

Some Results on Nonlinear Zigzag Functions

D. R. Stinson

Department of Combinatorics and Optimization
University of Waterloo
Waterloo Ontario, N2L 3G1
Canada

Abstract

Zigzag functions were defined by Brassard, Crépeau and Sántha [1] in connection with an application to the construction of oblivious transfers (a useful tool in cryptographic protocols). They proved that linear zigzag functions are equivalent to self-intersecting codes, which have been studied by several researchers.

In this paper, we begin an investigation of general (linear or non-linear) zigzag functions. In particular, we prove some bounds (i.e., necessary conditions for existence of zigzag functions) which generalize known bounds for linear zigzag functions.

1 Introduction and Definitions

Zigzag functions were defined by Brassard, Crépeau and Sántha [1]. We review basic concepts and definitions now.

Let \mathbb{F}_q denote the finite field with q elements. Suppose that $f : (\mathbb{F}_q)^n \rightarrow (\mathbb{F}_q)^m$, where $n > m$. Let $I \subseteq \{1, \dots, n\}$. We say that f is *unbiased* with respect to I if for all possible choices for $(x_i : i \in I) \in (\mathbb{F}_q)^{|I|}$, and for every $(y_1, \dots, y_m) \in (\mathbb{F}_q)^m$, there are exactly $q^{n-m-|I|}$ choices for $(x_i : i \in \{1, \dots, n\} \setminus I) \in (\mathbb{F}_q)^{n-|I|}$ such that $f(x_1, \dots, x_n) = (y_1, \dots, y_m)$.

We give an alternative way to define the unbiased concept, which is based on the approach used in [11] and [6] to study correlation-immune and resilient functions. For any $y \in (\mathbb{F}_q)^m$, define A_y to be the array whose rows consist of all the n -tuples in $f^{-1}(y)$. Index the columns of A_y by $1, \dots, n$, and for any $I \subseteq \{1, \dots, n\}$, let $A_y|_I$ denote the restriction of

A_y to the columns in I . Then f is unbiased with respect to I if and only if, for every $y \in (\mathbb{F}_q)^m$, $A_y|_I$ contains every $|I|$ -tuple exactly $q^{n-m-|I|}$ times.

Example 1.1 Consider the function

$$f(x_1, x_2, x_3, x_4) = (x_1 + x_2 + x_3, x_2 + x_3 + x_4),$$

where all arithmetic is performed in \mathbb{F}_2 . The arrays A_y ($y \in (\mathbb{F}_2)^2$) are as follows:

$A_{(0,0)}$	$A_{(0,1)}$	$A_{(1,0)}$	$A_{(1,1)}$
0 0 0 0	0 0 0 1	1 0 0 0	1 0 0 1
1 0 1 1	1 0 1 0	0 0 1 1	0 0 1 0
1 1 0 1	1 1 0 0	0 1 0 1	0 1 0 0
0 1 1 0	0 1 1 1	1 1 1 0	1 1 1 1

The sets I with respect to which f is unbiased are all sets of cardinality one, and all sets of cardinality two except $\{1, 4\}$.

Suppose that, for every $I \subseteq \{1, \dots, n\}$, f is unbiased with respect to at least one of I and $\{1, \dots, n\} \setminus I$. Then f is said to be a *zigzag function*, and denoted as an (n, m, q) -ZF.

Note that the example considered above is a $(4, 2, 2)$ -ZF.

Before proceeding, we observe that the case $m = 1$ is trivial: the function f defined as

$$f(x_1, \dots, x_n) = x_1 + \dots + x_n,$$

where addition is performed in \mathbb{F}_q , is an $(n, 1, q)$ -ZF for any n and q . Therefore, in the remainder of the paper, we will be interested only in (n, m, q) -ZF with $m \geq 2$.

The fundamental problem is to determine, for given q and m , the minimum n such that an (n, m, q) -ZF exists. This problem will be studied in later sections of this paper.

2 Linear Zigzag Functions

In this section, we discuss linear zigzag functions, concentrating on the connections between linear zigzag functions and linear codes.

