# *On Chebyshev radius of a set in Hamming space and the closest string problem*

**Citation:** Mazumdar, Arya, Yury Polyanskiy, and Barna Saha. "On Chebyshev Radius of a Set in Hamming Space and the Closest String Problem." 2013 IEEE International Symposium on Information Theory (July 2013).

**As Published:** http://dx.doi.org/10.1109/ISIT.2013.6620457

**Publisher:** Institute of Electrical and Electronics Engineers (IEEE)

**Persistent URL:** http://hdl.handle.net/1721.1/90564

**Version:** Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

# On Chebyshev Radius of a Set in Hamming Space and the Closest String Problem

Arya Mazumdar[†] Yury Polyanskiy[*] Barna Saha[‡]

*Abstract*—**The Chebyshev radius of a set in a metric space is defined to be the radius of the smallest ball containing the set. This quantity is closely related to the covering radius of the set and, in particular for Hamming set, is extensively studied in computational biology. This paper investigates some basic properties of radii of sets in $n$-dimensional Hamming space, provides a linear programing relaxation and gives tight bounds on the integrality gap. This results in a simple polynomial-time approximation algorithm that attains the performance of the best known such algorithms with shorter running time.**

## I. Introduction

A contribution of Sylvester [21] states[1]: "It is required to find the least circle which shall contain a given system of points in the plane." The first fast (linear in the number of points) algorithm for solving this problem was apparently found by Megiddo [18]. The corresponding problem in the $n$-dimensional Euclidean space is the smallest *bounding sphere* problem: compute the smallest $n$-sphere that encloses all of the points. Various and increasingly faster algorithms to solve this problem have been proposed in a number of papers, for example, [2], [10], [12], [20], [22], [24]. In this paper we address a similar question in the Hamming space, where the problem is interesting in its own right.

Consider a subset $S$ of size $m$ of the $n$-dimensional Hamming space, i.e., $S \subset \mathbb{F}_2^n$, and $|S| = m$. Define the following

$$\mathrm{r}_{\mathrm{cov}}(S) = \max_{x \in \mathbb{F}_2^n} \min_{y \in S} \mathrm{d}(y, x) \tag{1}$$

$$\mathrm{d}_{\min}(S) = \min_{x,y \in S, x \neq y} \mathrm{d}(x, y) \tag{2}$$

$$\mathrm{diam}(S) = \max_{x,y \in S; x \neq y} \mathrm{d}(x, y) \tag{3}$$

$$\mathrm{rad}(S) = \min_{x \in \mathbb{F}_2^n} \max_{y \in S} \mathrm{d}(y, x)\,, \tag{4}$$

where $\mathrm{d}(\cdot, \cdot)$ is the Hamming distance. These quantities are called *covering radius*, *minimum distance*, *diameter* and *radius* of $S$, respectively. Any point $x$ solving the optimization in (4) is a *Chebyshev center* of the set $S$. In words, $\mathrm{rad}(S)$ is the radius the smallest metric ball (or the circumscribed sphere) that fully contains $S$ and Chebyshev center is its center. Unlike Euclidean space, center in Hamming space may be non-unique as example of $\{(0,0), (1,1)\} \subset \mathbb{F}_2^2$ demonstrates.

The notion of radius and center comes up naturally in many problems. First, a simple identity

$$\mathrm{rad}(S) = n - \mathrm{r}_{\mathrm{cov}}(S) \tag{5}$$

relates it to one of the fundamental parameters of any code. Second, computing the center/radius of a set is an essential operation for the optimal decoder in combinatorial joint source-channel coding [11]. Third, the center usually yields a good representative of a set. In computational biology, this is known as the *closest-string* problem (see, [14] and the references therein) and in the theoretical computer science literature, often shows up as the *Hamming center* problem [7].

Our contributions are as follows. We study some basic properties of the center and radius of a set. Namely, we show existence of sets in Hamming space that have radii approximately half of the diameter. This rigorously establishes that the (easily computable) diameter can not serve as a good approximation of the radius. We show the radius and minimum distance follow a simple relation. A linear programming relaxation is given and its integrality gap is analyzed. Finally, we provide a polynomial-time approximation scheme (PTAS) that estimates $\mathrm{rad}(S)$ to within $(1 + \epsilon)$ error in $m^{O(\frac{1}{\epsilon^2})}$ time.

