Error-Correcting Codes for List Decoding

Peter Elias, *Fellow, IEEE*

Abstract—**In the list-of-L decoding of a block code the receiver of a noisy sequence lists L possible transmitted messages, and is in error only if the correct message is not on the list. This paper considers (n, e, L) codes, which correct all sets of e or fewer errors in a block of n bits under list-of-L decoding. New geometric relations between the number of errors corrected under list-of-1 decoding and the (larger) number corrected under list-of-L decoding of the same code lead to new lower bounds on the maximum rate of (n, e, L) codes. They show that a jammer who can change a fixed fraction $p < 1/2$ of the bits in an n-bit linear block code cannot prevent reliable communication at a positive rate using list-of-L decoding for sufficiently large n and an $L \leq n$. The new bounds are stronger for small n but weaker for fixed e/n in the limit of large n and L than known random coding bounds.**

Index Terms—**List decoding, error correcting codes, correcting nearly $n/2$ errors.**

I. History and Literature

IN THE LIST-OF-L decoding of a block code the receiver of a noisy sequence of n symbols lists not one but L possible transmitted messages. The decoding is in error only if the correct message is not on the list.

List decoding was introduced independently by Elias and by Wozencraft, who both used random coding arguments to explore the average decoding error probability of block codes at low rates for the binary symmetric channel [1], [2]. It was used by Shannon, Gallager, and Berlekamp in exploring low rate average error bounds for general discrete memoryless channels [3]; their results are cited in the texts [4]–[6]. Ahlswede used list decoding with and without noiseless feedback in finding the capacities of a broader class of channels, some with memory, in [7], and Csiszar and Körner discussed those and related topics in [8].

This paper, like [9] and [10], gives results for worst-case, not average, error-correcting behavior. It defines (n, e, L) codes as codes which, under list-of-L decoding, correct all sets of e or fewer errors per block of n binary symbols. It bounds the maximal rates of such codes both for fixed n, e, and L and in the limit of large n with e/n and L fixed. The limiting rates are matched to a model of a jammer who knows the code in use and the message being sent and can change a fixed fraction e/n of the symbols in each block, whatever block size the transmitter and receiver choose to use.

Finding best codes for fixed n, e, and L and maximal limiting rates for fixed e/n and L as n increases are generalizations of principal tasks in the theory of error-correcting codes, as presented e.g., in the encyclopedic work [13], which considers the case $L = 1$. Finding the limiting rate for fixed e/n is like a zero-error capacity problem since if some error pattern within the constraints on the jammer causes a list-of-L decoding error, the jammer will see to it that that error pattern occurs. However, it is not exactly a zero-error problem since the set of error patterns the jammer can introduce increases with the block size; so, the jammer is not a channel in the ordinary sense.

The performance of a different but related jamming model under list decoding is analyzed in [11], [12]. There the jammer also knows the code in use and the message being sent but the block size is $N = kn$ and the jammer can only change e bits in each n-bit subblock, while the transmitter and receiver do their coding and list-of-L decoding only after all k subblocks have been received. Maximal rates for fixed n, k, e, and L reduce to those of the first model at $k = 1$, but a different limit is of interest: the limit of large k with fixed e, n, and L. In that case the limit is precisely the zero-error capacity of a certain noisy channel under list-of-L decoding.

Zyablov and Pinsker in [9] and Blinovskii in [10] give random coding lower bounds to the asymptotic rates attainable by (n, e, L) codes with fixed e/n and L, averaging in [9] over an expurgated set of linear codes and in [10] over an expurgated set of all codes. Blinovskii's bounds on maximal rate are considerably stronger for finite L: the two agree in the limit of large L. The present paper uses geometric rather than random coding arguments. Its bounds on code sizes and rates are tighter for small n but weaker in the limit of large n and L with fixed e/n than those in [9], [10]. However, when the jammer can change almost half of the transmitted symbols, the random coding bounds require much larger lists to guarantee reliable communication with a linear code at a positive rate than do the bounds forthcoming.

Both [12] and this paper were stimulated by Levin, who reawakened my early interest in this subject matter by asking questions about channel capacity under jamming related to his cryptographic research [14], and who noted the implication just mentioned, that positive-rate linear codes with small lists can correct nearly $n/2$ errors [15]. Other recent work dealing with zero-error capacity appears in [16], which treats some problems in asymptotic combinatorics as zero-error capacity problems. The authors mention unpublished work by Körner and Marton that includes an analysis of perfect hashing as a zero-error capacity problem under list-of-L decoding.

II. Results

Section II defines general and linear codes, list-of-L decoding, (n, e, L) codes and the rate $R_L(\gamma)$ of an (n, e, L) code γ under list-of-L decoding, and gives some immediate consequences of those definitions. Section IV generalizes the Hamming and Gilbert–Varshamov bounds to list-of-L decoding in the case of finite n and L (Propositions 4 and 7), defines the asymptotic rate $r_L(p)$ attainable by $(n, \lfloor np \rfloor, L)$

Manuscript received February 14, 1989; revised June 10, 1990. This work was supported in part by DARPA Contract N00014-89-J-1988.

The author is with the Department of Electrical Engineering and Computer Science and the Laboratory for Computer Science at MIT, 77 Massachusetts Avenue, Cambridge, MA 02139.

IEEE Log Number 9040133.

codes for arbitrarily large n, fixed L and $0 \le p \le 1/2$, and $r_\infty(p)$, the limit of $r_L(p)$ as $L \to \infty$, and in Proposition 8 shows that the asymptotic Hamming and Gilbert–Varshamov bounds on $r_L(p)$ in (12), well known for $L = 1$, hold for all integers L and at $L = \infty$.

The principal new results appear in Section V. Proposition 9 notes that a code which corrects e errors out of n using list-of-1 decoding corrects more errors using list-of-L decoding for $L > 1$. Proposition 10 uses essentially known bounds to define a trading relation between list size and error correcting capability. Propositions 11 and 12 show that if $2 \le L \le n$, $0 \le p_1 \le 1/2$ and np_1, np_L, and np_∞ are all integers related by

$$p_\infty = p_1(1 - p_1), \qquad p_L = p_\infty L/(L-1), \tag{1}$$

then an $(n, np_L, 1)$ code is also an (n, np_1, L) code, and an $(n, np_\infty, 1)$ code is also an (n, np_1, n) code. Appendix A explores the behavior under list-of-L decoding of some codes known to correct e errors under list-of-1 decoding: Proposition 17 notes in part that an $(n, e, 1)$ code is also an $(n, e+1, \lfloor (n+1)/(e+1) \rfloor)$ code and Propositions 18–20 apply that result to the Hamming and Golay error correcting codes. Some open questions are mentioned at the end of Appendix A.

The relations (1) apply not only to known codes which correct e errors out of n under list-of-1 decoding but to codes which satisfy the Gilbert–Varshamov bound at $L = 1$, which are not known but are known to exist for arbitrarily large n. The maximal rate attainable under list-of-L decoding must be at least as great as the rate of such codes. Proposition 13 shows by that argument, among other things, the bounds

$$1 - H(p) \ge r_\infty(p) \ge 1 - H(2pq) > 0, \qquad 0 \le p < 1/2 \tag{2}$$

for the limiting rates of both general and linear codes which correct $\lfloor np \rfloor$ errors out of n in the limits of large n and L.

