The Non-Alignment Properties of BPSK Sequences with Good Peak Factor

Abstract

A technique is presented to identify maximal sets of BPSK sequences which have a continuous (DFT) spectral power profile with low worst-case peak. These maximal sets are identified by avoiding all sequences with very low or high weight. The technique is significantly refined by considering generalised 'weightings' using other bins of the $2N$-point DFT. Although the paper does not give a coding strategy it succeeds in identifying spectrally flat sequences with the set of sequences which have maximum distance from identifiable clusters of 'bad' sequences.

1 Introduction

This paper will show how the BPSK Peak Factor (PF) problem can be approximately recast as a problem of finding those binary codewords which are at least a prescribed set of distances away from a set of known binary vectors [3, 1, 2, 4, 5, 6]. By changing this set of prescribed distances we can roughly control the worst case Peak Factor possible. It would appear that the technique can find a large proportion of the binary sequences with PF $\leq$ a specified value, up to large blocklengths. In other words, there is only a small rate loss as the blocklength increases. We stress that, at the moment, we cannot propose a code **construction** that achieves these peak factor reductions. Instead we merely point out that a large subset of the binary sequences of a given blocklength with PF $\leq$ a prescribed value will have known minimum distances from a known set of binary vectors. The idea is based on a generalisation of 'constant-weight' sequences. Previous analysis has allowed us to generate the following tables, Tables 1 and 2, which demonstrate the continuous PF properties of constant-weight sequences:

1.1 Peak Factor Distribution of Constant Weight BPSK Codes

In the following, BPSK messages are grouped according to their (unipolar) weight. For each weight, the continuous PFs (i.e. using an infintely zero-padded DFT) of the messages with that weight are computed. The lowest and highest PF for each weight are then tabulated, as shown below. The rational nature of some (but not all) of these peak factors is indicated. Only half the weights are shown as the other half have a symmetrical distribution. It is important to notice that messages with the lowest and highest weights are always bad messages, i.e. they have a high PF. For instance, by eliminating 17 carrier BPSK messages with weight ≤ 4 and ≥ 13, we are guaranteed to eliminate only messages with PF ≥ 4.765. This is a very powerful 'first strike' for any PF limiting coding scheme. However, not all bad messages are eliminated in this way. Note that [3] used hamming weight as part of a metric to eliminate high PF messages.

For each blocklength, N, and each weight, d, the upper and lower PFs are tabulated. Table Key: N: Number of carriers, d: weight

N/d	0	1	2	3	4	5	6	7	8
4	4.000000	1.769800	4.000000						
	4.000000	1.769800	2.370370						
5	5.000000	2.112500	5.000000						
	5.000000	1.800000	2.101328						
6	6.000000	2.692787	3.317760	6.000000					
	6.000000	2.666667	2.053109	2.053109					
7	7.000000	3.571429	3.571429	7.000000					
	7.000000	3.571429	1.865810	1.375876					
8	8.000000	4.500000	3.420689	4.500000	8.000000				
	8.000000	4.500000	2.000000	1.660782	2.000000				
9	9.000000	5.444444	3.687000	5.444444	9.000000				
	9.000000	5.444444	2.777778	1.883913	2.026097				
10	10.000000	6.400000	4.118906	4.460938	6.400000	10.000000			
	10.000000	6.400000	3.600000	1.921601	1.921601	1.991558			
11	11.000000	7.363636	4.669122	4.769188	7.363636	11.000000			
	11.000000	7.363636	4.454545	2.272727	1.829646	1.314321			
12	12.000000	8.333333	5.366190	4.978308	5.730347	8.333333	12.000000		
	12.000000	8.333333	5.333333	3.000000	1.640376	1.740571	1.912673		
13	13.000000	9.307692	6.230769	5.310735	6.230769	9.307692	13.000000		
	13.000000	9.307692	6.230769	3.769231	1.923077	1.622804	1.824325		
14	14.000000	10.285714	7.142857	5.700241	6.130279	7.142857	10.285714	14.000000	
	14.000000	10.285714	7.142857	4.571429	2.571429	1.659606	1.659606	2.090919	
15	15.000000	11.266667	8.066667	6.156690	6.321273	8.066667	11.266667	15.000000	
	15.000000	11.266667	8.066667	5.400000	3.266667	1.666667	1.666667	1.851567	
16	16.000000	12.250000	9.000000	6.690924	6.615930	7.197406	9.000000	12.250000	16.000000
	16.000000	12.250000	9.000000	6.250000	4.000000	2.250000	1.712303	1.771833	1.855578
17	17.000000	13.235294	9.941176	7.315393	6.945385	7.497003	9.941176	13.235294	17.000000
	17.000000	13.235294	9.941176	7.117647	4.764706	2.882353	1.759558	1.771598	1.799987

