Spring 2007

Crossover property of the nonperiodic autocorrelation of quaternary sequences

Michael Pohl
University of Richmond

Crossover Property of the Nonperiodic Autocorrelation of Quaternary Sequences

Michael Pohl
Honors thesis[1]
Department of Mathematics & Computer Science
University of Richmond

[1]Under the direction of Dr. James A. Davis

The signatures below, by the thesis advisor, a departmental reader, and the honors coordinator for mathematics, certify that this thesis, prepared by Michael Pohl, has been approved, as to style and content.

5/23/07

(thesis advisor)

5/23/07

(reader)

5/23/07

(honors coordinator)

Abstract

Sequences with identical nonperiodic autocorrelation functions have recently been used to construct Golay sequences different than the Davis-Jedwab construction. In this thesis, we construct infinite families of quaternary sequences with identical nonperiodic autocorrelation functions. These results demonstrate that current constructions for quaternary families are not all encompassing and need further study.

Contents

1 Introduction

In today's digital world, devices with wireless capabilities are extremely popular. You can surf the net from your local coffee shop or park and call someone from your car. Three main problems need to be addressed with any wireless communications protocol: reliability, speed, and energy efficiency. Reliability in wireless devices addresses questions such as "Will the message get to a receiver" and "Will the message be same when it gets to the receiver." Speed refers to how much data one device can send to another per unit of time. With speed, there are many bottlenecks, including transmission rates, reception delay, and decoding rates. Energy efficiency is as important as ever with the advent of more advanced mobile wireless devices. These devices normally rely on battery power for all of their functionality, and often one of the biggest drains on the battery is radio communication. Thus reducing the amount of power for sending and receiving data has become very important.

These three objectives of wireless communication often require tradeoffs. For example, if an engineer wants more reliability, he or she will lose speed and efficiency. Therefore one of the major problems being addressed today in the communications field is to find ways to improve all three aspects of wireless communications simultaneously. This is the motivation behind this thesis.

1.1 Orthogonal Frequency-Domain Multiplexing

Orthogonal Frequency-Division Multiplexing (OFDM) is a modulation scheme that allows transmission of multiple bits in a single time unit. It uses the Fast Fourier Transform for efficient modulation and demodulation. OFDM is based on the concept of orthogonal functions. Suppose we have four bits to send to the same receiver. We could send these four bits serially, but we could speed up the process by sending four bits simultaneously on different carrier frequencies. For example, the first bit could be encoded as either $\cos\theta$ or $-\cos\theta$ to represent a 0 or 1 respectively. The second bit could be encoded as either $\cos 2\theta$ or $-\cos 2\theta$. The third bit could be encoded either as $\cos 3\theta$ or $-\cos 3\theta$. Finally the fourth bit could be encoded as either $\cos 4\theta$ or $-\cos 4\theta$. Think of these four bits being sent as a radio wave shaped like the sum of the four cosine waves. Observing that $\cos n\theta$ is orthogonal to $\cos m\theta$ when $n \neq m$ ($\int_0^{2\pi} \cos n\theta \cdot \cos m\theta = 0$), it is fairly simple to see that the original message could be recovered by breaking apart the received message into the intended bits.

A big issue for OFDM involves the potential for large spikes in the combined cosine waves. For example, if all four bits were zero, then at times 0 and 2π, the wave would peak at 4. If the peak power of a wireless transmission is bounded by a design or regulatory limit, then not all possible messages possible can be sent. Also, if there is no limit to the peaks by regulation or design, high peaks still drain battery more then a good peak-to-mean value.

1.2 Peak-to-Mean Envelope Power Ratio

In order to address battery efficiency we need the following definition.

Definition 1 *An M-ary sequence A of length n is the vector $A = \{a_0, a_1, \ldots, a_{n-1}\}$ where $a_i = \xi^j$, $\xi = e^{2\pi i/M}$, $0 \leq j \leq M-1$. The associated generating polynomial of A is $a(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1}x^{n-1}$.*

Peak-to-mean envelope power ratio (PMEPR) is the measure for how suitable a sequence is for use in OFDM transmission. It measures the ratio of the peak of the squared signal to the mean envelope power for a message. A technical definition of PMEPR can be found in [1]. We will restrict the allowable patterns for transmission in order to keep the PMEPR as low as possible. The importance of a sequence with low PMEPR is that it improves energy efficiency, and if we can guarantee low PMEPR, it is possible for engineers to design more efficient wireless communication systems.

1.3 Nonperiodic Autocorrelation Function

The nonperiodic autocorrelation function (NPAF) of a sequence, often called the aperiodic autocorrelation function, measures the self-similarity of a sequence at various shifts. We define the NPAF as follows:

Definition 2 *The nonperiodic autocorrelation function for a finite sequence of length* n $A = \{a_1, a_2, \ldots, a_n\}$ *is*

$$\mathrm{NPAF}(A) = \{\mathrm{NPAF}_1(A), \mathrm{NPAF}_2(A), \ldots, \mathrm{NPAF}_{n-1}(A)\},$$

where

$$\mathrm{NPAF}_k(A) = \sum_{i=0}^{n-k-1} \overline{a_i} a_{i+k},$$

and $\overline{a_i}$ *is the complex conjugate of* a_i.

Example 3 *Let* $A = (1, -1, -i, 1, 1)$; *then* $\mathrm{NPAF}_2(A)$ *can be computed as follows:*

$$\begin{array}{ccccccc} 1, & -1, & -i, & 1, & 1 & & \\ & & 1, & -1, & i, & 1, & 1 \\ \hline & & -i & -1 & +i & & \end{array}$$

Note the complex conjugate on the lower sequence. Thus the result is $\mathrm{NPAF}_2(A) = -i - 1 + i = -1$. *We get* $\mathrm{NPAF}(A) = \{2i, -1, 0, 1\}$.

In some papers ([1], [2]) one may find

$$\mathrm{NPAF}(A) = \{\mathrm{NPAF}_{-(n-1)}(A), \mathrm{NPAF}_{-(n-2)}(A), \ldots, \mathrm{NPAF}_0(A), \ldots, \mathrm{NPAF}_{n-1}(A)\}$$

but we do not do this here as it is guaranteed that $\mathrm{NPAF}_{-k}(A) = \overline{\mathrm{NPAF}_k(A)}$ and $\mathrm{NPAF}_0(A) = n$. Our definition does not include this repeated information.

1.4 Golay Sequences

We will focus attention on pairs of sequences that complement each other as in the following definition.