The lower bound in (2) is not best possible, but is tight enough to show that using large-list decoding against a jammer who can change at most np symbols per block of n, communication at a positive rate is possible as long as $p < 1/2$. This is different from list-of-1 decoding, for which it is well known that communication at a positive rate requires $p < 1/4$ (see [13], Chapter 17). Proposition 14 shows e.g., that if $\epsilon > 0$ then for both general and linear codes, $r_\infty((1-\epsilon)/2) \ge c\epsilon^4$, and that in both cases the rate under list-of-L decoding is positive if $L > 1/\epsilon^2$. Proposition 15 summarizes the random-coding lower bound given for linear codes in [9] and analogous results for general codes, which are proved in Appendix B. Proposition 16 shows that the random-coding results give the stronger lower bound $r_\infty((1-\epsilon)/2) \ge c\epsilon^2$ for both linear and general codes, and that for general codes the rate is positive if $L > c/\epsilon^2$, and that (as noted by Levin [15]) the random coding argument in [9] requires a list size *exponential* in $1/\epsilon^2$ to guarantee a positive rate. An open question about the typical performance of linear codes is mentioned at the end of Appendix B.

III. Binary Error-Correcting Codes and List Decoders

Let n and L be positive integers. Let e and d (with or without subscripts or arguments) denote nonnegative integers $\le n$.

Let $\mathscr{B} = \{0,1\}^n$, the set of n-dimensional binary vectors, under modulo 2 addition. Let $h(x, y)$ denote the Hamming distance between $x, y \in \mathscr{B}$. Let $\mathscr{S}_e^n(y) = \{z \in \mathscr{B} | h(z, y) \le e\}$, the sphere of radius e (e-sphere for short) about y, and let $V_e^n = |\mathscr{S}_e^n(y)| = 1 + \binom{n}{1} + \cdots + \binom{n}{e}$, the volume of an e-sphere.

A binary code γ of length n and size $M = M(\gamma)$ is a function $\gamma: \mathscr{M} \to \mathscr{B}$, which assigns an n-bit codeword $\gamma(i)$ to each integer message i in the message set $\mathscr{M} = \{0, 1, \cdots, M-1\}$. Let $d = d(\gamma)$ be the minimum of the distances $h(\gamma(i), \gamma(j))$ between codewords representing distinct messages $i, j \in \mathscr{M}$ ($d = 0$ if some codeword represents two messages).

For integer k some of the binary codes of size $M = 2^k$ and length n are linear. A linear code γ has a binary matrix G with k rows and n columns, whose jth row g_j is called the jth generator, $0 \le j \le k-1$. The set $C(\gamma) = \gamma\mathscr{M}$ of codewords is the row-space of G: message $i = i(0) + i(1)2^1 + \cdots + i(k-1)2^{k-1}$ has codeword $\gamma(i) = \sum_{j=0}^{k-1} i(j) g_j$; so $g_j = \gamma(2^j)$, and $\gamma(0)$ is the all-0 sequence.

A list decoder for γ is a function $\delta: \mathscr{B} \to 2^{\mathscr{M}}$, which assigns a subset $\delta(y)$ of $\mathscr{M}$ to each $y \in \mathscr{B}$. When message i is encoded and codeword $\gamma(i)$ is sent over a channel subject to bit errors and sequence y is received, the decoding $\delta(y)$ is correct if $i \in \delta(y)$ and otherwise is in error. δ corrects e errors with list size L for γ if

$$\forall\, i \in \mathscr{M} \text{ and } y \in \mathscr{S}_e^n(\gamma(i)), \qquad i \in \delta(y) \text{ and } |\delta(y)| \le L; \tag{3}$$

and γ is an (n, e, L) code (or is simply (n, e, L)) if it has a decoder that satisfies (3).

The rate of a code γ under list-of-1 decoding is $(1/n)\log M(\gamma)$ (logarithms are base 2 throughout). The rate of γ under list-of-L decoding is defined as

$$R_L(\gamma) = \begin{cases} 0, & M(\gamma) \le L \\ (1/n)\log(M(\gamma)/L), & M(\gamma) \ge L \end{cases} \tag{4}$$

bits per symbol since $\max\{\log(M(\gamma)), \log L\}$ bits are required after correct list decoding to decide which of the L (or $M(\gamma) < L$) messages on the list was actually sent. This rate definition is standard for $M(\gamma) \ge L$ [1,2,5]: for any e it assigns rate 0 to an (n, e, L) code if $L \ge M$, which is appropriate since the constant decoder $\delta(y) = \mathscr{M}$ is list-of-L and decodes such a code correctly with no channel. It is also convenient to define the (not necessarily integer) *effective size*:

$$M_L(\gamma) = 2^{nR_L(\gamma)} = \begin{cases} 1, & M(\gamma) \le L \\ M(\gamma)/L, & M(\gamma) \ge L \end{cases} \tag{5}$$

of the code γ for list-of-L decoding. By (3), $\delta(y) \subseteq \mathscr{M}$ and each message appears on at most 2^n lists; so $M(\gamma) \le 2^n L$, $M_L(\gamma) \le 2^n$ and

$$1 \ge R_1(\gamma) \ge R_L(\gamma) = \max\{R_1(\gamma) - (1/n)\log L, 0\}. \tag{6}$$

By the definition (3), an (n, e, L) code also corrects $\le e$ errors using lists of size $\ge L$. Proposition 1 records that fact for reference.

Proposition 1: γ is (n, e', L') if it is (n, e, L) and $e' \le e$, $L' \ge L$.

By Proposition 1 an $(n, e, 1)$ code γ of size $M_1(\gamma)$ is also an (n, e, L) code, of the same size but by (6) of smaller rate and effective size. To construct an (n, e, L) code γ' of the same

rate and effective size assign L messages to each codeword in γ, e.g., setting $\gamma'(j) = \gamma(j \bmod M(\gamma))$ for $0 \le j < LM(\gamma) - 1$. Since γ is $(n, e, 1)$, by (3) each e-sphere contains representations of ≤ 1 message in γ, and thus of $\le L$ messages in γ'; so γ' is (n, e, L) of effective size $M_L(\gamma') = M_1(\gamma)$ and rate $R_L(\gamma') = R_1(\gamma)$. If γ is linear, adding $\lfloor \log L \rfloor$ rows of 0's to its generator matrix produces a linear $(n, e, 2^{\lfloor \log L \rfloor})$ code γ'', which by Proposition 1 is also (n, e, L), with rate $R_L(\gamma'') = (1/n)\log(2^{\lfloor \log L \rfloor} M/L)$, proving Proposition 2.

Proposition 2: Let γ be an $(n, e, 1)$ code with size $M_1(\gamma)$ and rate $R_1(\gamma)$. Then

a) There is an (n, e, L) code γ' with the same codeword set, effective size $M_L(\gamma') = M_1(\gamma)$, and rate $R_L(\gamma') = R_1(\gamma)$.
b) If γ is linear, there is also a linear (n, e, L) code γ'' with size $M_L(\gamma'') = 2^{-\theta} M_1(\gamma)$ and rate $R_L(\gamma'') = R_1(\gamma) - \theta/n$, where $0 \le \theta = \log L - \lfloor \log L \rfloor < 1$.