An (n, m, q) -ZF, f , is said to be *linear* if there exists an $m \times n$ matrix M with entries from \mathbb{F}_q such that $f(x) = xM^T$ for all $x \in (\mathbb{F}_q)^n$. An m -dimensional subspace \mathcal{C} of $(\mathbb{F}_q)^n$ is said to be an $[n, m]$ q -ary code. An $m \times n$ matrix is said to be a *generating matrix* for \mathcal{C} if its rows form a basis for \mathcal{C} . An $(n - m) \times n$ matrix is said to be a *parity-check matrix* for \mathcal{C} if its rows form a basis for the *dual code*, \mathcal{C}^\perp (i.e., the orthogonal complement of \mathcal{C}).

Lemma 2.1 *Let M be a generating matrix for an $[n, m]$ q -ary code, \mathcal{C} , and let H be a parity-check matrix for \mathcal{C} . The function $f(x) = xM^T$ is unbiased with respect to $I \subseteq \{1, \dots, n\}$ if and only if the columns of H indexed by I are linearly independent.*

Proof. The result follows easily from the facts that $f^{-1}(0, \dots, 0) = \mathcal{C}^\perp$, and $f^{-1}(y_1, \dots, y_m)$ is a coset of \mathcal{C}^\perp , for any $(y_1, \dots, y_m) \in (\mathbb{F}_q)^m$. \square

Lemma 2.1 provides a fairly convenient way of checking to see if a given linear function is a zigzag function. We illustrate with an example.

Example 2.1 Suppose $n = 6$ and $m = 3$, and consider the function $f(x) = xM^T$, where

$$M = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

The parity-check matrix of the code generated by M is the matrix

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

All pairs of columns of H are linearly independent. The only sets of three columns of H that are linearly dependent are $\{1, 2, 4\}$, $\{1, 3, 6\}$, $\{2, 3, 5\}$ and $\{4, 5, 6\}$. It follows that f is a $(6, 3, 2)$ -ZF.

Let $x = (x_1, \dots, x_n) \in (\mathbb{F}_q)^n$. The *support* of x , denoted $\text{supp}(x)$, is defined as $\text{supp}(x) = \{i : x_i \neq 0\}$. We provide an alternative formulation of Lemma 2.1 using the idea of support.

Lemma 2.2 *Let M be a generating matrix for an $[n, m]$ q -ary code, \mathcal{C} . The function $f(x) = xM^T$ is unbiased with respect to $I \subseteq \{1, \dots, n\}$ if and only if there does not exist a non-zero codeword $y \in \mathcal{C}$ such that $\text{supp}(y) \subseteq I$.*

Proof. Let H be a parity-check matrix for \mathcal{C} . The result follows easily from Lemma 2.1, using the fact that there exists a non-zero codeword $y \in \mathcal{C}$ such that $\text{supp}(y) \subseteq I$ if and only if the columns of H indexed by I are linearly dependent. \square

An $[n, m]$ q -ary code \mathcal{C} is said to be *intersecting* if, for every $x, y \in \mathcal{C}$ such that $x, y \neq (0, \dots, 0)$, $\text{supp}(x) \cap \text{supp}(y) \neq \emptyset$. The following result, first proved in [1], says that linear zigzags are equivalent, in a strong sense, to intersecting codes. We provide a proof using Lemma 2.2.

Theorem 2.3 *Let M be an $m \times n$ matrix with entries from \mathbb{F}_q . The function $f(x) = xM^T$ is an (n, m, q) -ZF if and only if the linear code with generating matrix M is an intersecting $[n, m]$ q -ary code.*

Proof. Let \mathcal{C} be the code with generating matrix M . Suppose that \mathcal{C} is intersecting, and let $I \subseteq \{1, \dots, n\}$. Suppose that f is biased with respect to both I and $\{1, \dots, n\} \setminus I$. Then, from Lemma 2.2, there exist non-zero codewords $y, z \in \mathcal{C}$ such that $\text{supp}(y) \subseteq I$ and $\text{supp}(z) \subseteq \{1, \dots, n\} \setminus I$. This contradicts the fact that \mathcal{C} is intersecting, so we conclude that f is a zigzag function.