Table 1: PF Properties of Constant Weight BPSK Codes

N/d	0	1	2	3	4	5	6	7	8
4	4	$\frac{9+4\sqrt{3}}{9}$	4						
	4	$\frac{9+4\sqrt{3}}{9}$	$(\frac{4}{3})^3$						
5	5	$13^2/80$	5						
	5	$3^2/5$	2.101328*						
6	6	2.692787	$\frac{2^7 3^4}{5^5}$	6					
	6	$4^2/6$	2.053109*	2.053109*					
7	7	$5^2/7$	$5^2/7$	7					
	7	$5^2/7$	1.865810	1.375876					
8	8	$6^2/8$	3.420689	$6^2/8$	8				
	8	$6^2/8$	$4^2/8$	1.660782	$4^2/8$				
9	9	$7^2/9$	3.687000	$7^2/9$	9				
	9	$7^2/9$	$5^2/9$	1.883913	2.026097				
10	10	$8^2/10$	4.118906	4.460938	$8^2/10$	10			
	10	$8^2/10$	$6^2/10$	1.921601	1.921601	1.991558			
11	11	$9^2/11$	4.669122	4.769188	$9^2/11$	11			
	11	$9^2/11$	$7^2/11$	$5^2/11$	1.829646	1.314321			
12	12	$10^2/12$	5.366190	4.978308	5.730347	$10^2/12$	12		
	12	$10^2/12$	$8^2/12$	$6^2/12$	1.640376	1.740571	1.912673		
13	13	$11^2/13$	$9^2/13$	5.310735	$9^2/13$	$11^2/13$	13		
	13	$11^2/13$	$9^2/13$	$7^2/13$	$5^2/13$	1.622804	1.824325		
14	14	$12^2/14$	$10^2/14$	5.700241	6.130279	$10^2/14$	$12^2/14$	14	
	14	$12^2/14$	$10^2/14$	$8^2/14$	$6^2/14$	1.659606	1.659606	2.090919	
15	15	$13^2/15$	$11^2/15$	6.156690	6.321273	$11^2/15$	$13^2/15$	15	
	15	$13^2/15$	$11^2/15$	$9^2/15$	$7^2/15$	$5^2/15$	$5^2/15$	1.851567	
16	16	$14^2/16$	$12^2/16$	6.690924	6.615930	7.197406	$12^2/16$	$14^2/16$	16
	16	$14^2/16$	$12^2/16$	$10^2/16$	$8^2/16$	$6^2/16$	1.712303	1.771833	1.855578
17	17	$15^2/17$	$13^2/17$	7.315393	6.945385	7.497003	$13^2/17$	$15^2/17$	17
	17	$15^2/17$	$13^2/17$	$11^2/17$	$9^2/17$	$7^2/17$	1.759558	1.771598	1.799987

Table 2: PF Properties of Constant Weight BPSK Codes (with exact values)

*: The value approximated by 2.101328 can be represented exactly as,

$$\frac{2^6 3^2 5\alpha^4 + 2^5 3^3 \alpha^3 + 2^4 3.31\alpha^2 + 2^2 3^2 5\alpha + 5^3}{2^5 3^2 5\alpha^2}$$

where $\alpha = (\frac{3^2 + j\sqrt{3.223}}{2^5 3^2})$.

*: The value approximated by 2.053109 can be represented exactly as,

$$\frac{2^4 3.349 + 2^4 23\sqrt{2.23}}{5^5 3}$$

It is of interest to note that the tabulated collision points for $N = 6$ to 10 are always best-case PFs for a low or high weight codeword [1]. This is because such points refer to cases where the individual 'weight surplus' rotated carriers act in complete opposition to each other.

Note that for very low or high weights, the PF can be wholly defined:

N	4	5	6	7	8	9	10	11	12	13	14	15	16	17
d	0	0	0	1	1	1	1	1	1	2	2	2	2	2

Table 3: Threshold for Predictable Maximum Peak at Bin Zero

Note, also, the maximum weight where the maximum peak can possibly occur on bin zero:

N	4	5	6	7	8	9	10	11	12	13	14	15	16	17
d	0	1	1	1	2	2	2	3	3	4	4	5	5	5

Table 4: Threshold for Possible Maximum Peak at Bin Zero

1.2 Extending the Constant Weight Idea to Other Bins Using the Concept of 'Alignment'

The above observations for constant weight sequences can be interpreted thus: The weight 0 binary sequence has a maximum DFT bin on bin zero, being the sum of unity vectors. This sum is a maximum because the vectors are wholly **aligned**. Sequences which are distance 1 from this weight 0 sequence (i.e. of weight 1) have a bin zero which is computed from the sum of unity vectors which are all **aligned** apart from one vector. These binary vectors exhibit a high vector alignment on bin 0 and will therefore have high PF and can consequently be eliminated. As sequence blocklength, N, increases, the same argument can be used for weight 2,3,..etc sequences. The generalisation of this idea to bins other than 0 is the topic of this paper, and for the purpose of clarity the PF will be evaluated only over a $2N$-point DFT in this paper (i.e. we are approximately investigating the Merit Factor). The extension of the idea to the PF over larger blocklength DFTs is then straightforward.