We next briefly survey the algorithmic literature related to $\mathrm{rad}(S)$. In [6], the decision problem is shown to be NP-complete. Then in a number of papers increasingly better PTAS were proposed. In [7], the center problem was formulated as an integer linear program and a rounding algorithm for the corresponding linear programming relaxation was given that approximates the radius within a factor of $(\frac{4}{3} + \epsilon)$, $\epsilon > 0$. The authors in [7] also point out the difficulty of using independent randomized rounding technique [19], [23] for this problem when the optimal solution of the linear program is not so big. According to [7], in that "small" radius regime, employment of the Chernoff bound, the standard way of analyzing independent randomized rounding, leads to a significant deviation of $\sim \sqrt{n}$. This supposed shortcoming was overcome by [13] that performs more involved analysis by carefully studying the structural properties of the center problem. Specifically they show, that there exists an $l$-subset $L \subset S$, such that, for any $i \in [1, n]$ if all the $l$ vectors agree on the $i$th bit, then the $i$th bit of the computed center may be set to that same value without increasing the radius by much. Also, the number of positions where the vectors in $L$ agree is at least $n - l \cdot \mathrm{rad}(S)$. For the remaining

[†] Department of ECE, University of Minnesota, Minneapolis, MN, arya@umn.edu.

[*] Department of EECS, Massachusetts Inst. of Technology, Cambridge, MA 02139, yp@mit.edu.

[‡] AT&T Shannon Laboratories, Florham Park, NJ, barna@research.att.com

The work was partially supported by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-0939370, and a start-up grant from University of Minnesota.

[1] That is the entire content of [21].

$l \cdot \mathrm{rad}(S)$ positions, they determine the values of those bits in the center by doing independent randomized rounding on the LP-relaxation defined on those bits if $\mathrm{rad}(S) \geq \frac{1}{\epsilon^2}\log m$, and by exhaustive enumeration otherwise. In [13] $l$ is a polynomial in $\frac{1}{\epsilon}$. This leads to the first PTAS that runs in time $nm^{O(l/\epsilon^2)}$. Subsequently, the size of $l$ was reduced to $O(\log \frac{1}{\epsilon})$ in [16] and this improved the running time of the PTAS to $nm^{O(1/\epsilon^2 \log(1/\epsilon))}$ [1]. Finally, Ma and Sun proposed a fixed point tractable algorithm that computes radius in $O(m2^{O(\mathrm{rad}(S))})$ time [14]. Using this result along with [13] implies a running time of $nm^{O(1/\epsilon^2)}$ [14]. All of these results, and subsequent works [1], [14], crucially use the algorithm in [13] as subroutine which needs to solve $m^{O(\log \frac{1}{\epsilon})}$ linear programs even to guarantee an $(1+\epsilon)$ approximation when $\mathrm{rad}(S) \geq \frac{1}{\epsilon^2}\log m$. Our improvement is specifically on this front: we obtain the same approximation guarantee without any dependence on $\epsilon$ in the running time, by refining the analysis of [7].

Specifically, we show that a linear program of [7] approximates $\mathrm{rad}(S)$ within a factor of $1+\epsilon$ and runs in polynomial time in $m$ and $n$ (without any dependency on $\epsilon$) as long as the radius $\mathrm{rad}(S) \geq \frac{1}{\epsilon^2}\log m$. Thus, in our approach one needs to solve exactly one linear programming instance, as opposed to [13] where as many as $m^{O(\log \frac{1}{\epsilon})}$ different instances are to be solved. When the radius is strictly smaller than $\frac{1}{\epsilon^2}\log m$, we resort to the fixed parameter algorithm of [14], and overall complexity is dictated by this step, yielding the $m^{O(\frac{1}{\epsilon^2})}$ time.

The organization of the paper is as follows. In Section II we summarize some identities that the radius of a set must follow. Section III contains our main algorithms for estimating the center of a set and the analysis of their correctness.

## II. Properties of radius

One trivial property[2] of the radius is the following. Let $S_1 \subset \mathbb{F}_2^{n_1}$ and $S_2 \subset \mathbb{F}_2^{n_2}$, and define a set $S = S_1 \oplus S_2 \subset \mathbb{F}_2^{n_1+n_2} \cong \mathbb{F}_2^{n_1} \oplus \mathbb{F}_2^{n_2}$ of cardinality $|S_1| \cdot |S_2|$, then we have

$$\mathrm{rad}(S_1 \oplus S_2) = \mathrm{rad}(S_1) + \mathrm{rad}(S_2)\,. \tag{6}$$

