SPHERE PACKINGS CONSTRUCTED FROM BCH AND JUSTESEN CODES

N. J. A. SLOANE

Abstract. Bose–Chaudhuri–Hocquenghem and Justesen codes are used to pack equa spheres in n-dimensional Euclidean space with density Δ satisfying

$$\log_2 \Delta > -6n + o(n),$$

for all sufficiently large n of the form $m2^m$, where m is a power of 4. These appear to be the densest packings yet constructed in high dimensional space.

1. *Introduction.* The sphere packing problem is to pack equal spheres in n-dimensional Euclidean space E^n so as to maximize the density Δ, *i.e.*, the fraction of space covered by the spheres. It is known [**8**, pp. 4, 13] that the densest packings satisfy

$$-n < \log \Delta < -\frac{n}{2} + \log \frac{n+2}{2}.$$

(All logarithms in this paper are to the base 2.) Packings have been constructed from error-correcting codes in [**4**–**6**]. In [**5**–**6**], Bose–Chaudhuri–Hocquenghem (BCH) codes (see [**1**, Ch. 7]) were used to obtain non-lattice packings with

$$\log \Delta \sim -\frac{n}{2} \log \log n \text{ as } n \to \infty$$

when n is a power of 2. In the present paper we use BCH and Justesen [**2**] codes to construct non-lattice packings with

$$\log \Delta > -6n + o(n)$$

for all n of the form $n = m2^m$, where $m \geqslant 256$ is a power of 4.

2. *Definitions.* An (n, k, d) *linear code* over $GF(q)$ is a linear subspace of $GF(q)^n$ of dimension k, such that any two vectors (or *codewords*) in the subspace differ in at least d places. n is called the *block length* of the code, k the *dimension*, d the *minimum distance*, and $R = k/n$ the *rate*.

The *coordinate array* [**3**, §1.42] of a point in E^n having integer coordinates is formed by setting out in columns the values of the coordinates in the binary scale. The 1's row of the array comprises the 1's digits of the coordinates, and thus has 0's for even coordinates and 1's for odd coordinates. The 2's, 4's, 8's, ... rows similarly comprise the 2's, 4's, 8's, ... digits of the coordinates. Complementary notation is used for negative integers.

Other definitions of sphere packing terms will be found in [**8**], and of coding theory terms in [**1**, **7**].

3. *The Construction.* Let $\mathscr{C}_\beta$ be an $(n = m2^m = 2^{2a}, k_\beta, d_\beta = 4^\beta)$ linear binary code, for $0 \leqslant \beta \leqslant a$. A sphere packing in E^n is obtained by taking as centres all

points x with integer coordinates such that the $2^{a-\beta}$'s row of the coordinate array of x is in $\mathscr{C}_\beta$, for $\beta = 0, \ldots, a$. (This is Construction C of [6].)

If the first row in which two centres differ is the 2^i's row, then (i) if $i > a$ their (Euclidean) distance apart is at least 2^{a+1}, and (ii) if $0 \leqslant i \leqslant a$ they differ by at least 2^i in at least d_i coordinates, and so are at least $(d_i . 4^i)^{\frac{1}{2}} = 2^a$ apart. Thus we may take the radius of the spheres to be 2^{a-1}.

It is then easy to show [6, §6.3] that this packing has density Δ given by

$$\log \Delta = \sum_{\beta=0}^{a} k_\beta - \tfrac{1}{2} n \log (8n/e\pi) + O(\log n). \tag{1}$$

In general a non-lattice packing is obtained.

4. *The Codes.* We now specify the codes $\mathscr{C}_\beta$ to be used in the construction.

We divide the range of β into 3 parts: range 1,

$$0 \leqslant \beta < \beta_0 = \tfrac{1}{2} \log n - \tfrac{5}{8} \log \log n;$$

range 2, $\beta_0 \leqslant \beta < \beta_1 = \frac{1}{2} \log n - 5$; range 3, $\beta_1 \leqslant \beta \leqslant \frac{1}{2} \log n$.