Consider, say, bin 3 of a 2N-point DFT of a length N BPSK sequence: This is the sum of N unity vectors, where the first vector is the first vector element of the sequence rotated by 0 radians, the second vector is the second vector element of the sequence rotated by $\frac{3}{2N}2\pi$ radians, the third vector is the third vector element of the sequence rotated by $\frac{6}{2N}2\pi$ radians,...and so on. It is apparent that bin 3 of the $2N$-point DFT of the bipolar form of the binary sequence

$(\lfloor \frac{0}{N} \rfloor, \lfloor \frac{3}{N} \rfloor, \lfloor \frac{6}{N} \rfloor, \ldots, \lfloor \frac{(N-1)3}{N} \rfloor) \bmod 2$ will be the sum of N unity vectors which are all in the upper half of the complex plane, i.e. no vector points 'downwards'. In other words, bin 3 of the $2N$-point DFT of such a sequence will be the sum of N **'loosely-aligned'** unity vectors. Consequently it's PF will be quite high, especially for large N. Moreover, all binary sequences which are negacyclic shifts of the original sequence will only cause a phase rotation at bin 3 in the DFT domain and therefore leave the bin magnitude unaffected (if we had instead chosen to look at an even bin number then, instead of negacyclic shifts we would perform all combinations of negation and cyclic shifts of the sequence). We can identify these loosely-aligned vectors and their negacyclic or cyclic+negation family for each of the $2N$ bins of the $2N$-point DFT and associate with each bin family a worst-case PF and best-case PF (taken over a $2N$-point DFT in this paper, unlike the continuous PF for the results of the previous subsection). The bin family for bin 0 is simply $000\ldots00$ and $111\ldots11$, and for bin N of the $2N$-point DFT is $010101\ldots$ and $101010\ldots$. For other bins the bin family is bigger. As with bin 0 we can, for each bin, investigate all those sequences which are exactly weight 1 away from one or more sequences in the bin family, and a higher weight away from the rest of the sequences in the bin family. The DFT of these 'weight 1' sequences in these bin families will all be loosely-aligned on their respective bins apart from one vector, and will therefore also have pretty high PF, especially as N becomes large. We can further investigate the PF properties of all those sequences which are exactly weight 2 away from one or more sequences in the bin family, and a higher weight away from the rest of the sequences in the bin family.... and so on for weight 3...etc.

Thus we can examine the PF range for each bin family, and for binary sequences of a known distance from this bin family. Therefore, using a table of these PF ranges, sequences with good PF (and Merit Factor) can be identified which will be at least a known set of distances from each of the bin sequence families. Those with lower PF will have larger distance from each of these bin families.

The next section will present the above ideas in a more formal way.

2 The Peak Factor Problem Approximates to a Binary Distance Problem

We begin by presenting a few definitions followed by the main conjecture.

Definition 1 *Let $WNPF$ be the Peak Factor of a length N sequence, taken over a W-blocklength Discrete Fourier Transform (DFT).*

Consequently the NPF metric informs us of the periodic properties of the sequence, the odd bins of the $2N$PF metric will tell us about the negaperiodic properties of the sequence, and the $2N$PF metric will tell us about the aperiodic (and, approximately, the Merit Factor) properties of the sequence. The $4N$PF metric will tell us to a good accuracy about the continuous PF properties of the sequence, and the ∞PF metric will tell us exactly about the continuous PF properties of the sequence. In this paper, for simplicity of exposition, we will concentrate on the $2N$PF metric. Extending the idea to the $4N$PF, $8N$PF metric,...etc, will then be relatively straightforward.

Definition 2 *Let u_k be the length N binary sequence,*

$$u_k = (\lfloor \frac{0}{N} \rfloor, \lfloor \frac{k}{N} \rfloor, \lfloor \frac{2k}{N} \rfloor, \ldots, \lfloor \frac{(N-1)k}{N} \rfloor) \ mod \ 2$$

For instance, if $N = 6$, then $u_4 = 001001$, and $u_5 = 001010$.

Definition 3 *Let $\mathbf{U_k}$ be the set of length N binary sequences constructed as follows:*

If k is even,

$$\mathbf{U_k} = \{cyclic\ shifts(u_k) \bigcup negations(cyclic\ shifts(u_k))\}$$

If k is odd,

$$\mathbf{U_k} = \{negacyclic\ shifts(u_k)\}$$

For instance, if $N = 6$, then $\mathbf{U_4} = \{001001, 010010, 100100, 110110, 101101, 011011\}$, and

$$\mathbf{U_5} = \{001010, 010101, 101011, 010110, 101101, 011010, 110101, 101010, 010100, 101001, 010010, 100101\}$$

Definition 4 *Let $\mathbf{B_d}$ be the set of length N, weight d binary sequences. Then $\mathbf{D_{k,d}}$ is the set of length N binary sequences constructed as follows:*

$$\mathbf{D_{k,d}} = \bigcup_{v \in \mathbf{B_d}} \mathbf{U_k} \oplus v$$

For instance, if $N = 6$, then

$$\mathbf{D_{4,1}} =$$
$$\{000001, 000010, 000100, 001000, 001011, 001101, 010000, 010011, 010110, 011001, 011010, 011111,$$
$$100000, 100101, 100110, 101001, 101100, 101111, 110010, 110100, 110111, 111011, 111101, 111110\}$$

Note that, as evident in this example, a sequence may be constructed more than once using Definition 4, but it is only represented once in $\mathbf{D_{k,d}}$.