### A. Jung constant

It is evident that for any set $S \subset \mathbb{F}_2^n$,

$$\begin{aligned} \frac{1}{2}\mathrm{diam}(S) \leq \mathrm{rad}(S) \quad &\leq \quad \frac{J(\mathbb{F}_2^n)}{2}\mathrm{diam}(S) && (7)\\ &\leq \quad \mathrm{diam}(S)\,, && (8)\end{aligned}$$

where for any metric space $X$ we define its Jung constant to be

$$J(X) \equiv \sup_{S \subset X} \frac{2\,\mathrm{rad}(S)}{\mathrm{diam}(S)}\,.$$

Clearly both inequalities in (7) are tight, and clearly

$$1 \leq J(X) \leq 2\,.$$

The spaces with $J(X) = 1$ are called centrable, of which the primary examples are $(\mathbb{R}^n, \|\cdot\|_\infty)$. It turns out that asymptotically $J(\mathbb{F}_2^n) \to 2$ and thus (7) is (asymptotically) not an improvement of (8).

*Theorem 1:* For any $n$ for which there exists an $(n+1) \times (n+1)$ Hadamard matrix, we have

$$J(\mathbb{F}_2^{2n}) \geq \frac{2n}{n+1}\,. \tag{9}$$

In particular $\limsup_{n\to\infty} J(\mathbb{F}_2^n) = 2\,.$

*Proof:* Consider an isometric embedding of the set $\{-1, 0, 1\}$ into $\mathbb{F}_2^2$ where

$$-1 \to 00, \quad 0 \to 10, \quad 1 \to 11\,, \tag{10}$$

(this has a simple generalization for the embedding of $\{-d, \ldots, d\}$ into $\mathbb{F}_2^{2d}$). Naturally this extends into the isometric embedding of $\{-1, 0, 1\}^n$ with $\ell_1$ metric into $\mathbb{F}_2^{2n}$.

For any $n$, satisfying the conditions, a set of $n+1$ vectors in $\{-1, 0, 1\}^n$ is constructed in [5] such that the all-zero vector is its $\ell_1$-Chebyshev center and the ratio of the radius to diameter attains right hand side of (9). Since (10) is an isometry, the claim follows. ■

*Remark 1:* The smallest example of this construction is the set in $\mathbb{F}_2^6$ given by the rows in the matrix

$$\begin{pmatrix} 11 & 11 & 11\\ 11 & 00 & 00\\ 00 & 11 & 00\\ 00 & 00 & 11 \end{pmatrix}$$

whose diameter is 4, radius is 3 and a Chebyshev center is $(10, 10, 10)$ (direct argument convinces one that there is no vector at distance 2 from all of these vectors).

*Remark 2:* It is tempting to conjecture that

$$J(\mathbb{F}_2^n) \leq \frac{2n}{n+1}\,. \tag{11}$$

Indeed, consider an isometric embedding $\mathbb{F}_2^n \to (\mathbb{R}^n, \ell_1)$ and then apply the estimate of Bohnenblust [3]:

$$J(\mathbb{R}^n, \text{any norm}) \leq \frac{2n}{n+1}\,.$$

This, of course, does not imply (11) as Jung's constant may grow under isometric embeddings.

### B. Radius and minimum distance

In this section we show a relation between the minimum pairwise distance of a set to its radius. Define the following quantities:

$$A(n,d,w) = \max_{\substack{A \subseteq \mathbb{F}_2^n, \mathrm{d}_{\min}(A) \geq d\\ \forall x \in A, \mathrm{wt}(x) = w}} |A|\,, \tag{12}$$

$$B(n,d,w) = \max_{\substack{A \subseteq \mathbb{F}_2^n, \mathrm{d}_{\min}(A) \geq d\\ \forall x \in A, \mathrm{wt}(x) \leq w}} |A|\,, \tag{13}$$

where $\mathrm{wt}(\cdot)$ denotes the Hamming weight of a vector. The former is the maximum size of a constant weight code – a well-studied quantity [15].

An inequality that relates the radius and the minimum distance of a set is the following.

*Lemma 2:* For any $S \subset \mathbb{F}_2^n$,

$$|S| \leq B(n, \mathrm{d}_{\min}(S), \mathrm{rad}(S)). \tag{14}$$

[2] This happens due to an $\ell_1$ nature of the Hamming metric.

*Proof:* Enclose $S$ inside a ball of radius $\operatorname{rad}(S)$ and use definition (13). ■

A well-known bound on $A(n,d,w)$ is the Johnson Bound [9]. We can make slight changes in the proof of that bound such that it remains true for $B(n,d,w)$. Indeed, we can have,

*Lemma 3:* Suppose, $w \le n/2$. Then,

$$B(n,d,w) \le \frac{dn}{dn-2wn+2w^2}, \tag{15}$$

as long as the denominator is positive.

