Extensions to the Johnson bound

Venkatesan Guruswami* Madhu Sudan†

MIT Laboratory for Computer Science
200 Technology Square
Cambridge, MA 02139.
Email: venkat@theory.lcs.mit.edu, madhu@mit.edu.

February, 2001

Abstract

We present extensions of some recent geometric proofs of the well-known Johnson bound. Our extensions apply to arbitrary alphabets (while previous proofs were given only for the binary case). Our extensions yield a "weighted" version of the Johnson bound equally easily — the weighted version is of interest in light of the recent developments on soft-decision list decoding algorithms.

1 Introduction

We present extensions to the well-known Johnson bound in coding theory. The Johnson bound is a classical bound that provides an upper bound on the number of codewords in any Hamming ball of up to a certain radius (the radius till which the bound holds is a function of the minimum distance of the code). Such a bound is used in the Elias-Bassalygo upper bound on the dimension of codes with certain minimum distance, and is of interest to list decoding of codes.

Proofs of the Johnson bound seem to come in one of two flavors. The original proof and some of its derivates follows a linear algebra based argument [5, 6, 3, 10, 11], while more recent proofs, most notably [7, 4, 1] are more geometric. Our proof follows the latter spirit, extending these proofs to the case of general alphabets. (A more technical comparison of our proof with existing ones is given later, after outlining some formal definitions).

Moreover, we also prove a weighted version of the Johnson bound which is of interest to some questions raised by the recent investigations on soft-decision list decoding algorithms. Our result gives some improvements over the earlier results in this vein from [12], and furthermore sheds some light on the features of a weight vector that give the most information to a soft-decision list decoding algorithm.

*Supported in part by an IBM Graduate Fellowship.

†Supported in part by an MIT-NEC Research Initiation Award, a Sloan Foundation Fellowship and NSF Career Award CCR-9875511.

2 Our Results

2.1 Notation

We identify the elements of a q-ary alphabet with the integers $1, 2, \ldots, q$ in some canonical way. Let $[q] = \{1, 2, \ldots, q\}$. For $\mathbf{x}, \mathbf{y} \in [q]^n$, the Hamming distance between $\mathbf{x}$ and $\mathbf{y}$, denoted $\Delta(\mathbf{x}, \mathbf{y})$, is the number of positions where $\mathbf{x}$ and $\mathbf{y}$ differ. For $\mathbf{r} \in [q]^n$ and $0 \leq e \leq n$, the Hamming ball of radius e around $\mathbf{r}$ is defined by $B_q(\mathbf{r}, e) = \{\mathbf{x} \in [q]^n : \Delta(\mathbf{r}, \mathbf{x}) \leq e\}$.

The key quantity to study in our context is the following. Let $A'_q(n, d, e)$ denote the maximum number of points that may be placed in some ball $B_q(\mathbf{r}, e)$ such that all pairwise distances between the points are at least d. More formally, $A'_q(n, d, e) = \max\{|S| : S \subseteq B_q(\mathbf{r}, e)$ for some $\mathbf{r} \in [q]^n$ and $\forall \mathbf{x}, \mathbf{y} \in S,\ \Delta(\mathbf{x}, \mathbf{y}) \geq d\}$.[1] Clearly for any code $\mathcal{C} \subseteq [q]^n$ of minimum distance d, $A'_q(n, d, e)$ is an upper bound on the number of codewords of $\mathcal{C}$ that can lie in a Hamming ball of radius e.[2] Our objective, therefore, is to obtain an upper bound on the function $A'_q(n, d, e)$ and we do so below.

2.2 Our Main Result

Theorem 1 *Let $\mathcal{C}$ be any q-ary code of blocklength n and minimum distance $d = (1 - 1/q)(1 - \delta)n$ for some $0 < \delta < 1$. Let $e = (1 - 1/q)(1 - \gamma)n$ for some $0 < \gamma < 1$ and let $\mathbf{r} \in [q]^n$ be arbitrary. Then, provided $\gamma > \sqrt{\delta}$, we have*

$$|B_q(\mathbf{r}, e) \cap \mathcal{C}| \leq \min\{n(q-1), \frac{1-\delta}{\gamma^2 - \delta}\} \ .$$

Furthermore, for the case when $\gamma = \sqrt{\delta}$, we have $|B_q(\mathbf{r}, e) \cap \mathcal{C}| \leq 2n(q-1) - 1$.