Definition 5

$$\mathbf{E_{k,d}} = \bigcup_{e=0}^{d} \mathbf{D_{k,d}}$$

Note that a sequence may be included more than once using Definition 5, but it is only represented once in $\mathbf{E_{k,d}}$.

Definition 6

$$\mathbf{F_{k,d}} = \mathbf{E_{k,d}} - \mathbf{E_{k,d-1}}$$

Definition 7 *A length N binary sequence, s, is said to have a 'distance' d from $\mathbf{U_k}$ if it is not nearer than distance d to any sequence in $\mathbf{U_k}$, but is a distance d from one or more members of $\mathbf{U_k}$. Consequently $s \in \mathbf{F_{k,d}}$.*

The set $\mathbf{F_{k,d}}$ is the set of all binary sequences which are a distance d away from the bin family $\mathbf{U_k}$.

For small values of d the lowest $2N$PF of $\mathbf{F_{k,d}}$ becomes larger as d falls.

We can therefore state the following lemma:

Lemma 1 *There exists a value d_k such that all members of $\mathbf{E_{k,d_k}}$ have high $2NPF$.*

Definition 8

$$\mathbf{G} = \bigcup_{k=0}^{N-1} \mathbf{E_{k,d_k}}$$

$\mathbf{G}$ is the set of binary sequences which are distance d_k or less away from the bin family $\mathbf{U_k}$, for all k, $0 \le k < 2N$. $\mathbf{G}$ is defined for N distance values, d_k.

Note that $\mathbf{G}$ need only be defined for N values of k, not $2N$ values of k, due to symmetry between the first and second halves of the spectrum.

Definition 9 *Define the vector $R = (d_0, d_1, d_2, \ldots, d_{N-1})$ to be the vector of N distance values over which* $\mathbf{G}$ *is defined.*

Lemma 2 *There exists a vector R such that all members of* $\mathbf{G}$ *have high $2NPF$.*

Conversely, if $\mathbf{A}$ is the set of all 2^N binary vectors, then,

Definition 10

$$\mathbf{H} = \mathbf{A} - \mathbf{G}$$

Lemma 3 *There exists a vector R such that all members of* $\mathbf{H}$ *have low or quite low $2NPF$.*

R is the 'distance profile' of $\mathbf{H}$.

Lemma 3 is still a bit vague, but it seems that by using all $2N$ bins (or a well chosen subset) to constrain the distance of desirable codewords we can home in on a large number of sequences with good $2N$PF. In other words, if one can find a code construction satisfying the distance profile R, then one has found a very good $2N$PF code. This paper merely shows that such a code will be good. We do not currently have such a code construction.

The next section presents an example for length 16 binary sequences.

3 Example: Length 16 Binary Sequences

Here is the computed table of worst-case and best-case $2N$PFs for each set $\mathbf{F_{k,d}}$ for $0 \leq k < 2N$ and small values of d for the case $N = 16$.

The figure in square brackets is the number of sequences in $\mathbf{F_{k,d}}$ for the specified values of k and d. The figures below this in each box are the $2N$PF range from lowest $2N$PF - highest $2N$PF. The '$*$' indicates the smallest value of d for a given k such that $\mathbf{F_{k,d}} \subset \mathbf{D_{k,d}}$, i.e. where some sequences in $\mathbf{D_{k,d}}$ are also in $\mathbf{D_{k,d'}}$, $d' < d$. The results for different bins, k, are grouped according to identical results. For instance, the results for bins $1, 15, 17, 31$ are grouped together.