The proof of the above lemma is standard and we omit it here.

As a consequence of the above lemma, we have a corollary of Thm. 2.

*Corollary 4:* For any $S \subset \mathbb{F}_2^n$, if $\operatorname{rad}(S) \le n/2$, then

$$|S| \le \frac{n\,\mathrm{d}_{\min}(S)}{n\,\mathrm{d}_{\min}(S) - 2n\operatorname{rad}(S) + 2\operatorname{rad}(S)^2},$$

as long as the denominator is positive.

This trivially leads us to the following observation:

*Theorem 5:* For any $S \subset \mathbb{F}_2^n$, if $|S| > n\,\mathrm{d}_{\min}(S)$ then

$$\frac{\operatorname{rad}(S)}{n} \ge \frac{1}{2} - \frac{1}{2}\sqrt{1 - 2\frac{\mathrm{d}_{\min}(S)}{n}} \triangleq J\left(\frac{\mathrm{d}_{\min}(S)}{n}\right) \tag{16}$$

The term $J(\delta)$ is known as the Johnson radius [15].

## III. Computing the radius

In this section we show an algorithm to estimate the center closely and thus also approximating the radius of a set. Note that, it is easy to compute the diameter of the set $S$. Indeed, one just have to compute all possible pairwise distances. The complexity of this is $O(m^2 n)$. Hence, in time $O(m^2 n)$, it is possible to estimate the radius (and center) of the set within a factor of 2: simply output one of the points of $S$ as the center and the distance of the farthest point to this one as the radius.

A simple linear programming algorithm for computation of radius was proposed in [7]. We rephrase the formulation below and use a refined analysis of the rounding technique to obtain our result.

### A. An integer programming formulation

Consider map $g : \mathbb{F}_2 \to \mathbb{R}$: $\quad 0 \mapsto +1$ ; $1 \mapsto -1$. We can extend this map to $g : \mathbb{F}_2^n \to \{-1,+1\}^n \subset \mathbb{R}^n$ by mapping each coordinate according to $g$. For any $x \in \mathbb{F}_2^n$, let us write $\hat{x} \equiv g(x)$. Clearly, $\mathrm{d}(x,y) = \frac{1}{2}\|\hat{x}-\hat{y}\|_1$ , where $\|\cdot\|_1$ denotes the $\ell_1$ distance in $\mathbb{R}^n$. Observe now that if $a \in \{-1,1\}^n$ and $x \in [-1,1]^n$ then we have

$$\|a - x\|_1 = n - \langle a, x\rangle, \tag{17}$$

with $\langle\cdot,\cdot\rangle$ denoting the standard inner product in $\mathbb{R}^n$.

Suppose, the set $S = \{a_1,\dots,a_m\} \subset \mathbb{F}_2^n$ has radius $d$. Then there exists $x \in \mathbb{F}_2^n$, such that,

$$\mathrm{d}(x, a_i) \le d, \tag{18}$$

for $1 \le i \le m$. This implies,

$$\langle \hat{a}_i, \hat{x}\rangle \ge n - 2d, \quad 1 \le i \le m. \tag{19}$$

Note that,

$$\hat{x} \in \{-1,+1\}^n. \tag{20}$$

The smallest positive integer $d$ such that there exists a non-empty set of feasible solutions to the Equations (19) and (20), is the radius of the set. As $\operatorname{rad}(S) < n$, the complexity of finding the radius (and a center) is at most $n$ times (actually, at most $\log\operatorname{diam}(S)$ times) the complexity of solving the integer linear program given by Equations (19), (20). The coordinates of the center can be found to be $(1-\hat{x}(j))/2, 1 \le j \le n$, which is the inverse map of $g$.

### B. A linear programming relaxation

Let us relax the condition (20) to the following.

$$-1 \le \hat{x}(j) \le 1, \quad 1 \le j \le n. \tag{21}$$

If there exists a feasible $\hat{x}$ such that Equations (19), (21) are satisfied, then that $\hat{x}$ can be found by using an algorithm that solves linear programs. However, that $\hat{x}$ might be a non-integer point and therefore may not provide a valid center in the Hamming space.