Corollary 2 *Let q, n, d be arbitrary positive integers with $d < (1 - 1/q)n$. Let $e \geq 1$ be any integer that satisfies the condition*

$$e < \Big(1 - \frac{1}{q}\Big)\Big(1 - \sqrt{1 - \frac{q}{q-1} \cdot \frac{d}{n}}\Big)n \ . \tag{1}$$

Then we have

$$A'_q(n, d, e) \leq \min\{n(q-1), \frac{nd}{nd - 2e(n - \frac{qe}{2(q-1)})}\} \ . \tag{2}$$

Furthermore, if e equals the R.H.S of Condition (1) then $A'_q(n, d, e) \leq 2n(q-1) - 1$.

Comparison with Previous Bounds: The second upper bound on $A'_q(n, d, e)$ in (2) is the "classical" version of Johnson bound for the q-ary case (cf. [9]; proofs appear, for instance, in [10, 11]). The new aspect of our result is the $n(q-1)$ upper bound. For the case $q = 2$, this result was known. Specifically, Elias [3] proved that if d is *odd*, then $A'_2(n, d, e) \leq n$ as long as e satisfies

[1] We use the notation $A'_q(n, d, e)$ instead of the apparently more natural choice $A_q(n, d, e)$ because the notation $A_q(n, d, e)$ in coding theory literature normally refers to the maximum number of points that may be placed **on** the surface of (instead of within) the ball $B_q(\mathbf{r}, e)$ with pairwise distances at least d. To avoid confusion with this standard terminology, we use $A'_q(n, d, e)$ instead. We clearly have $A_q(n, d, e) \leq A'_q(n, d, e)$, and thus any upper bound we derive on $A'_q(n, d, e)$ also applies to $A_q(n, d, e)$.

[2] The minimum distance of a code $\mathcal{C}$ is defined as the minimum Hamming distance between two distinct elements of $\mathcal{C}$.

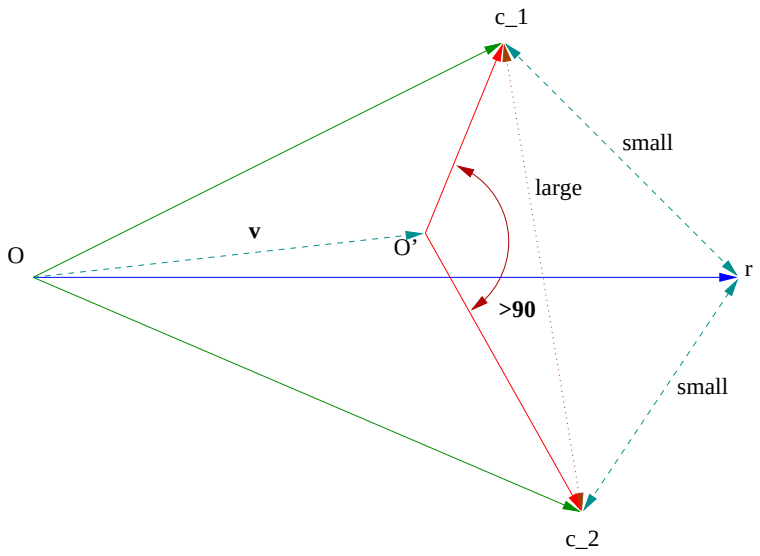

Figure 1: Geometric picture behind proof of Theorem 1

Condition (1). For even d, however, $A'_2(n,d,e) = O(n^2)$ was the best known bound that was made explicit till the recent work of Agrell, Vardy and Zeger [1], who showed that $A'_2(n,d,e) \leq n$ for all d, e that satisfy Condition (1). [3]

For the case of $q > 2$, to the best of our knowledge, the only known upper bound on $A'_q(n,d,e)$ seems to be the second bound in Equation (2) and our upper bound of $n(q-1)$ seems to be new.