Definition 4 *A Golay sequence A is a sequence such that there exists a sequence B so that $\mathrm{NPAF}_k(A) + \mathrm{NPAF}_k(B) = 0$ for all k. The sequences A and B are called a Golay pair.*

Example 5 *The sequence $A = \{1, 1, 1, -1\}$ has $NPAF(A) = \{1, 0, -1\}$. The sequence $B = \{1, 1, -1, 1\}$ has $\mathrm{NPAF}(B) = \{-1, 0, 1\}$. Since $\mathrm{NPAF}_k(A) + \mathrm{NPAF}_k(B) = 0$ for all k, A and B are a Golay pair.*

Example 6 *The sequence $A = \{1, 1, 1, -1, i, i, -i, i\}$ has $NPAF(A) = \{-i, 0, 3i, 0, i, 0, i\}$. The sequence $B = \{1, 1, 1, -1, -i, -i, i, -i\}$ has $\mathrm{NPAF}(B) = \{i, 0, -3i, 0, -i, 0, -i\}$. Since $\mathrm{NPAF}_k(A) + \mathrm{NPAF}_k(B) = 0$ for all k, A and B are a Golay pair.*

It was first noticed in [6] that Golay sequences have low PMEPR. In that paper, Popovic gives a simple proof that all Golay sequences have a PMEPR of at most 2 (as compared to the maximum possible PMEPR of n), which is a very nice value, especially for large n. Thus, if we restrict allowable sequences to Golay pairs, then battery efficiency is guaranteed.

1.4.1 Davis-Jedwab Construction

In [1], Davis and Jedwab construct length 2^n Golay sequences for an alphabet of size 2^M (hereafter referred to the DJ construction). Their construction creates families of sequences in $\mathbb{Z}_{2^M}$, which could be encoded for transmission. For more details on this construction, see [1] or [2]. Despite this construction, Davis and Jedwab were unable to prove that their construction created all Golay sequences. Recently, it has been shown that this algorithm does not account for all Golay sequences for non-binary alphabets.

1.4.2 More Golay Sequences

In [4], six years after the DJ construction paper was published, Li and Chu showed by exhaustive search that there were 1024 quaternary Golay sequences of length 16 that were not part of the DJ construction. Their paper found 16 characteristic sequences from which the rest of the 1024 new quaternary Golay sequences of length 16 can be constructed, but Li and Chu give no theoretical explanation for the existence of these sequences. This result, however, does indicate that we still do not completely understand Golay sequences.

About half a year later, Fiedler and Jedwab explained the existence of these sequences in [2]. In order to understand their explanation, we need to first describe when we do expect two sequences to have identical NPAFs.

1.5 Whitehead

In 1978, Whitehead [3] published a paper that determined an upper-bound for the number of distinct NPAFs for binary sequences, and also indicated where this upper bound is not sharp. He identifies families of sequences that have the same NPAF. His paper did not focus on Golay sequences, so all binary sequences were considered.

Many of Whitehead's arguments about these sequences utilize the generating polynomial of a sequence. The following lemma demonstrates the importance of the generating polynomial in relation to the NPAF.

Lemma 7 *If $A = \{a_0, a_1, \ldots, a_{n-1}\}$, where A is a M-ary sequence and $a(x)$ is its generating polynomial, then* $a(x)\overline{a(x^{-1})} = n + \sum_{k=1}^{n-1} \mathrm{NPAF}_k(A)x^k + \overline{\mathrm{NPAF}_k(A)}x^{-k}$.

Proof:

$$\begin{aligned}
a(x)\overline{a(x^{-1})} &= (a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1}x^{n-1}) \\
&\quad (\overline{a_0} + \overline{a_1}x^{-1} + \overline{a_2}x^{-2} + \cdots + \overline{a_{n-1}}x^{-(n-1)}) \\
&= (a_0\overline{a_{n-1}})x^{-(n-1)} + (a_0\overline{a_{n-2}} + a_1\overline{a_{n-1}})x^{-(n-2)} + \cdots + \\
&\quad (a_0\overline{a_0} + a_1\overline{a_1} + \cdots + a_{n-1}\overline{a_{n-1}}) + \cdots +
\end{aligned}$$

$$
\begin{aligned}
&(\overline{a_0}a_{n-2} + \overline{a_1}a_{n-1})x^{n-2} + (\overline{a_0}a_{n-1})x^{n-1} \\
= \; & \overline{\mathrm{NPAF}_{n-1}(A)}x^{-(n-1)} + \overline{\mathrm{NPAF}_{n-2}(A)}x^{-(n-2)} + \cdots + n + \cdots + \\
& \mathrm{NPAF}_{n-2}(A)x^{n-2} + \mathrm{NPAF}_{n-1}(A)x^{n-1} \\
= \; & n + \sum_{k=1}^{n-1} \mathrm{NPAF}_k(A)x^k + \overline{\mathrm{NPAF}_k(A)}x^{-k}.
\end{aligned}
$$

□

Whitehead only proves the above result for binary sequences, but the generalization stated above is obvious. Lemma 7 gives us this important corollary:

Corollary 8 *Let A and B be M-ary sequences of length n and let $a(x)$ and $b(x)$ be their respective generating polynomials. Then $a(x)\overline{a(x^{-1})} = b(x)\overline{b(x^{-1})}$ if and only if* $\mathrm{NPAF}(A) = \mathrm{NPAF}(B)$.

Because of the result of Corollary 8, if $a(x)$ is the generating polynomial of a sequence A, the function $a(x)\overline{a(x^{-1})}$ is called the *generating function* for the $\mathrm{NPAF}(A)$.

Definition 9 *The reversal of a sequence $A = \{a_0, a_1, ..., a_{n-1}\}$ is $R(A) = \{a_{n-1}, a_{n-1}, ..., a_0\}$ and the negation of a sequence $A = \{a_0, a_1, ..., a_{n-1}\}$ is $N(A) = \{-a_0, -a_1, \ldots, -a_{n-1}\}$.*

We note that $N(R(A)) = R(N(A))$: for sequence $A = a_0, a_1, \ldots, a_{n-1}$, $N(R(A)) = N(a_{n-1}, a_{n-1}, \ldots, a_0) = -a_{n-1}, -a_{n-1}, \ldots, -a_0 = R(-a_0, -a_1, \ldots, -a_{n-1}) = R(N(A))$. Also, given A with generating polynomial $a(x)$, the generating polynomial of $N(A)$ is $-a_0 + (-a_1)x + \cdots + (-a_{n-1})x^{n-1}) = -a(x)$, and the generating polynomial for $R(A)$ is $a_{n-1} + a_{n-2}x + \cdots + a_0x^{n-1} = x^{n-1}a(x^{-1})$.

From these facts, Whitehead shows:

Lemma 10 *For any finite binary sequence A,* $\mathrm{NPAF}(A) = \mathrm{NPAF}(N(A)) = \mathrm{NPAF}(R(A)) = \mathrm{NPAF}(N(R(A)))$.

