when flipping an even coin $n_2 - 1$ times is at least the probability, $B(n_1, t) \cdot 2^{-n_1}$, of that event when flipping the coin $n_1 - 1$ times.

Taking logarithms, we obtain

$$\begin{aligned}\rho(\mathcal{C}(n_1 \times n_2, t)) &\le -\log_2(1 - 2n_1 B(n_2, t) \cdot 2^{-n_2}) \\ &\le -\log_2\left(1 - 4n_1 \cdot n_2^{t-1} \cdot 2^{-n_2}\right) \\ &= -\log_2\left(1 - 4 \cdot 2^{-\Delta(n_1 \times n_2, t)}\right)\end{aligned}$$

thus proving (1).

REFERENCES

[1] S. Al-Bassam and B. Bose, "On balanced codes," *IEEE Trans. Inform. Theory*, vol. 36, pp. 406–408, 1990.

[2] J. J. Ashley, M. Blaum, and B. H. Marcus, "Report on coding techniques for holographic storage," IBM Res. Rep. RJ 10013, 1996.

[3] J. J. Ashley and B. H. Marcus, "Two-dimensional low-pass filtering codes," *IEEE Trans. Commun.*, vol. 46, pp. 724–727, 1998.

[4] R. Berger, "The undecidability of the Domino problem," *Mem. Amer. Math. Soc.*, vol. 66, 1966.

[5] D. Brady and D. Psaltis, "Control of volume holograms," *J. Opt. Soc. Am. A*, vol. 9, pp. 1167–1182, 1992.

[6] N. Calkin and H. S. Wilf, "The number of independent sets in a grid graph," *SIAM J. Discrete Math.*, vol. 11, pp. 54–60, 1997.

[7] T. Etzion, "Cascading methods for runlength-limited arrays," *IEEE Trans. Inform. Theory*, vol. 43, pp. 319–324, 1997.

[8] T. Etzion and V. K. Wei, "On two-dimensional runlength-limited codes," *IEEE Int. Workshop on Information Theory*, Salvador, Brazil, 1991.

[9] S. Forchhammer and J. Justesen, "Entropy bounds for constrained two-dimensional random fields," *IEEE Trans. Inform. Theory*, vol. 45, pp. 118–127, 1999.

[10] J. F. Heanue, M. C. Bashaw, and L. Hesselink, "Volume holographic storage and retrieval of digital data," *Science*, vol. 265, pp. 749–752, 1994.

[11] ——, "Channel codes for digital holographic data storage," *J. Opt. Soc. Am. A*, vol. 12, 1995.

[12] D. E. Knuth, "Efficient balanced codes," *IEEE Trans. Inform. Theory*, vol. 32, pp. 51–53, 1986.

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

[14] M. Marcellin and H. Weber, "Two-dimensional modulation codes," *IEEE J. Selected Areas Commun.*, vol. 10, pp. 254–266, 1992.

[15] B. H. Marcus, R. M. Roth, and P. H. Siegel, "Constrained systems and coding for recording channels," in *Handbook of Coding Theory*, V. S. Pless and W. C. Huffman, Eds. Amsterdam, The Netherlands: Elsevier, pp. 1635–1764.

[16] B. H. Marcus, P. H. Siegel, and J. K. Wolf, "Finite-state modulation codes for data storage," *IEEE J. Select. Areas Comm.*, vol. 10, pp. 5–37, 1992.

[17] E. K. Orcutt and M. W. Marcellin, "Enumerable multi-track (d, k) block codes," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1738–1744, 1993.

[18] ——, "Redundant multi-track (d, k) codes," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1744–1750, 1993.

[19] P. Pavlidis, J. Swartz, and Y. P. Wang, "Fundamentals of bar code information theory," *Computers*, vol. 23, pp. 74–86, 1990.

[20] D. Psaltis and F. Mok, *Holographic Memories, Scientific American*, vol. 273, no. 5, pp. 70–76, 1995.

[21] E. Ordentlich and R. M. Roth, "On the redundancy of two-dimensional balanced codes," Hewlett-Packard Lab. Tech. Rep. HPL-97-143, 1997; also in *Proc. Int. Symp. Information Theory*, Cambridge, U.K., Aug. 1998, p. 113.