Proof Idea: The proof follows a "geometric" approach. We identify elements of $[q]^n$ with vectors in $\mathbb{R}^{nq}$ by replacing the symbol i $(1 \leq i \leq q)$ by the unit vector of length q with a 1 in position i. This allows us to embed the codewords and the "received" word $\mathbf{r}$ into $\mathbb{R}^{nq}$. Next, by appropriately shifting the set of vectors corresponding to the codewords that are close to $\mathbf{r}$, we get a set of vectors such that the inner product of any two distinct vectors from this set is non-positive. By a standard geometric upper bound on the cardinality of such a set of vectors, we get the required upper bound on the number of codewords that are "close" to $\mathbf{r}$.

A idea extends proofs for the binary case, given by [4, 7, 1], where an appropriate embedding of the binary codewords in $\mathbb{R}^n$ and an appropriate shifting of vectors was used to establish "Johnson-style" bounds by appealing to bounds on spherical codes, i.e., bounds on the cardinality of a set of unit vectors in real space with a specified minimum angle between any pair of vectors. It may be noted that the generaliztion to the case of general alphabets is not canonical. Of the several potential approaches, our proof hits upon the right path.

Proof of Theorem 1: Assume without loss of generality that $\mathbf{r} = \langle q, q, \ldots, q \rangle$, i.e is the symbol q repeated n times. Let $C_1, C_2, \ldots, C_m$ be all the codewords of $\mathcal{C}$ that lie within $B_q(\mathbf{r}, e)$ where $e = (1 - 1/q)(1 - \gamma)n$. Our goal is to get an upper bound on m provided γ is large enough. The bound $m \leq \frac{1-\delta}{\gamma^2 - \delta}$ is well known, but we reprove it here since we can deduce it for almost free from the technique we use to establish the upper bound $m \leq nq$.

[3] Actually, Agrell *et al.* claim their result only for $A_2(n,d,e)$, but their proof will work for the case of $A'_2(n,d,e)$ as well.

We associate a vector in $\mathbb{R}^{nq}$ with $\mathbf{r}$ and with each codeword C_i. (Actually these vectors will all lie in an $n(q-1)$-dimensional subspace of $\mathbb{R}^{nq}$, and this will used later in the proof, but it is easiest to specify these vectors as embedded in $\mathbb{R}^{nq}$.) Each vector is to be viewed as having n blocks each having q components (the n blocks correspond to the n codeword positions). For $1 \leq l \leq q$, denote by $\hat{e}_l$ the q-dimensional unit vector with 1 in the lth position and 0 elsewhere. For $1 \leq i \leq m$, the vector $\mathbf{c_i}$ associated with the codeword C_i has in its jth block the components of the vector $\hat{e}_{C_i[j]}$ ($C_i[j]$ is the jth symbol of C_i, treated as an integer between 1 and q). The vector associated with the received word $\mathbf{r}$, which we also denote $\mathbf{r}$ by abuse of notation, is defined similarly. Let $\mathbf{1} \in \mathbb{R}^{nq}$ be the all 1's vector. Now define $\mathbf{v} = \alpha\mathbf{r} + \frac{(1-\alpha)}{q}\mathbf{1}$ for a parameter $0 \leq \alpha \leq 1$ to be specified later in the proof. Note that each $\mathbf{c_i}$ and $\mathbf{v}$ all lie in the space defined by the intersection of the n "hyperplanes" $\{\ \mathcal{H}'_j : \ \sum_{\ell=1}^{q} x_{j,\ell} = 1\ \}$ for $1 \leq j \leq n$. Hence the vectors $(\mathbf{c_i} - \mathbf{v})$, for $1 \leq i \leq m$, all lie in $\mathcal{H} = \bigcap_{j=1}^{n} \mathcal{H}_j$ where $\mathcal{H}_j = \{\mathbf{x} \in \mathbb{R}^{nq} : \sum_{\ell=1}^{q} x_{j,\ell} = 0\}$. It is easy to see that $\mathcal{H}$ is an $n(q-1)$-dimensional subspace of $\mathbb{R}^{nq}$. We thus conclude that the vectors $(\mathbf{c_i} - \mathbf{v})$, $1 \leq i \leq m$, all lie in an $n(q-1)$-dimensional space.

