R}^L$.

Indeed,
\[
g(\vec{t}) = \sum_{i=1}^{L} \vec{t}(i)^2 - \frac{1}{L} \left( \sum_{i=1}^{L} \vec{t}(i) \right)^2
= \left(1 - \frac{1}{L}\right) \sum_{i=1}^{L} \vec{t}(i)^2 - \frac{2}{L} \sum_{i,j \in [L], i < j} \vec{t}(i)\vec{t}(j)
= \vec{t}^\top A\vec{t},
\]

where
\[
A := I_L - \frac{1}{L} J_L \in \mathbb{R}^{L \times L}
\]
and $I_L$ denotes the $L \times L$ identity matrix and $J_L$ denotes the $L \times L$ all-one matrix. Therefore we can write Equation (31) as
\[
\Pr \left[ \sum_{j=1}^{n} \mathbf{x}_j^\top A\mathbf{x}_j \leq LnN \right],
\]

where $\mathbf{x}_j := [\mathbf{x}_1(j), \cdots, \mathbf{x}_L(j)] \in \mathbb{R}^L$ and $\mathbf{x}_j(j) \stackrel{i.i.d.}{\sim} \text{Unif}(\mathcal{I})$ for each $1 \leq i \leq L$ and $1 \leq j \leq n$.

### A. Large Deviation Principle

Since $\mathbf{x}_j^\top A\mathbf{x}_j$ is independent for each $1 \leq j \leq n$, we can apply the large deviation principle (Theorem 5) to get the asymptotic behaviour of Equation (36). Specifically,
\[
\frac{1}{n} \ln \left[ \text{Equation (36)} \right] \to -\max_{\lambda \geq 0} \left\{ \lambda LN - \ln \mathbb{E} \left[ e^{\lambda \mathbf{x}^\top A\mathbf{x}} \right] \right\}
= -\max_{\lambda \geq 0} \left\{ -\lambda LN - \ln \mathbb{E} \left[ e^{-\lambda \mathbf{x}^\top A\mathbf{x}} \right] \right\},
\]

where $\mathbf{x} \sim \text{Unif}^{\otimes L}(\mathcal{I})$.

We need to compute the following integral:
\[
\mathbb{E} \left[ e^{-\lambda \mathbf{x}^\top A\mathbf{x}} \right] = \frac{1}{(2K)^L} \int_{\mathcal{I}^L} e^{-\lambda \mathbf{x}^\top A\mathbf{x}} d\mathbf{x}
= \frac{1}{(2K)^L} \int_{[-1,1]^L} e^{-\lambda K^2 \vec{t}^\top A\vec{t}} K^L d\vec{t}
= \frac{1}{2^L} \int_{[-1,1]^L} e^{-\lambda K^2 \vec{t}^\top A\vec{t}} d\vec{t},
\]

where $\lambda \geq 0$.

Note that $A \in \mathbb{R}^{L \times L}$ has rank $L-1$ and therefore is singular.

In fact, $A$ has eigendecomposition $A = PDP^{-1}$ where $D := \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{L \times L}$ consists of the eigenvalues of $A$ as its diagonal entries. However, $P$ is not orthogonal. One can orthogonalize it using the Gram–Schmidt process which gives us an orthogonal matrix $U \in \mathbb{R}^{L \times L}$. We claim that $U$ is given by Equation (39), shown at the bottom of the next page.

The above $U$ gives us the Singular Value Decomposition of $A$ which is $A = UDU^\top$. Note that $U^\top = U^{-1}$ by orthogonality of $U$ and the diagonalization of $A$ is given by $D = U^{-1}AU^{-\top} = U^\top AU$. Under the change of variable $\vec{t} = U\vec{y}$, the quadratic form $\vec{t}^\top A\vec{t}$ becomes a diagonal form $\vec{y}^\top D\vec{y}$ and the RHS of Equation (38) becomes
\[
\frac{1}{2^L} \int_{[-1,1]^L} \exp \left( -K^2 \lambda \vec{t}^\top A\vec{t} \right) d\vec{t}
= \frac{1}{2^L} \int_{U[-1,1]^L} \exp \left( -K^2 \lambda (U\vec{y})^\top A(U\vec{y}) \right) \cdot |\det(U)| d\vec{y}
= \frac{1}{2^L} \int_{U[-1,1]^L} \exp \left( -K^2 \lambda \vec{y}^\top (U^\top AU) \vec{y} \right) d\vec{y}
= \frac{1}{2^L} \int_{U[-1,1]^L} \exp \left( -K^2 \lambda \sum_{i=1}^{L-1} \vec{t}(i)^2 \right) d\vec{t},
\]

Equation (40) is by Lemma 3. In Equation (41), we use the facts that $U^{-1} = U^\top$ and $|\det(U)| = 1$.