[22] D. Psaltis, M. A. Neifeld, A. Yamamura, and S. Kobayashi, "Optical memory disks in optical information processing," *Appl. Optics*, vol. 29, pp. 2038–2057, 1990.

[23] R. M. Robinson, "Undecidability and nonperiodicity for tilings of the plane," *Inventiones Math.*, vol. 12, pp. 177–209, 1971.

[24] K. A. S. Immink, *Coding Techniques for Digital Recorders*. New York: Prentice-Hall, 1991.

[25] G. T. Sincerbox, "Holographic storage re-visited," *Current Trends in Optics*, C. Dainty, Ed. New York: Academic, 1994.

[26] L. G. Tallini, R. M. Capocelli, and B. Bose, "Design of some new balanced codes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 790–802, 1996.

[27] A. Vardy, M. Blaum, P. H. Siegel, and G. T. Sincerbox, "Conservative arrays: Multidimensional modulation codes for holographic recording," *IEEE Trans. Inform. Theory*, vol. 42, pp. 227–230, 1996.

[28] H. S. Wilf, "The problem of the kings," *Electronic J. Combinatorics*, vol. 2, no. R3, 1995.

An Improved Upper Bound of the Rate of Euclidean Superimposed Codes

Zoltán Füredi and Miklós Ruszinkó

***Abstract*—A family of n-dimensional unit norm vectors is an *Euclidean superimposed code* if the sums of any two distinct at most m-tuples of vectors are separated by a certain minimum Euclidean distance d. Ericson and Györfi [8] proved that the rate of such a code is between $(\log m)/4m$ and $(\log m)/m$ for m large enough. In this paper—improving the above long-standing best upper bound for the rate—it is shown that the rate is always at most $(\log m)/2m$, i.e., the size of a possible superimposed code is at most the root of the size given in [8]. We also generalize these codes to other normed vector spaces.**

***Index Terms*—Codes, growth rate, superimposed geometric codes.**

I. SUPERIMPOSED CODES

Binary superimposed codes were introduced by Kautz and Singleton [18]. They studied binary codewords with the property that the disjunctions (bitwise OR) of any pair of distinct at most m-tuples of codewords have to be different. Later, this question has been investigated in several papers on multiple-access communication (see, e.g., [1], [4], [5], [14]–[16]). The same problem has been posed—in different terms—by Erdős [6] and Frankl and Füredi [7] in combinatorics, by Sós [22] in combinatorial number theory, and by Hwang [11] and Sós [12] in group testing. One can find an easy proof of the best known upper bound of these codes in the papers by Füredi [9] and Ruszinkó [21]. In the paper of Füredi and Ruszinkó [10], the connection of these codes to the big distance ones is shown.

In 1988, Ericson and Györfi [8] introduced a new class of superimposed codes for Euclidean channels. Roughly speaking, a family of n-dimensional unit norm vectors is an *Euclidean superimposed code,* if the sums of any two distinct at most m-tuples of vectors are separated by a certain minimum Euclidean distance d. Similarly

Manuscript received December 17, 1997; revised August 25, 1998. This work was supported in part by the Hungarian National Science Foundation under Grant OTKA 016389, National Security Agency under Grant MDA904-98-I-0022, and OTKA under Grants T 016414 and F 014919.

Z. Füredi is with the Department of Mathematics, University of Illinois, Urbana, IL 61801 USA (e-mail: z-furedi@math.uiuc.edu) and the Mathematical Institute of the Hungarian Academy of Sciences, Budapest 1364, Hungary (e-mail: furedi@math-inst.hu).

M. Ruszinkó is with the Department of Mathematical Sciences, Carnegie-Mellon University, Pittsburgh, PA 15213 USA (e-mail: ruszinko@andrew.cmu.edu) and the Computer and Automation Research Institute of the Hungarian Academy of Sciences, Budapest 1518, Hungary (e-mail: ruszinko@lutra.sztaki.hu).