$k \quad \backslash \quad d$	0	1	2	3	4	5	6	7	8
$\{0, 16\}$	[2] 16.0	[32] 12.25	[240] 9.00	[1120] 6.25 − 6.51	[3640] 4.0 − 6.51	[8736] 2.25 − 6.51	[16016] 1.71 − 9.00	[22880] 1.60 − 12.25	[12870] 1.74 − 16.0
$\{1, 15, 17, 31\}$	[32] 6.51 − 16.0	[416]* 4.22 − 12.25	[2432] 2.54 − 9.00	[8320] 1.71 − 6.57	[17920] 1.60 − 6.83	[23296] 1.60 − 9.00	[12608] 1.71 − 12.25	[512] 2.33 − 16.00	– –
$\{2, 14, 18, 30\}$	[16] 6.57	[224] 4.30 − 6.51	[1424]* 2.78 − 6.51	[5376] 1.75 − 6.51	[13120] 1.71 − 6.57	[20608] 1.60 − 6.51	[17952] 1.71 − 9.00	[6560] 2.25 − 12.25	[256] 3.0 − 16.0
$\{3, 13, 19, 29\}$	[32] 6.51	[416]* 4.22 − 6.32	[2432] 2.54 − 6.83	[8320] 1.71 − 9.00	[17920] 1.60 − 12.25	[23296] 1.71 − 16.0	[12608] 1.74 − 16.0	[512] 2.33 − 9.00	– –
$\{5, 11, 21, 27\}$	[32] 6.51	[416]* 4.22 − 6.32	[2432] 2.54 − 6.83	[8320] 1.71 − 9.00	[17920] 1.60 − 12.25	[23296] 1.60 − 16.0	[12608] 1.81 − 16.0	[512] 2.33 − 12.25	– –
$\{4, 12, 20, 28\}$	[8] 6.83	[128] 4.66 − 6.08	[912] 2.85 − 6.51	[3840]* 2.25 − 6.51	[10472] 1.71 − 6.83	[18816] 1.60 − 6.51	[20080] 1.71 − 9.00	[9984] 1.75 − 12.25	[1296] 2.21 − 16.0
$\{6, 10, 22, 26\}$	[16] 6.57	[224] 4.30 − 6.32	[1424]* 2.78 − 6.51	[5376] 1.75 − 6.51	[13120] 1.71 − 6.57	[20608] 1.60 − 6.51	[17952] 1.71 − 9.00	[6560] 2.25 − 12.25	[256] 3.00 − 16.00
$\{7, 9, 23, 25\}$	[32] 6.51	[416]* 4.22 − 6.02	[2432] 2.54 − 6.32	[8320] 1.81 − 8.00	[17920] 1.60 − 9.00	[23296] 1.71 − 12.25	[12608] 1.60 − 16.0	[512] 2.33 − 6.51	– –
$\{8, 24\}$	[4] 8.00	[64] 6.25	[480] 4.5 − 6.02	[2240] 3.25 − 6.51	[7000] 2.00 − 6.51	[14784]* 1.60 − 6.51	[20384] 1.71 − 9.00	[15680] 1.75 − 12.25	[4900] 2.00 − 16.00

Table 5: $2N$PF Properties of $\mathbf{F_{k,d}}$ Sets for $N = 16$

We can use this table to define a codeset with specified distances from the bin families, $\mathbf{U_k}$, such that the codeset comprises sequences with low $2N$PF. Table 6 shows the $2N$PF results for sequence sets $\mathbf{H}$ for various distance profiles, out of a total message space of $2^{16} = 65536$.

The distance profile, R, is abbreviated in this table to the sequence $a, b, c, d, e, f, g, h, i$ where $d_0, d_{16} = a$, $d_1, d_{15}, d_{17}, d_{31} = b$, $d_2, d_{14}, d_{18}, d_{30} = c$, $d_3, d_{13}, d_{19}, d_{29} = d$, $d_4, d_{12}, d_{20}, d_{28} = e$, $d_5, d_{11}, d_{21}, d_{27}, = f$, $d_6, d_{10}, d_{22}, d_{26} = g$, $d_7, d_9, d_{23}, d_{25} = h$, $d_8, d_{24} = i$. The tail distance profile is similarly abbreviate for the same d_i subsets on the line underneath for each example. The last example in the table represents $d_0 - d_{16}$ directly. The notation $|\mathbf{H}(\mathbf{x})|$ means 'the number of sequences in $\mathbf{H}$ which have a $2N$PF $\leq x$'. $|\mathbf{A}(\mathbf{x})|$ is similarly defined.