### B. Laplace’s Method and Proof of Lemma 10

To compute Equation (42), we note that the integral is degenerate along the direction of the last coordinate $\vec{t}(L)$. Since the integral domain is bounded, the integral is still finite. We first integrate out $\vec{t}(L)$ and get an $(L-1)$-dimensional integral w.r.t. $\vec{t}(1), \cdots, \vec{t}(L-1)$. To this end, observe that for $\vec{t} \in U[-1,1]^L$, the last component $\vec{t}(L)$ is a function of $\vec{t}(1), \cdots, \vec{t}(L-1)$ and it can take any value of the last coordinate of $U[-1,1]^L$. Therefore the range of $\vec{t}(L)$ can be written as $[g_1(\vec{t}(1), \cdots, \vec{t}(L-1)), g_2(\vec{t}(1), \cdots, \vec{t}(L-1))]$ where $g_1(\cdot)$ and $g_2(\cdot)$ are piecewise linear continuous functions given by $U$. We now integrate out $\vec{t}(L)$ and get
\[
\int_{U[-1,1]^L} \int_{[g_1(\vec{t}(1), \cdots, \vec{t}(L-1)), g_2(\vec{t}(1), \cdots, \vec{t}(L-1))] \cap \{ \vec{t}(L) \}} e^{-K^2 \lambda \sum_{i=1}^{L-1} \vec{t}(i)^2} (g_2(\vec{t}(1), \cdots, \vec{t}(L-1)) - g_1(\vec{t}(1), \cdots, \vec{t}(L-1))) d\vec{t}(1), \cdots, d\vec{t}(L-1),
\]

where $(U[-1,1]^L)_{\cap \{ \vec{t}(L) \}} \subset \mathbb{R}^{L-1}$ denotes the set obtained by restricting each vector in $U[-1,1]^L \subset \mathbb{R}^L$ to the first $L-1$ coordinates $(\vec{t}(1), \cdots, \vec{t}(L-1))$. Note that the quadratic function $f(\vec{t}(1), \cdots, \vec{t}(L-1)) := \lambda \sum_{i=1}^{L-1} \vec{t}_i^2$ is nonnegative and attains its unique minimum (which is zero) at $[\vec{t}(1), \cdots, \vec{t}(L-1)] = [0, \cdots, 0]$ which is in
the interior of $U^T [-1,1]^L$. Therefore, by Laplace’s method (Theorem 4), Equation (43) converges to
\[
\left(\frac{2\pi}{K^2}\right)^{\frac{L-1}{2}} \frac{g_2(0,\ldots,0) - g_1(0,\ldots,0)}{\det((\text{Hess } f)(0,\ldots,0))}
\]
as $K \to \infty$. Since $f = 2M_{L-1} > 0$, we have
\[
\left(\frac{2\pi}{K^2}\right)^{\frac{L-1}{2}} \frac{g_2(0,\ldots,0) - g_1(0,\ldots,0)}{\sqrt{(2\lambda)^{L-1}}}
\] {(K^2\lambda)}^{\frac{L-1}{2}} (g_2(0,\ldots,0) - g_1(0,\ldots,0)).
\]
Note that $g_2(0,\ldots,0) - g_1(0,\ldots,0)$ is nothing but the length of the range of the last coordinate $t_{L-1}$ of vectors in $U^T [-1,1]^L$. Since any vector $\bar{f} \in U^T [-1,1]^L$ can be written as $U^T \bar{u}$ for some $\bar{u} \in [-1,1]^L$, the length of the range of the last coordinate of $\bar{f}$ is twice the $\ell_2$-norm of the last row of $U^T$, i.e., the last column of $U$. From Equation (39), it is not hard to see that
\[
g_2(0,\ldots,0) - g_1(0,\ldots,0) = 2 \cdot L \cdot 1 = 2 \sqrt{L}. \tag{44}
\]
Finally, we get that Equation (43) (asymptotically) equals
\[
\left(\frac{\pi}{K^2\lambda}\right)^{\frac{L-1}{2}} \cdot 2\sqrt{L}. \tag{45}
\]
Recall that $E(K)$ is the error exponent corresponding to Equation (31). Plugging Equation (45) back to Equation (42) and then back to Equation (37), we have
\[
\begin{align*}
E(K) &= \max_{\lambda \geq 0} \left\{ -\lambda LN - \ln \left( \frac{2\pi}{K^2\lambda} \right)^{\frac{L-1}{2}} \cdot 2\sqrt{L} \right\} \\
&= \max_{\lambda \geq 0} \left\{ -\lambda LN - \ln \left( \left(\frac{\pi}{2}\right)^{\frac{L-1}{2}} \sqrt{L} \right) \right\} \\
&= \max_{\lambda \geq 0} \left\{ -LN\lambda + L - \frac{1}{2} \ln \lambda + (L-1) \ln(2K) \right. \\
&\quad \left. - \frac{L-1}{2} \ln \pi - \frac{1}{2} \ln L \right\} \\
&= \max_{\lambda \geq 0} \left\{ -LN\lambda + L - \frac{1}{2} \ln \lambda + (L-1) \ln(2K) \right\} \\
&= \max_{\lambda \geq 0} \left\{ -LN\lambda + L - \frac{1}{2} \ln \lambda + (L-1) \ln(2K) \right\}
\end{align*}
\]