There is no analog to Proposition 2 that allows reducing rather than increasing L at no cost in $R_L(\gamma)$. In fact, in general increasing L increases rate. Stronger results than Proposition 2 are given in Section V.

Proposition 3 notes for reference some obvious geometric properties of (n, e, L) codes.

Proposition 3:

a) A code γ is (n, e, L) iff no e-sphere about any $y \in \mathscr{B}$ contains codewords that between them represent more than L messages, that is, iff the decoder δ defined for $y \in \mathscr{B}$ by

$$\delta(y) = \{i | \gamma(i) \in \mathscr{S}_e^n(y)\} = \{i | y \in \mathscr{S}_e^n(\gamma(i))\} \quad (7)$$

satisfies

$$|\delta(y)| \le L, \qquad \text{for all } y \in \mathscr{B}. \quad (8)$$

b) If γ is (n, e, L) the δ in (7) is minimal, in the sense that a decoder δ' corrects e errors with list size L for γ iff, for all $y \in \mathscr{B}$, $\delta'(y) \supseteq \delta(y)$ and $|\delta'(y)| \le L$.

Proposition 3 b) is used without proof by the authors of [9], [10]. The (obvious) proof is the observation that (3) holds if (7) and (8) do, and that (7) and (8) fail only if more than L messages lie within e of some y. But then any list $\delta'(y)$ containing only L of them is in error when an $i \in \delta(y) - \delta'(y)$ is selected, $\gamma(i)$ is sent and y is received, so (3) fails. Therefore the decoder in (7) is e-error correcting and minimal for γ. Proposition 3 b) is immediate.

Substituting in (8) the first expression for $\delta(y)$ in (7) gives the defining geometric property cited in Proposition 3 a). Substituting the second expression gives another property, more commonly used at $L = 1$: γ is $(n, e, 1)$ iff e-spheres about codewords representing different messages are disjoint; so the construction of an $(n, e, 1)$ code is a sphere-packing problem. For general L the second property is more complex: γ is (n, e, L) iff no point in $\mathscr{B}$ is covered more than L times by the M e-spheres centered on the codewords that represent each message. A third property equivalent at $L = 1$ has to do with minimum distance: the condition $d(\gamma) \ge 2e + 1$ is necessary and sufficient for γ to be $(n, e, 1)$. Minimum distance, unlike error-correction capability, seems to have no natural generalization to $L > 1$. However, the classical upper (Hamming) and lower (Gilbert–Varshamov) bounds on rate, whose proofs for $L = 1$ use distance properties, generalize nicely to $L > 1$.

IV. The Generalized Hamming and Gilbert–Varshamov Bounds

The Hamming, or volume, bound is given, e.g., as Theorem 6, Chapter 1 in [13]. Let γ be (n, e, L), and let T be the number of occurrences of all messages on all lists. By Proposition 3 each message appears on at least V_e^n lists, and there are 2^n lists of at most L members each; so $M(\gamma)V_e^n \le T \le L2^n$. That with the rate definition (4) proves the following.

Proposition 4—The generalized Hamming upper bound: If γ is (n, e, L) then $M_L(\gamma) \le 2^n / V_e^n$, $R_L(\gamma) \le 1 - (1/n)\log V_e^n$.

Two implications of Proposition 4 are worth noting. First, finding maximal-rate codes for given n and e is easy when L is sufficiently large. By Proposition 3, if $L = V_e^n$ then for each x in $\mathscr{B}$ the decoder $\delta(x) = S_e^n(x)$ is list-of-L and corrects all sets of e or fewer errors. Assigning k messages rather than 1 to each codeword and increasing the list length from L to kL, as in the proof of Proposition 2, keeps the effective size constant while increasing L, and proves the following.

Proposition 5: If $L = kV_e^n$ for some positive integer k then there is an (n, e, L) code that attains the Hamming bound.

Second, at $L = 1$ obvious geometrical arguments show that a code with $n = 2e + 1$ has at most two messages, and that a code that corrects $n/2$ errors or more has one message and rate 0. For $L > 1$ the Hamming bound and the symmetry properties of the binomial coefficients imply the first of these properties and a generalization of the second, proving the following.

Proposition 6: Let γ be (n, e, L). If $n = 2e + 1$ then $M_L(\gamma) \le 2$. If $n \le 2e$ then $M_L(\gamma) < 2$.

Comment: The bound $M_L(\gamma) = 2$ is attained at $n = 2e + 1$ for any L by using each of the codewords $0^n, 1^n$ to represent L messages. $M_L(\gamma) < 2$ does not imply $M_L(\gamma) = 1$ since for $L > 1$ the effective size in (5) need not be an integer. Proposition 5 produces many examples of codes with $e < n < 2e + 1$ and effective size between 1 and 2. Transmitting a bit of information in such a case requires encoding each message into a sequence of subblocks of size n, i.e., changing the model discussed in Section I to $k > 1$ and analyzing a zero-error channel capacity problem as in [12].

The classical Gilbert and Varshamov lower bounds for $L = 1$ are extended in Proposition 7 to general L.

Proposition 7—The generalized Gilbert–Varshamov lower bound: Let $2e \le n$. Then

a) there is a linear $(n, e, 1)$ code γ of size $M_1(\gamma) \ge 2^{n-1} / V_{2e-1}^{n-1} \ge 2^n V_{2e}^n$;
b) there is a linear (n, e, L) code γ' with $M_L(\gamma') = 2^{-\theta} M_1(\gamma)$, $R_L(\gamma') = R_1(\gamma) - \theta/n$, where $0 \le \theta = \log L - \lfloor \log L \rfloor < 1$.

Proof: The first inequality in a) is Varshamov's, and follows from Theorem 12, Chapter 1 in [13]. The Pascal triangle construction of the binomial coefficients shows that $2V_{2e-1}^{n-1} \le V_{2e}^n$ with equality at $2e = n$, proving the second inequality, which is Gilbert's. Part b) follows by Proposition 2. □

The asymptotic versions of the classic bounds for codes which correct $\lfloor np \rfloor$ errors out of n in the limit of large n may also be generalized.

Let $\Gamma(n,e,L)$ be the set of all (n,e,L) codes and $\Gamma'(n,e,L)$ the linear subset. Define the maximal rates

$$\rho(n,e,L) = \max_{\gamma \in \Gamma(n,e,L)} R_L(\gamma),$$
$$\rho'(n,e,L) = \max_{\gamma \in \Gamma'(n,e,L)} R_L(\gamma). \quad (9)$$

The limit of $\rho(n,\lfloor np \rfloor, L)$ as $n \to \infty$ is not known to exist, but the sequence of rates is bounded above by 1 in (6), so it has a limit superior. Let

$$r_L(p) = \limsup_{n\to\infty} \rho(n,\lfloor np \rfloor, L), \qquad r_\infty(p) = \lim_{L\to\infty} r_L(p), \quad (10)$$

and similarly $r'_L(p), r'_\infty(p)$. As $L \to \infty$ the limits $r_\infty(p), r'_\infty(p)$ exist and are nonincreasing in p and bounded by 0 and 1 since, by Proposition 1 and (6), $\rho(n,\lfloor np \rfloor, L), \rho'(n,\lfloor np \rfloor, L)$ have those properties.