Communicated by A. Barg, Associate Editor for Coding Theory.

Publisher Item Identifier S 0018-9448(99)01415-7.

to the previous ones, these codes are motivated by multiple access communication as follows.

Suppose that T users share a common channel. To each of them a real-valued n-dimensional unit norm vector is associated. The ith user transmits its vector $\mathbf{x_i} = (x_i^1, x_i^2, \cdots, x_i^n)$ $(i = 1, 2, \cdots, T)$—i.e., in each slot a real is contained, while the total block length is n—if it is active, otherwise not. It is assumed that the transmission is bit and block synchronized. The destination of the messages is a single receiver, which, in an optimal case, gets the sum of the vectors

$$\mathbf{y} = \sum_{\forall i \text{ active}} \mathbf{x_i}$$

associated to the active users. Moreover, suppose that at most m users are active simultaneously. The problem is how to choose the vectors associated to the single users to assure that the receiver may recognize the set of the active ones even in the case if the sequence $\mathbf{y}$ is further contaminated by some (less than $d/2$) disturbance.

The central problem is here to determine the maximum number of real unit norm vectors $T(n, m, d)$ assuring the above identification, given the block length n, maximum number of simultaneously active users m, and disturbance $<d/2$. Ericson and Györfi [8] showed that for the rate of such a code, the following inequalities hold for $n \gg m$ large enough, $d \le 1$:

$$\frac{\log m}{4m} \le \frac{\log T(n, m, d)}{n} \le \frac{\log m}{m}. \tag{1}$$

Observe that this does not depend on d, since d is constant. Anderson [2], [3] slightly improved the bounds given by Ericson and Györfi, but his attempts based on methods elaborated by Kabatjanskiĭ and Levenštein [17] gave only polynomial improvements, and thus he did not improve the bounds on the *rate* of the code. In this paper, we improve the upper bound of the rate—given by Ericson and Györfi—by a factor of two, thus giving an exponential improvement of this long-standing open problem.

Note that instead of unit norm vectors, at most unit norm vectors can be considered, whereby this approach gives the same rate, since adding one additional dimension, which does not change the rate, one can easily make of at most unit vectors unit ones.

In the next section, the mathematical formulation is given, which is followed by the proof of the new upper bound. Finally, the last section contains some remarks on geometric codes in other normed spaces.

II. Euclidean Codes

Let $\mathcal{C}$ be a finite set of unit norm vectors in $\mathbf{R^n}$ (a *spherical code*). For a subset $\mathcal{A}$ of $\mathcal{C}$—following the notation of [8]—$f(\mathcal{A})$ stands for the sum of vectors $\mathbf{x} \in \mathcal{A}$, $f(\mathcal{A}) = \sum_{\mathbf{x} \in \mathcal{A}} \mathbf{x}$. Let m and T be positive integers, $m \le T$, and let d be a real number, $0 < d \le 1$. The finite set $\mathcal{C}$ of unit norm vectors in $\mathbf{R^n}$ is an *Euclidean superimposed code* with parameters (n, m, T, d) if $|\mathcal{C}| = T$ and for two arbitrary distinct subsets $\mathcal{A}$ and $\mathcal{B}$ of $\mathcal{C}$ with $0 \le |\mathcal{A}|, |\mathcal{B}| \le m$ the Euclidean distance of the vectors $f(\mathcal{A})$ and $f(\mathcal{B})$ is at least d. More precisely, denote by $\mathcal{C}^m$ the set of the sums of all at most m-tuples of vectors of $\mathcal{C}$, i.e.,

$$\mathcal{C}^m = \{f(\mathcal{A}) \colon \mathcal{A} \subseteq \mathcal{C},\, |\mathcal{A}| \le m\}.$$