\begin{bmatrix}
-\frac{1}{\sqrt{4 \times 2}} & -\frac{1}{\sqrt{4 \times 3}} & -\frac{1}{\sqrt{4 \times 4}} & \cdots & -\frac{1}{\sqrt{(L-2) \times (L-1)}} & -\frac{1}{\sqrt{(L-1) \times L}} & 1 \\
\sqrt{\frac{2}{5}} & \sqrt{\frac{3}{5}} & -\frac{1}{\sqrt{4 \times 4}} & : & \cdots & : & : \\
\sqrt{\frac{2}{3}} & \sqrt{\frac{3}{4}} & -\frac{1}{\sqrt{4 \times 5}} & : & \cdots & : & : \\
\sqrt{\frac{2}{5}} & -\frac{1}{\sqrt{4 \times 3}} & -\frac{1}{\sqrt{4 \times 4}} & \cdots & -\frac{1}{\sqrt{(L-2) \times (L-1)}} & -\frac{1}{\sqrt{(L-1) \times L}} & 1 \\
\sqrt{\frac{2}{3}} & -\frac{1}{\sqrt{4 \times 3}} & -\frac{1}{\sqrt{4 \times 4}} & \cdots & -\frac{1}{\sqrt{(L-2) \times (L-1)}} & -\frac{1}{\sqrt{(L-1) \times L}} & 1 \\
\end{bmatrix} \in \mathbb{R}^{L \times L}. \tag{39}
\]
After expurgating out one codeword from each bad list, we get an \((N, L - 1)\)-average-radius multiple packing \(C_K\) of size at least \(\frac{1}{2} \mathbb{E}[M]\) and the density \(\frac{1}{n} \ln \frac{|C_K|}{|A|}\) is therefore at least

\[
\frac{1}{n} \ln \frac{\frac{1}{2} \mathbb{E}[M]}{|A|} = \frac{1}{n} \ln \frac{\frac{1}{2} N}{|A|} = \frac{1}{n} \ln \frac{\lambda_n}{2}.
\]

(52)

Substituting Equation (51) here, we get the following lower bound on the density

\[
\frac{1}{n} \ln \frac{L - 1}{2\pi eNL} - \frac{1}{2(L - 1)} \ln L + o(1),
\]

(53)

as promised in Lemma 9.

D. Unbounded Packing and Proof of Theorem 8

The above derivation shows the existence of a finite codebook \(C \cap [-K, K]^n\) in which all \(L\)-tuple of points have radius at least \(\sqrt{nN}\). To obtain an unbounded \((N, L - 1)\)-packing, let us take \(K = n^2\) and define \(C_n := C \cap [-n^2, n^2]^n\) and

\[
C_\infty := C_n + (n^2 + n^{0.6})\mathbb{Z}^n.
\]

In words, \(C_\infty\) is obtained by tiling \(\mathbb{R}^n\) using translations of \(C_n\) and leaving a gap of width \(2 \cdot n^{0.6}\) between adjacent copies of \(C_n\). The NLD of \(C_\infty\) is essentially the same as that of \(C \cap [-n^2, n^2]^n\) which is given by Equation (53). Indeed, since \(C_\infty\) is periodic, we have

\[
R(C_\infty) = \frac{1}{n} \ln \frac{|C_n|}{|[-(n^2 + n^{0.6}), (n^2 + n^{0.6})]|} = \frac{1}{n} \ln \frac{|C_n|}{|[-n^2, n^2]|} + \frac{1}{n} \ln \frac{(2n^2)^n}{(2(n^2 + n^{0.6}))^n} \to \infty R(C_n).
\]

