On the weight distribution of some countably infinite sequences of Justesen codes and their asymptotic relative minimum distance

Toshiyuki KOHNSU
Dept. of Engineering Informatics
Osaka Electro-Communication University
Neyagawa, Osaka 572-8530, Japan
kohnsu@kns.osakac.ac.jp

Kin-ichiroh TOKIWA
Dept. of Electrical Eng. and Electronics
Osaka Sangyo University
Daito, Osaka 574-8530, Japan
tokiwa@elec.osaka-sangyo.ac.jp

Toshihisa NISHIJIMA
Dept. of Digital Media Sciences
Hosei University
Koganei, Tokyo 184-8584, Japan
tnishio@hosei.ac.jp

Abstract — Justesen first constructed asymptotically good concatenated codes such that the outer code is Reed-Solomon (RS) code and the inner code is Wozencraft’s ensemble of randomly shifted codes. Kolev gave the weight distributions of Justesen codes in some cases. In this paper, we give the weight distributions of Justesen codes obtained by using some other RS codes as the outer code and examine their relative minimum distance.

I. INTRODUCTION

Justesen first constructed asymptotically good concatenated codes such that the outer code is Reed-Solomon (RS) code and the inner code is Wozencraft’s ensemble of randomly shifted codes[1]. Kolev gave the weight distributions and the possible weights of Justesen codes without proof in some cases[2]. In this paper, we give the weight distributions of Justesen codes obtained by using some other RS codes as the outer code and examine their relative minimum distance. One of the weight distributions obtained by Kolev is proved.

II. PRELIMINARIES

A. Justesen codes

Justesen codes are examples of concatenated codes. The outer code is an \((N,K,D)\)\(^1\) RS code over \(GF(2^m)\), where \(N = 2^m - 1\), \(D = 2^m - K\). Let \(\mathcal{F}_{m,K,b}\) denote the \((N,K)\) RS code over \(GF(2^m)\) with the generator polynomial

\[
G(x) = \prod_{i=0}^{b+N-K-1} (x - \alpha^i),
\]

where \(\alpha\) is a primitive element of \(GF(2^m)\) and \(b\) is an integer. Then we have a codeword polynomial

\[
C(x) = U(x)G(x) = c_0 + c_1 x + \cdots + c_{N-1} x^{N-1}, \quad c_i \in GF(2^m),
\]

where \(U(x)\) is an information polynomial

\[
U(x) = u_0 + u_1 x + \cdots + u_{K-1} x^{K-1}, \quad u_i \in GF(2^m).
\]

In this paper, we associate the codeword polynomial \(C(x)\) with its codeword vector representation

\[
C = (c_0, c_1, \ldots, c_{N-1}).
\]

Each component \(c_i\) of \(C\) is encoded into \((c_i, \alpha^i c_i)\), where \(\alpha^i c_i\) is a check symbol appended by an \(i\)-th inner encoder. As a result, we have a vector

\[
J = (c_0, c_0; c_1, \alpha c_1; \ldots; c_{N-1}, \alpha^{N-1} c_{N-1})
\]

over \(GF(2^m)\). Replacing each component of \(J\) with its corresponding binary \(m\)-tuple, we obtain a binary codeword \(j\) of a \((2mN, mK)\) Justesen code \(\mathcal{F}_{m,K,b}\).

The vector representation

\[
B = (c_0, \alpha c_1, \ldots, \alpha^{N-1} c_{N-1})
\]

of check symbols appended by inner encoders has the associated polynomial

\[
B(x) = c_0 + \alpha c_1 x + \cdots + \alpha^{N-1} c_{N-1} x^{N-1} = C(\alpha x).
\]

It should be noted that \(B(x)\) can be interpreted as a codeword polynomial of RS code with generator polynomial \(G(\alpha x)\) and information polynomial \(U(\alpha x)\).

A simple lower bound on the minimum distance of Justesen codes is described in [3]. The ratio of minimum distance to length, that is, the relative minimum distance of Justesen code is given by

\[
\frac{d}{2mN} = \frac{d}{2m(2^m - 1)},
\]

where \(d\) is the minimum distance of Justesen code.