For $p \le 1/2$ and $q = 1-p$, the inequalities

$$H(p) - (1/2n)\log(8npq) \le (1/n)\log V^n_{\lfloor np \rfloor} \le H(p) \quad (11)$$

hold, where H is the entropy function, $H(p) = -p\log p - q\log q$. (See, e.g., [13], Chapter 10, Corollary 9.) Using (11) in Propositions 4 and 7 extends the classic asymptotic bounds on rate to arbitrary L. Proposition 8 gives the results.

Proposition 8: Let $p \le 1/2$, $q = 1-p$. Then $r_L(p)$ and $r'_L(p)$ are nondecreasing in L and nonincreasing in p and converge as L increases to limit functions $r_\infty(p), r'_\infty(p)$ that are nonincreasing in p, with $0 \le r'_L(p) \le r_L(p) \le 1$, and for $1 \le L \le \infty$,

$$1 - H(p) \ge r_L(p) \ge r'_L(p) \ge \begin{cases} 1 - H(2p), & 0 \le p \le 1/4 \\ 0, & 1/4 \le p \le 1/2 \end{cases}. \quad (12)$$

V. List Error-Correction Capabilities of $(n,e,1)$ Codes

The principal new results follow from the fact that an $(n,e,1)$ code, which has minimum distance $d = 2e+1$ and corrects e errors using list-of-1 decoding, corrects more errors when it is used with list-of-L decoding. This fact leads to stronger versions of Propositions 7 and 8.

Let $A(n,d,e)$ denote the maximum number of points that may be placed in one e-sphere with pairwise distances $\ge d$. If a code γ has minimum distance d then at most $A(n,d,e)$ of its codewords lie in any e-sphere. It follows by Proposition 3 that γ is an $(n,e,A(n,d,e))$ code, proving Proposition 9—an obvious but apparently unnoticed result.

Proposition 9: A code γ of length n and minimum distance d is an $(n,e,A(n,d,e))$ code.

If γ has minimum distance $d > 2e$ then any e-sphere contains at most one codeword. Proposition 10 records this and other bounds to $A(n,d,e)$.

Proposition 10:

a) If $n \ge d \ge 2e+1$ then $A(n,d,e) = 1$.

b) If $e > 0$ and $n \ge 2e \ge d > 2e(n-e)/n$ then $n \ge 2$ and

$$2 \le A(n,d,e) \le \left\lfloor \frac{nd}{nd - 2e(n-e)} \right\rfloor \le \max\{2, n(n-1)/2\}. \quad (13)$$

c) If d is odd then $A(n,d,e) = A(n+1,d+1,e)$. Therefore for odd d, if $e > 0$ and $n+1 \ge 2e \ge d+1 > 2e(n+1-e/(n+1))$, then $n \ge 1$ and

$$A(n,d,e) \le \left\lfloor \frac{(n+1)(d+1)}{(n+1)(d+1) - 2e(n+1-e)} \right\rfloor \le \{\max 2, n(n+1)/2\}. \quad (14)$$

d) If d is odd and the conditions in the first line of b) hold then $A(n,d,e) \le n$.

Proof: The leftmost inequality in (13) is obvious since e.g., the two sequences $0^{n-e}1^e$ and 1^e0^{n-e} are in $S^n_e(0)$ and have distance $d = 2e$ if $2e \le n$. The first published proof of the second inequality in (13) is by Johnson in [17]. It shows that the rightmost term in (13) bounds the number of codewords of length n with pairwise distances $\ge d$ *on the surface* $\mathscr{S}^n_e(\mathbf{0}) - \mathscr{S}^n_{e-1}(\mathbf{0})$ of the e-sphere $\mathscr{S}^n_e(\mathbf{0})$. A more accessible proof of that result is in [13]. A simple and complete proof of the (also known) result that the second inequality in (13) bounds the number of codewords in the whole e-sphere $\mathscr{S}^n_e(\mathbf{0})$ seems not to be available, but is quite short.

Let $\gamma(0), \gamma(1), \cdots$ be A codewords with an average of αn 1's and $(1-\alpha)n$ 0's each and pairwise distances $\ge d$, with $\alpha \le e/n \le 1/2$. Then

$$A(A-1)d/2 \le \sum_{0 \le i < j < A} h(\gamma(i), \gamma(j)) \le nA^2\alpha(1-\alpha) \le nA^2(e/n)(1-e/n). \quad (15)$$

The first inequality in (15) holds since each of $A(A-1)/2$ summands is $\ge d$. By the convexity $\cap$ of $\alpha(1-\alpha)$ the second inequality is strict unless each column has just $A\alpha$ 1's and $A(1-\alpha)$ 0's, contributing $A^2\alpha(1-\alpha)$ to the sum. The third expression increases with $\alpha < 1/2$, completing the proof of (15). The second inequality in (13) follows on multiplying the first and last terms in (15) by $2/Ad$ and solving for A.

The third inequality in (13) obviously holds for $d \le (n-1)/2$, since then the numerator is $nd \le n(n-1)/2$ while the denominator is always ≥ 1. For $d = n/2$ the maximum finite ratio in the bound is attained at $e = (n-2)/2$, which gives $A \le n^2/4 \le n(n-1)/2$ for $n \ge 2$. For $d = (n+1)/2$ the maximizing $e = (n-1)/2$, giving $A \le n$, which is 2 at $n = 2$ and is $\le n(n-1)/2$ for $n > 2$. And for $d = (n+k)/2$, $k \ge 2$, the ratio is bounded by its value at $e = n/2$, which gives $A \le (n+k)/k \le 1 + n/2$, also 2 at $n = 2$ and $\le n(n-1)/2$ for $n > 2$, completing the proof of (13) and b).

When all words are on the surface of $\mathscr{S}^n_e(\mathbf{0})$ only even distances occur, so (14) is vacuous. (13) holds for all d, but (14) is tighter for odd d, and may not have been published. The inequalities in (14) and the conditions on d,e,n are the inequalities in (13) and the conditions on d,e,n given there with d and n increased by 1. Thus it remains only to prove the asserted equality in c).

To prove the equality note first that $A(n,d,e) \ge A(n+1,d+1,e)$ for any d. For construct C in $\mathscr{S}^n_e(\mathbf{0})$ by dropping the last bit of a maximal set C' in $\mathscr{S}^{n+1}_e(\mathbf{0})$ of length $n+1$, distance $d+1$ and size $A(n+1,d+1,e)$. Clearly C has minimum distance $\ge d$ and lies in or on $\mathscr{S}^n_e(\mathbf{0})$.