In range 1 we take $\mathscr{C}_\beta$ to be the extended BCH code of length n and minimum distance 4^β which has the largest number 2^{k_β} of codewords. Then it follows immediately from [6, §6.7] that

$$\sum_{0 \leqslant \beta < \beta_0} k_\beta > \tfrac{1}{2} n \log n - \tfrac{5}{8} n \log \log n + o(n). \tag{2}$$

In range 3 we take $\mathscr{C}_\beta$ to consist just of the zero codeword, so $k_\beta = 0$.

5. *Justesen Codes.* In range 2 we take $\mathscr{C}_\beta$ to be a shortened Justesen code. Since the original Justesen codes have block lengths which depend in a complicated way on the rate, and our codes $\mathscr{C}_\beta$ must all have the same block length n, some care is needed in the construction.

In what follows $m \geqslant 256$ is a fixed power of 4, and $n = m2^m$.

For any integer s, $0 < s < m$, we construct a Justesen code $\mathscr{J}_s$ as follows. Let $r = r(s) = m/(2m - s)$, $K = K(s) = [r2^m]$.

First let α_s be an $(N = 2^m, K, N - K)$ extended Reed–Solomon code [1, p. 310] over $GF(2^m)$, in which the last coordinate of each codeword is an overall parity check. Let α be a primitive element of $GF(2^m)$.

We form the matrix

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a_0 \ldots a_{N-1} \\ b_0 \ldots b_{N-1} \end{pmatrix}$$

over $GF(2^m)$, where $a \in \alpha_s$ and $b_i = \alpha^i a_i$. A fixed basis for $GF(2^m)$ over $GF(2)$ is chosen, and each element of the matrix is replaced by the corresponding binary column vector of length m. Then the last s rows of the new matrix are deleted.

The resulting binary $(2m - s) \times 2^m$ matrices, for all $a \in \alpha_s$, considered now as vectors of length $n' = n'(s) = (2m - s) 2^m$, form an $(n', k' = mK, d)$ linear binary Justesen [2] code which we denote by $\mathscr{J}_s$. The rate of this code is

$$R = R(s) = k'/n' = rK2^{-m},$$

and the minimum distance d will be bounded in the next section.

Finally let the shortened Justesen code $\mathscr{K}_s$ be obtained from $\mathscr{J}_s$ as follows. k' binary symbols in each codeword of $\mathscr{J}_s$ can be chosen arbitrarily. If a fixed set of $n' - n$ of them are set equal to zero, the $(n, k = k' - n' + n, d)$ linear binary code $\mathscr{K}_s$ is obtained.

6. *The Minimum Distance of the Justesen Codes.* The aim of this section is to establish a lower bound on the minimum distance d of the Justesen code $\mathscr{J}_s$ (and so of $\mathscr{K}_s$) when s is a large fraction of m.

Notation. $H(x) = -x \log x - (1-x) \log (1-x)$ denotes the binary entropy function, and $x = H^{-1}(y)$ denotes the smaller of the two values of x for which $y = H(x)$.

Let $\delta = 1 - r = (m-s)/(2m-s)$ and $\gamma = 1 - R/r$.

THEOREM 1. *If*

$$0 \cdot 8m < s < m(1 - (\log\log n)^{\frac{1}{2}} (\log n)^{-\frac{1}{2}}), \tag{3}$$

then the ratio of minimum distance to block length of the Justesen code $\mathscr{J}_s$ satisfies

$$\frac{d}{n'} > H^{-1}(\delta)(\delta - \varepsilon_1) \tag{4}$$

where

$$0 < \varepsilon_1 < 5 \log\log n/\log n. \tag{5}$$

Theorem 1 is a refinement of Theorem 2 of [2]. The proof depends on two lemmas, the first of which is elementary.