The idea behind the rest of the proof is the following. We will pick α so that the nq-dimensional vectors $(\mathbf{c_i} - \mathbf{v})$, for $1 \leq i \leq m$, have all pairwise dot products less than 0. Geometrically speaking, we shift the origin O to O' where $OO' = \mathbf{v}$, and require that relative to the new origin the vectors corresponding to the codewords have pairwise angles which are greater than 90 degrees (see Figure 1). By a simple geometric fact (stated in Lemma 3 below), it will then follow that the number of codewords m is at most the dimension $n(q-1)$ of the space in which these vectors all lie.

For $1 \leq i \leq m$, let $e_i = \Delta(\mathbf{r}, C_i)$. Note that $e_i \leq e$ for every i. Now

$$\langle \mathbf{c_i}, \mathbf{v} \rangle = \alpha\langle \mathbf{c_i}, \mathbf{r} \rangle + \frac{(1-\alpha)}{q}\langle \mathbf{c_i}, \mathbf{1} \rangle = \alpha(n - e_i) + (1-\alpha)\frac{n}{q} \geq \alpha(n-e) + (1-\alpha)\frac{n}{q} \tag{3}$$

$$\langle \mathbf{v}, \mathbf{v} \rangle = \alpha^2 n + 2(1-\alpha)\alpha\frac{n}{q} + (1-\alpha)^2\frac{n}{q} = \frac{n}{q} + \alpha^2\Big(1 - \frac{1}{q}\Big)n \tag{4}$$

$$\langle \mathbf{c_i}, \mathbf{c_j} \rangle = n - \Delta(C_i, C_j) \leq n - d\ . \tag{5}$$

Using (3), (4) and (5), we get for $i \neq j$

$$\langle \mathbf{c_i} - \mathbf{v}, \mathbf{c_j} - \mathbf{v} \rangle \leq 2\alpha e - d + \Big(1 - \frac{1}{q}\Big)(1-\alpha)^2 n \tag{6}$$

which using $e = (1-1/q)(1-\gamma)n$ and $d = (1-1/q)(1-\delta)n$ simplifies to

$$\langle \mathbf{c_i} - \mathbf{v}, \mathbf{c_j} - \mathbf{v} \rangle \leq \Big(1 - \frac{1}{q}\Big)n\Big(\delta + \alpha^2 - 2\alpha\gamma\Big) \tag{7}$$

Thus as long as $\gamma > \frac{1}{2}\Big(\frac{\delta}{\alpha} + \alpha\Big)$ we will have all pairwise dot products to be negative just as we wanted. We pick α to minimize $(\frac{\delta}{\alpha} + \alpha)$, or in other words we set $\alpha = \sqrt{\delta}$. Now as long as $\gamma > \sqrt{\delta}$, we will have $\langle \mathbf{c_i} - \mathbf{v}, \mathbf{c_j} - \mathbf{v} \rangle < 0$ for all $1 \leq i < j \leq m$. To complete the proof, we note that (for the choice $\alpha = \sqrt{\delta}$), for every $1 \leq i \leq m$, $\langle \mathbf{c_i} - \mathbf{v}, \mathbf{v} \rangle \geq (1-1/q)n\sqrt{\delta}(\gamma - \sqrt{\delta}) > 0$ provided $\gamma > \sqrt{\delta}$. Now applying Part (iii) of Lemma 3, with the setting $\mathbf{v_i} = \mathbf{c_i} - \mathbf{v}$ and $\mathbf{u} = \mathbf{v}_{|\mathcal{H}}$, the projection of $\mathbf{v}$ onto the subspace $\mathcal{H}$, implies that $m \leq n(q-1)$ (recall that the vectors $(\mathbf{c_i} - \mathbf{v})$, $1 \leq i \leq m$, all lie in $\mathcal{H}$ and $\dim(\mathcal{H}) = n(q-1)$).