Conversely, suppose that f is a zigzag function. Let $y \in \mathcal{C}$, $y \neq (0, \dots, 0)$. By Lemma 2.2, f is biased with respect to $\text{supp}(y)$. Therefore f is unbiased with respect to $\{1, \dots, n\} \setminus \text{supp}(y)$. Hence, from Lemma 2.2, there does not exist a non-zero codeword $z \in \mathcal{C}$ such that $\text{supp}(z) \subseteq \{1, \dots, n\} \setminus \text{supp}(y)$. Therefore \mathcal{C} is an intersecting code. \square

Various interesting results on linear zigzag functions can be obtained by invoking known results on intersecting codes, which are studied in [7, 9, 3, 10, 4, 1], for example. The best (asymptotic) necessary condition for the existence of linear zigzag functions follows from the observation that an intersecting $[n, m]$ q -ary code has minimum distance at least m , together with the MRRW bound ([8, Ch. 17]). We will return to this later.

3 Nonlinear Zigzag Functions

We begin an investigation of arbitrary (i.e., linear or nonlinear) zigzag functions. Suppose that $f : (\mathbb{F}_q)^n \rightarrow (\mathbb{F}_q)^m$.

Lemma 3.1 *If f is unbiased with respect to I , then $|I| \leq n - m$.*

Proof. This follows immediately from the fact that $q^{n-m-|I|}$ must be an integer. \square

Lemma 3.2 *If an (n, m, q) -ZF exists, then $n \geq 2m - 1$.*

Proof. Apply Lemma 3.1 with $|I| = \lfloor \frac{n}{2} \rfloor$. Then $m \leq \lceil \frac{n}{2} \rceil$, and hence $n \geq 2m - 1$. \square

The following is our main necessary existence condition for the existence of zigzag functions.

Theorem 3.3 *If f is an (n, m, q) -ZF, then f is unbiased with respect to I for all I such that $|I| = m - 1$.*

Proof. Let $|I| = m - 1$. Since f is a zigzag function, f is unbiased with respect to at least one of I and $\{1, \dots, n\} \setminus I$. Since $n - |I| > n - m$, f is biased with respect to $\{1, \dots, n\} \setminus I$, by Lemma 3.1. Therefore, f must be unbiased with respect to I . \square

Suppose that $1 \leq t \leq k$, and $v \geq 2$. An *orthogonal array* $OA_\lambda(t, k, v)$ is a $\lambda v^t \times k$ array A of v symbols, such that within any t columns of A , every possible t -tuple of symbols occurs in exactly λ rows of A . An orthogonal array is *simple* if it does not contain two identical rows. A *large set* of orthogonal arrays $OA_\lambda(t, k, v)$, denoted $LOA_\lambda(t, k, v)$, is a set of v^{k-t}/λ simple $OA_\lambda(t, k, v)$, such that every possible k -tuple occurs as a row in exactly one of the orthogonal arrays in the set.

The following is an immediate corollary of Theorem 3.3.

Corollary 3.4 *If there is an (n, m, q) -ZF with $m \geq 2$, then there is an $LOA_{q^{n-2m+1}}(m - 1, n, q)$, and hence an $OA_{q^{n-2m+1}}(m - 1, n, q)$.*

3.1 The case $n = 2m - 1$

In the special case when $n = 2m - 1$, we can say more. The following theorem is proved in the same way as [11, Theorem 2.1].

Theorem 3.5 *An $(2m - 1, m, q)$ -ZF with $m \geq 2$ exists if and only if an $\text{LOA}_1(m - 1, 2m - 1, q)$ exists.*

Proof. Suppose first that a $(2m - 1, m, q)$ -ZF exists. Applying Corollary 3.4, we see that an $\text{LOA}_1(m - 1, 2m - 1, q)$ exists.