Proof: Let $a(x)$ be the generating polynomial of A. First, $a(x)a(x^{-1}) = -a(x)(-a(x^{-1}))$, which implies $\text{NPAF}(A) = \text{NPAF}(N(A))$. Next, $a(x)a(x^{-1}) = (x^{n-1}x^{-(n-1)})a(x)a(x^{-1}) = x^{n-1}a(x^{-1})x^{-(n-1)}a(x)$, which implies $\text{NPAF}(A) = \text{NPAF}(R(A))$. Finally, since $\text{NPAF}(A) = \text{NPAF}(N(A))$ and $\text{NPAF}(A) = \text{NPAF}(R(A))$, we have $\text{NPAF}(A) = \text{NPAF}(N(R(A)))$. Thus $\text{NPAF}(A) = \text{NPAF}(N(A)) = \text{NPAF}(R(A)) = \text{NPAF}(N(R(A)))$.
$\square$

Since $N(A)$ can never equal A, Lemma 10 implies that if $A \neq R(A)$ and $R(A) \neq N(A)$, there are four distinct sequences that have identical NPAFs: $A, N(A), R(A)$, and $N(R(A))$. If $A = R(A)$, there are two distinct sequences which have identical NPAFs: A and $N(A)$. Finally, if $N(A) = R(A)$, there are two distinct sequences that have identical NPAFs: A and $N(A)$.

1.6 Extending To M-ary Sequences

Since we are interested in more than just binary sequences, it seems logical to try to extend Whitehead's work to M-ary sequences. Most of his results extend naturally. We are interested in finite sequences for an M-ary alphabet using the M-th roots of unity $\{\xi^0, \xi^1, ..., \xi^{M-1}\}$, where $\xi = \exp(i\pi/M)$.

Recalling that negation in binary was simply a scalar multiplication of the sequence by -1, it is logical to extend this to all values in $\{\xi^0, \xi^1, ..., \xi^{M-1}\}$. Given a sequence from the Mth roots of unity $A = \{a_0, a_1, \ldots, a_{n-1}\}$ and $c \in \{\xi^0, \xi^1, ..., \xi^{M-1}\}$, define scalar multiplication to be $cA = \{ca_0, ca_1, \ldots, ca_{n-1}\}$.

Lemma 11 *Let A and B be length n $M-$ary sequences. If $A \in \{cB | c \in \{\xi^0, \xi^1, ..., \xi^{M-1}\}\} \cup \{c\overline{R(B)} | c \in \{\xi^0, \xi^1, ..., \xi^{M-1}\}\}$, then* $\text{NPAF}(A) = \text{NPAF}(B)$.

Proof: Let $b(x)$ be the generating polynomial of the sequence B. Suppose $A = cB$ for some $c \in \{\xi^0, \xi^1, ..., \xi^{M-1}\}$. Observe that the generating polynomial of

A is simply $cb(x)$. Thus $cb(x)\overline{cb(x^{-1})} = c\bar{c}b(x)\overline{b(x^{-1})} = b(x)\overline{b(x^{-1})}$. This implies NPAF($A$) = NPAF($B$) by Corollary 8. Now suppose $A = c\overline{R(B)}$ for some $c \in \{\xi^0, \xi^1, ..., \xi^{M-1}\}$. The generating polynomial of A is $cx^{n-1}\overline{b(x^{-1})}$. Therefore $cx^{n-1}\overline{b(x^{-1})}\bar{c}x^{-(n-1)}b(x) = c\bar{c}x^{n-1}x^{-(n-1)}b(x)\overline{b(x^{-1})} = b(x)\overline{b(x^{-1})}$, which implies NPAF($A$) = NPAF($B$), completing the proof. $\square$

Lemma 11 shows that any two sequences are expected to have the same NPAF is one sequence is equal to a scalar multiple or the other or a scalar multiple of the reversal conjugate of the other.

1.7 Crossover

We are now prepared to define crossover. In Lemma 11, we showed which sequences are guaranteed to have the same NPAF.

Definition 12 *Two sequences A and B are defined to be crossover if* NPAF(A) = NPAF(B), *yet* $A \notin \{cB | c \in \{\xi^0, \xi^1, ..., \xi^{M-1}\}\} \cup \{c\overline{R(B)} | c \in \{\xi^0, \xi^1, ..., \xi^{M-1}\}\}$.

Davis and Jedwab observed a pair of length 8 quaternary Golay sequences that were crossover in [1]. In [2], Fiedler and Jedwab show that if any length n Golay pair crossover exists, then that pair can be used to construct Golay pairs in length $2n$ that are not accounted for by the DJ construction. Putting these two facts together, Fiedler and Jedwab explain the Li and Chu Golay sequences.

This new understanding of Golay sequences based on crossover motivates the rest of the thesis.

2 Finding and Predicting Crossover

In view of the fact that the newly discovered Golay sequences in [4] arise because of crossover, the next step is to try to find and explain where crossover occurs. To do this, we follow a technique used in [3] by Whitehead. In that paper, he put an upper bound on the number of distinct NPAFs that occur for different length binary sequences. He then used computer search to find where this upper bound is not sharp, and using this, he explains why the upper bound might not be sharp. Following this technique, we search for crossover in larger alphabets.

2.1 Binary Upper Bound

From Lemma 10, Whitehead creates a function f which gives an upper bound on the number of distinct NPAFs of length n:

Theorem 13 *An upper bound for the number of distinct NPAFs for all binary sequences of length n is*

$$f(n) = \begin{cases} \frac{2^n - 2^{(n+1)/2}}{4} + 2^{(n-1)/2} & \text{if } n \text{ odd;} \\ \frac{2^n - 2^{n/2+1}}{4} + 2^{n/2} & \text{if } n \text{ even.} \end{cases}$$

Proof: Assume n odd for sequence A. Note that $N(A) \neq R(A)$, since the middle values in the sequences are forced to have opposite signs. We claim there are $2^{(n+1)/2}$ different sequences such that the sequence is equal to its own reversal. To see this, let $A = \{a_0, a_1, \ldots, a_{n-1}\}$. For $A = R(A)$, $a_0 = a_{n-1}$, $a_1 = a_{n-2}$, $\ldots a_{(n-3)/2} = a_{(n+1)/2}$, or in other words, $A = \{a_0, a_1, \ldots, a_{(n-3)/2}, a_{(n-1)/2}, a_{(n-3)/2}, \ldots, a_1, a_0\}$. This implies that there are $(n+1)/2$ free choices, each of which can assume one of two values $(+1, -1)$. Thus there are $2^{(n+1)/2}$ sequences that equal their reversal. Therefore, counting distinct NPAFs where a sequence equals its own reversal, there are at most $2^{(n-1)/2}$ NPAFs, as NPAF(A) = NPAF$(N(A))$. Since there are 2^n sequences of length n and $2^{(n+1)/2}$ that equal their reversals, there must be $(2^n - 2^{(n+1)/2})$ sequences of length n that do not equal their reversals. Since they do not equal their reversals, we know that $(2^n - 2^{(n+1)/2})$ sequences yield at most $(2^n - 2^{(n+1)/2})/4$ distinct NPAFs. These facts verify our results about $f(n)$ for n odd.