Distance Profile	$2N$PF Range of $\mathbf{H}$	$\|\mathbf{H}(\mathbf{x})\|$	$\|\mathbf{A}(\mathbf{x})\|$
$3, 0, 0, 0, 0, 0, 0, 0, 1$ $-, -, -, -, -, -, -, -, -$	$1.60 - 6.32$	$\|\mathbf{H}\| = 62376$	$\|\mathbf{A}(\mathbf{6.32})\| = 64672$
$5, 2, 2, 2, 2, 2, 2, 2, 4$ $-, -, -, -, -, -, -, -, -$	$1.60 - 4.54$	$\|\mathbf{H}\| = 16012$ $\|\mathbf{H}(\mathbf{2.85})\| = 9072$ $\|\mathbf{H}(\mathbf{1.99})\| = 1072$	$\|\mathbf{A}(\mathbf{4.54})\| = 57896$ $\|\mathbf{A}(\mathbf{2.85})\| = 19688$ $\|\mathbf{A}(\mathbf{1.99})\| = 1072$
$5, 2, 2, 2, 3, 2, 2, 2, 4$ $-, -, -, -, -, -, -, -, -$	$1.60 - 4.54$	$\|\mathbf{H}\| = 13700$ $\|\mathbf{H}(\mathbf{2.77})\| = 7612$ $\|\mathbf{H}(\mathbf{1.99})\| = 1072$	$\|\mathbf{A}(\mathbf{4.54})\| = 57896$ $\|\mathbf{A}(\mathbf{2.77})\| = 17556$ $\|\mathbf{A}(\mathbf{1.99})\| = 1072$
$5, 2, 2, 2, 3, 2, 2, 2, 4$ $-, 7, 7, 7, 8, 7, 7, 7, 8$	$1.60 - 3.89$	$\|\mathbf{H}\| = 10944$ $\|\mathbf{H}(\mathbf{2.77})\| = 6284$ $\|\mathbf{H}(\mathbf{1.99})\| = 1072$	$\|\mathbf{A}(\mathbf{3.89})\| = 42788$ $\|\mathbf{A}(\mathbf{2.77})\| = 17556$ $\|\mathbf{A}(\mathbf{1.99})\| = 1072$
$5, 2, 3, 2, 3, 2, 2, 2, 4$ $-, -, -, -, -, -, -, -, -$	$1.60 - 3.92$	$\|\mathbf{H}\| = 8228$ $\|\mathbf{H}(\mathbf{2.53})\| = 3240$ $\|\mathbf{H}(\mathbf{1.99})\| = 840$	$\|\mathbf{A}(\mathbf{3.92})\| = 44108$ $\|\mathbf{A}(\mathbf{2.53})\| = 11580$ $\|\mathbf{A}(\mathbf{1.99})\| = 1072$
$5, 2, 3, 2, 3, 2, 2, 2, 4$ $-, 7, 7, 7, 8, 7, 7, 7, 8$	$1.60 - 3.89$	$\|\mathbf{H}\| = 7888$ $\|\mathbf{H}(\mathbf{2.53})\| = 2912$ $\|\mathbf{H}(\mathbf{1.99})\| = 896$	$\|\mathbf{A}(\mathbf{3.92})\| = 44108$ $\|\mathbf{A}(\mathbf{2.53})\| = 11580$ $\|\mathbf{A}(\mathbf{1.99})\| = 1072$
$6, 3, 3, 3, 4, 3, 3, 3, 4$ $-, -, -, -, -, -, -, -, -$	$1.60 - 2.24$	$\|\mathbf{H}\| = 16$ $\|\mathbf{H}(\mathbf{1.60})\| = 8$	$\|\mathbf{A}(\mathbf{2.24})\| = 2248$ $\|\mathbf{A}(\mathbf{1.60})\| = 8$
$6, 3, 3, 3, 4, 3, 3, 3, 4$ $8, 6, 6, 5, 6, 6, 6, 7, 6$	1.60	$\|\mathbf{H}\| = 8$ $\|\mathbf{H}(\mathbf{1.60})\| = 8$	$\|\mathbf{A}(\mathbf{1.60})\| = 8$ $\|\mathbf{A}(\mathbf{1.60})\| = 8$
Bins 0 - 16: $6, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 5$ $8, -, -, -, -, -, -, -, -, -, -, -, -, -, -, -, -$	1.60	$\|\mathbf{H}\| = 8$ $\|\mathbf{H}(\mathbf{1.60})\| = 8$	$\|\mathbf{A}(\mathbf{1.60})\| = 8$ $\|\mathbf{A}(\mathbf{1.60})\| = 8$

Table 6: $2N$PF Properties of $\mathbf{G}$ Sets for $N = 16$

Table 6 indicates how the distance profile can control the $2N$PF range of $\mathbf{H}$. In the case of $3, 0, 0, 0, 0, 0, 0, 1$ in the table, the distance criteria successfully eliminates most messages with $2N\text{PF} > 6.32$. In the case of $5, 2, 2, 2, 3, 2, 2, 4$ in the table, although there a number of sequences with moderately high $2N$PF, **all** messages with $2N\text{PF} \leq 1.99$ are included. Also, the case of $6, 3, 3, 3, 4, 3, 3, 4$ captures 16 messages, 8 of which have the best possible $2N$PF (and very good Merit Factor) of 1.60. The other 8 have a $2N$PF of 2.24. By adding additional 'tail' constraints on each bin we can further improve results. In particular, for $6, 3, 3, 3, 4, 3, 3, 4$ with the tail constraints $8, 6, 6, 5, 6, 6, 6, 7, 6$, we eliminate all but the set of 8 messages with optimal $2N$PF. The final entry in the table shows that these constraints can be limited to the first 16 bins, and eased somewhat, and yet we are still able to capture only the 8 optimum $2N$PF messages.

The above results suggest that the distance profile constraints are quite successful in limiting

$2N$PF. They also suggest that the bin families, $\mathbf{U_k}$, could perhaps be augmented in some way so as to eliminate further unwanted messages.

For comparison purposes we should note that the Golay-Davis-Jedwab (GDJ) 'complementary sequence' approach will yield 384 length 16 binary sequences with ∞PF ≤ 2.0. It is hoped that the method outlined in this paper (or a slight improvement of it) will yield much higher rates, particularly as N increases. However, unlike the GDJ approach, we do not yet have a code construction. This paper has only succeeded in proposing **binary** code constraints.

We can gain further insight by looking at the PF taken only over each bin in question (7), as opposed the $2N$PF taken over all $2N$ bins for each bin in question (5). We thus see that messages which are a large distance away from the bin families (i.e. around 7 or 8) will have very low PF value at bin k. This in turn forces a moderate rise in the best case $2N$PF at these distances, as was previously shown in Table 5.

The figure in square brackets is the number of sequences in $\mathbf{F_{k,d}}$ for the specified values of k and d. The figures below this in each box is the PF range fromlowest PF - highest PF at bin k only. The results for different bins, k, are grouped according to identical results. For instance, the results for bins $2, 6, 10, 14, 18, 22, 26, 30$ are grouped together.