Now, suppose that an $\text{LOA}_1(m - 1, 2m - 1, q)$ exists. Denote the q^m arrays in the large set as A_y , $y \in (\mathbb{F}_q)^m$, and think of each A_y as being composed of a set of $(2m - 1)$ -tuples from $(\mathbb{F}_q)^{2m-1}$. Now define a function $f : (\mathbb{F}_q)^{2m-1} \rightarrow (\mathbb{F}_q)^m$ by the rule

$$f(x_1, \dots, x_{2m-1}) = (y_1, \dots, y_m) \Leftrightarrow (x_1, \dots, x_{2m-1}) \in A_{(y_1, \dots, y_m)}.$$

The resulting function is unbiased with respect to any set of size $m - 1$. Since $n = 2m - 1$, it follows that f is a $(2m - 1, m, q)$ -ZF. \square

Remark A $(2m - 1, m, q)$ -ZF is equivalent to a $(2m - 1, m, m - 1, q)$ -resilient function (see [6, Theorem 5.2]).

We can use the above theorem to obtain necessary conditions for the existence of zigzag functions with $n = 2m - 1$.

Corollary 3.6 *If $q \leq m - 1$ and $m \geq 2$, then there does not exist a $(2m - 1, m, q)$ -ZF.*

Proof. The Bush bound for orthogonal arrays (see, for example, [5, p. 180]), states that if an $\text{OA}_1(t, k, q)$ exists with $q \leq t$, then $k \leq t + 1$. Hence, if a $(2m - 1, m, q)$ -ZF exists with $q \leq m - 1$, then $2m - 1 \leq m$, or $m \leq 1$. Since $m \geq 2$, we have a contradiction. \square

The following existence result is given in [10, Theorem 1]. It uses Reed-Solomon codes.

Theorem 3.7 *For every prime power q and every integer $m \geq 2$ such that $q \geq 2m - 2$, there exists a (linear) $(2m - 1, m, q)$ -ZF.*

3.2 A General Bound

In general, we can use Corollary 3.4 to obtain bounds for zigzag functions that are stronger than the bound of Lemma 3.2. Applying the classical Rao bound for orthogonal arrays (see, for example, [5, p. 180]), the following is obtained.

Theorem 3.8 *Suppose there exists an (n, m, q) -ZF. Then*

$$q^{n-m} \geq \begin{cases} \sum_{i=0}^{\frac{m-1}{2}} \binom{n}{i} (q-1)^i & \text{if } m \text{ is odd} \\ \sum_{i=0}^{\frac{m-2}{2}} \binom{n}{i} (q-1)^i + \binom{n-1}{\frac{m-2}{2}} (q-1)^{\frac{m}{2}} & \text{if } m \text{ is even.} \end{cases}$$

We give an example to illustrate Theorem 3.8.

Example 3.1 Suppose that $n = 16$, $m = 7$ and $q = 2$. Theorem 3.8 tells us that a $(16, 7, 2)$ -ZF exists only if

$$2^{16-7} \geq \sum_{i=0}^3 \binom{16}{i}.$$

Since

$$2^{16-7} = 512$$

and

$$\sum_{i=0}^3 \binom{16}{i} = 1 + 16 + 120 + 560 = 697,$$

we have shown that a $(16, 7, 2)$ -ZF does not exist.

4 Bounds on Binary Zigzag Functions

Define $n^*(m)$ to be the minimum n such that an $(n, m, 2)$ -ZF exists, and define $n_L^*(m)$ to be the minimum n such that a linear $(n, m, 2)$ -ZF exists. We provide a table summarizing known results on $n_L^*(m)$ and $n^*(m)$. Of course it is always the case that $n_L^*(m) \geq n^*(m)$. The values in Table 1 in the column “linear zigzag” are all taken from [1].

Table 1: Results on the minimum n such that an $(n, m, 2)$ -ZF exists

m	linear zigzag	arbitrary zigzag	
	[1]	Theorem 3.8	exact values
2	$n_L^* = 3$	$n^* \geq 3$	$n^* = 3$
3	$n_L^* = 6$	$n^* \geq 6$	$n^* = 6$
4	$n_L^* = 9$	$n^* \geq 8$	$n^* = 9$
5	$n_L^* = 13$	$n^* \geq 12$	
6	$n_L^* = 15$	$n^* \geq 14$	
7	$n_L^* \leq 21$	$n^* \geq 17$	
8	$n_L^* \leq 25$	$n^* \geq 19$	
9	$n_L^* \leq 29$	$n^* \geq 23$	

4.1 The Nonexistence of an $(8, 4, 2)$ -ZF

The case $m = 4$ is an interesting one. It was previously known that a (linear) $(9, 4, 2)$ -ZF exists, and no linear $(8, 4, 2)$ -ZF exists. Theorem 3.8 does not rule out the existence of a (non-linear) $(8, 4, 2)$ -ZF (it cannot do so since an $OA_2(3, 8, 2)$ does in fact exist). In this section, we prove that no $(8, 4, 2)$ -ZF exists.