Assume n even for sequence A. First we count the number of sequences that equal their own reversal. For A to equal its reversal, it must be the case that $A = \{a_0, a_1, \ldots, a_{n/2-1}, a_{n/2-1}, \ldots, a_1, a_0\}$. Thus there are $n/2$ free choices that can assume either of the two values $(+1, -1)$, and hence there are $2^{n/2}$ distinct sequences that equal their reversal. These account for $2^{n/2-1}$ possibly distinct NPAFs. Next, we need to count the sequences that have their negation equal to their reversal. Similar to reversal, for $N(A) = R(A)$, $A = \{a_0, a_1, \ldots, a_{n/2-1}, -a_{n/2-1}, \ldots, -a_1, -a_0\}$, giving $n/2$ free choices that can assume either one of two values $(+1, -1)$. Thus there are $2^{n/2}$ distinct sequences satisfying $N(A) = R(A)$, accounting for another $2^{n/2-1}$ possibly distinct NPAFs. Since there are 2^n sequences and $2^{n/2+1}$ sequences that either satisfy $A = R(A)$ or $R(A) = N(A)$, there must be $(2^n - 2^{n/2+1})$ sequences that satisfy $A \neq R(A)$ and $R(A) \neq N(A)$. This yields at most $(2^n - 2^{n/2+1})/4$ distinct NPAFs. $\square$

Since we now have f, an upper bound for the number of distinct NPAFs we can expect for different length binary sequences, the logical question to ask is "Is this upper bound sharp? If not, how good is this upper bound?" To answer this, Whitehead exhaustively searched the first 15 values of n. We extended this to 23 by brute search. The results are shown in Table 1.

Some very interesting results are noticeable in this table: the one length 17 crossover and the 67 length 21 crossovers. These are unexpected, as seventeen is prime and all of Whitehead's theorems about crossover involve composite lengths. The length 17 pair is as follows: $\{1, 1, 1, 1, 1, -1, -1, -1, 1, 1, -1, 1, 1, -1, 1, -1, 1\}$ and $\{1, 1, 1, -1, -1, -1, 1, -1, 1, 1, 1, 1, -1, 1, 1, -1, 1\}$. Both of these sequences have NPAF $\{0, 1, 0, -3, 0, -3, 0, 5, 0, 1, 0, 1, 0, 1, 0, 1\}$. Noting that all odd offsets had value 0, we realized that these sequences were both interleavings of length nine symmetric sequences with length eight skew-symmetric sequences. Because of this,

Table 1: Comparison of f(n) to the true number of distinct NPAFs

n = length of binary sequence	f(n) (upper bound on number of distinct NPAFs)	Actual number of distinct NPAFs	Error = f(n) - Actual
1	1	1	0
2	2	2	0
3	3	3	0
4	6	6	0
5	10	10	0
6	20	20	0
7	36	36	0
8	72	72	0
9	136	135	1
10	272	272	0
11	528	528	0
12	1056	1048	8
13	2080	2080	0
14	4160	4160	0
15	8256	8242	14
16	16512	16500	12
17	32896	32895	1
18	65792	65750	42
19	131328	131328	0
20	262656	262612	44
21	524800	524733	67
22	1049600	1049600	0
23	2098176	2098176	0

we decided to test all length $2n+1$ symmetric sequences interleaved with all length $2n$ skew-symmetric sequences to try and find crossover at values unreachable by brute force. This included length 29, 37, and 41. No crossovers of these types occurred at these lengths, and no further explanation for this length 17 sequence pair has been achieved. Also, for length 21 the number of unexpected crossovers being 67 is somewhat of an anomaly. Sadly, because of time, we were unable to further explore these results.

2.1.1 Tensor Product

To explain where this upper bound fails, Whitehead noted that it failed when $n = n_1 n_2$ where n_1 and n_2 were both integers greater than 2. He then looked at the tensor product of two sequences. The tensor product of two sequences $A = \{a_0, a_1, \ldots, a_{n-1}\}$ and $B = \{b_0, b_1, \ldots, b_{m-1}\}$ is defined to be

$$A * B = \{a_0b_0, a_1b_0, \ldots, a_{n-1}b_0, a_0b_1, \ldots, a_{n-1}b_1, \ldots, a_0b_{m-1}, a_1b_{m-1}, \ldots, a_{n-1}b_{m-1}\}.$$

From the tensor product equation $A * B = C$, we want to find the generating polynomial for C, which we denote by $c(x)$. Let A's generating function be $a(x)$ and B's generating function be $b(x)$. Then

$$\begin{aligned} c(x) &= a_0b_0 + a_1b_0x + a_2b_0x^2 + \cdots + a_{n-1}b_{m-1}x^{mn-1} \\ &= (a_0 + a_1x^1 + a_2x^2 + \cdots + a_{n-1}x^{n-1}) \\ &\quad (b_0 + b_1x^n + b_2x^{2n} + \cdots + b_{m-1}x^{(m-1)n}) \\ &= a(x)b(x^n). \end{aligned}$$

This gives us a way of looking at the function $a(x)b(x^n)a(x^{-1})b(x^{-n})$ as the corresponding to the NPAF$(A*B)$, and led Whitehead to prove the following lemma for binary, which we extend here to M-ary:

Lemma 14 *Given finite sequences A, B, C, and D of the M-th roots of unity such that* $\mathrm{NPAF}(A) = \mathrm{NPAF}(B)$ *and* $\mathrm{NPAF}(C) = \mathrm{NPAF}(D)$, $\mathrm{NPAF}(A*C) = \mathrm{NPAF}(B*D)$.

<u>**Proof:**</u> Suppose we have two sequences A, B, of length n, and two sequences C, D, of length m, such that $\mathrm{NPAF}(A) = \mathrm{NPAF}(B)$ and $\mathrm{NPAF}(C) = \mathrm{NPAF}(D)$. Let the generating polynomials of A, B, C, and D be $a(x), b(x), c(x)$, and $d(x)$ respectively. These assumptions imply that $a(x)\overline{a(x^{-1})} = b(x)\overline{b(x^{-1})}$ and $c(x)\overline{c(x^{-1})} = d(x)\overline{d(x^{-1})}$. The generating function of $A * C$ is $a(x)c(x^n)$ and $B * D$ is $b(x)d(x^n)$. Thus, $a(x)c(x^n)\overline{a(x^{-1})c(x^{-n})} = a(x)\overline{a(x^{-1})}c(x^n)\overline{c(x^{-n})} = b(x)\overline{b(x^{-1})}d(x^n)\overline{d(x^{-n})} = b(x)d(x^n)\overline{b(x^{-1})d(x^{-n})}$, implying that $\mathrm{NPAF}(A * C) = \mathrm{NPAF}(B * D)$ by Corollary 8. $\square$