LEMMA 1. *For* $0 < y < \frac{1}{2}$,

$$y/\left(\log\frac{1}{y} + \log\log\frac{1}{y} + 4\right) < H^{-1}(y) < y/\left(\log\frac{1}{y} + \log\log\frac{1}{y}\right).$$

LEMMA 2. *The total number of* 1's, W, *in* $N_0 = [\gamma 2^{m-s}]$ *distinct nonzero binary* $(2m-s)$*-tuples satisfies*

$$W > (2m-s)H^{-1}(\delta)N_0(1 - \varepsilon_2/\delta)$$

where $0 < \varepsilon_2 < 4{\cdot}2 \log\log n/\log n$.

This refines the lemma of [2].

Proof. The idea is to choose a fraction t in such a way that the total number, N_1, say, of binary $(2m-s)$-tuples of weight $\leqslant [t(2m-s)]$,

$$N_1 = \sum_{i=1}^{[t(2m-s)]} \binom{2m-s}{i},$$

is less than N_0. Then the remaining $N_0 - N_1 (2m-s)$-tuples have weight $> t(2m-s)$, and so

$$W > t(2m-s)(N_0 - N_1).$$

(The weight of a vector is the number of its nonzero components.) A good choice for t is

$$t = H^{-1}(\delta) - \psi,$$

where

$$\psi = \frac{\log m + \log\log(1/\delta)}{(2m - s)\log(1/\delta)}.$$

Then

$$\begin{aligned} W &> (2m - s)(H^{-1}(\delta) - \psi)(N_0 - N_1) \\ &> (2m - s)H^{-1}(\delta)N_0(1 - \varepsilon_3) \end{aligned}$$

where

$$\varepsilon_3 = \frac{\psi}{H^{-1}(\delta)} + \frac{N_1}{N_0}.$$

In order to estimate ε_3 we need some bounds. From $n = m2^m$ follows, for $m \geqslant 16$,

$$\log n - \log\log n < m < \log n - \tfrac{1}{2}\log\log n.$$

From (3) follow

$$\tfrac{1}{2}(\log\log n)^{\frac{1}{2}}(\log n)^{-\frac{5}{8}} < \delta < \tfrac{1}{6}, \tag{6a}$$

$$2{\cdot}58 < \log(1/\delta) < \tfrac{5}{8}\log\log n, \tag{6b}$$

$$1{\cdot}36 < \log\log(1/\delta) < \log\log\log n. \tag{6c}$$

For any fixed s satisfying (3), it follows from their definitions that

$$\delta \leqslant \gamma < \delta + 2^{-m}, \tag{7a}$$

$$r(r - 2^{-m}) < R \leqslant r^2, \tag{7b}$$

and from Lemma 1 that

$$\delta\Big/\left(\log\frac{1}{\delta} + \log\log\frac{1}{\delta} + 4\right) < H^{-1}(\delta) < \delta\Big/\left(\log\frac{1}{\delta} + \log\log\frac{1}{\delta}\right).$$

From the standard bounds [7] on binomial coefficients,

$$N_1 < \frac{1 - t}{1 - 2t}\binom{2m - s}{[t(2m - s)]} < \frac{1 - t}{1 - 2t}2^{(2m-s)H(t)}.$$

From the mean value theorem, for some θ between 0 and 1,

$$\begin{aligned} H(t) &= H(t_0) - \psi H'(t_0 - \theta\psi), \text{ where } t_0 = H^{-1}(\delta), \\ &= \delta - \psi\log\frac{1 - t_0 + \theta\psi}{t_0 - \theta\psi} \\ &< \delta - \psi\log\frac{1 - t_0}{t_0}. \end{aligned}$$

Also

$$\frac{1}{N_0} < \frac{1}{\gamma 2^{\delta(2m-s)}}\left(1 + \frac{2}{N_0}\right).$$

Therefore, for the second term in ε_3,

$$\frac{N_1}{N_0} < \frac{1}{\gamma}\cdot\frac{1 - t}{1 - 2t}\cdot\left(1 + \frac{2}{N_0}\right).2^{-(2m-s)(\delta - H(t))} < \frac{2}{\delta}2^{-(2m-s)\psi\log((1-t_0)/t_0)}.$$