Note next that $A(n,d,e) \le A(n+1,d+1,e)$ for odd $d = 2j+1$. For let C be a set of size $A(n,d,e)$ in $\mathscr{S}^n_e(\mathbf{0})$ with minimum distance d. Add to each codeword $\boldsymbol{x} \in C$ of weight $w(\boldsymbol{x}) \le e$ an $(n+1)$st bit which is the parity of $w(\boldsymbol{x}) + e$. The result is a codeword $\boldsymbol{x}'$ in a set C' in $\mathscr{S}^{n+1}_e(\mathbf{0})$, which has minimum distance $d+1$. If $w(\boldsymbol{x}) = e$ then the new bit is 0,

and thus the new weight is also e, while if $w(\boldsymbol{x}) < e$ then $w(\boldsymbol{x}') \le e$, and thus $\boldsymbol{x}' \in \mathscr{S}_e^{n+1}(\mathbf{0})$. And if two words $\boldsymbol{x}, \boldsymbol{y}$ in $\mathscr{C}$ have odd distance $d = 2j+1$ their weights must have opposite parity, so they differ in the new bit and $\boldsymbol{x}', \boldsymbol{y}'$ have Hamming distance $2j+2$ in C', while if they have even distance $d \ge 2j+1$ then $d \ge 2j+2$, in C and thus in C'.

To prove d), note the condition $nd > 2e(n-e)$. Thus the denominator in (13) is ≥ 1, so the denominator in (14) is $\ge 1+1+n+d-2e$, which is $\ge 2+d$ since $n \ge 2e$. Thus the ratio in (14) is $\le (n+1)(d+1)/(d+2) \le (n+1)n/(n+1) = n$ since $n \ge 2e \ge d+1$ for odd d, completing the proof of the proposition. □

The bounds in Proposition 10 are not always the best possible. For the maximal number of points *on the surface* of an e-sphere at distances $\ge d$, [13] gives other bounds tighter in some cases and a table of values or of upper and lower bounds for $n \le 23$ and $d \le 10$.

Some of the implications of Propositions 9 and 10 for known error-correcting codes are given in Appendix A. Proposition 11, which also follows directly from Propositions 9 and 10, is useful for asymptotic analysis with L, p_L and p fixed and $e_L \approx np_L$, $e \approx np$ as $n \to \infty$.

Proposition 11: Let $L \ge 2$ and let p_L and p satisfy

$$p_L = p(1-p)L/(L-1), \qquad 0 \le p \le 1/2,$$
$$p = \frac{1}{2}\left(1 - \sqrt{1 - 4p_L(1-1/L)}\right), \qquad 0 \le p_L \le L/4(L-1). \tag{16}$$

If np_L and np are integers (e.g., rational $p = j/k$ and $n = ik^2(L-1)$) then each $(n, np_L, 1)$ code is also an (n, np, L) code. More generally for any n let

$$e_L(n) = \lfloor (n+1)p_L \rfloor, \qquad e(n) = \lfloor (n+1)p \rfloor. \tag{17}$$

Then if a code is $(n, e_L(n), 1)$, it is also $(n, e(n), L)$.

Proof: If $e_L(n) \ge e(n)$ then the result follows from Proposition 1. Thus assume $e_L(n) < e(n)$, $e_L(n)+1 \le e(n)$. In the general case with e_L and e given by (17) let γ be an $(n, e_L(n), 1)$ code. In (14) let $d+1 = 2e_L(n)+2$ and $e = e(n)$. Then the conditions in Proposition 10 c) are satisfied. For first, since $p \le 1/2$, $n+1 \ge 2\lfloor (n+1)/2 \rfloor \ge 2\lfloor (n+1)p \rfloor = 2e(n) = 2e \ge 2e_L(n)+2 = d+1$. And second, rewriting the ratio that bounds $A(n,d,e)$ in (14) using these definitions gives the first inequality in (18): replacing $e_L(n)+1$ by its lower bound $(n+1)p_L$ and $e(n)$ by its upper bound $(n+1)p$ from (17) can only increase that ratio (when $p \le 1/2$), and gives the second inequality in (18). Using the relation (16) for p_L in terms of p shows the denominator in the increased ratio to be positive, and the increased ratio to equal L:

$$A(n,d,e) \le 1 \Big/ \left\{ 1 - \frac{e(n)}{n+1}\left(1 - \frac{e(n)}{n+1}\right) \Big/ \left(\frac{e_L(n)+1}{n+1}\right)\right\}$$
$$\le 1/\{1 - p(1-p)/p_L\} = L. \tag{18}$$

Thus, by Propositions 1 and 9, γ is $(n, e(n), L)$. In the integer case $e_L(n) = \lfloor (n+1)p_L \rfloor = np_L$ and $e(n) = \lfloor (n+1)p \rfloor = np$ since p_L and p are < 1. □

Keeping p fixed as L increases in (16) gives the simpler relation (19) between p_∞ and p. Proposition 12 shows that for finite n an infinite L is not required to satisfy (19): keeping $L = n$ as n increases will do.

Proposition 12: Let p and p_∞ satisfy

$$p_\infty = p(1-p), \qquad 0 \le p \le 1/2,$$
$$p = \frac{1}{2}\left(1 - \sqrt{1-4p_\infty}\right), \qquad 0 \le p_\infty \le 1/4. \tag{19}$$

If np and np_∞ are integers (e.g., rational $p = j/k$ and $n = ik^2$) and a code is $(n, np_\infty, 1)$, then it is also (n, np, n). More generally, for any n let

$$e_\infty(n) = \lceil np_\infty \rceil, \qquad e(n) = \lfloor np \rfloor. \tag{20}$$

Then if a code is $(n, e_\infty(n), 1)$, it is also $(n, e(n), n)$.

Proof: In the general case, what needs proving is that the condition $d > 2e(1 - e/n)$ in Proposition 10 b) holds, where $d = 2e_\infty(n)+1$ and $e = e(n)$. Then the result follows from Proposition 10 d) and Proposition 9. But from (20), $d \ge 2np_\infty + 1$ while for $p \le 1/2$, $2e(1-e/n) \le 2np(1-p)$, which is $2np_\infty$ by (19). The integer case follows since then $e_\infty = np_\infty$ and $e = np$. □

For particular p, p_L and p_∞, Propositions 11 and 12 can be used to find the list-of-L behavior of known codes and code families. In addition, these propositions may be used to extrapolate to $L > 1$ the behavior of the codes (unknown for large n, but known to exist) which satisfy the Gilbert–Varshamov lower bound in Proposition 8 at $L = 1$. This gives better lower bounds to the maximal rates $r_L(p), r_L'(p)$ in (10) for $L > 1$ than Proposition 8 gives directly, for large n and in the limit as n increases and $e \approx np$.

Proposition 13:

a) Let $0 \le p < 1/2$, $q = 1-p$, and $1/(1-4pq) \le L < \infty$. Then

$$1 - H(p) \ge r_L'(p) \ge r_1'(pqL/(L-1))$$
$$\ge 1 - H(2pqL/(L-1)) > 0 \tag{21}$$

and, in the limit of large L,

$$1 - H(p) \ge r_\infty'(p) \ge r_1'(pq) \ge 1 - H(2pq) > 0. \tag{22}$$

The results also hold when primes are deleted from r_1', r_L' and r_∞'.

b) Let $0 \le p < 1/4$. Then

$$1 - H\left(\left(1 - \sqrt{1-4p}\right)/2\right) \ge r_\infty\left(\left(1 - \sqrt{1-4p}\right)/2\right) \ge r_1(p)$$
$$\ge 1 - H(2p) > 0. \tag{23}$$

The results also hold when the functions r_1 and r_∞ are primed.

The primed and unprimed versions of (22) and (23) are both of interest since by Proposition 8 lower bounds are stronger on primed, and upper bounds on unprimed, rate functions.