Using Lemma 14, one can easily see situations where this tensor product yields sequences that have the same NPAF, but do not fall under expectations from Lemma 11. Consider the following example. Looking at the sequences $A = \{1, 1, -1\}$ and $B = \{1, -1, -1\}$, we first note that $\mathrm{NPAF}(A) = \{0, -1\} = \mathrm{NPAF}(B)$. Applying Lemma 14, we get $\mathrm{NPAF}(A * A) = \mathrm{NPAF}(A * B)$. We can verify this by noting that $\{1, 1, -1\} * \{1, 1, -1\} = \{1, 1, -1, 1, 1, -1, -1, -1, 1\}$ and $\{1, 1, -1\} * \{1, -1, -1\} = \{1, 1, -1, -1, -1, 1, -1, -1, 1\}$, and $\mathrm{NPAF}(\{1, 1, -1, 1, 1, -1, -1, -1, 1\}) = \{-3, 0, 1, 0, -3, 0, 1\} = \mathrm{NPAF}(\{1, 1, -1, -1, -1, 1, -1, -1, 1\})$. Finally, observe that $\{1, 1, -1, 1, 1, -1, -1, -1, 1\}$ is not the negation, reversal, or negation of the reversal of $\{1, 1, -1, -1, -1, 1, -1, -1, 1\}$, and thus this is the one duplicate NPAF value of length 9 not explained by Lemma 10.

Whitehead's paper gives a recursive argument explaining why it is necessary that $n = n_1 n_2$ with n_1 and n_2 integers greater than two. In other words, he proves why there is no crossover for length 4, 6, 8, 10, 14, or 22 as is seen in Table 1.

2.1.2 Other Situations When The Upper Bound Fails

Whitehead provides another interesting proof that shows why even if one were able to count the tensor product crossovers, obtaining a sharp upper bound would be difficult. His proof is as follows:

Lemma 15 *Given binary sequences $A_1, ..., A_k$ of length n with generating polynomials $a_1(x), \ldots, a_k(x)$ and binary sequences T and U of length m with generating polynomials $t(x)$ and $u(x)$ where* $\mathrm{NPAF}(T) = \mathrm{NPAF}(U)$, *then* $\mathrm{NPAF}(A_1 * T|A_2 * T|\ldots|A_k * T) = \mathrm{NPAF}(A_1 * U|A_2 * U|\ldots|A_k * U)$, *where the operation $|$ is concatenation.*

Proof: Assume we have binary sequences $A_1, ..., A_k$ of length n with generating polynomials $a_1(x), \ldots, a_k(x)$ as well as binary sequences T and U of length m with generating polynomials $t(x)$ and $u(x)$ where $\mathrm{NPAF}(T) = \mathrm{NPAF}(U)$. The generating function for the NPAF of the first sequence is

$$\left[\sum_{i=1}^{k} a_i(x)x^{(i-1)mn}\right] t(x^n)t(x^{-n}) \left[\sum_{i=1}^{k} a_i(x^{-1})x^{-(i-1)mn}\right]$$

while the generating of the second sequence is

$$\left[\sum_{i=1}^{k} a_i(x)x^{(i-1)mn}\right] u(x^n)u(x^{-n}) \left[\sum_{i=1}^{k} a_i(x^{-1})x^{-(i-1)mn}\right].$$

Since we assumed $\mathrm{NPAF}(T) = \mathrm{NPAF}(U)$, we know $u(x)u(x^{-1}) = t(x)t(x^{-1})$, and thus these two sequences have the same NPAF. □

Lemmas 14 and 15 shed light on the reasons why it is difficult to come up with a count that would give a sharp upper bound on the number of distinct NPAFs. These two proofs, though, give us two different ways to look for occurrences of crossover.

2.2 M-ary Upper Bound

As in the binary alphabet case, we find an upper bound on the number of distinct NPAF values expected noting Lemma 11.

Theorem 16 *An upper bound on the number of distinct NPAFs for all M-ary sequences of length n is*

$$f(n,M) = \begin{cases} \frac{M^n - M^{(n+1)/2}}{2M} + M^{(n-1)/2} & \textit{if } n \textit{ odd;} \\ \frac{M^n - M^{n/2+1}}{2M} + M^{n/2} & \textit{if } n \textit{ even.} \end{cases}$$

Proof: Again, we will proceed by cases. First assume n is odd. We first can count the sequences that equal their reversal conjugate times a scalar multiple (including 1). For this to be true for a sequence A and multiple c, we note that $A = \{a_0, a_1, \ldots, a_{(n-3)/2}, a_{(n-1)/2}, c\overline{a_{(n-3)/2}}, \ldots, c\overline{a_1}, c\overline{a_0}\}$ must hold with $a_{(n-1)/2} = c\overline{a_{(n-1)/2}}$. Thus $c = a^2_{(n-1)/2}$. So we have M choices for $a_{(n-1)/2}$, which fixes c. Next we note that there are M free choices each of the for the first $(n-1)/2$ values, giving no more free choices for the rest of the sequence. Thus there are $M^{(n+1)/2}$ sequences that equal their reversal conjugate times a scalar multiple, yielding $M^{(n-1)/2}$ distinct NPAF values after removing duplicates for the M scalar multiples of each sequence. This leaves $M^n - M^{(n+1)/2}$ sequences which do not equal their reversal conjugates times a scalar, and thus yield $(M^n - M^{(n+1)/2})/2M$ distinct NPAF values. Thus for odd n, there are $(M^n - M^{(n+1)/2})/2M + M^{(n-1)/2}$ distinct NPAFs possible. Notice with $M = 2$, this is $(2^n - 2^{(n+1)/2})/4 + 2^{(n-1)/2}$, which is same as the odd case from Theorem 13.

Now assume n is even. We first count the sequences that equal their reversal conjugates times a scalar multiple (including 1). For this to be true for a sequence A and multiple c, we note that with $A = \{a_0, a_1, \ldots, a_{n-1}\}$, $\overline{cR(A)} = \{c\overline{a_{n-1}}, c\overline{a_{n-2}}, \ldots, c\overline{a_0}\}$. Thus $c\overline{a_i} = a_{n-i-1}$ for all i. Thus we have M choices for c and M free choices for each of the first $n/2$ values of A. The last half of A is constrained by these free choices. Thus there are $M^{n/2+1}$ sequences such that A is equal to its reversal conjugate times a scalar multiple. These sequences yield at most $M^{n/2}$ distinct NPAFs after removing the M duplicates for scalar multiples of

Table 2: Comparison of f(n, 4) to the true number of distinct NPAFs

n = length of quaternary sequence	f(n) (upper bound on number of distinct NPAFs)	Actual number of distinct NPAFs	Error = f(n) - Actual
1	1	1	0
2	4	4	0
3	10	10	0
4	40	40	0
5	136	136	0
6	544	540	4
7	2080	2080	0
8	8320	8308	12
9	32896	32776	120
10	131584	131492	92
11	524800	524800	0
12	2099200	2097166	2034
13	8390656	8390656	0
14	33562624	33561832	792

each sequence. This leaves $M^n - M^{n/2+1}$ sequences which do not equal their reversal conjugates, and thus yield $(M^n - M^{n/2+1})/2M$ distinct NPAF values. Thus for even n, there are $(M^n - M^{n/2+1})/2M + M^{(n/2}$ distinct NPAFs possible. Notice with $M = 2$, this is $(2^n - 2^{n/2+1})/4 + 2^{n/2}$, which is same as the odd case from Theorem 13.