Set

$$d_E(\mathcal{C}^m) = \min_{\substack{\mathcal{A} \ne \mathcal{B} \\ 0 \le |\mathcal{A}|,\, |\mathcal{B}| \le m \\ \mathcal{A},\, \mathcal{B} \subseteq \mathcal{C}}} \|f(\mathcal{A}) - f(\mathcal{B})\|$$

where

$$\|\mathbf{x}\| = \left(\sum_{i=1}^n x_i^2\right)^{1/2}$$

is the usual Euclidean norm. Equivalently, the set $\mathcal{C}$ of finite unit norm vectors in $\mathbf{R^n}$ is an *Euclidean superimposed code* with parameters (n, m, T, d) if $d_E(\mathcal{C}^m) \ge d$. Given n, m, and d, let $T(n, m, d)$ denote the maximum size of an *Euclidean superimposed code,* i.e.,

$$T(n, m, d) = \max\{T \colon \mathcal{C}(n, m, T, d) \ne \emptyset\}.$$

Similarly, let

$$\mathcal{C}^{m'} = \{f(\mathcal{A}) \colon \mathcal{A} \subseteq \mathcal{C},\, |\mathcal{A}| = m\}$$

and $T'(n, m, d)$ denote the maximum size of a spherical code $\mathcal{C}$ satisfying $d_E(\mathcal{C}^{m'}) \ge d$. Obviously, $T(n, m, d) \le T'(n, m, d)$.

As it was shown in [8], $T(n, m, d)$ increases exponentially in n. Therefore, due to coding theoretic traditions

$$R(m, d) = \limsup_{n \to \infty} \frac{\log T(n, m, d)}{n}$$

is the exponent of the growth. It is also called the *rate* of the code. Here and throughout the paper $\log$ stands for the logarithm in base 2, $\mathbf{v_i v_j}$ is the scalar product of vectors $\mathbf{v_i}$ and $\mathbf{v_j}$ (i.e., $\mathbf{v}^2$ is the scalar product of the vector $\mathbf{v}$ with itself), and $\binom{\mathcal{C}}{m}$ stands for all m-subsets of $\mathcal{C}$.

III. The Improved Upper Bound

The main idea of the proof of the upper bound given in [8] is a sphere packing argument. Observe that if $\mathcal{C}$ is a spherical code, then all vectors of $\mathcal{C}^m$ are within a ball of radius m. From this fact one can get the Ericson–Györfi upper bound as follows. The fraction of the volume of this ball and the volume of a ball of radius $d/2$ is an upper bound for the size $|\mathcal{C}^m|$, since if $\mathcal{C}$ is an *Euclidean superimposed code*—by the definition—$\mathcal{C}^m$ is of minimum distance d. Our main idea is to show that for any spherical code $\mathcal{C}$ almost all members of $\mathcal{C}^{m'}$ are within a ball of radius of magnitude $m^{1/2}$. From this the improvement on the rate of an *Euclidean superimposed code* immediately follows.

Lemma 3.1: For any spherical code $\mathcal{C}$ of size T' and integer $m \le T'$ the inequality

$$\sum_{\mathcal{A} \in \binom{\mathcal{C}}{m}} \|f(\mathcal{A}) - m\mathbf{c}\|^2 \le m \binom{T'}{m}$$

holds, where $\mathbf{c}$ is the (average) vector $(1/T') \sum_{\mathbf{v} \in \mathcal{C}} \mathbf{v}$.

Proof:

$$\begin{aligned}\sum_{\mathcal{A} \in \binom{\mathcal{C}}{m}} &\|f(\mathcal{A}) - m\mathbf{c}\|^2 \\ &= \sum_{\mathcal{A} \in \binom{\mathcal{C}}{m}} \big((f(\mathcal{A}))^2 + m^2\mathbf{c}^2 - 2m\mathbf{c}f(\mathcal{A})\big).\end{aligned} \tag{2}$$