Define the LP relaxed radius as

$$r_{LP} = \frac{n}{2} - \frac{1}{2}\max_{z\in[-1,1]^n}\min_i \langle z, \hat{a}_i\rangle, \tag{22}$$

and let $z^*$ be a maximizer in (22).

*Theorem 6:* We have

$$\begin{aligned} r_{LP} \le \operatorname{rad}(S) &\le r_{LP} + \frac{1}{3}\ln|S| + \sqrt{\frac{1}{9}\ln^2|S| + 2V\ln|S|} \\ &\le r_{LP} + \frac{2}{3}\ln|S| + \sqrt{8 r_{LP}\ln|S|} \end{aligned} \tag{23}$$

where $V = n - \|z^*\|_2^2 \le 4r_{LP}$.

*Remark 3:* In view of (17), $r_{LP}$ is simply half of the Chebyshev radius of the set $S$ in $(\mathbb{R}, \ell_1)$. Thus the theorem compares the radius in Hamming space vs. radius in the ambient $\ell_1$ space. The result is useful for two reasons. First, in general $r_{LP}$ and $\operatorname{rad}(S)$ can be very different: e.g. for $S = \mathbb{F}_2^n$ we have $r_{LP} = \frac{n}{2}$ while $\operatorname{rad}(S) = n$. Theorem shows that order-$n$ gaps, however, are only possible for exponentially large sets. Second, unlike $\operatorname{rad}(S)$ relaxation $r_{LP}$ is easy to compute as a solution of a linear program.

*Remark 4:* At the same time, the bound in (23) is not tight for small sets, e.g. $|S| = O(1)$ and $n \to \infty$. In fact, it can be shown that for $|S| > 1$

$$\operatorname{rad}(S) \le r_{LP} + \frac{2^{|S|-1} - 2}{2}$$

To see this, let $m = |S|$. By translation, assume one element of $S$ to 0. Then all $n$ coordinates can be split into $2^{m-1}$ groups, such that elements of $S \setminus \{0\}$ are constant inside each group. Two of the groups are where all elements of $S$ have zeros and ones only, and can be ignored. Clearly, the optimal solution ($\ell_1$-center) in (22) maybe assumed to be constant inside each group. Then its value can be approximated by a zero-one vector to within $\pm\frac{1}{2}$. Thus, the resulting vector approximates the distance to *each* of the elements in $S$ to within $\frac{2^{m-1}-2}{2}$.

Proof of Theorem 6 is similar to that of Theorem 9, and we omit it here.

### *C. The randomized rounding: RANDRAD*

The input of the algorithm is the set $S = \{a_1, \dots, a_m\} \subset \mathbb{F}_2^n$. Suppose $d$ is the smallest positive integer such that Equations (19), (21), have a feasible point. Let, $z \in \mathbb{R}^n$ be a feasible point for that $d$ (can be found by solving a linear program). Construct a random vector $y \in \mathbb{F}_2^n$ from $z$ in the following way.

$$y(j) = \begin{cases} 0 & \text{with probability } \frac{1+z(i)}{2} \\ 1 & \text{with probability } \frac{1-z(i)}{2}. \end{cases}$$

Each coordinate of $y$ is independent. Output $y$ as the center and $\max_{1 \le i \le m} \mathrm{d}(y, a_i)$ as the radius.

The algorithm runs in polynomial time.

*Remark 5:* This method of rounding linear programming solutions to estimate the corresponding integer programming solution is called *independent randomized rounding* [19] and is standard in the literature of approximation algorithms [23]. This rounding technique was used in [7]. In conjunction with the Hoeffding's inequality [8], we have the following result, also present in [7].

*Theorem 7:* Suppose $S \subset \mathbb{F}_2^n$ and $|S| = m$. The randomized polynomial time algorithm RANDRAD($S$) outputs estimate $y$ of the center of $S$ such that for any $d' > 0$,

$$\Pr\Big(\max_{1 \le i \le m} d(y, a_i) \le \mathrm{rad}(S) + d'\Big) \ge 1 - m \exp\Big(-\frac{2d'^2}{n}\Big).$$

The radius is thus approximated with high probability within an additive term $d' = \sqrt{n \ln m}$. Hence whenever the true radius is large RANDRAD produces a quite accurate estimate with high probability.

Note that there is an additive error term in the approximation of the radius that grows with $\sqrt{n}$. When the true radius of set is $\omega(\sqrt{n})$ this error can be withstood. But when the radius is $O(\sqrt{n})$ this error term becomes the dominating factor. In the following section we show a way to overcome this.