We now prove that if $\gamma > \sqrt{\delta}$, then $m \leq \frac{1-\delta}{\gamma^2 - \delta}$. For this we set $\alpha = \gamma$. Now from Equation (7) we have $\langle \mathbf{c_i} - \mathbf{v}, \mathbf{c_j} - \mathbf{v} \rangle \leq (1-1/q)n(\delta - \gamma^2)$. Thus if $\gamma > \sqrt{\delta}$, we have $\langle \mathbf{c_i} - \mathbf{v}, \mathbf{c_j} - \mathbf{v} \rangle < 0$. Now for

each i, $1 \leq i \leq m$, we have $\|\mathbf{c_i}-\mathbf{v}\|^2 = \langle \mathbf{c_i}-\mathbf{v}, \mathbf{c_i}-\mathbf{v}\rangle \leq 2\alpha e + (1-1/q)(1-\alpha)^2 n = (1-1/q)n(1-\gamma^2)$ for the choice $\alpha = \gamma$. Denoting by $\mathbf{w_i}$ the unit vector $\frac{\mathbf{c_i}-\mathbf{v}}{\|\mathbf{c_i}-\mathbf{v}\|}$, we thus have

$$\langle \mathbf{w_i}, \mathbf{w_j}\rangle \leq -\frac{\gamma^2-\delta}{1-\gamma^2} \tag{8}$$

for $1 \leq i < j \leq m$. By a well-known geometric fact (see Lemma 4 for the simple proof), it follows that the number of such vectors, m, is at most $(1+\frac{1-\gamma^2}{\gamma^2-\delta}) = \frac{1-\delta}{\gamma^2-\delta}$, as desired.

To handle the case when $\gamma = \sqrt{\delta}$, we have $\langle \mathbf{c_i}-\mathbf{v}, \mathbf{c_j}-\mathbf{v}\rangle \leq 0$ for all $1 \leq i < j \leq m$, and also $\langle \mathbf{c_i}-\mathbf{v}, \mathbf{v}\rangle \geq 0$ for each $i = 1, 2, \ldots, m$. Now applying Part (ii) of Lemma 3, we get $m \leq 2n(q-1)-1$. □

2.3 Geometric Lemmas

We now state the geometric facts that were used in the above proof.

Lemma 3 *Let $\mathbf{v_1}, \ldots, \mathbf{v_m}$ be non-zero vectors in $\mathbb{R}^N$ such that $\langle \mathbf{v_i}, \mathbf{v_j}\rangle \leq 0$ for all $1 \leq i < j \leq m$. Then the following hold:*

(i) $m \leq 2N$.

(ii) Suppose that there exists a non-zero $\mathbf{u} \in \mathbb{R}^N$ such that $\langle \mathbf{u}, \mathbf{v_i}\rangle \geq 0$ for $i = 1, 2, \ldots, m$. Then $m \leq 2N-1$.

(iii) Suppose there exists an $\mathbf{u} \in \mathbb{R}^N$ such that $\langle \mathbf{u}, \mathbf{v_i}\rangle > 0$ for $i = 1, 2, \ldots, m$. Then $m \leq N$.

A proof of Part (i) of the above lemma can be found, for instance, in [2, Chapter 10, page 71]. The proofs of the other two parts are similar. For completeness, we present a self-contained proof of the above lemma in Appendix A.