Proof: Assume rational p. Consider the increasing sequence $n_i = ik^2(L-1)$ and let γ_i be linear and maximal in $\Gamma(n_i, n_i p_L, 1)$, with rate $R_1(\gamma_i) = \rho'(n_i, n_i p_L, 1)$ and the maximal limiting rate $r_1'(p_L) \ge 1 - H(2p_L)$ by Proposition 8. Using (16) gives the lower bound in (21); using (19) gives the lower bound in (22). By Proposition 11 γ_i is also in $\Gamma(n_i, n_i p, L)$, and by (6) its rate $R_L(\gamma_i) = R_1(\gamma_i) - (\log L)/n_i$; so its limiting L-rate approaches its limiting 1-rate $r_1'(p_L)$ as n_i increases. The rate of the best linear $(n_i, n_i p, L)$ code is no smaller, so $r_L'(p) \ge r_1'(p_L)$. That, with the upper bound to $r_L'(p)$ for finite and infinite L from Proposition 8, completes

the proof of (21) and (22) for rational p and linear codes. For general codes, let γ_i be maximal without the linearity constraint. By Proposition 8 the upper and lower bounds still hold, proving the assertion following (21). The relations between p and p_∞ in (19) convert the unprimed version of (22) to (23). The extension of all results to real p follows by the continuity of the bounds and the monotonicity of $r_L(p)$ and $r'_L(p)$ in p from Proposition 8. □

The inequality between first and third terms in (23) is the asymptotic Elias upper bound to r_1 [4], [5], [13], which is not the best possible (see e.g., Chapter 17.7 in [13]) but, unlike the Hamming bound in Proposition 4, shows that a jammer who can change one fourth of the transmitted symbols can guarantee a vanishing rate $1 - H(1/2) = 0$ of transmission under list-of-1 decoding. The inequality between the second and fourth terms in (22) is essentially an inverse of that bound (with primes) via the relation (19) between p_∞ and p_1. It gives a lower bound to $r'_\infty(p)$, which is not the best possible (see Proposition 15) but which, unlike the Gilbert–Varshamov bound in Proposition 7, remains strictly positive for $0 < p < 1/2$. Thus if the communicators use list decoding with sufficiently large lists, a jammer must be able to change half or more of their transmitted symbols to guarantee a vanishing rate of transmission. More precisely, communicating at a positive rate requires $L > 1/\epsilon^2$, and the rate for large L behaves like ϵ^4.

Proposition 14: Let $p = (1-\epsilon)/2$, $L = \alpha/\epsilon^2 \geq 2$, $0 < \epsilon \leq 1$, and $c = (\log e)/2 = 0.721\cdots$. Then

$$r'_L(p) \geq \begin{cases} 0, & \alpha \leq 1 \\ 1 - H\left(\frac{1}{2}\left(1 - \frac{\epsilon^2 L - 1}{L-1}\right)\right) > c\left(\frac{\epsilon^2 L - 1}{L-1}\right)^2 > c\epsilon^4(1-1/\alpha)^2, & \alpha > 1 \end{cases} \tag{24}$$

and

$$r'_\infty(p) \geq 1 - H((1-\epsilon^2)/2) > c\epsilon^4. \tag{25}$$

Proof: Equation (25) and the first inequality in (24) follow from (21), the fact that $2p(1-p) = (1-\epsilon^2)/2$ and a Taylor expansion of $1 - H((1-\delta)/2)$ about $\delta = 0$, where it has zero first derivative and minimal second derivative, giving the lower bound in

$$\delta^2 \geq 1 - H((1-\delta)/2) \geq c\delta^2. \tag{26}$$

The second inequality in (24) follows from the first by substituting $L = \alpha/\epsilon^2$, multiplying through by ϵ^2 and using $\epsilon > 0$. The upper bound in (26) is also useful later. It follows from the fact that there is equality on the left when δ is 0 or 1 and a comparison of second derivatives between those limits. □

The random coding argument in [9] proves that in the linear case the best lower bound to $r_\infty(p)$ is in fact the upper bound in (12). Proposition 15 summarizes the results proved in [9] for linear codes and gives analogous results for general codes. The upper and lower bounds for finite L found by Blinovskii in [10] using ingenious new random coding arguments are tighter but not easily summarized.

Proposition 15: Let $L \geq 2$ and $J = \lceil \log(L+1) \rceil$. Then

$$r_L(p) \geq f_L(p) = 1 - (1 + 1/L)H(p),$$
$$r'_L(p) \geq f'_L(p) = 1 - (1 + 1/J)H(p), \tag{27}$$
$$r_\infty(p) = r'_\infty(p) = 1 - H(p) \geq r_L(p) \geq r'_L(p). \tag{28}$$

Appendix B proves Proposition 15 in the linear and nonlinear cases. Proposition 16 shows that, for p near $1/2$, the random coding lower bound to $r_\infty((1-\epsilon)/2)$ is $c\epsilon^2$ rather than the bound $c\epsilon^4$ from Proposition 14, but that (as noticed by Levin [15]) guaranteeing positive rate for linear codes requires much larger lists.

Proposition 16: Let $p = (1-\epsilon)/2$, $0 < \epsilon \leq 1$, and let $c_0 = (\log e)/2$. Let K be a positive integer and let $c_1 = \epsilon^2/(K+1)$.

a) If $c_1 < c_0$, then

$$f_L(p) \geq (c_0 - c_1)\epsilon^2, \qquad f'_{L'}(p) \geq (c_0 - c_1)\epsilon^2 \tag{29}$$

for all integers $L \geq K$ and $L' \geq 2^{K-1}$.

b) If $c_1 \geq 1$, then

$$f_L(p) \leq 0, \qquad f'_{L'}(p) \leq 0 \tag{30}$$

for all integers $L \leq K$ and $L' \leq 2^K - 1$.

c) In the limit,

$$\epsilon^2 \geq r_\infty(p) = r'_\infty(p) \geq c\epsilon^2. \tag{31}$$

Proof: From (27), $f_L(p) = [(1 - H(p)) - 1/(L+1)](1 + 1/L)$, and similarly for f' with L replaced by J. Dropping the $1/L$ and using the lower bound to H in (26) with $\delta = \epsilon$ and the definition of c_1 with $L = K$ gives the first inequality in (29). The second follows from a parallel treatment of $f'_{L'}$ replacing L by $J' = \lceil \log(L'+1) \rceil$, and using $J' \geq K$ iff $L' \geq 2^{K-1}$, proving a). Using the upper bound from (26) and the fact that $J' \leq K$ iff $L' \leq 2^K - 1$ proves b). Part c) follows directly from (28) and (26). □

Appendix A
Examples of List Error-Correcting Codes

Taking $d = 2e_1 + 1$ in Propositions 9 and 10 proves Proposition 17. It is in convenient form for use in the construction of (n, e, L) codes from $(n, e_1, 1)$ codes for finite n, e_1, and e.

Proposition 17:

a) Let γ be $(n, e_1, 1)$ and $e \geq e_1 + 1$. Then γ is also (n, e, L), where

$$L = \left\lfloor \frac{(n+1)(e_1+1)}{e^2 - (n+1)(e - e_1 - 1)} \right\rfloor, \tag{32}$$

if the denominator in (32) is positive.

b) In particular, if γ is $(n, e, 1)$ it is $(n, (e+1), \lfloor (n+1)/(e+1) \rfloor)$.