Moreover, we claim that \(C_\infty\) is an \((N, L - 1)\)-packing. To see this, take any \(L \subset C_\infty\). If \(L \subset C_n + \frac{z}{n}\) for some \(z \in (n^2 + n^{0.6})\mathbb{Z}^n\), then \(\text{rad}(L) \geq \sqrt{nN}\) by the guarantee of \(C_n\). Otherwise, there exist two points \(x_1, x_2 \in L\) such that \(x_1 \in C_n + \frac{z_1}{n}\) and \(x_2 \in C_n + \frac{z_2}{n}\) for two distinct \(z_1 \neq z_2 \in (n^2 + n^{0.6})\mathbb{Z}^n\). Then

\[
\text{rad}(L) \geq \frac{1}{2} \|x_1 - x_2\|_\infty \geq n^{0.6} \geq \sqrt{nN}\.
\]

Therefore, we obtain an \((N, L - 1)\)-packing \(C_\infty\) of NLD asymptotically equal to Equation (53). The proof of Theorem 8 is complete.

E. Alternate Approaches for Bounding Equation (36)

We managed to compute the exact asymptotics (up to lower order terms in the exponent) of the tail probability given by Equation (36). The way we did so is by applying the large deviation principle and analyzing the moment generating function of the random quadratic form of interest. From the perspective of concentration of measure, the tail probability we computed could potentially be bounded using different approaches:

1. Gaussian integral w.r.t. a general (non-identity) degenerate covariance matrix
2. Concentration of the uniform measure on a solid cube (a useful trick for which is to push it forward to the Gaussian measure and applying Lipschitz concentration [62]);
3. The (standard) Gaussian measure of a parallelepiped defined by the linear transformation \(U\);
4. The probability that a Gaussian (with zero mean and general covariance matrix) lies in a cube;
5. Hanson–Wright inequality for quadratic forms in subgaussian random vectors [63] which, in our case, are uniform vectors in a solid cube.

We tried all the above techniques. However, they do not seem to yield the correct exponent, at least in their vanilla forms, though they may give certain exponentially decaying bounds.

Therefore, we feel that Equation (36) is an interesting example for which standard concentration tools do not directly yield the optimal bound.

F. Connections to [14]

The paper [14] analyzed the list-decodability of expurgated PPPs and arrived at the same bound (Equation (30)) as ours, and in fact the current paper was inspired by [14].

However, there were some gaps in the proof of [14] that we were not able to resolve. In the paper, it was shown that for every sufficiently large \(K > 0\), there exists an (infinite) codebook \(C\) obtained by expurgating a PPP such that every \(L\)-tuple of points in \(C \cap [-K, K]^n\) has radius at least \(\sqrt{nN}\). However, we could not find a rigorous way to pass to the limit as \(K \to \infty\) and argue that \(C\) itself is an infinite point set is an \((N, L - 1)\)-multiple packing. Indeed, it is claimed in [64, Note 5] that \(C \cap [-K, K]^n\) may not converge as \(K \to \infty\).

Although our proof also involves analyzing the tail probability of the average squared radius (Equation (31)), our techniques are different, as outlined below.

Let \(C\) be a PPP (without expurgation yet) and \(K \in \mathbb{R}\). Let \(x_1, \ldots, x_L \in C \cap [-K, K]^n\) be an \(L\)-list. Recall that they are independent and uniformly distributed in \([-K, K]^n\). Let \(x = \frac{1}{L} \sum_{i=1}^L x_i\) denote the centroid of the list. To compute Equation (31), [14] claimed that we could use an orthogonal transformation \(x_i \mapsto u_i, (1 \leq i \leq L)\) to the list so that \(\frac{1}{L} \|u_1\|_2 = \|u\|_2\), where \(u := \frac{1}{L} \sum_{i=1}^L u_i\). This is in contrast to our approach. However, it should be noted that an orthogonal transformation only reflects and/or rotates the list, but does not translate it.