Lemma 4 *Let $\varepsilon > 0$ be a positive real and let $\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_m}$ be m unit vectors such that $\langle \mathbf{w_i}, \mathbf{w_j}\rangle \leq -\varepsilon$ for all $1 \leq i < j \leq m$. Then $m \leq 1 + \frac{1}{\varepsilon}$.*

Proof: We have

$$0 \leq \langle \sum_{i=1}^{m} \mathbf{w_i}, \sum_{i=1}^{m} \mathbf{w_i}\rangle = \sum_{i=1}^{m} \langle \mathbf{w_i}, \mathbf{w_i}\rangle + 2 \sum_{1 \leq i < j \leq m} \langle \mathbf{w_i}, \mathbf{w_j}\rangle \leq m - m(m-1)\varepsilon \ ,$$

which gives $m \leq 1 + 1/\varepsilon$. □

2.4 Generalization in Presence of Weights

For applications to "soft-decision" list decoding algorithms, it is of interest to prove a version of the Johnson bound in the presence of weights on codeword symbols. A version of such a bound appears for instance in [12]. Here we state the weighted version of the Johnson bound that follows from our proof technique. The bound in Part (i) of the theorem generalizes the result of Corollary 2. The result from Part (ii) applies under a more general condition than Condition (1) (or even Condition (9)), but the upper bound is itself is slightly weaker (since it is $(nq-1)$ instead of $n(q-1)$).

Theorem 5 *Let $\mathcal{C} \subseteq [q]^n$ be a code of blocklength n and minimum distance d. Let $\{w_{i,j} : 1 \leq i \leq n; 1 \leq j \leq q\}$ be an arbitrary set of non-negative real weights. Define $W_i = \sum_{j=1}^{q} w_{i,j}$ and $W_i^{(2)} = \sum_{j=1}^{q} w_{i,j}^2$, $W_{\rm tot} = \sum_{i,j} w_{i,j}$, and $W_{\rm tot}^{(2)} = \sum_{i,j} w_{i,j}^2$. Then:*

(i) The number of codewords $C \in \mathcal{C}$ that satisfy

$$\sum_{i=1}^{n} \frac{w_{i,C_i}}{W_i} > \frac{n}{q} + \sqrt{\Big(n(1-\frac{1}{q}) - d\Big)\Big(\sum_{i=1}^{n} \frac{W_i^{(2)}}{W_i^2} - \frac{n}{q}\Big)} \ . \tag{9}$$

is at most $n(q-1)$.

(ii) The number of codewords $C \in \mathcal{C}$ that satisfy

$$\sum_{i=1}^{n} w_{i,C_i} > \frac{W_{\rm tot}}{q} + \sqrt{\Big(n(1-\frac{1}{q}) - d\Big)\Big(W_{\rm tot}^{(2)} - \frac{(W_{\rm tot})^2}{nq}\Big)} \tag{10}$$

is at most $(nq-1)$.

(iii) For any integer $L \geq 2$, the number of codewords $C \in \mathcal{C}$ that satisfy

$$\sum_{i=1}^{n} w_{i,C_i} \geq \frac{W_{\rm tot}}{q} + \sqrt{\Big(n(1-\frac{1}{q}) - d + \frac{d}{L}\Big)\Big(W_{\rm tot}^{(2)} - \frac{(W_{\rm tot})^2}{nq}\Big)} \tag{11}$$

is at most L.

Proof: We do not give a full proof here, rather we indicate the only changes that must be made to the proof of Theorem 1 in order to prove our claim. For Part (i), the only modification required in the proof of Theorem 1 is to pick $\mathbf{r}$ so that its (i,j)'th component, for $1 \leq i \leq n$ and $1 \leq j \leq q$, equals $\frac{w_{i,j}}{W_i}$. The vector $\mathbf{v}$ is defined as before to be $\alpha \mathbf{r} + \frac{(1-\alpha)}{q}\mathbf{1}$ for

$$\alpha = \sqrt{\frac{n(1-1/q) - d}{\sum_i \frac{W_i^{(2)}}{W_i^2} - n/q}} \ .$$

Once once again all the vectors $(\mathbf{c_i} - \mathbf{v})$ lie in an $n(q-1)$-dimensional subspace of $\mathbb{R}^{nq}$. It can be proved as in the proof of Theorem 1 that these vectors have pairwise non-positive dot products, which gives the desired $n(q-1)$ upper bound on the number of codewords.