Putting both the odd and even results together, we get $f(n, M)$. $\square$

Again the question to ask is how good is this upper bound. In extending from binary to M-ary sequences, we know the bound will not be tight, but will the first unexpected crossover happen before length 9? We know that the Golay pairs at length 8 will occur by the Fiedler and Jedwab observation, but are they the only ones? In Table 2, the results for quaternary sequences of length up to 14 are shown.

As expected, the Golay pairs of interest that initiated this search for crossover occurred. Unexpectedly, though, there were four different length six crossovers, as

Table 3: The four crossover in quaternary length six sequences

Sequence A	Sequence B	NPAF	Diff. Sequence C
1, 1, 1, -i, i, -i	1, i, 1, i, 1, -i	-i, 2, -i, 0, -i	1, i, 1, -1, -i, 1
1, 1, -1, -1, 1, -1	1, i, -1, 1, -i, -1	-1, -2, 1, 0, -1	1, i, 1, -1, -i, 1
1, 1, 1, i, -i, i	1, i, -1, i, -1, i	i, 2, i, 0, i	1, i, -1, 1, -i, 1
1, 1, 1, -1, -1, 1	1, i, -1, -1, i, 1	1, -2, -1, 0, 1	1, i, -1, 1, -i, 1

well as ten other sequences in length eight and 92 in length ten. These were unexpected because they were not tensor products. We know this because Whitehead showed that crossover due to Lemma 14 only occur for length $n = n_1 n_2$ where n_1 and n_2 are integers greater than 2. Also unexpected were many sequences for length nine and twelve that were not tensors.

3 Crossover Patterns

To begin an investigation into where these crossover values are occurring in quaternary, it seemed logical to begin with the length six sequences. These sequences, shown as sequence A and sequence B in Table 3, are not tensors. Though no simple pattern has been found, in the table we show a difference sequence, which we call sequence C.

Definition 17 *Given M-ary sequences $A = \{a_0, a_1, \ldots, a_{n-1}\}$ and $B = \{b_0, b_1, \ldots, b_{n-1}\}$ where* $\mathrm{NPAF}(A) = \mathrm{NPAF}(B)$, *then the difference sequence $C = \{b_0/a_0, b_1/a_1, \ldots, b_{n-1}/a_{n-1}\}$.*

Example 18 *Given $A = \{1, 1, 1, -i, i, -i\}$ and $B = \{1, i, 1, i, 1, -i\}$, then the difference sequence $C = \{1/1, i/1, 1/1, i/-i, 1/i, -i/-i\} = \{1, i, 1, -1, -i, 1\}$.*

Also for reference, the twelve crossover pairs of length eight are shown in Table 4.

Table 4: The twelve crossover in quaternary length eight sequences

Sequence A	Sequence B	NPAF	Diff. Sequence C
1, 1, 1, 1, -i, -i, i, -i	1, i, i, 1, i, 1, 1, -i	2-i, 2-2i, 2-i, -2i, -i, 0, -i	1, i, i, 1, -1, i, -i, 1
1, 1, 1, -1, -1, -1, 1, -1	1, i, i, -1, 1, -i, -i, -1	1, 0, -3, 0, -1, 0, -1	1, i, i, 1, -1, i, -i, 1
1, 1, 1, 1, i, i, -i, i	1, i, i, -1, i, -1, -1, i	2+i, 2+2i, 2+i, 2i, i, 0, i	1, i, i, -1, 1, i, -i, 1
1, 1, -i, -i, 1, -1, i, -i	1, i, -1, i, 1, i, 1, -i	-i, 0, -3i, 0, i, 0, -i	1, i, -i, -1, 1, -i, -i, 1
1, 1, -i, -i, i, i, 1, -1	1, i, -1, i, i, 1, -i, -1	1-2i, -2-2i, -1+2i, 2i, 1, 0, -1	1, i, -i, -1, 1, -i, -i, 1
1, 1, i, i, 1, -1, -i, i	1, i, -1, i, -1, -i, -1, i	i, 0, 3i, 0, -i, 0, i	1, i, i, 1, -1, i, -i, 1
1, 1, -i, i, i, -i, -1, 1	1, i, -1, i, -i, -1, i, 1	-1-2i, -2+2i, 1+2i, -2i, -1, 0, 1	1, i, -i, 1, -1, -i, -i, 1
1, 1, i, -i, -i, i, -1, 1	1, i, -1, -i, i, -1, i, 1	-1+2i, -2-2i, 1-2i, 2i, -1, 0, 1	1, i, i, 1, -1, i, -i, 1
1, 1, i, i, -i, -i, 1, -1	1, i, -1, -i, -i, 1, -i, -1	1+2i, -2+2i, -1-2i, -2i, 1, 0, -1	1, i, i, -1, 1, i, -i, 1
1, 1, 1, -1, 1, 1, -1, 1	1, i, -i, 1, 1, -i, i, 1	-1, 0, 3, 0, 1, 0, 1	1, i, -i, -1, 1, -i, -i, 1
1, 1, 1, -1, -i, i, -i, i	1, i, -i, 1, -i, 1, -1, i	-2+i, 2-2i, -2+i, -2i, i, 0, i	1, i, -i, -1, 1, -i, -i, 1
1, 1, 1, -1, i, -i, i, -i	1, i, -i, -1, -i, -1, 1, -i	-2-i, 2+2i, -2-i, 2i, -i, 0, -i	1, i, -i, 1, -1, -i, -i, 1

3.1 The Difference Sequences

The reason we included the difference sequences in the table was because they have some very interesting patterns. All pairs of crossover that could were not caused by tensor products had common properties in the difference sequences. For all lengths, we noticed that the difference sequences were symmetrical up to negation. That is, given a difference sequence $C = \{c_0, ..., c_{n-1}\}$, we observed that each element c_j must be such that either $c_j = c_{n-1-j}$ or $c_j = -c_{n-1-j}$. Also, if the sequence was of length $n = n_1 n_2$ where $n_1 \leq n_2$, splitting the sequence into n_1 consecutive parts of length n_2, each part of the sequence was symmetric up to negation, and each part was equal up to negation. So far, we have not come up with a formal proof of either of these.