B. Kolev’s results

Kolev gave the weight distributions of \(\mathcal{F}_{m,1,1}\) and \(\mathcal{F}_{m,2,1}\), and the possible weights of codewords of several Justesen codes without proof in [2]. Let \(A_{m,K,b}^w\) denote the number of codewords of weight \(w\) in \(\mathcal{F}_{m,K,b}\). The weight distributions of \(\mathcal{F}_{m,1,1}\) and \(\mathcal{F}_{m,2,1}\) are as follows:

**Theorem 1 ([2]):** The weight distribution of \(\mathcal{F}_{m,1,1}\) is given by

\[
A_{m,1,1}^{m,1,1} = 1
\]

\[
A_{m,2m-1+2^m-1}^{m,1,1} = \binom{m}{t}
\]

for \(t = 1, 2, \ldots, m\).

**Theorem 2 ([2]):** The weight distribution of \(\mathcal{F}_{m,2,1}\) is given by

\[
A_{m,2m}^{m,2,1} = A_{m,2m}^{m,1,1} + 2^m - 1
\]

\[
A_{m,2m-t}^{m,2,1} = A_{m,2m-t}^{m,1,1} + (2^m - m - 1) \binom{m}{t}
\]

\[
A_{m,2m-2^m-1-t}^{m,2,1} = A_{m,2m-2^m-1-t}^{m,1,1} + m \binom{m}{t}
\]

1 The \((N,K,D)\) code denotes the code of length \(N\), dimension \(K\) and minimum distance \(D\).
\[ A_w^{m,2,1} = A_w^{m,1,1}, \]
\[ w \neq m2^m, m2^m - t, m2^m - 2^{m-1} - t, \]
\[ (13) \]
for \( t = 1, 2, \ldots, m. \)

### III. Main Results

In this section, we show the weight distributions of Justesen codes obtained by using some other RS codes as the outer code. In addition, we discuss the relative minimum distance of Justesen codes.

#### A. The case of \( K = 1 \)

**Theorem 3:** The weight distribution of \( J_{m,1,2} \) is given by

\[ A_0^{m,1,2} = 1 \]
\[ A_{m2^m-1}^{m,1,2} = \binom{m}{t} \]
\[ (14) \]
\[ (15) \]
for \( t = 1, 2, \ldots, m. \)

**Proof:** The generator polynomial of \( R_{m,1,2} \) is

\[ G(x) = \prod_{i=2}^{N} (x - \alpha^i) = \frac{x^N - 1}{x - \alpha} = \alpha^{N-1} + \alpha^{N-2} x + \cdots + \alpha x^{N-2} + x^{N-1}, \]
\[ (16) \]
and a sequence of check symbols appended by inner encoders is a codeword of RS code whose generator polynomial is

\[ G(\alpha x) = \frac{x^N - 1}{\alpha(x - 1)} = \alpha^{-1} (1 + x + x^2 + \cdots + x^{N-1}). \]
\[ (17) \]
Each of all nonzero codewords of \( R_{m,1,2} \) has all the distinct nonzero elements of \( GF(2^m) \) as its components and the binary weight of its binary expanded codeword is \( m2^m-1 \) for any nonzero information symbols. Since the sequence of check symbols generated by inner encoders has all the same elements of \( GF(2^m) \), its binary weight is \( (2^m - 1) t \), where \( t \) is binary weight of the element of \( GF(2^m) \). As the number of the sequences whose binary weight is \( (2^m - 1) t \) is \( \binom{m}{t} \), Theorem 3 holds. \( \square \)