For Parts (ii) and (iii), we pick $\mathbf{r}$ so that its (i,j)'th component for $1 \leq i \leq n$ and $1 \leq j \leq q$, equals $\frac{n w_{i,j}}{W_{\rm tot}}$, and the rest of the proof follows that of Theorem 1. Note that $W_{\rm tot}/q$ is the expected value of $\sum_i w_{i,r_i}$ for a random vector $\mathbf{r} \in [q]^n$, and $(W_{\rm tot}^{(2)} - \frac{(W_{\rm tot})^2}{nq})$ is proportional to the variance of the $w_{i,j}$'s. Thus, the above theorem states that the number of codewords which have weighted agreement bounded away from the expectation by a certain number of standard deviations is small. The upper bound of $(nq-1)$ (instead of $n(q-1)$) in Part (ii) of above theorem arises since we are only able to ensure that the vectors $(\mathbf{c_i} - \mathbf{v})$ all lie in an $(nq-1)$-dimensional subspace (namely that defined by $\sum_{i,j} x_{i,j} = 0$), and not an $n(q-1)$-dimensional subspace as in Part (i). $\Box$

A bound similar to Theorem 5 can also be worked out for the case when the different codeword symbols belong to different alphabets (say the ith symbol belongs to an alphabet of size q_i). Such a bound is of interest for certain codes like the Chinese Remainder Code [12].

Acknowledgments

We would like to thank an anonymous referee for crucial pointers to the works [1, 4, 7] and numerous other suggestions.

References

[1] E. Agrell, A. Vardy and K. Zeger. Upper bounds for constant-weight codes. *IEEE Transactions on Information Theory*, **46** (2000), pp. 2373-2395.

[2] B. Bollobás. *Combinatorics*, Cambridge University Press, Cambridge, U.K, 1986.

[3] P. Elias. Error-correcting codes for List decoding. *IEEE Trans. Info. Theory*, **37** (1), pp. 5-12, 1991.

[4] T. Ericson and V. Zinoviev. Spherical codes generated by binary partitions of symmetric pointsets. *IEEE Trans. on Information Theory*, **41** (1995), pp. 107-129.

[5] S. M. Johnson. A new upper bound for error-correcting codes. *IEEE Trans. on Info. Theory*, **8** (1962), pp. 203-207.

[6] S. M. Johnson. Improved asymptotic bounds for error-correcting codes. *IEEE Trans. on Info. Theory*, **9** (1963), pp. 198-205.

[7] V. I. Levenshtein. Universal bounds for codes and designs. Chapter 6 in *Handbook of Coding Theory*, V. S. Pless and W. C. Huffman (eds), Elsevier, 1998, pp. 499-648.

[8] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*, North-Holland, Amsterdam, 1981.

[9] *Handbook of Coding Theory*, Volume I, V. S. Pless and W. C. Huffman, Editors, North-Holland, 1998.

[10] O. Goldreich, R. Rubinfeld and M. Sudan. Learning polynomials with queries: the highly noisy case. *Proc. of FOCS 95.*

[11] V. Guruswami and M. Sudan. List decoding algorithms for certain concatenated codes. *Proc. of the 32nd Annual ACM Symposium on Theory of Computing*, May 2000, pp. 181-190.

[12] V. Guruswami, A. Sahai and M. Sudan. "Soft-decision" decoding of Chinese Remainder Codes. *Proc. of the 41st IEEE Symposium on Foundations of Computer Science*, November 2000, pp. 159-168.

[13] M. Sudan. List Decoding: Algorithms and Applications. *SIGACT News*, **31** (2000), pp. 16-27.