$k \quad \backslash \quad d$	0	1	2	3	4	5	6	7	8
$\{0, 16\}$	[2] 16.0	[32] 12.25	[240] 9.00	[1120] 6.25	[3640] 4.0	[8736] 2.25	[16016] 1.00	[22880] 0.25	[12870] 0
$\{1, 3, 5, 7,$ $9, 11, 13, 15,$ $17, 19, 21, 23,$ $25, 27, 29, 31\}$	[32] 6.51	[416]* $4.22 - 6.02$	[2432] $2.42 - 5.11$	[8320] $1.18 - 3.92$	[17920] $0.36 - 2.64$	[23296] $0.05 - 1.46$	[12608] $0.0008 - 0.55$	[512] $0.0004 - 0.06$	– –
$\{2, 6, 10, 14,$ $18, 22, 26, 30\}$	[16] 6.57	[224] $4.30 - 6.32$	[1424]* $2.50 - 5.61$	[5376] $1.20 - 4.54$	[13120] $0.36 - 3.28$	[20608] $0.08 - 2.03$	[17952] $0.002 - 0.96$	[6560] $0.002 - 0.25$	[256] 0
$\{4, 12, 20, 28\}$	[8] 6.83	[128] $4.66 - 6.08$	[912] $2.85 - 5.83$	[3840]* $1.54 - 4.37$	[10472] $0.59 - 3.41$	[18816] $0.13 - 1.96$	[20080] $0.03 - 1.0$	[9984] $0.04 - 0.25$	[1296] 0
$\{8, 24\}$	[4] 8.00	[64] 6.25	[480] $4.5 - 5.0$	[2240] $3.25 - 4.25$	[7000] $2.00 - 4.00$	[14784]* $1.25 - 2.25$	[20384] $0.5 - 1.0$	[15680] 0.25	[4900] 0

Table 7: Bin k PF Properties of $\mathbf{F_{k,d}}$ Sets for $N = 16$

4 Example: Length 17 Binary Sequences

A brief 'spot' check (guess) of the length 17 case came up with the following result:

This profile obtains the lowest possible $2N$PF class of 1.57.

Distance Profile	$2N$PF Range of $\mathbf{H}$	$\lvert\mathbf{H(x)}\rvert$	$\lvert\mathbf{A(x)}\rvert$
$R = (5,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,$ $5,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3)$	$1.57 - 2.87$	$\lvert\mathbf{H}\rvert = 4052$	$\lvert\mathbf{A(2.87)}\rvert = 30128$

Table 8: $2N$PF Properties of $\mathbf{G}$ Sets for $N = 17$

5 Computing the $2N$PF Using the L-point DFT

One can also, more generally, include the bin families for the L bins of the L-point DFT and constrain the required sequences for the $2N$PF accordingly. These may include the cyclic and negacyclic bin families discussed so far, (if $2N|L$), plus further bin families. However, whereas the bin families for the cyclic and negacyclic cases are obviously generated, the situation is a bit more complicated for the new bin families. For example, consider bin 3 of a 32-point DFT for a length 8 sequence. Then the sequence,

$$u_3 = (\lceil\frac{0}{16}\rceil, \lceil\frac{3}{16}\rceil, \lceil\frac{6}{16}\rceil, \lceil\frac{9}{16}\rceil, \lceil\frac{12}{16}\rceil, \lceil\frac{15}{16}\rceil, \lceil\frac{18}{16}\rceil, \lceil\frac{21}{16}\rceil) \bmod 2$$

will have all vectors in the top half of the complex plane (i.e. loosely aligned) at bin 3. More generally, all vectors of the form,

$$u_{3,s} = (\lceil\frac{0+s}{16}\rceil, \lceil\frac{3+s}{16}\rceil, \lceil\frac{6+s}{16}\rceil, \lceil\frac{9+s}{16}\rceil, \lceil\frac{12+s}{16}\rceil, \lceil\frac{15+s}{16}\rceil, \lceil\frac{18+s}{16}\rceil, \lceil\frac{21+s}{16}\rceil) \bmod 2$$

will be loosely aligned, with all vectors at bin 3 in some half of the complex plane. This will give the following set of sequences:

00000011, 10000011, 10000001, 11000001, 11000000, 11100000, 11110000, 11111000,
11111100, 01111100, 01111110, 00111110, 00111111, 00011111, 00001111, 00000111

Note that this family is neither cyclic nor negacyclic shift.

In this way, we arrive at the following definitions, (which are completely general and can be used to construct the cyclic, negacyclic, and all bin cases).

Definition 11 *Let $u_{k,s}$ be the length N binary sequence,*

$$u_{k,s} = (\lfloor\frac{2(0+s)}{L}\rfloor, \lfloor\frac{2(k+s)}{L}\rfloor, \lfloor\frac{2(2k+s)}{L}\rfloor, \ldots, \lfloor\frac{2((N-1)k+s)}{L}\rfloor) \ mod\ 2$$

Definition 12 *Let $\mathbf{U_k}$ be the set of length N binary sequences constructed as follows:*

$$\mathbf{U_k} = \bigcup_{s=0}^{L-1} u_{k,s}$$

The distance definitions defined in previous sections hold in exactly the same way, however we have now generalised the problem to identify further bin families from which flat sequences should be distant. We therefore have a more detailed distance profile. A desirable aim will be to identify a smallest set of distance families from which optimal sequences should have a known distance.