**Theorem 4:** Let \( N \) and \( b - 1 \) be relatively prime. And let \( N \) and \( b - 2 \) be also relatively prime. The weight distribution of \( J_{m,1,b} \) is given by

\[ A_0^{m,1,b} = 1 \]
\[ A_{m2^m}^{m,1,b} = 2^m - 1. \]
\[ (18) \]
\[ (19) \]
\[ \text{Proof:} \quad \text{The generator polynomial of } R_{m,1,b} \]
\[ G(x) = \frac{x^N - 1}{x - \alpha^{b-1}} = \alpha^{(N-1)(b-1)} + \alpha^{(N-2)(b-1)} x + \cdots + \alpha x^{(N-2)} + x^{N-1}, \]
\[ (20) \]
and a sequence of check symbols appended by inner encoders is a codeword of RS code whose generator polynomial is

\[ G(\alpha x) = \frac{x^N - 1}{\alpha(x - \alpha^{b-2})} = \alpha^{-1} \left( \alpha^{(N-1)(b-2)} + \alpha^{(N-2)(b-2)} x + \cdots + \alpha^{b-2} x^{N-2} + x^{N-1} \right). \]
\[ (21) \]
Since \( N \) and \( b - 1 \) is relatively prime, each of all nonzero codewords of \( R_{m,1,b} \) has all the distinct nonzero elements of \( GF(2^m) \) for any nonzero information symbols. The sequence of check symbols generated by inner encoders also has all the distinct any nonzero elements of \( GF(2^m) \), because \( N \) and \( b - 2 \) is relatively prime. Since each binary weight is \( m2^m-1 \), the nonzero codewords of \( J_{m,1,b} \) have binary weight \( m2^m \) for any nonzero information symbols. Thus, we obtain Theorem 4.

In the case of \( b = 3 \), then \( N \) and \( b - 1 = N - 1 \) is relatively prime for any \( N \), that is, for any \( m \). Moreover, \( N \) and \( b - 2 = N - 2 \) are also relatively prime for any \( N \). So we have the following corollary.

**Corollary 1:** The weight distribution of \( J_{m,1,3} \) is given by

\[ A_0^{m,1,3} = 1 \]
\[ A_{m2^m}^{m,1,3} = 2^m - 1. \]
\[ (22) \]
\[ (23) \]
**Proof:** The weight distribution of \( J_{m,1,0} \) is given by

\[ A_0^{m,1,0} = 1 \]
\[ A_{m2^m}^{m,1,0} = 2^m - 1. \]
\[ (24) \]
\[ (25) \]
**Proof:** The weight distribution of \( J_{m,1,0} \) is given by

\[ (26) \]
\[ (27) \]
for \( t = 1, 2, \ldots, m \), and

\[ A_{m2^m-s}^{m,2,2} = \sum_{\mu = \max(1, s - m)}^{\min(s+1,m)} \binom{m}{\mu} \binom{m}{s - \mu} \]
\[ (28) \]
for \( s = 2, 3, \ldots, 2m \).

**Proof:** The generator polynomial of \( R_{m,2,2} \) is

\[ G(x) = \prod_{i=2}^{N-1} (x - \alpha^i) = \frac{x^N - 1}{(x - 1)(x - \alpha)} \]
\[ (29) \]
and a sequence of check symbols appended by inner encoders is a codeword of RS code whose generator polynomial is
\[ G(ax) = \frac{x^N - 1}{\alpha^2(x - \alpha^{-1})(x - 1)}. \] (30)

When \( U_1(x) = x + \alpha \) and \( U_2(x) = ax + \alpha \), we have codewords
\[ C_1(x) = 1 + x + x^2 + \cdots + x^{N-1} \] (31)
\[ C_2(x) = 1 + \alpha x + \alpha^2 x^2 + \cdots + \alpha^{N-1} x^{N-1} \] (32)
of \( \mathcal{R}_{m,2,2} \), respectively, and they are a basis of \( \mathcal{R}_{m,2,2} \).
Then \( C = \beta C_1 + \gamma C_2 \) for \( \beta, \gamma \in GF(2^m) \) is a codeword of \( \mathcal{R}_{m,2,2} \).