A Proof of Lemma 3

Proof: We first prove (iii). Suppose for contradiction that $m \geq N + 1$. Then since the vectors $\mathbf{v_1}, \ldots, \mathbf{v_m}$ all lie in $\mathbb{R}^N$, they must be linearly dependent. Let $S \subseteq [m]$ be a non-empty set of *minimum* size for which a relation of the form $\sum_{i \in S} a_i \mathbf{v_i} = \mathbf{0}$ holds with each $a_i \neq 0$. We claim that the a_i's must all be positive or all be negative. Indeed, if not, by collecting terms with positive

a_i's on one side and those with negative a_i's on the other, we will have an equation of the form $\sum_{i\in T^+} a_i\mathbf{v_i} = \sum_{j\in T^-} b_j\mathbf{v_j} = \mathbf{w}$ (for some vector $\mathbf{w}$) where T^+ and T^- are *disjoint* non-empty sets with $T^+ \cup T^- = S$, and all $a_i, b_j > 0$. By the minimality of S, $\mathbf{w} \neq \mathbf{0}$ and hence $\langle \mathbf{w}, \mathbf{w}\rangle > 0$. On the other hand $\langle \mathbf{w}, \mathbf{w}\rangle = \langle \sum_{i\in T^+} a_i\mathbf{v_i}, \sum_{j\in T^-} b_j\mathbf{v_j}\rangle = \sum_{i,j} a_ib_j\langle \mathbf{v_i}, \mathbf{v_j}\rangle \leq 0$ since $a_ib_j > 0$ and $\langle \mathbf{v_i}, \mathbf{v_j}\rangle \leq 0$ for each $i \in T^+$ and $j \in T^-$. This contradiction shows that we may assume that $a_i > 0$ for all $i \in S$.

Now $\sum_{i=1}^s a_i\mathbf{v_i} = \mathbf{0}$, so that $\sum_{i=1}^s a_i\langle \mathbf{u}, \mathbf{v_i}\rangle = 0$. But this is impossible since for each i we have $a_i > 0$ and $\langle \mathbf{u}, \mathbf{v_i}\rangle > 0$. We have thus arrived at a contradiction, and therefore we must have $m \leq N$.

To prove (ii), we use induction on N. The statement clearly holds for $N = 1$. For $N > 1$, we proceed exactly as above. If $m \leq N$, we have nothing to prove, so assume $m > N$ so that $\mathbf{v_1}, \ldots, \mathbf{v_m}$ are linearly independent, and as above, let $S \subseteq [m]$ be a non-empty set of minimum size for which a relation of the form $\sum_{i\in S} a_i\mathbf{v_i} = \mathbf{0}$ holds with each $a_i \neq 0$. Assume for definiteness that $S = \{1, 2, \ldots, s\}$. We thus have the linear dependence $\sum_{i=1}^s a_i\mathbf{v_i} = \mathbf{0}$ with each $a_i > 0$, and since this is a minimum sized linear dependence, $\mathbf{v_1}, \ldots, \mathbf{v_s}$ must span a subspace W of $\mathbb{R}^N$ of dimension $(s-1)$.

Since $\sum_{i=1}^s a_i\mathbf{v_i} = \mathbf{0}$, we have $\sum_{i=1}^s a_i\langle \mathbf{v_i}, \mathbf{v}_\ell\rangle = 0$ for each $\ell = s+1, \ldots, m$. Since $a_i > 0$ for $1 \leq i \leq s$ and $\langle \mathbf{v_i}, \mathbf{v}_\ell\rangle \leq 0$, it must be therefore be the case that $\mathbf{v_i}$ is orthogonal to $\mathbf{v}_\ell$ for all i, ℓ with $1 \leq i \leq s$ and $s < \ell \leq m$. A similar argument shows $\mathbf{u}$ is orthogonal to $\mathbf{v_i}$ for each $i = 1, 2, \ldots, s$. Thus the vectors $\mathbf{v_{s+1}}, \ldots, \mathbf{v_m}$ and $\mathbf{u}$ all lie in $W^\perp$ which has dimension equal to $(N - s + 1)$. Since $s > 1$, the induction hypothesis applied to these vectors implies that $m - s \leq 2(N - s + 1) - 1$, or in other words $m \leq 2N - s + 1 \leq 2N - 1$, as desired.

Finally (i) follows immediately from (ii). Indeed, apply (ii) with vectors $\mathbf{v_1}, \ldots, \mathbf{v_{m-1}}$ and $-\mathbf{v_m}$ playing the role of $\mathbf{u}$. This implies $m - 1 \leq 2N - 1$, or in other words $m \leq 2N$. □