6 Distance Profiles for Optimal Sequence Sets

The first table is for $2N$PF properties, the second table is for NPF properties.

N	DFT size, w	Distance Profile	$2N$PF Range of $\mathbf{H}$	$\lvert\mathbf{H}(\mathbf{x})\rvert$	$\lvert\mathbf{A}(\mathbf{x})\rvert$	comments
6	12	1001001	2.0	4	12	ideal ∞PF class
6	12	1000001	2.0	12	12	ideal $2N$PF class
7	14	20000002	1.14	4	4	ideal $2N$PF and ∞PF class
8	32	2.00000.2.00000.2 Tail .2...1.....1...2.	1.65	16	16	ideal $2N$PF class
8	32	2.00000.2.00000.2 Tail .2....2...2....2.	1.65	8	16	ideal ∞PF class
9	18	2001001002 Tail 4..4..4..4	1.72	24	24	ideal $2N$PF class
10	40	...1..11.111.11..1... Tail ..3.4...4...4...4.3..	1.85	8	72	ideal ∞PF class
11	22	3.11111111.4 Tail .3333333333.	1.09	4	4	ideal $2N$PF and ∞PF class

Table 9: $2N$PF Properties of $\mathbf{G}$ Sets for Various N

N	DFT size, w	Distance Profile	NPF Range of $\mathbf{H}$	$\lvert\mathbf{H}(\mathbf{x})\rvert$	$\lvert\mathbf{A}(\mathbf{x})\rvert$	comments
6	6	1001	2.0($PACF$ = 2.00)	24	24	ideal NPF class
7	7	2000	1.14($PACF$ = 1.00)	28	28	ideal NPF class
8	8	20202 Tail .3.3.	1.5($PACF$ = 4.00)	32	32	ideal NPF class
9	9	200100 Tail 4..4..	1.72($PACF$ = 3.00)	108	108	ideal NPF class
10	10	211112 Tail .4444.	1.85($PACF$ = 2.00)	320	320	ideal NPF class
11	11	...... Tail .3.433	1.09($PACF$ = 1.00)	44	44	ideal NPF class

Table 10: NPF Properties of $\mathbf{G}$ Sets for Various N

7 Conclusion and Further Work

This paper has identified the problem of finding sequences which are spectrally flat with a discrete distance problem, where all acceptable sequences have sufficient distance from a set of clusters of high spectral peak sequences. Further work should consider the following:

- Recast the periodic autocorrelation problem using only the $\mathbf{U_k}$ families when k is even. This is a subproblem of the problem under discussion in this paper.
- Recast the negaperiodic autocorrelation problem using only the $\mathbf{U_k}$ families when k is odd. This is a subproblem of the problem under discussion in this paper.
- Derive equations for the offline computation of tables like Table 5, at least for small distances, d. I think this is quite easy to do.
- Use the Merit Factor metric instead of the $2N$PF metric. The ranges will be slightly different.
- Augment the bin families $\mathbf{U_k}$ somehow, or introduce extra bin families, so as to eradicate the remaining sequences with moderately high $2N$PF.
- Develop code constructions to generate codesets $\mathbf{H}$ which satisfy the distance profile.
- Develop code constructions to generate codesets $\mathbf{H}$ which satisfy the distance profile **and** such that the minimum distance of $\mathbf{H}$ is high, (i.e. $\mathbf{H}$ possesses intrinsic error correction properties).
- Extend the ideas to WNPF where $W = 4, 8$,...etc.

References

[1] M.G.Parker,S.J.Shepherd,S.K.Barton, “Final Report of EPSRC Grant GR/K48914: Multi-Function Coding for Minimisation of Peak Envelope Power and Error Control in Multitone Modulation Systems”, *University of Bradford Internal Report No 601, '97*

[2] S.J.Shepherd,J.Orriss,S.K.Barton, ”Asymptotic Limits in Peak Envelope Power Reduction by Redundant Coding in QPSK Multi-Carrier Modulation”, *IEEE Trans on Communications*, Vol 46, No 1, pp 5-10, Jan '98

[3] P.W.J.Van Eetvelt,G.Wade,M.Tomlinson, "Peak to Average Power Reduction for OFDM Schemes by Selective Scrambling", *Electronics Letters*, Vol 32, No 21, pp 1963-1964, 10th Oct, '96

[4] J.A.Davis,J.Jedwab, "Peak-to-mean Power Control in OFDM, Golay Complementary Sequences and Reed-Muller Codes," *HP Laboratories Technical Report*, HP Laboratories Bristol, HPL-97-158, Dec '97

[5] K.G.Paterson, "Generalised Reed-Muller Codes and Power Control in OFDM Modulation," *HP Technical Report, HPL-98-57* March '98

[6] K.G.Paterson,V.Tarokh, "On the Existence and Construction of Good Codes with Low Peak-to-Average Power Ratios," *HP Technical Report*, HPL-1999-51, '99