From our understanding of the paper [14], the ideas can be interpreted as follows. Since the average squared radius is invariant under rigid transformations (i.e., translations, rotations, reflections and their combination) and a homogeneous PPP is stationary and isotropic, [14] attempts to transform the list rigidly so that the resulting average squared radius admits a simpler expression and the list is still independent and uniformly distributed in the cube. However, it appears that such a rigid transformation does not exist. We instead use a different, much simpler approach by first constructing a finite codebook and then tiling this. The high-level construction is similar to [5] which was originally studied for the problem.
of reliable communication over additive-white Gaussian noise channels.

XI. Open Questions

We end the paper with several intriguing open questions.

1) The problem of packing spheres in $\ell_p$ space was also addressed in the literature [65], [66], [67], [68]. Recently, there was an exponential improvement on the optimal packing density in $\ell_2$ space [69] relying on the Kabatiansky–Levenshtein bound [13]. It is worth exploring the $\ell_p$ version of the multiple packing problem. One obstacle here is that the $\ell_p$ average radius does not admit a closed form expression unlike the $p = 2$ case.

2) In this paper, we treat (regular) list-decoding and average-radius list-decoding as two different notions and obtain bounds for the latter (which automatically lower bounds the former). It follows from our bounds that the largest multiple packing density under these two notions coincide as $L \to \infty$.

However, as far as we know, it is unknown whether the largest multiple packing density under standard and average-radius list-decoding is the same for any finite $L$.

Acknowledgment

The author Yihan Zhang would like to thank Jiajin Li for making the observation given by (20). He would also like to thank Nir Aharon and Ely Porat for several helpful conversations throughout this project, and Alexander Barg for insightful comments on the manuscript.

References

[1] Y. Zhang and S. Vatedka, “List-decodability of Poisson point processes,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), 2022, pp. 2559–2564.

[2] J. H. Conway and N. J. A. Sloane, Sphere Packings, Lattices and Groups, vol. 290. Cham, Switzerland: Springer, 2013.

[3] P. Elias, “List decoding for noisy channels,” Massachusetts Inst. Tech. Res. Lab. Electron., Cambridge, MA, USA, Tech. Rep., 335, 1957.

[4] J. M. Wozencraft, “List decoding,” Quart. Prog. Rep., vol. 48, pp. 90–95, Jan. 1958.

[5] G. Poltyrev, “On coding without restrictions for the AWGN channel,” IEEE Trans. Inf. Theory, vol. 40, no. 2, pp. 409–417, Mar. 1994.

[6] M. Langberg, “Private codes or succinct random codes that are (almost) perfect,” in Proc. 45th Annu. IEEE Symp. Found. Comput. Sci., Jun., 2004, pp. 325–334.

[7] A. D. Sarwate, “Robust and adaptive communication under uncertain interference,” Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Univ. California, Berkeley, CA, USA, Jul. 2008. [Online]. Available: http://www2.eecs.berkeley.edu/Pubs/TechRpts/2008/ERCS-2008-86.html

[8] S. Bhattacharya, A. J. Budkuley, and S. Jaggi, “Shared randomness in arbitrarily varying channels,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Jul. 2019, pp. 627–631.

[9] Y. Zhang, S. Vatedka, S. Jaggi, and A. D. Sarwate, “Quadratically constrained myopic adversarial channels,” IEEE Trans. Inf. Theory, vol. 68, no. 8, pp. 4001–4014, Aug. 2022.

[10] Y. Zhang, S. Vatedka, and S. Jaggi, “Quadratically constrained two-way adversarial channels,” 2020, arXiv:2001.02575.

[11] J. Kepler. Strena Seu De Nive Sexangula (the Six-Cornered Snowflake), vol. 1611. G. Tampach, Ed., Frankfurt, Germany: Oxford, 1966.

[12] H. Minkowski, Geometrie der Zahlen, B. Teubner, Ed. Leipzig, Germany: Teubner, 1910.

[13] 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.
[46] A. Ben-Aroya, D. Doron, and A. Ta-Shma, “Near-optimal erasure list-decodable codes,” in Proc. 35th Comput. Complex. Conf. (CCC), Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2020, pp. 1–27.

[47] V. Guruswami, B. Haeupler, and A. Shahrasbi, “Optimaly resilient codes for list-decoding from insertions and deletions,” in Proc. 52nd Annu. ACM SIGACT Symp. Theory Comput., Jun. 2020, pp. 524–537.

[48] N. Polyanskii and Y. Zhang, “Codes for the Z-channel,” 2021, arXiv:2105.01427.