Now $(1 - t_0)/t_0 > \dfrac{1}{\delta} \log \dfrac{1}{\delta}$, so

$$\frac{N_1}{N_0} < \frac{2}{\delta} 2^{-\{\log m + \log \log (1/\delta)\}\{1 + \log \log (1/\delta)/\log (1/\delta)\}}$$

$$= \frac{2}{\delta} \left(m \log \frac{1}{\delta}\right)^{-\{1 + \log \log (1/\delta)/\log (1/\delta)\}}.$$

For the first term in ε_2,

$$\frac{\psi}{H^{-1}(\delta)} < \frac{\{\log m + \log \log (1/\delta)\}\{\log (1/\delta) + \log \log (1/\delta) + 4\}}{m\delta \log (1/\delta)}$$

$$< \frac{4}{m\delta}\left(\log m + \log \log \frac{1}{\delta}\right),$$

which is much larger than the bound on N_1/N_0. Therefore

$$\varepsilon_3 < \frac{4 \cdot 1}{m\delta}\left(\log m + \log \log \frac{1}{\delta}\right) < \frac{(4 \cdot 2) \log \log n}{\delta \log n},$$

which establishes the lemma.

Proof of Theorem 1. The minimum distance d of the Justesen code $\mathscr{J}_s$ is estimated as follows. From the definition of α_s, at least $N - K$ out of the first $N - 1$ components of each nonzero codeword of α_s are nonzero. For these nonzero components a_i, the binary columns of length $2m$ corresponding to $\begin{pmatrix} a_i \\ \alpha_i a_i \end{pmatrix}$ are obviously all distinct. However, after s rows of the matrix have been deleted, each nonzero $(2m - s)$-tuple may be repeated up to 2^s times. The worst case occurs when the Justesen codeword contains 2^s repetitions of each of the $[2^{-s}(N - K)]$ distinct lowest weight nonzero $(2m - s)$-tuples.

Now $2^{-s}(N - K) = \left(1 - \dfrac{R}{r}\right) 2^{m-s} = \gamma 2^{m-s}$, so Lemma 2 may be applied directly to give, with the help of (7a),

$$\frac{d}{n'} \geqslant \frac{2^s W}{n'} > H^{-1}(\delta)\delta\left(1 - \frac{\varepsilon_2}{\delta}\right)\left(1 - \frac{1}{\gamma 2^{m-s}}\right)$$

$$> H^{-1}(\delta)\left(\delta - \varepsilon_2 - \frac{1}{\gamma 2^{m-s}}\right)$$

$$> H^{-1}(\delta)(\delta - \varepsilon_1)$$

where ε_1 satisfies (5). This completes the proof of Theorem 1.

7. *The rate of the Justesen codes.*

THEOREM 2. *For any β in range 2, there is an $s = s_\beta$ satisfying* (3) *such that the corresponding $(n_\beta' = (2m - s_\beta) 2^m, k_\beta', d_\beta)$ Justesen code $\mathscr{J}_{s_\beta}$ has minimum distance*

$d_\beta > 4^\beta$. *Furthermore the rate of this code satisfies*

$$\frac{k_\beta'}{n_\beta'} > 1 - Q(4^\beta/n) - 2\varepsilon_1,$$

where ε_1 *satisfies* (5), *and*

$$Q(y) = \left\{2y\left(\log\frac{1}{y} + 4\log\log\frac{1}{y}\right)\right\}^{\frac{1}{2}}.$$

Proof. Provided s satisfies (3), it follows from Theorem 1 and Lemma 1 that for the $(n' = (2m - s)\,2^m, k', d)$ Justesen code $\mathscr{J}_s$,

$$\frac{d}{n} > \frac{d}{n'} > \frac{\delta(\delta - \varepsilon_1)}{\log(1/\delta) + \log\log(1/\delta) + 4}. \tag{8}$$

We first note that as s runs through the values specified by (3), δ runs through (6a), and the right hand side of (8) runs from

$$\tfrac{2}{5}(\log n)^{-\frac{5}{4}} \text{ to } \tfrac{1}{288},$$

Now as β runs through range 2, $d/n = 4^\beta/n$ runs from

$$(\log n)^{-\frac{5}{4}} \text{ to } \tfrac{1}{1024},$$

which is included in the above range. Therefore we can find codes with the desired minimum distance with s_β satisfying (3).

