Arithmetic Correlations and Walsh Transforms

Mark Goresky* Andrew Klapper†

Abstract

In this paper we introduce an *arithmetic Walsh transform*. It is a with-carry analog, based on modular arithmetic, of the usual Walsh transform of Boolean functions. We first develop some tools for analyzing arithmetic Walsh transforms. We then prove that the mapping from a Boolean function to its arithmetic Walsh transform is injective. We then compute the average arithmetic Walsh transforms and the arithmetic Walsh transforms of affine functions.

Keywords: Walsh transform, correlation functions, p-adic numbers.

1 Definitions

A *Boolean function* is a function

$$f : V_n = \mathbb{F}_2^n \to \mathbb{F}_2$$

for some positive integer n. Here $\mathbb{F}_2 = \{0, 1\}$ is the field with 2 elements. We define addition on the set of Boolean functions termwise, $(f + g)(a) = f(a) + g(a)$. The *imbalance* of a Boolean function is the real number $Z(a)$ defined by

$$Z(f) = \sum_{a \in V_n} (-1)^{f(a)}.$$

*School of Mathematics, Institute for Advanced Study

†Dept. of Computer Science, 779A Anderson Hall, University of Kentucky, Lexington, KY, 40506-0046. www.cs.uky.edu/∼klapper. This material is based upon work supported by the National Science Foundation under Grants No. CCF-0514660 and CCF-0914828. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

If $a \in V_n$, then the *shift* of f by a is the real valued function $f_a : V_n \to \mathbb{R}$ defined by

$$f_a(b) = f(a + b).$$

The *cross-correlation* of two Boolean functions f and g is the real valued function $C_{f,g} : V_n \to \mathbb{R}$ defined by

$$C_{f,g}(b) = Z(f + g_b).$$

The *autocorrelation* of f is $A_f(b) = C_{f,f}(b)$. Let $a \cdot b$ denote the inner product of a and b. For any $a \in V_n$, let $T_a(b) = a \cdot b$, $a, b \in V_n$, so that T_a is a linear function. The *Walsh Transform* of f is the real valued function $\widehat{f} : V_n \to \mathbb{R}$ defined by

$$\widehat{f}(a) = Z(f + T_a) = C_{f,T_a}.$$

The Walsh transform plays a central role in the study of the nonlinearity of functions, a study which is central to understanding the cryptographic security of block and stream cipher.

We define an arithmetic analog of the Walsh transform by replacing the termwise sum (which is the same as the difference) of functions by the *with carry* difference. This takes some work since the carries naturally take us outside the domain of the Boolean function. Let $\mathbb{N} = \{0, 1, 2, \cdots\}$ denote the natural numbers. We extend the Boolean function f to $\mathbf{f} : \mathbb{N}^n \to \mathbb{F}_2$ by setting

$$\mathbf{f}(a_1, \cdots, a_n) = f(a_1 \pmod 2, \cdots, a_n \pmod 2).$$

The set P_n of such extensions of Boolean functions is a subset of the set R_n of Boolean valued functions on $\mathbb{N}^n$. It is exactly the set of elements of R_n that are periodic with period 2 in all directions. That is, for every $a, b \in V_n$ we have $\mathbf{f}(a + 2b) = \mathbf{f}(a)$.

In general in this paper we denote Boolean functions by lower case letters and elements of R_n by boldface lowercase letters. The extension of a Boolean function to R_n is denoted by the boldface version of the letter denoting the Boolean function. Vectors in $\mathbb{N}^n$ are denoted by lowercase letters from the beginning of the alphabet. We denote the inner product of two integer vectors a and b by $a \cdot b$. We denote the reduction of an integer x modulo 2 by $[x]_2$. Thus the $\mathbb{F}_2$-inner product of two binary vectors a and b is $[a \cdot b]_2$

We now define an unusual algebraic structure on the set R_n. To understand this definition it is helpful to recall the definition of the 2-adic numbers (in fact R_1 is exactly the 2-adic numbers). A 2-adic number is a formal expression

$$\mathbf{f} = \sum_{i=0}^{\infty} f_i 2^i,$$

where $f_i \in \mathbb{F}_2$. We can identify this 2-adic number with the function on $\mathbb{N}$ that maps i to f_i. We denote the set of 2-adic numbers by $\mathbb{Z}_2$. There is a well defined algebraic structure on the set of 2-adic numbers that makes it a ring. It is based on doing addition and multiplication with carry. Specifically, we say that