### *D. Refinement of the rounding and analysis*

Let us propose a refined rounding algorithm of the linear programming solution. This rounding uses a mix of deterministic and randomized rounding. We change the algorithm RANDRAD as follows. Assume the input of the algorithm is the set $S = \{a_1, \dots, a_m\} \subset \mathbb{F}_2^n$, and $d$ is the smallest positive integer such that Equations (19), (21), have a feasible solution. Let, $z \in \mathbb{R}^n$ be a feasible point for that $d$. Construct a random vector $y \in \mathbb{F}_2^n$ from $z$ in the following way. Say, $0 < b < 1$, whose value will be decided in the next section.

$$y(i) = \begin{cases} 0 & \text{when } z(i) \ge b \\ 1 & \text{when } z(i) \le -b. \end{cases}$$

When $i$ is such that $-b < z_i < b$, set independently

$$y(i) = \begin{cases} 0 & \text{with probability } \frac{1+z(i)}{2} \\ 1 & \text{with probability } \frac{1-z(i)}{2}. \end{cases}$$

Output $y$ as the center and $\max_{1 \le i \le m} \mathrm{d}(y, a_i)$ as the radius.

The algorithm, called RANDRAD-REF, runs in polynomial time. Notice that, before resorting to the randomize rounding we deterministically round some points that are very close to integer values. Hence the previous rounding can be thought of a special case of this algorithm. This gives us an extra parameter to optimize our algorithm with. Indeed, by computing the diameter of the set first we can be sure of the range of the radius and then set $b = 1$ above when diameter is $\omega(\sqrt{n})$.

The following can be said regarding the performance of RANDRAD-REF.

*Theorem 8:* Suppose $S \subset \mathbb{F}_2^n$ and $|S| = m$. For any $0 < \epsilon < \frac{1}{2}$ and $b = 1 - 2\epsilon$, the algorithm RANDRAD–REF($S$) estimates the center and radius $d_R$, such that $d_R \ge \mathrm{rad}(S)$ and with probability at least $1 - \delta$, $d_R$ is at most

$$\frac{\mathrm{rad}(S)}{1-\epsilon} + \sqrt{\frac{\mathrm{rad}(S)}{2\epsilon} \ln \frac{m}{\delta}}.$$

It is apparent that the additive error is now proportional to only $\sqrt{\mathrm{rad}(S)}$ and does not depend on $n$ at all. Hence for relatively smaller values of the radius this algorithm gives a much better approximation guarantee than the previous naive rounding algorithm.

*Proof of Theorem 8:* Remember $d$ is the smallest positive integer such that Equations (19), (21), have a feasible solution. Obviously

$$d \le \mathrm{rad}(S). \tag{24}$$

Suppose, $I \subseteq \{1, \dots, n\}$ be the set such that $-b < z_i < b$. Let, $\bar{I} = \{1, \dots, n\} \setminus I$, and for any vector $w$ and any index set $K$, $w(K)$ denotes the projection of $w$ on to $K$.

Now, $\mathrm{d}(y, a_i) = \mathrm{d}(y(I), a_i(I)) + \mathrm{d}(y(\bar{I}), a_i(\bar{I}))$. Also, $\langle \hat{a}_i, z \rangle = \sum_{j=1}^{n} (1 - 2|a_i(j) - \frac{1-z(j)}{2}|) \ge n - 2d$, which means, if $d_1 = \sum_{j \in I} |a_i(j) - \frac{1-z(j)}{2}|$ and $d_2 = \sum_{j \in \bar{I}} |a_i(j) - \frac{1-z(j)}{2}|$, then $d_1 + d_2 \le d$. Each element in the summation of the expression for $d_1$ is greater than $\frac{1-b}{2}$ and less than $\frac{1+b}{2}$. Hence,

$$\frac{2d_1}{1+b} \le |I| \le \frac{2d_1}{1-b}. \tag{25}$$

Now, because for any $j \in \bar{I}$, $|a_i(j) - y(j)| \le \frac{2}{1+b}|a_i(j) - \frac{1-z(j)}{2}|$,

$$\mathrm{d}(y(\bar{I}), a_i(\bar{I})) \le \frac{2d_2}{1+b}. \tag{26}$$

Now just as in the case of RANDRAD, the random variable $\mathrm{d}(y(I), a_i(I))$ is concentrated around its mean, $d_1$. Indeed, using Hoeffding's inequality [8],

$$\Pr(\mathrm{d}(y(I), a_i(I)) \ge d_1 + \lambda) \le \exp(-\frac{2\lambda^2}{|I|}).$$