Next we calculate s_β. For $y > 0$ let $\delta = \delta_a(y)$ be the unique solution of

$$y = \frac{\delta(\delta - \varepsilon_1)}{\log(1/\delta) + \log\log(1/\delta) + 4}. \tag{9}$$

From (8), (9), if the code $\mathscr{J}_s$ is such that $\delta \geqslant \delta_a(4^\beta/n)$, then $d > 4^\beta$. If it were not for the fact that s_β must be an integer, we would take $s_\beta = m(1 - 2\delta_\beta)/(1-\delta_\beta)$ where $\delta_\beta = \delta_a(4^\beta/n)$.

But s_β must be an integer, and since $\delta = (m - s)/(2m-s)$, an increase in s by 1 corresponds to a decrease in δ by $m(2m - s)^{-2} < m^{-1}$. Therefore let

$$\delta_b(y) = \delta_a(y) + m^{-1}, \tag{10}$$

$$s_\beta = \left[m \,.\, \frac{1 - 2\delta_b(4^\beta/n)}{1 - 2\delta_b(4^\beta/n)}\right] + 1.$$

The corresponding value of δ is

$$\delta_\beta = \frac{m - s_\beta}{2m - s_\beta};$$

(10) implies that $\delta_b(4^\beta/n) \geqslant \delta_\beta \geqslant \delta_a(4^\beta/n)$, and so $d > 4^\beta$.

An upper bound on $\delta_a(y)$ is obtained as follows. From (9) and (6) we find that

$$\log\frac{1}{y} > 2\log\frac{1}{\delta_a} + \log\log\frac{1}{\delta_a}, \tag{11a}$$

$$\log\log\frac{1}{y} > \log\log\frac{1}{\delta_a}. \tag{11b}$$

Also (9) implies

$$\delta_a^2 - \varepsilon_1 \delta_a - y\left(\log \frac{1}{\delta_a} + \log\log \frac{1}{\delta_a} + 4\right) = 0,$$

and so from (11),

$$\delta_a < \frac{1}{2}\varepsilon_1 + \left\{\frac{1}{4}\varepsilon_1^2 + \frac{1}{2}y\left(\log \frac{1}{y} + \log\log \frac{1}{y} + 8\right)\right\}^{\frac{1}{2}}$$

$$< \frac{1}{2}\varepsilon_1 + \frac{1}{2}Q(y). \tag{12}$$

Finally we calculate the rate of $\mathscr{I}_{s_\beta}$. From (7b) and (12),

$$\begin{aligned} R &> r(r - 2^{-m}) = (1 - \delta_\beta)(1 - \delta_\beta - 2^{-m}) \\ &> (1 - \delta_a - m^{-1})(1 - \delta_a - m^{-1} - 2^{-m}) \\ &> 1 - \varepsilon_1 - Q(4^\beta/n) - 2m^{-1} - 2^{-m} \\ &> 1 - 2\varepsilon_1 - Q(4^\beta/n), \end{aligned}$$

which completes the proof of the theorem.

8. *The Density of the Sphere Packing.* We can now complete the construction of the sphere packing, taking $\mathscr{C}_\beta = \mathscr{K}_{s_\beta}$ for β in range 2, where $\mathscr{K}_{s_\beta}$ is obtained by shortening the Justesen code $\mathscr{I}_{s_\beta}$ as described at the end of §5.