The second term of (2) is clearly $\binom{T'}{m} m^2 \mathbf{c^2}$, while the third one is

$$-2m\mathbf{c} \sum_{\mathcal{A} \in \binom{\mathcal{C}}{m}} f(\mathcal{A}) = -2m\mathbf{c} \binom{T'-1}{m-1} T\mathbf{c} = -2\binom{T'}{m} m^2 \mathbf{c}^2$$

since in the sum $\sum_{\mathcal{A} \in \binom{\mathcal{C}}{m}} f(\mathcal{A})$ every vector of the spherical code is summed up with multiplicity $\binom{T'-1}{m-1}$, i.e., this is the number of distinct m-tuples in which a given vector is contained. The first term of (2) can be estimated as follows:

$$\begin{aligned}\sum_{\mathcal{A} \in \binom{\mathcal{C}}{m}} &(f(\mathcal{A}))^2 \\ &= \sum_{\mathcal{A} \in \binom{\mathcal{C}}{m}} \left(\sum_{\mathbf{v} \in \mathcal{A}} \mathbf{v}\right)^2 = \sum_{\mathcal{A} \in \binom{\mathcal{C}}{m}} \left(m + 2 \sum_{\substack{1 \le i < j \le m \\ \mathbf{v_i},\, \mathbf{v_j} \in \mathcal{A}}} \mathbf{v_i v_j}\right)\end{aligned} \tag{3}$$

$$= m\binom{T'}{m} + 2\binom{T'-2}{m-2} \sum_{\substack{1\le i<j\le T' \\ \mathbf{v_i},\mathbf{v_j}\in\mathcal{C}}} \mathbf{v_i}\mathbf{v_j} \tag{4}$$

$$= m\binom{T'}{m} + 2\binom{T'-2}{m-2} \sum_{\substack{1\le i<j\le T' \\ \mathbf{v_i},\mathbf{v_j}\in\mathcal{C}}} \mathbf{v_i}\mathbf{v_j} + T'\binom{T'-2}{m-2} - T'\binom{T'-2}{m-2} \tag{5}$$

$$= m\binom{T'}{m} + 2\binom{T'-2}{m-2} \sum_{\substack{1\le i<j\le T' \\ \mathbf{v_i},\mathbf{v_j}\in\mathcal{C}}} \mathbf{v_i}\mathbf{v_j} + \binom{T'-2}{m-2}\sum_{\mathbf{v}\in\mathcal{C}}\mathbf{v}^2 - T'\binom{T'-2}{m-2} \tag{6}$$

$$= m\binom{T'}{m} + \binom{T'-2}{m-2}\left(\sum_{\mathbf{v}\in\mathcal{C}}\mathbf{v}\right)^2 - T'\binom{T'-2}{m-2} \tag{7}$$

$$= m\binom{T'}{m} + T'^2\mathbf{c}^2\binom{T'-2}{m-2} - T'\binom{T'-2}{m-2}. \tag{8}$$

In (3), it is used that all vectors are of unit norm, while (4) follows from the fact that a pair of vectors is contained in exactly $\binom{T'-2}{m-2}$ m-tuples, thus every product $\mathbf{v_i}\mathbf{v_j}$ occurs with this multiplicity. In (6), it is used again, that the vectors are unit norm ones, thus $\sum_{\mathbf{v}\in\mathcal{C}}\mathbf{v}^2 = T'$. An easy calculation gives (7), while (8) follows from the definition $\sum_{\mathbf{v}\in\mathcal{C}}\mathbf{v} = T'\mathbf{c}$.

Continuing (2) by the above computation we get

$$\sum_{\mathcal{A}\in\binom{\mathcal{C}}{m}} \|f(\mathcal{A}) - m\mathbf{c}\|^2 = m\binom{T'}{m} + T'^2\mathbf{c}^2\binom{T'-2}{m-2} - T'\binom{T'-2}{m-2} - m^2\mathbf{c}^2\binom{T'}{m}$$
$$= m\binom{T'}{m} - T'\binom{T'-2}{m-2} + T'^2\mathbf{c}^2\binom{T'-2}{m-2} - m^2\mathbf{c}^2\frac{T'(T'-1)}{m(m-1)}\binom{T'-2}{m-2}.$$

From $m \le T'$, it follows that $-m^2(T'(T'-1)/m(m-1)) \le -T'^2$. Therefore

$$\sum_{\mathcal{A}\in\binom{\mathcal{C}}{m}} \|f(\mathcal{A}) - m\mathbf{c}\|^2 \le m\binom{T'}{m} - T'\binom{T'-2}{m-2} + T'^2\mathbf{c}^2\binom{T'-2}{m-2} - T'^2\mathbf{c}^2\binom{T'-2}{m-2} \le m\binom{T'}{m}$$

which gives the desired result. □

Now we are ready to prove the new upper bound on the rate of *Euclidean superimposed codes.*

Theorem 3.2:

$$R(m, d) = \limsup_{n\to\infty} \frac{\log T(n, m, d)}{n} \le \frac{\log m}{2m}(1 + o(1))$$

where d is a constant and $o(1)$ is a function tending to zero as m tends to infinity.