[49] Y. Zhang, A. J. Budkuley, and S. Jaggi, “Generalized list decoding,” in Proc. 11th Innov. Theor. Comput. Sci. Conf. (ITCS), in Leibniz International Proceedings in Informatics (LIPIcs), vol. 151, T. Vidick, Ed. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2020, p. 51. [Online]. Available: https://drops.dagstuhl.de/opus/volltexte/2020/11736

[50] B. L. Hughes, “The smallest list for the arbitrarily varying channel,” IEEE Trans. Inf. Theory, vol. 43, no. 3, pp. 803–815, May 1997.

[51] A. D. Sarwate and M. Gastpar, “List-decoding for the arbitrarily varying channel under state constraints,” IEEE Trans. Inf. Theory, vol. 58, no. 3, pp. 1372–1384, Mar. 2012.

[52] Y. Zhang, S. Jaggi, and A. J. Budkuley, “Tight list-sizes for oblivious AVCs under constraints,” 2020, arXiv:2009.03788.

[53] F. Hosseinigoki and O. Kosut, “List-decoding capacity of the Gaussian arbitrarily-varying channel,” Entropy, vol. 21, no. 6, p. 575, 2019, doi: 10.3390/e21060575.

[54] Y. Zhang and S. Vatedka, “Multiple packing: Lower bounds via error exponents,” 2022, arXiv:2211.04408.

[55] N. M. Blachman and L. Few, “Multiple packing of spherical caps,” Mathematika, vol. 10, no. 1, pp. 84–88, 1963.

[56] V. Blinovsky, “Multiple packing of the Euclidean sphere,” IEEE Trans. Inf. Theory, vol. 45, no. 4, pp. 1334–1337, May 1999.

[57] H. Groemer, “Existenzsätze für lagerungen im Euklidischen Raum,” Mathematische Zeitschrift, vol. 81, no. 3, pp. 260–278, 1963, doi: 10.1007/BF01111546

[58] H. Cohn and N. Elkies, “New upper bounds on sphere packings,” Advances in Mathematics, vol. 157, no. 2, pp. 689–714, 2003, doi: 10.1006/aima.2003.1576.

[59] G. Kuperberg, “Notions of denseness,” Ann. Math., vol. 157, no. 2, pp. 689–714, 2003, doi: 10.17323/1609-4514-2001-1-1-73-89.

[60] R. A. Rankin, “The closest packing of spherical caps in n dimensions,” Proc. Glasgow Math. Assoc., vol. 2, no. 3, pp. 139–144, 1955.

[61] E. Spence, “Packing of spheres in LP,” Glasgow Math. J., vol. 11, no. 1, pp. 72–80, 1970.

[62] K. Ball, “Inequalities and sphere-packing in l_p,” Isr. J. Math., vol. 58, no. 2, pp. 243–256, 1987.

[63] A. Samorodnitsky. (2013). A Bound on L1 Codes. [Online]. Available: https://www.cs.huji.ac.il/~sales/papers/L1_codes.pdf

[64] S. Shlosman and M. Tsfasman, “Random lattices and random sphere packings: Typical properties,” Moscow Math. J., vol. 1, no. 1, pp. 73–89, 2001, doi: 10.17323/1609-4514-2001-1-1-73-89.

[65] R. A. Rankin, “The closest packing of spherical caps in n dimensions,” Proc. Glasgow Math. Assoc., vol. 2, no. 3, pp. 139–144, 1955.

[66] E. Spence, “Packing of spheres in LP,” Glasgow Math. J., vol. 11, no. 1, pp. 72–80, 1970.

[67] K. Ball, “Inequalities and sphere-packing in l_p,” Isr. J. Math., vol. 58, no. 2, pp. 243–256, 1987.

[68] A. Samorodnitsky. (2013). A Bound on L1 Codes. [Online]. Available: https://www.cs.huji.ac.il/~sales/papers/L1_codes.pdf

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 coding theory, information theory, and statistics theory (in no particular order).

Shashank Vatedka (Member, IEEE) received the M.Sc. (Engineering) and Ph.D. degrees from the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru, India, 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, SAR. From 2018 to 2019, he was a Post-Doctoral Fellow at Télécom Paris, France. Since October 2019, he has been an Assistant Professor with the Department of Electrical Engineering, IIT 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 Kailini Medal for Best Ph.D. Thesis from the Department of Electrical Communication Engineering, IISc Bengaluru. He also received the Best Paper Award-Honorable Mention from SPCCOM 2020, the Best Poster Award from the 2021 Stanford Compression Workshop, and the Best Paper Award (Communications Track) from NCC 2023.