Comment: Proposition 17 b) shows that the Hamming $(2^k - 1, 1, 1)$ code γ_k of size $M_1(\gamma_k) = 2^{2^k - 1 - k}$ is also a $(2^k - 1, 2, 2^{k-1})$ code of effective size $M_{2^{k-1}}(\gamma_k) = 2^{2^k - 2k}$. If a $(2^k - 1, 2, 1)$ code γ'_k of the same effective size were available, it would be preferable since it is more versatile: by Proposition 2 it could also be used as an $(2^k - 1, 2, 2^{k-1})$ code of the same effective size by mapping 2^{k-1} messages into each codeword. It turns out that there is such γ'_k for even k but that for odd k all $(2^k - 1, 2, 1)$ codes are strictly smaller in size than $2^{2^k - 2k}$.

Proposition 18: The Hamming $(2^k-1,1,1)$ code γ_k of size $M_1(\gamma_k)=2^{2^k-1-k}$ is also a $(2^k-1,2,2^{k-1})$ code of effective size $M_{2^{k-1}}(\gamma_k)=2^{2^k-2k}$. For odd k, the largest $(2^k-1,2,1)$ code γ_k' is strictly smaller, i.e., $M_1(\gamma_k')<2^{2^k-2k}$. For even k, there are $(2^k-1,2,1)$ codes of size 2^{2^k-2k} but no larger.

Proof: The first assertion has been proved. The Hamming bound of Proposition 4 gives

$$M \le 2^{2^k-1}/\left[1+(2^k-1)+(2^k-1)(2^k-2)/2\right] = 2^{2^k-2k}/(1-2^{-k}+2^{-2k+1}), \quad (33)$$

not quite tight enough to show that there is no larger code for any k. A refinement of the Hamming bound due to Johnson (Corollary 14, Chapter 17 in [13]) proves the second assertion, however. (I am obliged to Neil Sloane for pointing out the applicability of this result to my purposes [18].) In our terminology that result gives the upper bound

$$2^n\Big/\left[V_e^n+\binom{n}{e}\phi((n-e)/(e+1))/\lfloor n/(e+1)\rfloor\right], \quad (34)$$

to the size of an $(n,e,1)$ code, where $\phi(x)=x-\lfloor x\rfloor$ gives the fractional part of x.

Let $e=2$ and $n=2^k-1$ in (33). Since 2^j is not divisible by 3, either 2^j-1 or 2^j+1 is; so, for $k=2j$,

$$(2^j-1)(2^j+1)=2^{2j}-1=2^k-1\equiv 0 \bmod 3. \quad (35)$$

Equation (35) gives $\phi((2^{2j}-3)/3)/\lfloor(2^{2j}-1)/3\rfloor=1/(2^{2j}-1)$; so (34), for $n=2^k-1=2^{2j}-1$, gives

$$2^{2^k-1}\Big/\left[1+(2^k-1)+\binom{2^k-1}{2}(1+1/(2^k-1))\right]=2^{2^k-2k}. \quad (36)$$

This proves that there are no larger $(2^k-1,2,1)$ codes. The (nonlinear) Preparata codes attain the bound (see [13], Chapter 15). For odd $k=2j+1$, multiplying by 2 in (34) gives $2^k-2\equiv 0 \bmod 3$; so

$$\phi((2^{2j+1}-3)/3)/\lfloor(2^{2j+1}-1)/3\rfloor=2/(2^{2j+1}-2),$$

and (34) gives

$$2^{2^k-1}\Big/\left(1+(2^k-1)+\binom{2^k-1}{2}(1+2/(2^k-2))\right) = 2^{2^k-2k}/(1+2^{-k}), \quad (37)$$

which proves the second assertion in the proposition: for odd k no $(2^k-1,2,1)$ code is as large as 2^{2^k-2k}. It follows that no linear code is larger than 2^{2^k-2k-1} for odd k. Verhoeff's table [19] shows that linear $(2^k-1,2,1)$ codes of size 2^{2^k-2k-1} exist for all $k<8$, and that linear codes of size 2^{2^k-2k} do not exist for $k=4$; the question seems to be open for $k=6$, and perhaps for some even $k\ge 8$. □

Proposition 18 says all there is to say about using Proposition 17 with $e_1=1$ to convert $(n,1,1)$ codes into (n,e,L) codes of effective size $M_L\ge 2$. Proposition 18 has dealt with the case $e=2$, and there are no cases with $e\ge 3$.

Proposition 19: Let γ be an $(n,1,1)$ code, which can be shown by Proposition 17, to also be an (n,e,L) code, with $e\ge 3$. Then $M_L(\gamma)<2$, $R_L(\gamma)<1/n$.

Proof: Assume first that γ is an $(n,1,1)$ code which Proposition 17 shows to also be (n,e,L) with $e\ge 3$ and $M_L(\gamma)\ge 2$. Then Proposition 6 holds and the denominator in (32) is ≥ 1, giving the two inequalities

$$2e+2\le n+1\le (e^2-1)/(e-2). \quad (38)$$

Dropping the center term in (38) gives $2(e+1)\le(e^2-1)/(e-2)$, $e\le 3$; so, by the hypothesis, $e=3$. Substituting $e=3$ in (38) gives the unique value $n=7$. Those two values in (32) give $L=16$. By Proposition 4, the largest $(7,1,1)$ code is the Hamming code of size $2^7/(1+7)=16$; so the effective size $M_L(\gamma)=M_1(\gamma)/L\le 1$, which contradicts the hypothesis. □

The Golay $(23,3,1)$ code is an example of the use of Proposition 17 with $e_1>1$.

Proposition 20: The Golay $(23,3,1)$ code is also a $(23,4,6)$ code, of effective size $2^{11}/3=682\ 2/3$.

Comment: This proposition also follows directly from Proposition 17 b) with $L=24/4=6$. The Hamming bound on effective size is $2^{23}/V_4^{23}=2^{23}/10903=769.385\cdots$, and the Johnson bound is $747.697\cdots$, but there are tighter bounds in this case: [13] bounds above the size of the (unknown) best $(23,4,1)$ code by 280, and [19] gives 128 as the size of the best linear $(23,4,1)$ code. Proposition 17 a) shows that the Golay code is also a $(23,5,96)$ code, of effective size $2^{12}/96=2^7/3=42\ 2/3$, but that is less interesting since a $(23,5,1)$ code of size 48 is given in [13], though the best linear $(23,5,1)$ code only has size 32 [19]. The Proposition does not apply to the Golay code for $e=6$ since the denominator is then negative.

Open problems: Practically all questions about the error correction capabilities of particular $(n,e,1)$ codes under list decoding remain open. Here are a few.

1) By Proposition 17, all $(n,e_1,1)$ codes are (n,e,L) codes for the appropriate L. For what $(n,e,1)$ codes and values of L are the resulting codeword sets significantly better in effective size, i.e., packed significantly more densely, than the best (or best known) $(n,e,1)$ code?
2) Is there an (n,e,L) code of maximal size which is not an optimal $(n,e_1,1)$ code, i.e., a proof that maximizing minimum distance increases the minimum number of codewords in some e-sphere?
3) Is it true that in some of the known good $(n,e,1)$ codes no e-sphere contains the maximum number of codewords at distance d allowed by Proposition 10, so that the bounds in Propositions 11, 12, and 17 can be improved?