On the other hand, \( U_1(ax) = ax + \alpha \) and \( U_2(ax) = \alpha^2 x + \alpha \) give two sequences of check symbols as follows:
\[ C_1(ax) = (ax + \alpha) \frac{x^N - 1}{\alpha^2(x - \alpha^{-1})(x - 1)} \]
\[ = 1 + ax + \alpha^2 x^2 + \cdots + \alpha^{N-1} x^{N-1} \]
\[ = B_1(x) \] (33)
\[ C_2(ax) = (\alpha^2 x + \alpha) \frac{x^N - 1}{\alpha^2(x - \alpha^{-1})(x - 1)} \]
\[ = 1 + x + x^2 + \cdots + x^{N-1} \]
\[ = B_2(x). \] (34)

They are a basis of RS code with generator polynomial \( G(ax) \). Since they are check symbols obtained from \( U_1(x), U_2(x) \) by inner encoders, \( B = \beta B_1 + \gamma B_2 \) and \( C = \beta C_1 + \gamma C_2 \) for \( \beta, \gamma \).

In order to obtain weight distribution, we consider four cases such that (i) \( \beta = 0, \gamma = 0 \), (ii) \( \beta \neq 0, \gamma = 0 \), (iii) \( \beta = 0, \gamma \neq 0 \), (iv) \( \beta \neq 0, \gamma \neq 0 \). The weight distribution of each case is as follows:

(i) The binary weight of the codeword of \( \mathcal{J}_{m,2,1} \) is 0.
(ii) The binary weight of \( C \) is \( (2^m - 1) t \), where \( t \) is the binary weight of \( B \), and the binary weight of \( B \) is \( m 2^{m-1} \) for any \( \beta \). The number of codewords in this case is \( \binom{m}{t} \) for \( t = 1, 2, \ldots, m \).
(iii) The binary weights of \( C \) and \( B \) are opposite to the case (ii) and the number of codewords is the same as the case (ii).
(iv) When the binary representation of \( \gamma C_2 \) is represented as a binary \( mxN \) matrix by a basis of \( GF(2^m) \), each row is \( m \)-sequence and its binary weight is \( 2^{m-1} \). The row vector with weight \( 2^{m-1} \) added to all 1’s vector makes the vector with weight \( N - 2^{m-1} = 2^{m-1} - 1 \). When binary representation of \( \beta \) has binary weight \( t \), rows of binary representation of \( \gamma C_2 \) is added to all 1’s vector. So the binary weight of \( C \) is \( (2^{m-1} - 1) t + 2^{m-1} (m - t) = m 2^{m-1} - t \). The binary weight of \( B \) is also obtained as \( (2^{m-1} - 1) t' + 2^{m-1} (m - t') = m 2^{m-1} - t' \), where \( t' \) is the binary weight of binary representation of \( \gamma \) in the same way. Since \( t \) and \( t' \) are independent, we have the weight distribution
\[ A_{m,2,2}^{m,2} = \binom{m}{t} \binom{m}{t'} \] (35)
for \( t = 1, 2, \ldots, m \) and for \( t' = 1, 2, \ldots, m \). Let \( s = t + t' \), then we have (28).

As the result, we have (26) from (i), (27) from (ii) and (iii), and (28) from (iv).

As shown in Theorem 2, Kolev gave the weight distribution of \( \mathcal{J}_{m,2,1} \) as a recurrence formula from \( \mathcal{J}_{m,1,1} \) without proof. We show the weight distribution of \( \mathcal{J}_{m,2,1} \) explicitly and prove it.

**Theorem 6:** The weight distribution of \( \mathcal{J}_{m,2,1} \) is given by
\[ A_0^{m,2,1} = 1 \] (36)
\[ A_{m,2,1}^{m} = 2^m - 1 \] (37)
\[ A_m^{m,2} = (2^m - m - 1) \binom{m}{t} \] (38)
\[ A_{m,2}^{m,2} = (2^m - m - 1) t \binom{m}{t} \] (39)
\[ A_{m,2}^{m,2} = (2^m - m - 1) t \binom{m}{t} \] (40)
for \( t = 1, 2, \ldots, m \).