Also of interest, but still unproven, is that difference sequences come in at least pairs. As can be seen in Table 3, the four pairs all had different NPAFs, yet there were only two pairs of difference sequences that occur. This pattern seems to continue for length 8 where all occur in quadruples or pairs, as there are only four difference sequences for the twelve sequences. We believe this is more than a coincidence. Any proof of this fact could shed light on where crossover occurs.

3.2 Golay Sets

We noticed that summing up the NPAFs for all four pairs of crossover sequences of length 6 (Table 3) yields the all zero sequence. Sets of this nature have special properties and a special name, Golay sets (note Golay pairs are just a Golay set of order 2) [5]. In Table 4, one sees that the 8 pairs of crossover that are not Golay sequences create a Golay octuple. This seems to hold for larger lengths, though no proof of this has been found. The occurrence of Golay sets at nearly every place crossover occurs is very intriguing, and any explanation for this should shed light on the problem of explaining crossover in general.

3.3 Infinite Families of Crossover Sequences

When examining our computer search data for patterns, we observed that each of the length 6 crossover sequences seemed to extend naturally to a very similar length 10 crossover sequence, and the length 10s similarly extended to length 14s. For example, the pair $\{1,1,1,-i,i,-i\}$ and $\{1,i,1,i,1,-i\}$ extends to the length 10 crossover pair $\{1,1,1,1,1,-i,i,i,-i,-i\}$ and $\{1,1,i,1,1,i,i,1,-i,-i\}$, and these extend to the length 14 crossover pair $\{1,1,1,1,1,1,1,-i,i,i,i,-i,-i,-i\}$ and $\{1,1,1,i,1,1,1,i,i,i,1,-i,-i,-i\}$.

In order to describe families of sequences with crossover, we need a convent notation that will allow us to do the NPAF computations.

Definition 19 *The sequence $\{a_0u_0, a_1u_1, \ldots, a_{n-1}u_{n-1}\}$, where $a_i \in \mathbb{N}$ and $u_i \in \{\xi^0, \xi^1, \ldots, \xi^{M-1}\}$, corresponds to the length $\sum_{i=0}^{n-1} a_i$ sequence whose first a_0 terms are u_0, the next a_1 terms are u_1, and so on until the last a_{n-1} terms are u_{n-1}.*

For example, the sequence $\{1,2i,-3,-4i\}$ is simply $\{1,i,i,-1,-1,-1,-i,-i,-i,-i\}$.

Theorem 20 *If* $A_n = \{(n+1), -(n+1), n, -n\}$ *and* $B_n = \{n, i, -n, n, -i, -n\}$, *then* $\text{NPAF}(A_n) = \text{NPAF}(B_n)$.

<u>**Proof:**</u> We need to show that $\text{NPAF}_k(A_n) = \text{NPAF}_k(B_n)$ for all $1 \leq k \leq 4n+1$. We break this up into the cases.

Case 1: $3n+2 \leq k \leq 4n+1$. Let $j = 4n+2-k$. $\text{NPAF}_k(A_n)$ is calculated by

$$\begin{array}{ccccccc}
(n+1), & -(n+1), & n, & (n-j), & -j & & \\
 & & & & j, & (n+1-j), & \dots \\
\hline
 & & & & -j & &
\end{array}$$

and thus $\text{NPAF}_k(A_n) = -j$. $\text{NPAF}_k(B_n)$ is calculated by

$$\begin{array}{ccccccccc}
n, & i, & -n, & n, & -i, & -(n-j), & -j & & \\
 & & & & & & j, & (n-j), & \dots \\
\hline
 & & & & & & -j & &
\end{array}$$

and thus $\text{NPAF}_k(B_n) = -j = \text{NPAF}_k(A_n)$.

Case 2: $k = 3n+1$. $\text{NPAF}_k(A_n)$ is calculated by

$$\begin{array}{ccccccc}
(n+1), & -(n+1), & (n-1), & 1, & -n & & \\
 & & & 1, & n, & -(n+1), & \dots \\
\hline
 & & & 1 & -n & &
\end{array}$$

and thus $\text{NPAF}_k(A_n) = 1-n$. $\text{NPAF}_k(B_n)$ is calculated by

$$\begin{array}{ccccccccc}
n, & i, & -n, & n, & -i, & -(n-1), & -1 & & \\
 & & & & 1, & (n-1), & -i, & -n, & \dots \\
\hline
 & & & & -i & +1-n & +i & &
\end{array}$$

and thus $\text{NPAF}_k(B_n) = -i + 1 - n + i = 1 - n = \text{NPAF}_k(A_n)$.

Case 3: $2n+2 \leq k \leq 3n$. Let $j = 3n+1-k$. $\text{NPAF}_k(A_n)$ is calculated by

$$\begin{array}{cccccccc}
(n+1), & -(n+1), & (n-j-1), & j+1, & -(n-j), & -(j) & & \\
 & & & j+1, & (n-j), & -(j), & -(n+1-j+1), & \dots \\
\hline
 & & & j+1 & -n+j & +j & &
\end{array}$$

and thus $\text{NPAF}_k(A_n) = j+1-n+j+j = 3j-n+1$. $\text{NPAF}_k(B_n)$ is calculated by

$$\begin{array}{cccccccccccc}
n, & i, & -n, & 1, & (n-1-j), & j, & -i, & -(n-j-1), & -1, & -j & & \\
 & & & & & j, & 1, & (n-j-1), & -i, & -j, & -(n-j), & \dots \\
\hline
 & & & & & j & -i & -n+j+1 & +i & +j & &
\end{array}$$

and thus $\text{NPAF}_k(B_n) = j - i - n + j + 1 + i + j = 3j - n + 1 = \text{NPAF}_k(A_n)$.

Case 4: $k = 2n+1$. $\mathrm{NPAF}_k(A_n)$ is calculated by

$$\begin{array}{ccccccc}
(n+1), & -n, & -1, & n, & -n & & \\
 & & 1, & n, & -n, & -1, & \ldots \\
\hline
 & & -1 & +n & +n & &
\end{array}$$

and thus $\mathrm{NPAF}_k(A_n) = -1+n+n = 2n-1$. $\mathrm{NPAF}_k(B_n)$ is calculated by

$$\begin{array}{cccccc}
n, & i, & -n, & n, & -i, & -n \\
 & & & n, & -i, & -n, \ \ldots \\
\hline
 & & & n & -1 & +n
\end{array}$$

and thus $\mathrm{NPAF}_k(B_n) = n-1+n = 2n-1 = \mathrm{NPAF}_k(A_n)$.

Case 5: $n+1 \le k \le 2n$. Let $j = 2n+1-k$. $\mathrm{NPAF}_k(A_n)$ is calculated by

$$\begin{array}{ccccccccc}
(n+1), & -(n+1-j-1), & -(j+1), & (n-j), & j, & -(n+1-j), & -(j-1) & & \\
 & & (j+1), & (n-j), & -j, & -(n+1-j), & (j-1), & (n-j+1), & \ldots \\
\hline
 & & -j-1 & +n-j & -j & +n+1-j & -j+1 & &
\end{array}$$

and thus $\mathrm{NPAF}_k(A_n) = -j-1+n-j-j+n+1-j-j+1 = 2n-5j+1$.