Suppose that f is an $(8, 4, 2)$ -ZF. We will study the array $A_{(0,0)}$, as defined in the Introduction. Without loss of generality, we will assume that $(0, \dots, 0)$ is one of the rows of $A_{(0,0)}$. By viewing $A_{(0,0)}$ as a set of 8-tuples, we can think of it as an $(8, 16, d)$ code, which we will call \mathcal{C} , where d denotes the minimum Hamming distance of \mathcal{C} . We consider two cases, depending on the value of d .

4.1.1 Case 1: $d \geq 4$

The punctured code is a $(7, 16, 3)$ code, which is perfect. (It has the same parameters as a Hamming code, but recall that we are not assuming that it is linear.) This code is distance invariant, and its weight and distance distributions coincide (see [8, Ch. 6]). It can be shown that, in \mathcal{C} , every codeword is at distance eight from exactly one codeword and at distance four from all other codewords.

The function f cannot be unbiased with respect to all sets of cardinality

four, for then $A_{(0,0)}$ would be an $OA_1(4, 8, 2)$, which can be shown not to exist by the Rao bound. Let us suppose without loss of generality that f is biased with respect to $\{1, 2, 3, 4\}$, and therefore unbiased with respect to $\{5, 6, 7, 8\}$. It follows that there are two codewords in \mathcal{C} which agree in the first four coordinates. These codewords must have distance four, so they look like

$$\begin{aligned} & (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) \quad \text{and} \\ & (x_1, x_2, x_3, x_4, 1 - x_5, 1 - x_6, 1 - x_7, 1 - x_8). \end{aligned}$$

Since each codeword is at distance eight from exactly one codeword, we obtain two more codewords:

$$\begin{aligned} & (1 - x_1, 1 - x_2, 1 - x_3, 1 - x_4, 1 - x_5, 1 - x_6, 1 - x_7, 1 - x_8) \quad \text{and} \\ & (1 - x_1, 1 - x_2, 1 - x_3, 1 - x_4, x_5, x_6, x_7, x_8). \end{aligned}$$

However, we have now found codewords that agree in the last four coordinates, which contradicts the fact that f is unbiased with respect to $\{5, 6, 7, 8\}$. We conclude that $d \geq 4$ cannot occur.

4.1.2 Case 2: $d \leq 3$

In this case, there must be two codewords that agree in at least five coordinates. So, without loss of generality, we have two codewords that look like

$$(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) \quad \text{and} \quad (y_1, y_2, y_3, x_4, x_5, x_6, x_7, x_8).$$

It follows that f is biased with respect to any 4-subset of $\{4, 5, 6, 7, 8\}$. Therefore, f is unbiased with respect to the following sets: $\{1, 2, 3, 4\}$, $\{1, 2, 3, 5\}$, $\{1, 2, 3, 6\}$, $\{1, 2, 3, 7\}$ and $\{1, 2, 3, 8\}$. As well, f is unbiased with respect to any set of size three.

We prove that this is impossible by using the same idea that is used to prove the Rao bound. Let $A'_{(0,0)}$ be constructed from $A_{(0,0)}$ by replacing every entry x by $(-1)^x$. Thus $A'_{(0,0)}$ has entries ± 1 . Denote the entry in row i and column j of $A'_{(0,0)}$ by $a_{i,j}$ ($1 \leq i \leq 16, 1 \leq j \leq 8$).