Proof: Take an arbitrary *Euclidean superimposed code* $\mathcal{C}(n, m, T, d)$ and denote—similarly to the above lemma—by $\mathbf{c}$ the (average) vector $(1/T)\sum_{\mathbf{v}\in\mathcal{C}}\mathbf{v}$. Let ξ be a random variable with probability distribution $\mathbf{P}(\xi = \|f(\mathcal{A}) - m\mathbf{c}\|) = 1/\binom{T}{m}$ ($\mathcal{A} \subseteq \mathcal{C}$, $|\mathcal{A}| = m$). By Lemma 3.1 (and Jensen's inequality), the expected distance $\mathbf{E}(\xi) \le m^{1/2}$. Thus, by Markov's inequality, $\mathbf{P}(\xi \ge \lambda m^{1/2}) \le 1/\lambda$. This means that for any constant $\lambda > 1$, at least the $(1 - 1/\lambda)$ fraction of all sums of the m-tuples of $\mathcal{C}$ lies within the n-dimensional ball of radius $\lambda m^{1/2}$ centered about the point $m\mathbf{c}$.

But $\mathcal{C}$ is an *Euclidean superimposed code,* which means that those vectors (of sums of m-tuples) have distance at least d from each other. Applying to these vectors the sphere packing argument, we get

$$\frac{\lambda - 1}{\lambda}\binom{T}{m} \le \left(\frac{\lambda m^{1/2} + d/2}{d/2}\right)^n = \left(1 + \frac{2\lambda m^{1/2}}{d}\right)^n \tag{9}$$

from which (since λ is a constant)

$$R(m, d) \le \frac{1}{m}\log\left(1 + \frac{2\lambda m^{1/2}}{d}\right) \tag{10}$$

immediately follows. For large m, (10) is of the form

$$R(m, d) \le \frac{\log m}{2m}(1 + o(1))$$

which gives the desired result.

IV. Superimposed Codes in Other Normed Spaces

Let $\mathcal{N} = (X, \|\cdot\|)$ be a finite-dimensional (n-dimensional) normed vector space, and let $B(\mathbf{c}, r)$ denote the closed ball with center $\mathbf{c}$ and radius $r > 0$. We also use B for the unit ball $B(\mathbf{0}, 1)$ of $\mathcal{N}$. In general, this B may also be considered as an arbitrary n-dimensional symmetric convex body in $\mathbf{R}^n$, the symmetry being with respect to the origin. One might also be interested in the growth rate of *superimposed codes* in the more general normed vector space $\mathcal{N}$, where the norm is defined by an arbitrary n-dimensional central symmetric convex body. Similarly to the Euclidean norm case, this means the following.

Let $\mathcal{C}$ be a finite set of (at most) unit norm vectors in $\mathcal{N}$. It is called a *superimposed code in* $\mathcal{N}$ with parameters (n, m, T, d) if $|\mathcal{C}| = T$ and for two arbitrary distinct subsets $\mathcal{A}$ and $\mathcal{B}$ of $\mathcal{C}$ with $0 \le |\mathcal{A}|, |\mathcal{B}| \le m$ the $\mathcal{N}$ distance of the vectors $f(\mathcal{A})$ and $f(\mathcal{B})$ is at least d. That is

$$d_{\mathcal{N}}(\mathcal{C}^m) := \min_{\substack{\mathcal{A}\ne\mathcal{B} \\ 0\le|\mathcal{A}|,|\mathcal{B}|\le m \\ \mathcal{A},\mathcal{B}\subseteq\mathcal{C}}} \|f(\mathcal{A}) - f(\mathcal{B})\|_{\mathcal{N}} \ge d.$$

As before, for given n, m, and d, let $T_{\mathcal{N}}(n, m, d)$ denote the maximum size of such a code.