$\mathrm{NPAF}_k(B_n)$ is calculated by

$$\begin{array}{cccccccccccc}
n, & i, & -(n-j), & -j, & (n-j), & 1, & (j-1), & -i, & -(n-j), & -j & & \\
 & & & j, & (n-j), & -i, & -(j-1), & -1, & -(n-j), & j, & (n-j), & \ldots \\
\hline
 & & & -j & +n-j & -i & -j+1 & i & +n-j & -j & &
\end{array}$$

and thus $\mathrm{NPAF}_k(B_n) = -j+n-j-i-j+1+i+n-j-j = 2n-5j+1 = \mathrm{NPAF}_k(A_n)$.

Case 6: $k = n$. $\mathrm{NPAF}_k(A_n)$ is calculated by

$$\begin{array}{ccccccc}
n, & 1, & -n, & -1, & n, & -n & \\
 & 1, & n, & -1, & -n, & n, & -n \\
\hline
 & 1 & -n & +1 & -n & -n &
\end{array}$$

and thus $\mathrm{NPAF}_k(A_n) = 1-n+1-n-n = 2-3n$. $\mathrm{NPAF}_k(B_n)$ is calculated by

$$\begin{array}{ccccccccc}
n, & i, & -(n-1), & -1, & n, & -i, & -(n-1), & -1 & \\
 & 1, & (n-1), & -i, & -n, & 1, & (n-1), & i, & -n \\
\hline
 & i & -n+1 & +i & -n & -i & -n+1 & -i &
\end{array}$$

and thus $\mathrm{NPAF}_k(B_n) = i-n+1+i-n-i-n+1-i = 2-3n = \mathrm{NPAF}_k(A_n)$.

Case 7: $1 \le k \le n-1$. $n+1 \le k \le 2n$. Let $j = n-k$. $\mathrm{NPAF}_k(A_n)$ is calculated by

$$\begin{array}{cccccccccc}
(n-j), & j, & 1, & -(n-j), & -(j+1), & (n-j), & j, & -(n-j), & -j & \\
 & j, & 1, & (n-j), & -(j+1), & -(n-j), & j, & (n-j), & -j, & -(n-j) \\
\hline
 & j & +1 & -n+j & j+1 & -n+j & +j & -n+j & +j &
\end{array}$$

and thus $\text{NPAF}_k(A_n) = j+1-n+j-j-1-n+j+j-n+j-j = 7j-3n+2$.

$\text{NPAF}_k(B_n)$ is calculated by

$$\begin{array}{ccccccccccccc}
(n-j), & j, & i, & -(n-j-1), & -1, & -j, & (n-j), & j, & -i, & -(n-j-1), & -1, & -j & \\
 & j, & 1, & (n-j-1), & -i, & -j, & -(n-j), & j, & 1, & (n-j-1), & i, & -j, & -(n-j) \\
\hline
 & j & +i & -n+j+1 & +i & +j & -n+j & +j & -i & -n+j+1 & -i & +j &
\end{array}$$

and thus $\text{NPAF}_k(B_n) = j+i-n+j+1+i+j-n+j+j-i-n+j+1-i+j = 7j-3n+2 = \text{NPAF}_k(A_n)$.

Thus, for all k, $\text{NPAF}_k(A_n) = \text{NPAF}_k(B_n)$ which implies that $\text{NPAF}(A_n) = \text{NPAF}(B_n)$. $\square$

From Theorem 20 we know one length 14 crossover is $\{4,-4,3,-3\}$ and $\{3,i,-3,3,-i,-3\}$ with NPAF equal to $\{7,0,-7,-8,-3,2,5,4,1,-2,-3,-2,-1\}$, which was also found in the computer search. This theorem gives us an infinite family of pairs of crossover sequences of length $4n+2$. Moreover, as can easily be observed, these sequences are not tensor products of smaller sequences. Interestingly, we can generalize two of the other length 6 crossover sequence pairs as members of similar families that can be shown to exist using the same method as Theorem 20. These two families are the pair $\{(2n+1),-i,ni,-ni\}$ and $\{n,i,n,ni,1,-ni\}$ as well as the pair $\{(2n+1),i,-ni,ni\}$ and $\{n,i,-n,ni,-1,ni\}$. Also, the example $\{(n+3),-(n+3),n,-n\}$ and $\{n,3i,-n,n,-3i,-n\}$ is a family starting at length 10, demonstrating that will be many more infinite families of crossover pairs.

4 Conclusion

In this thesis, we have constructed a new infinite family of quaternary sequences with identical NPAFs. We have indicated that many more of these families exist. Ultimately, by understanding crossover better, we hope future research will uncover more Golay sequences.

- First and most pressing, find a general explanation for crossover. If this can be done, we could compare this construction to the DJ construction and possibly find more Golay pairs.

- Second, explain the binary length 17 crossover pair. Whitehead's explanations for crossover pairs all involved composites, yet 17 is prime.
- Third, illustrate the reasons for the 67 binary crossover sequence pairs of length 21. These could exist because of some currently unknown reason for crossover.
- Fourth, the crossover sequences found by our computer search seem to occur in Golay sets. We have no idea why.
- Fifth, preliminary findings indicate the existence of many infinite families of crossover pairs of length $4n$. These are an excellent place to look for new Golay sequences.
- Finally, extend our computer search to different alphabets, like octary, and try to find families of crossover that are not extensions from binary or quaternary. To do this, smarter brute force techniques would need to be implemented. This could lead to new families of Golay sequences for different alphabets.

References

[1] James A. Davis and Jonathan Jedwab. Peak-to-mean power control in ODFM, Golay complementary sequences, and Reed-Muller codes. *IEEE Transactions on Information Theory*, 45(7):2397–2417, 1999.

[2] Frank Fiedler and Jonathan Jedwab. How do more Golay sequences arise? *IEEE Transactions on Information Theory*, 52(9):4261–4266, 2006.

[3] Earl Glen Whitehead Jr. Autocorrelation of (+1, -1) sequences. *Combinatorial Mathematics*, 686:329–336, 1978.

[4] Y. Li and W. B. Chu. More Golay sequences. *IEEE Transactions on Information Theory*, 51(3):1141–1145, 2005.

[5] K. G. Paterson. Generalized Reed-Muller codes and power control in OFDM modulation. *IEEE Transactions on Communication*, 39:1031–1033, 1991.

[6] B. M. Popovic. Synthesis of power efficient multitone signals with flat amplitude spectrum. *IEEE Transactions on Communication*, 39:1031–1033, 1991.