But this implies,

$$\max_{1 \le i \le m} \Pr(\mathrm{d}(y, a_i) \ge \frac{2d_2}{1+b} + d_1 + \lambda) \le m \exp(-\frac{2\lambda^2}{|I|}). \tag{27}$$

Hence, with probability at least $1 - \delta$, the estimated radius $d_R$ is at most,

$$\begin{aligned} d_R &\le \frac{2d_2}{1+b} + d_1 + \sqrt{\frac{|I|}{2} \ln \frac{m}{\delta}} \\ &\le \frac{2d_2}{1+b} + d_1 + \sqrt{\frac{d_1}{1-b} \ln \frac{m}{\delta}} \end{aligned}$$

$$\begin{aligned}&\leq \quad d + d_2\left(\frac{2}{1+b} - 1\right) + \sqrt{\frac{d}{1-b}\ln\frac{m}{\delta}}\\ &\leq \quad \frac{2d}{1+b} + \sqrt{\frac{d}{1-b}\ln\frac{m}{\delta}}\end{aligned}$$

Substituting $\epsilon = \frac{1-b}{2}$ proves the theorem. ■

*Remark 6:* The optimal value of $\epsilon$ that gives the lowest possible bound on the estimated radius can be found by solving the equation $\frac{(1-\epsilon)^2}{\epsilon^{3/2}} = 2\sqrt{\frac{2d}{\ln m/\delta}}$. In our theorem, $\epsilon$ denotes the trade-off we want between the multiplicative approximation error and the additive error.

The two-stage rounding algorithm RANDRAD-REF, reduces the amount of random bits required from the straightforward rounding. It is an immediate corollary of the above theorem that, if $\mathrm{rad}(S) = \Omega(\frac{\log m}{\epsilon^3})$, with probability $1 - o(1)$, RANDRAD-REF estimates the radius within a factor of $(1+\epsilon)$ and runs in polynomial time, independent of $\epsilon$. Instead of using RANDRAD-REF, we could have used a better concentration inequality than the Hoeffding's bound in the analysis of RANDRAD. Indeed, the above rounding comes from the intuition that the variables with low variance better be deterministically rounded to the nearest integer. By using the Bernstein inequality, we can have the following result.

*Theorem 9:* Suppose $S \subset \mathbb{F}_2^n$ and $|S| = m$. The randomized polynomial time algorithm RANDRAD($S$) outputs estimate $y$ of the center of $S$ such that for any $\epsilon > 0$,

$$\Pr\Big(\max_{1\leq i\leq m} d(y, a_i) > \mathrm{rad}(S)(1+\epsilon)\Big) \leq m\exp\Big(-\frac{\mathrm{rad}(S)\epsilon^2}{2(1+\frac{\epsilon}{3})}\Big).$$

*Proof:* Suppose, $a \in S$ and $z$ is the solution of the linear program of (19), (21). Let, $X_i = \hat{a}(i)\hat{y}(i)$ where,

$$\hat{y}(j) = \begin{cases} 1 & \text{with probability } \frac{1+z(i)}{2} \\ -1 & \text{with probability } \frac{1-z(i)}{2}. \end{cases}$$

Also, let, $X = \sum_{i=1}^{n} X_i$. We have, $X = n - 2\mathrm{d}(a, y)$, and $\mathbb{E}X \geq n - 2d$, where $d$ is the radius of the set $S$. Define, $Y_i = \frac{\mathbb{E}X_i - X_i}{2}$. We have, $Y_j \leq 1$ and $\mathbb{E}Y_j = 0$ for $0 \leq j \leq n$. Now,

$$\begin{aligned}\sigma^2 \equiv \frac{1}{n}\sum_{i=1}^{n}\mathrm{var}Y_i &= \frac{1}{4n}\sum_{i=1}^{n}(1 - (\mathbb{E}X_i)^2)\\ &\leq \frac{1}{4} - \frac{1}{4}\Big(\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}X_i\Big)^2\\ &\leq \frac{1}{4} - \frac{1}{4}\Big(\frac{n-2d}{n}\Big)^2 \leq \frac{d}{n}.\end{aligned}$$

Using Bernstein inequality [4, Thm. 3], we have,

$$\Pr\Big(\sum_{i=1}^{n} Y_i > d\epsilon\Big) \leq \exp\Big(-\frac{d^2\epsilon^2}{2(\sigma^2 n + d\epsilon/3)}\Big).$$

Using union bound the theorem is proved. ■

As a consequence of the above theorem, when $\mathrm{rad}(S) \geq \frac{3\log m}{\epsilon^2}$, with probability $1 - o(1)$, RANDRAD estimates the radius within a factor of $(1+\epsilon)$ and runs in polynomial time, independent of $\epsilon$. On the other hand when $\mathrm{rad}(S) < \frac{3\log m}{\epsilon^2}$, we can use the fixed parameter algorithm of [14], and have an algorithm that runs in time $O(m^{O(\frac{1}{\epsilon^2})})$.