We are able to extend the bounds of inequality (1) for all finite-dimensional normed spaces $\mathcal{N}$ in a somewhat weaker form.

Theorem 4.1:

$$\frac{\log T_{\mathcal{N}}(n, m, d)}{n} = \Theta\left(\frac{\log m}{m}\right).$$

More precisely, there is an absolute constant C (independent of m, n, $d \le 1$, and of the space $\mathcal{N}$) such that

$$\frac{\log T_{\mathcal{N}}(n, m, d)}{n} \ge C\frac{\log m}{m} \tag{11}$$

as $n \gg m$. Moreover, for all m, n, and $\mathcal{N}$, one has

$$\frac{\log T_{\mathcal{N}}(n, m, d)}{n} \le \frac{\log m}{m} + O\left(\frac{1}{m}\right) + O\left(\frac{\log n}{n}\right). \tag{12}$$

Here, O and Θ are used in the conventional sense, i.e., for sequences $f(m)$ and $g(m)$, $f(m) = O(g(m))$ if $f(m) \le cg(m)$ holds for some constant $c > 0$ and every m, and $f(m) = \Theta(g(m))$ if both $f(m) = O(g(m))$ and $g(m) = O(f(m))$ hold.

Proof: (Sketch) To prove the lower bound (11), use the following theorem of Milman [19]. For every $\varepsilon > 0$, there exists a positive constant $\psi(\varepsilon) > 0$, such that one can find a projection of a section of B [say, $\Pi_{F_2}(F_1 \cap B)$ with $F_2 \subset F_1 \subset \mathbf{R}^n$] which is $(1+\varepsilon)$ equivalent to an ellipsoid and has a dimension at least $\psi(\varepsilon)n$. Here, the ψ is independent from the convex body B, but, of course, the choice of the subspaces F_1 and F_2 varies with B. (For more background on this topic and proofs, see the excellent book of Pisier [20].) An ellipsoid is affine invariant to the Euclidean ball, so taking an *Euclidean superimposed code* $\mathcal{C}$ of maximum size in the subspace F_2—by the affine invariant transformation mapping the unit ball to the ellipsoid—we will get a superimposed code with the same parameters with respect to the distance defined by the ellipsoid. Project $\mathcal{C}$ back to $F_1 \cap B$, and by Milman's theorem obtain a superimposed code in $\mathcal{N}$ with parameters $(n, m, |\mathcal{C}|, d/(1+\varepsilon))$.

The upper bound in (12) easily follows from the volume bound of Ericson and Györfi

$$\binom{T}{m} \le \left(\frac{m + d/2}{d/2}\right)^n \tag{13}$$

which is true for every space $\mathcal{N}$ and every n, m, and d. □

At the present, we are not able to sharpen the upper bound of (1) for normed spaces other than the Euclidean. The argument in the previous section substantially utilized the properties of the scalar product, which other spaces lack. However, one can slightly improve (13) for *small* dimension, i.e., for $n < m$, although this does not say anything about the growth rate when $n \to \infty$. Fritz John's [13] classical result says that for every symmetric convex body B centered about the origin there is an ellipsoid D such that $D \subset B \subset \sqrt{n}D$. This implies that every normed space is (affine) $\sqrt{n}$ equivalent to $\mathbf{R}^n$, so our modified volume bound (9) gives

$$\frac{\lambda - 1}{\lambda}\binom{T}{m} \le \left(\frac{\lambda m^{1/2} + d/(2\sqrt{n})}{d/(2\sqrt{n})}\right)^n = \left(1 + \frac{2\lambda m^{1/2} n^{1/2}}{d}\right)^n.$$

It would be interesting to find better bounds, especially for the maximum norm ℓ_∞^n, where B is the (hyper)cube.

Acknowledgment

The authors would like to thank T. Ericson and L. Györfi for fruitful discussions.