For a set $I \subseteq \{1, \dots, 8\}$, define a vector $\phi(I) = (\phi(I)_1, \dots, \phi(I)_{16}) \in \mathbb{R}^{16}$ by the rule

$$\phi(I)_i = \prod_{j \in I} a_{i,j}.$$

By convention, we will define

$$\phi(\emptyset) = (1, \dots, 1).$$

Let $\langle x, y \rangle$ denote the usual inner product of two vectors $x, y \in \mathbb{R}^{16}$.

Now, it is easy to see that

$$\langle \phi(I), \phi(J) \rangle = 0$$

provided that $I \neq J$ and f is unbiased with respect to $I \cup J$. Let

$$\mathcal{J} = \{I \subseteq \{1, \dots, 8\} : |I| \leq 1\} \cup \{\{1, i\} : 2 \leq i \leq 8\} \cup \{\{2, 3\}\}.$$

By the observation made above, the vectors in \mathcal{J} are mutually orthogonal, and hence they are linearly independent. However, $|\mathcal{J}| = 9 + 7 + 1 = 17$, so we have found 17 linearly independent vectors in \mathbb{R}^{16} , which is a contradiction. We conclude that $d \leq 3$ cannot occur.

Combining the two cases, we obtain the main result of this section.

Theorem 4.1 *There does not exist an $(8, 4, 2)$ -ZF.*

4.2 An Asymptotic Bound

Our last result is an asymptotic lower bound on $n^*(m)$, obtained by appealing to Corollary 3.4 in conjunction with the MRRW bound for orthogonal arrays [2, Theorem 3.5]. The result is the following.

Theorem 4.2

$$\liminf_{m \rightarrow \infty} \frac{n^*(m)}{m} > 3.5277.$$

Theorem 4.2 extends the result proved in the linear case in [7] (this is the result mentioned at the end of Section 2) to the nonlinear case. As in the linear case, the constant 3.5277 is computed as $1/\delta$, where $\delta \approx .28347$ is the solution to the equation

$$\delta = h\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right),$$

where h is the binary entropy function defined as

$$h(x) = -x \log_2 x - (1-x) \log_2(1-x).$$

5 Summary

In [1], the authors ask if there is any parameter situation where there exists a nonlinear zigzag function, but no linear zigzag function exists. Although we have not been able to answer this question, we have generalized several bounds on linear zigzag functions to the general (nonlinear) case.

Acknowledgements

This research was carried out while the author was on sabbatical with the Department of Computer Science at the University of Manitoba. He would like to thank the department for its hospitality.

D. R. Stinson's research is supported by NSERC grants IRC # 216431-96 and RGPIN # 203114-98.

References

- [1] G. Brassard, C. Crépeau and M. Sántha. Oblivious transfers and intersecting codes. *IEEE Transactions on Information Theory* **42** (1996), 1769–1780.
- [2] J. Bierbrauer, K. Gopalakrishnan and D. R. Stinson. A note on the duality of linear programming bounds for orthogonal arrays and codes. *Bulletin of the ICA* **22** (1998), 17–24.
- [3] G. D. Cohen and A. Lempel. Linear intersecting codes. *Discrete Mathematics* **56** (1985), 35–43.
- [4] G. D. Cohen and G. Zemor. Intersecting codes and independent families. *IEEE Transactions on Information Theory* **40** (1994), 1872–1881.
- [5] C. J. Colbourn and J. H. Dinitz, eds. *The CRC Handbook of Combinatorial Designs*, CRC Press, Inc., 1996.
- [6] K. Gopalakrishnan and D. R. Stinson. Three characterizations of non-binary correlation-immune and resilient functions. *Designs, Codes and Cryptography* **5** (1995), 241–251.

- [7] G. Katona and J. Srivastava. Minimal 2-coverings of finite affine spaces based on $GF(2)$. *Journal of Statistical Planning and Inference* **8** (1983), 375–388.
- [8] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*, North-Holland, 1977.
- [9] D. Miklós. Linear binary codes with intersection properties. *Discrete Applied Mathematics* **9** (1984), 187–196.
- [10] N. J. A. Sloane. Covering arrays and intersecting codes. *Journal of Combinatorial Designs* **1** (1993), 51–63.
- [11] D. R. Stinson. Resilient functions and large sets of orthogonal arrays. *Congressus Numerantium* **92** (1993), 105–110.