Appendix B
Proof of Random Coding Lower Bounds

Proof: In Proposition 15, (28) follows from (27) (by taking the limit) and Proposition 8. Given n let $e=\lfloor np\rfloor$, and let $\mathscr{U}(n,M)$ denote the set of all codes of length n and size M. To prove (27) consider the 2^n e-spheres about each point in $\mathscr{B}$. The fraction of the codes in $\mathscr{U}(n,2M)$ whose first $L+1$ codewords all lie in the first of those spheres is $(V_e^n/2^n)^{L+1}$; the expected number of the $\binom{2M}{L+1}$ $(L+1)$-tuples of codewords which lie in one of the 2^n e-spheres is just

$$2^n(V_e^n/2^n)^{L+1}\binom{2M}{L+1}<2^n(2MV_e^n/2^n)^{L+1}/(L+1)! \quad (39)$$

per code, using $\binom{2M}{L+1} < (2M)^{L+1}/(L+1)!$. Set M equal to the upper bound in (39); so that that expected number is $\leq M$. Then in each $C(\gamma_n')$, removing one codeword from each $(L+1)$-tuple that lies in an e-sphere reduces the size of that code by an amount whose expectation is $\leq M$. The resulting subcodes of the codes in $\mathscr{U}(n, 2M)$ are (n, e, L) by Proposition 3 and have an expected number $\geq M$ of messages left; so, some (n, e, L) subcode γ_n has at least M messages and effective size $\geq M/L$.

Thus setting the bound in (39) equal to M, dividing by M, and solving for (M/L) gives a lower bound to the effective sizes of (n, e, L) codes by (4):

$$M_L = M/L$$

$$\geq 2^{n(1-(1+1/L)(1/n)\log V_e^n)}\left(2(L+1)!/(2L)^{L+1}\right)^{1/L}, \quad (40)$$

from which the bounds on the unprimed rates in the proposition follow on taking logarithms, dividing by n, using (11), and taking the limit as $n \to \infty$.

In the linear case, following Zyablov and Pinsker [9], define a J-tuple of integer messages to be *independent* if the n-bit binary representations of the J integers are linearly independent when considered as vectors in $\mathscr{B}$. For independent J-tuples, all 2^{nJ} assignments of codewords in $\mathscr{B}$ to those J messages occur with equal frequency in the set of all linear codes of length n. By Proposition 3, a linear code γ of size $M = 2^k$ is (n, e, L) iff no one of the V_e^n e-spheres about points in $\mathscr{S}_e^n(0)$ contains codewords representing L nonzero messages since when message 0 is coded into the all-zero codeword $\mathbf{0}$ and is sent and any sequence $y \in \mathscr{S}_e^n(\mathbf{0})$ is received, 0 will be on the receiver's list $\delta(y)$. Every set of L nonzero messages has a linearly independent subset of size at least J, so γ is certainly (n, e, L) if none of the V_e^n e-spheres with centers in $\mathscr{S}_e^n(0)$ contains codewords representing J independent messages.

The fraction of the codebooks in which a particular J-tuple of independent messages falls in a particular e-sphere is just $(V_e^n/2^n)^J$. There are $M - 2^j = 2^k - 2^j$ nonzero choices for the $(j+1)$st message in such a J-tuple whose binary representation must not be one of the 2^j linear combinations of the representations of the first j codewords, and order of choice does not matter. Thus the total number of independent J-tuples is

$$(2^k - 2^0)(2^k - 2^1) \cdots (2^k - 2^{J-1})/J! < M^J/J! \quad (41)$$

and the expected number of *all* J-tuples which fall in *some* e-sphere per code is bounded above by

$$V_e^n\left(MV_e^n/2^n\right)^J/J! < 2^{-nJ(1-(1+1/L)(1/n)\log V_e^n - R)}/J!, \quad (42)$$

using $M = L2^{nR}$. There is at least one (n, e, L) code of rate R if the lefthand side of (42) is ≤ 1. Setting its upper bound at 1, taking logarithms, dividing by n, using (11) again, and taking the limit of large n gives Zyablov and Pinsker's bound on r_L' in (27). □

Open problems: The most obvious problem left open by the random coding results is whether the requirement noted in Proposition 16 of very large lists for typical linear codes is real or (as seems more likely) is an artifact of the proof techniques.

References

[1] P. Elias, "List decoding for noisy channels," *Wescon Convention Record*, Part 2, Institute of Radio Engineers (now IEEE), pp. 94–104, 1957.

[2] J. M. Wozencraft, "List decoding," *Quarterly Progress Report*, vol. 48, pp. 90–95, Research Laboratory of Electronics, MIT, Jan. 15, 1958.

[3] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels," *Inform. Contr.*, vol. 10, pp. 65–103 (Part I), pp. 522–552 (Part II), 1967.

[4] E. R. Berlekamp, *Algebraic Coding Theory*. New York: McGraw Hill, 1968.

[5] R. G. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley, 1968.

[6] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw Hill, 1979.

[7] R. Ahlswede, "Channel capacities for list codes," *J. Appl. Probability*, vol. 10, pp. 824–836, 1973.

[8] I. Csiszar and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic Press, 1981, p. 196.

[9] V. V. Zyablov and M. S. Pinsker, "List cascade decoding," in *Prob. Inform. Transm.*, vol. 17, no. 4, pp. 29–34 (in Russian), Oct.–Dec. 1981; pp. 236–240 (in English), 1982.

[10] V. M. Blinovskii, "Bounds for codes in the case of list encoding of finite volume," *Prob. Inform. Transm.*, vol. 22, no. 1, pp. 11–25 (in Russian), Jan.–Mar. 1986; pp. 7–19 (in English), 1986.

[11] P. Elias, "Zero error capacity for list detection," *Quarterly Progress Report*, vol. 48, pp. 88–90, Research Laboratory of Electronics, MIT, Jan. 15, 1958.

[12] ___, "Zero error capacity under list decoding," *IEEE Trans. Inform. Theory*, vol. 34, no. 5, pp. 1070–1074, Sept. 1988.

[13] J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam: North Holland, 1977.

[14] L. Levin, "One-way functions and pseudorandom generators," *Combinatorica*, vol. 7, no. 4, pp. 357–363, 1987.

[15] ___, personal communication, Sept. 1988.

[16] G. Cohen, J. Körner, and G. Simonyi, "Zero-error capacities and very different sequences (preliminary version)," in *Sequences: Combinatorics, Compression, Security, and Transmission*, R. M. Capocelli, Ed. New York: Springer-Verlag, 1990, pp. 144–155.

[17] S. M. Johnson, "A new upper bound for error-correcting codes," *IEEE Trans. Inform. Theory*, vol. IT-8, no. 3, pp. 203–207, Apr. 1962.

[18] Neil J. A. Sloane, personal communication, May 1988.

[19] T. Verhoeff, "An updated table of minimum-distance bounds for binary linear codes," *IEEE Trans. Inform. Theory*, vol. IT-33, no. 5, pp. 665–680, Sept. 1987.