References

[1] Q. A. Nguyen and T. Zeisel, "Bounds on constant weight binary superimposed codes," *Prob. Contr. Inform. Theory,* vol. 17, pp. 223–230, 1988.

[2] P.-O. Anderson, "Superimposed codes for the Euclidean channel," Ph.D. dissertation, Linköping Studies in Science and Technology, Linköping, Sweden.

[3] ——, "A construction of superimposed codes for the Euclidean channel," in *Proc. 1994 IEEE Int. Symp. Inform. Theory,* Trondheim, Norway, June 27–July 1, 1994, p. 283.

[4] A. G. Dyachkov and V. V. Rykov, "Bounds on the length of disjunctive codes," *Problemy Peredaci Informacii,* vol. 18, no. 3, pp. 7–13, 1982.

[5] ——, "A survey of superimposed codes theory," *Prob. Contr. Inform. Theory,* vol. 12, no. 4, pp. 1–13, 1983.

[6] P. Erdős, P. Frankl, and Z. Füredi, "Families of finite sets in which no set is covered by the union of two others," *J. Comb. Theory, Series A,* vol. 33, no. 2, pp. 158–166, 1982.

[7] ——, "Families of finite sets in which no set is covered by the union of r others," *Israel J. Math.,* vol. 51, no. 1/2, pp. 79–89, 1985.

[8] T. Ericson and L. Györfi, "Superimposed codes in R^n," *IEEE Trans. Inform. Theory,* vol. 34, pp. 877–880, July 1988.

[9] Z. Füredi, "A note on r-cover-free families," *J. Comb. Theory, Series A,* vol. 73, pp. 172–173, 1996.

[10] Z. Füredi and M. Ruszinkó, "Superimposed codes are almost big distance ones," in *Proc. 1997 IEEE Int. Symp. Inform. Theory,* Ulm, Germany, June 29–July 4, 1997, p. 118.

[11] F. K. Hwang, "A method for detecting all defective members in a population by group testing," *J. Amer. Stat. Assoc.,* vol. 67, no. 339, pp. 605–608, 1972.

[12] F. K. Hwang and V. T. Sós, "Non adaptive hypergeometric group testing," *Studia Sci. Math. Hungar.,* vol. 22, pp. 257–263, 1987.

[13] F. John, "Extremum problems with inequalities as subsidiary conditions," in *Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948.* New York: Interscience, 1948, pp. 187–204.

[14] S. M. Johnson, "On the upper bounds for unrestricted binary error-correcting codes," *IEEE Trans. Inform. Theory,* vol. IT-17, no. 4, pp. 466–478, 1971.

[15] ——, "Improved asymptotic bounds for error-correcting codes," *IEEE Trans. Inform. Theory,* vol. IT-9, no. 4, pp. 198–205, 1963.

[16] ——, "A new upper bound for error-correcting codes," *IRE Trans. Inform. Theory,* vol. IT-8, pp. 203–207, 1962.

[17] G. A. Kabatjanskiĭ and V. I. Levenštein, "Bounds for packings on a sphere and in a space," (Russian) *Problemy Peredaci Informacii,* vol. 14, pp. 3–25, 1978.

[18] W. H. Kautz, and R. C. Singleton, "Nonrandom binary superimposed codes," *IEEE Trans. Inform. Theory,* vol. IT-10, pp. 363–377, Oct. 1964.

[19] V. D. Milman, "Almost Euclidean quotient spaces of subspaces of finite dimensional normed spaces," in *Proc. Amer. Math. Soc.,* 1985, vol. 94, pp. 445–449.

[20] G. Pisier, "The volume of convex bodies and Banach space geometry," *Cambridge Tracts in Mathematics 9.* Cambridge, U.K.: Cambridge, Univ. Press, 1989.

[21] M. Ruszinkó, "On the upper bound of the size of the r-cover-free families," *J. Comb. Theory, Series A,* vol. 66, no. 2, pp. 302–310, May 1994.

[22] V. T. Sós, "An additive problem in different structures," in *Proc. 2nd Int. Conf. in Graph Theory, Combinatorics, Algorithms, and Applications,* San Francisco Univ., San Francisco, CA, July 1989, pp. 486–510.