In conclusion, using our results, polynomial time algorithm for estimating covering radius arbitrarily close, can also be proposed. One interesting direction of research would be to come up with approximation algorithms to compute radius/covering radius of *linear codes*, in time that is polynomial in the dimension of the space. It was shown in [17] that this problem is NP-hard when an exact solution is needed.

## References

[1] A. Andoni, P. Indyk, and M. Patrascu. On the optimality of the dimensionality reduction method. In *Foundations of Computer Science, 2006. FOCS'06. 47th Annual IEEE Symposium on*, pages 449–458. IEEE, 2006.

[2] M. Bădoiu, S. Har-Peled, and P. Indyk. Approximate clustering via core-sets. In *Proceedings of the thirty-fourth annual ACM symposium on Theory of computing*, pages 250–257. ACM, 2002.

[3] H. F. Bohnenblust. Convex regions and projections in Minkowski spaces. *Ann. Math.*, 39(2):301–308, 1938.

[4] S. Boucheron, G. Lugosi, and O. Bousquet. Concentration inequalities. In *Advanced Lectures on Machine Learning*, pages 208–240. Springer, 2004.

[5] V. P. Dolnikov. On Jung's constant in $\ell_1^n$. *Matem. Zametki*, 42(4):519–526, 1987.

[6] M. Frances and A. Litman. On covering problems of codes. *Theory of Computing Systems*, 30(2):113–119, 1997.

[7] L. Gasieniec, J. Jansson, and A. Lingas. Efficient approximation algorithms for the hamming center problem. In *Proceedings of the tenth annual ACM-SIAM symposium on Discrete algorithms*, pages 905–906. Society for Industrial and Applied Mathematics, 1999.

[8] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[9] S. Johnson. A new upper bound for error-correcting codes. *IRE Transactions on Information Theory*, 8(3):203–207, 1962.

[10] M. K. K. Fischer, B. Gartner. Fast smallest-enclosing-ball computation in high dimensions. In *European Symposium on Algorithms*, pages 630–641. Springer, LNCS Vol. 2832, 2003.

[11] Y. Kochman, A. Mazumdar, and Y. Polyanskiy. The adversarial joint source-channel problem. In *Proc. 2012 IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, MA, July 2012.

[12] P. Kumar, J. Mitchell, and A. Yildirim. Computing core-sets and approximate smallest enclosing hyperspheres in high dimensions. In *5th Workshop on Algorithm Engineering and Experiments*, 2003.

[13] M. Li, B. Ma, and L. Wang. On the closest string and substring problems. *Journal of the ACM (JACM)*, 49(2):157–171, 2002.

[14] B. Ma and X. Sun. More efficient algorithms for closest string and substring problems. In *Research in Computational Molecular Biology*, pages 396–409. Springer, 2008.

[15] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. North-Holland, 1997.

[16] D. Marx. The closest substring problem with small distances. In *FOCS*, pages 63–72, 2005.

[17] A. McLoughlin. The complexity of computing the covering radius of a code. *Information Theory, IEEE Transactions on*, 30(6):800–804, 1984.

[18] N. Megiddo. Linear-time algorithms for linear programming in $\mathbb{R}^3$ and related problems. *SIAM J. Comp.*, 12(4):759–776, 1983.

[19] P. Raghavan and C. D. Thompson. Randomized rounding: a technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7(4):365–374, 1987.

[20] J. Ritter. An efficient bounding sphere. In A. S. Glassner, editor, *Graphics Gems*. Academic Press, Boston, MA, 1990.

[21] J. Sylvester. A question in the geometry of situation. *Quart. J. Pure Appl. Math.*, 1:79–79, 1857.

[22] B. Tian. Bouncing bubble: A fast algorithm for minimal enclosing ball problem. 2012.

[23] V. V. Vazirani. *Approximation algorithms*. Springer, 2001.

[24] E. Welzl. Smallest enclosing disks (balls and ellipsoids). *New results and new trends in computer science*, pages 359–370, 1991.