**Proof:** The generator polynomial of \( \mathcal{R}_{m,2,1} \) is
\[ G(x) = \prod_{i=1}^{N-2} (x - \alpha^i) \]
\[ = \frac{x^N - 1}{(x - \alpha^{-1})(x - 1)} \] (41)
and a sequence of check symbols appended by inner encoders is a codeword of RS code whose generator polynomial is
\[ G(ax) = \frac{x^N - 1}{\alpha^2(x - \alpha^{-2})(x - \alpha^{-1})}. \] (42)

When \( U_1(x) = x + \alpha^{-1} \) and \( U_2(x) = \alpha^{-1} x + \alpha^{-1} \), we have codewords
\[ C_1(x) = 1 + x + x^2 + \cdots + x^{N-1} \] (43)
\[ C_2(x) = 1 + ax + \alpha^2 x^2 + \cdots + \alpha^{N-1} x^{N-1} \] (44)
of \( \mathcal{R}_{m,1,2} \), respectively. Then \( C = \beta C_1 + \gamma C_2 \) for \( \beta, \gamma \in GF(2^m) \) is a codeword of \( \mathcal{R}_{m,1,2} \).
On the other hand, \( U_1(ax) = ax + \alpha^{-1} \) and \( U_2(ax) = x + \alpha^{-1} \) give two sequences of check symbols as follows:
\[ C_1(ax) = (ax + \alpha^{-1}) \frac{x^N - 1}{\alpha^2(x - \alpha^{-2})(x - \alpha^{-1})} \]
\[ = \frac{x^N - 1}{\alpha(x - \alpha^{-1})} \]
\[ = 1 + ax + \alpha^2 x^2 + \cdots + \alpha^{N-1} x^{N-1} \]
\[ = B_1(x) \] (45)
\[ C_2(ax) = (x + \alpha^{-1}) \frac{x^N - 1}{\alpha^2(x - \alpha^{-2})(x - \alpha^{-1})} \]
\[ = \frac{x^N - 1}{\alpha^2(x - \alpha^{-2})} \]
\[ = 1 + \alpha^2 x + \alpha^4 x^2 + \cdots + \alpha^{N-2} x^{N-1} \]
\[ = B_2(x). \] (46)

They are a basis of RS code with generator polynomial \( G(ax) \) and \( B = \beta B_1 + \gamma B_2 \) for \( C = \beta C_1 + \gamma C_2 \).
The binary weight of $B$ can be obtained from the binary weight distribution of $\mathcal{F}_{m,2,2}$. Imamura et al. [5] showed the binary weight distribution of $\mathcal{F}_{m,2,2}$. They showed $\mathcal{F}_{m,2,2}$ is divided into three cases by using the complementary basis $\lambda_1, \lambda_2, \ldots, \lambda_m$ corresponding to a basis of $G(F(2^m))$ such that (i) $\gamma = \beta = 0$, (ii) $\gamma = \beta^2 \lambda_i \neq 0$, (iii) $\gamma \neq \beta^2 \lambda_i$ for $i = 1, 2, \ldots, m$. In these cases, the binary weight are $0, (m - 1)2^{m-1}, m2^{m-1}$, and the numbers of corresponding codewords are $1, m(2^m - 1), (2^m + 1 - m)(2^m - 1)$, respectively.

We can obtain the weight distribution of $\mathcal{F}_{m,2,1}$ by taking account of the binary weight of $C$ as follows:
(i) The binary weight of $C$ is 0, so we obtain (36).
(ii) Since the mapping $\delta \rightarrow \delta^2, \delta \in G(F(2^m))$ is an isomorphism from $G(F(2^m))$ onto itself, the mapping nonzero $\beta \rightarrow \beta^2 \lambda_i$ for arbitrary $\lambda_i$ is also an isomorphism from nonzero elements of $G(F(2^m))$ onto itself. We have binary weight of $C$ as $m2^{m-1} - t$ by similar methods in the case (iv) of Proof of Theorem 5, where $t$ is the binary weight of $\beta$. Since the number of $\beta$ with weight $t$ is $\binom{m}{t}$ and it is true for each $\lambda_i$, $i = 1, 2, \ldots, m$. Therefore the number of codewords with the weight $(m - 1)2^{m-1} + m2^{m-1} - t = m2^m - 2^m - t$ of $\mathcal{F}_{m,2,1}$ is $m \binom{m}{t}$, then we have
\begin{equation}
30.
\end{equation}
To obtain the weight distribution of the other codewords, we divide the case (iii) into three cases, that is, (iii-a) $\gamma \neq \beta^2 \lambda_i, \beta \neq 0, \gamma \neq 0$, (iii-b) $\beta = 0, \gamma \neq 0$, (iii-c) $\beta \neq 0, \gamma = 0$.