THEOREM 3. *Let $m \geqslant 256$ be a power of* 4, *and $n = m2^m$. The sphere packing in E^n obtained by using BCH codes in range* 1, *shortened Justesen codes in range* 2, *and the zero code in range* 3, *has density Δ satisfying*

$$\log \Delta > -6n + o(n).$$

Proof. The contribution to (1) from the BCH codes is given by (2). The contribution from the Justesen codes is, from Theorem 2,

$$\begin{aligned} \sum_{\beta_0 \leqslant \beta < \beta_1} k_\beta &= \sum(k_\beta' - n_\beta' + n) \\ &> \sum\{n - n'(Q(4^\beta/n) + 2\varepsilon_1)\} \\ &> n(\tfrac{5}{8}\log\log n - 5) - 2n\sum_2 + o(n), \end{aligned} \tag{13}$$

where

$$\sum_2 = \sum_{\beta_0 \leqslant \beta < \beta_1}\left\{2.\frac{4^\beta}{n}\left(\log\frac{n}{4^\beta} + 4\log\log\frac{n}{4^\beta}\right)\right\}^{\frac{1}{2}}$$

$$= \sqrt{\left(\frac{2}{n}\right)} \sum_{\beta_0 \leqslant \beta < \beta_1} 2^\beta\left(\log\frac{n}{4^\beta} + 4\log\log\frac{n}{4^\beta}\right)^{\frac{1}{2}}.$$

Let T_β be the β-th term in the latter sum. A straightforward calculation shows that $T_\beta/T_{\beta+1}$ is an increasing function of β in range 2. Therefore

$$T_\beta/T_{\beta+1} < T_{\beta_1-1}/T_{\beta_1} < 0{\cdot}6;$$

$\sum T_\beta < (0{\cdot}6)\, T_{\beta_1}/(1 - 0{\cdot}6) < \frac{1}{4}\sqrt{n}$; and $\sum_2 < 2^{-\frac{3}{2}}$. The theorem then follows from (1), (2) and (13).

9. *A Remark.* The coefficient 6 in the statement of Theorem 3 could be reduced by a more careful analysis. However, the Hamming or sphere-packing bound for error-correcting codes [7, p. 52] implies that the density Δ of *any* packing in E^n obtained from Construction C of [6] is bounded by

$$\log \Delta < n\left(\tfrac{1}{2}\log\frac{\pi e}{4} - \sum_{i=0}^{\infty} H\left(\frac{1}{4^i}\right)\right) + o(n)$$

$$= -\,0{\cdot}7702\ldots n + o(n).$$

References

1. E. R. Berlekamp, *Algebraic Coding Theory* (McGraw–Hill, New York, 1968).
2. J. Justesen, " A class of constructive asymptotically good algebraic codes ", *IEEE Trans. Information Theory*, 18 (1972), 652–656.
3. J. Leech, " Notes on Sphere Packings ", *Canadian J. Math.*, 19 (1967), 251–267.
4. ———, and N. J. A. Sloane, " New Sphere Packings in Dimensions 9–15 ", *Bull. Amer. Math. Soc.*, 76 (1970), 1006–1010.
5. ———, and ———, " New Sphere Packings in More than Thirty-Two Dimensions ", *Proc. 2nd Chapel Hill Conf. on a Combinatorial Mathematics and Its Applications* (University of North Carolina at Chapel Hill, 1970), pp. 345–355.
6. ———, and ———, " Sphere Packings and Error-Correcting Codes ", *Canadian J. Math.*, 23 (1971), 718–745.
7. W. W. Peterson, *Error-Correcting Codes* (M.I.T. Press, Cambridge, Mass., 1961), pp. 245–247.
8. C. A. Rogers, *Packings and Coverings* (Cambridge University Press, Cambridge, 1964).

10E30: *Number Theory; Geometry of Numbers; Lattice packing.*
94A10: *Information and Communication; Coding theory.*

Bell Telephone Laboratories, Inc.,
Murray Hill, New Jersey, U.S.A.

(*Received on the* 14*th of April*, 1972.)