(iii-a) The binary weight of $B$ is $m2^{m-1}$ and the binary weight of $C$ is $m2^{m-1} - t$ in the same way as the case (ii). Since the number of nonzero $\gamma$ satisfying $\gamma \neq \beta \lambda_i$ for fixed nonzero $\beta$ is $(2^m - 1 - m)$, the number of codewords with the weight $(2^m - 1 - t)2^m + (2^m - 1 - m)t$ of $C$ is $(2^m - m - 1) \binom{m}{t}$. So we have (38).

(iii-b) Since the binary weight of $C$ is also $m2^{m-1}$, we have (37).

(iii-c) Since the binary weight of $C$ is $(2^m - 1)t$, we have (40).

As the result, Theorem 6 holds.

C. The asymptotic relative minimum distance
On the relative minimum distance of the above mentioned Justesen codes, we have the following theorems.

**Theorem 7:** Let $N$ and $b-1$ be relatively prime. And let $N$ and $2-b$ be also relatively prime. The relative minimum distance of $\mathcal{F}_{m,1,0}$ asymptotically approaches 0.5 as the rate approaches zero.

**Proof:** In this case, the weight of nonzero codewords of $\mathcal{F}_{m,1,b}$ is only $m2^m$. Then relative minimum distance is $\frac{m2^m}{2m(2^m - 1)}$ and approaches 0.5 as $m \rightarrow \infty$, that is, length $2mN \rightarrow \infty$. The rate of $\mathcal{F}_{m,1,b}$ is $\frac{1}{2m(2^m - 1)}$ and approaches zero as $m \rightarrow \infty$.

**Theorem 8:** The relative minimum distance of $\mathcal{F}_{m,1,1}, \mathcal{F}_{m,1,2}$, $\mathcal{F}_{m,2,1}$ and $\mathcal{F}_{m,2,2}$ asymptotically approaches 0.25 as the rate approaches zero.

**Proof:** The minimum weight of nonzero codewords is $m2^{m-1} + 2^m - 1$ for large $m$ in these cases. Then relative minimum distance is $\frac{m2^{m-1} + 2^m - 1}{2m(2^m - 1) - 1}$ and approaches 0.25 as $m \rightarrow \infty$. The rate is $\frac{2m(2^m - 1)}{2m(2^m - 1)}$ or $\frac{2}{2m(2^m - 1)}$ and approaches zero as $m \rightarrow \infty$.

The values of the relative minimum distance for these codes as the rate approaches zero are obviously greater than 0.11, the lower bound at the rate of zero, which is given by Justesen[1]. In particular, a sequence of the codes in Theorem 7 asymptotically meets the Varshamov-Gilbert bound.

IV. Conclusions
In this paper, we showed the weight distributions of some Justesen codes in some cases and examined their relative minimum distance.

Difference of generator polynomials of RS codes as the outer code gives the different weight distributions. The weight distributions of $\mathcal{F}_{m,1,b}$ for $b = 0, 1, 2, 3$ and $\mathcal{F}_{m,2,b}$ for $b = 1, 2$ were derived for any $m$.

Each component over $G(F(2^m))$ is expanded into the corresponding binary $m$-tuple by a basis. It is worthy of notice that the results of the weight distributions of Justesen codes in this paper do not depend on difference of bases.

Further studies will be required to obtain the weight distributions or the minimum distance of the other Justesen codes theoretically.

**References**