$$\sum_{i=0}^{\infty} f_i 2^i + \sum_{i=0}^{\infty} g_i 2^i = \sum_{i=0}^{\infty} h_i 2^i$$

if there are integers $d_0, d_1, d_2, \cdots$ so that $d_0 = 0$ and for all $i \geq 0$ we have $f_i + g_i + d_i = h_i + 2d_{i+1}$. Similarly, we say that

$$\sum_{i=0}^{\infty} f_i 2^i \cdot \sum_{i=0}^{\infty} g_i 2^i = \sum_{i=0}^{\infty} h_i 2^i$$

if there are integers $d_0, d_1, d_2, \cdots$ so that $d_0 = 0$ and for all $i \geq 1$ we have $f_i g_0 + f_{i-1} g_1 + \cdots + h_0 g_i + d_i = h_i + 2d_{i+1}$. The d_is are the carries. The algebra of 2-adic numbers has been studied for more than 100 years [1, 4] and recently the authors and others have used this algebra extensively in the study of fast generation of pseudorandom sequences [2, 3].

We can identify R_1 with the set of 2-adic numbers: a function $\mathbf{f} \in R_1$ is identified with the 2-adic number

$$\sum_{a=0}^{\infty} \mathbf{f}(a) 2^a.$$

To generalize this notion to multiple variables, we want a multiple term analog of the 2-adic number in much the same way that we generalize power series in one variable to power series in several variables. Our new structure will have several "2s". To distinguish them from the ordinary integer 2, we denote them by $t_1, \cdots, t_2$. Then a multi-2-adic number is a formal expression

$$\sum_{a=(a_1,\cdots,a_n)\in\mathbb{N}^n} f_a t_1^{a_1} \cdots t_n^{a_n},$$

with $f_a \in \{0, 1\}$. For convenience we use the following standard notation: if $a = (a_1, \cdots, a_n) \in \mathbb{N}^n$ then $t^a = t_1^{a_1} \cdots t_n^{a_n}$, and if $b = (b_1, \cdots, b_n) \in \mathbb{N}^n$ then $a > b$ means that $a_i > b_i$ for all i $(1 \leq i \leq n)$. Write $\bar{1} = (1, \cdots, 1)$ and similarly for $\bar{0}$, etc.

We can identify a Boolean function $\mathbf{f} \in R_n$ with a multi-2-adic number by setting $f_a = \mathbf{f}(a)$. To think of this geometrically, each lattice point $(a_1, a_2, \cdots, a_n) \in \mathbb{N}^n$ corresponds to a monomial $t_1^{a_1} t_2^{a_2} \cdots t_n^{a_n}$ and the multi-2-adic number $\sum_{a\in\mathbb{N}^n} f_a t^a$ can be identified with the collection of lattice points $a \in \mathbb{N}^n$ such that $f_a = 1$ as in Figure 1 for $n = 2$.

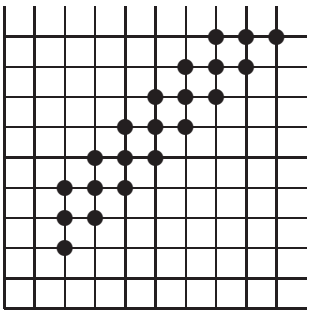

Figure 1: $t_1^2 t_2^2 (1 + t_2 + t_2^2) \sum_{n=0}^{\infty} (t_1 t_2)^n$

When we do arithmetic, we want a coefficient equal to 2 to induce a carry to "the next place in each variable". That is,

$$2t_1^{a_1} t_2^{a_2} \cdots t_n^{a_n} = t_1^{a_1+1} t_2^{a_2+1} \cdots t_n^{a_n+1}.$$

Accordingly we define an addition operation by saying that

$$\sum_{a \in \mathbb{N}^n} f_a t^a + \sum_{a \in \mathbb{N}^n} g_a t^a = \sum_{a \in \mathbb{N}^n} h_a t^a \tag{1}$$

if there exist integers $\{d_a : a \in \mathbb{N}^n\}$ so that $d_a = 0$ if any component of a is zero, and for all $a \in \mathbb{N}^n$, we have

$$f_a + g_a + d_a = h_a + 2d_{a+\overline{1}}.$$

In other words, addition is just 2-adic addition along the diagonals $D_a = \{a+c(1,1,\cdots,1) : c \in \mathbb{N}\}$, where $a \in \mathbb{N}$.

Define a multiplication operation by saying that

$$\sum_{a \in \mathbb{N}^n} f_a t^a \cdot \sum_{a \in \mathbb{N}^n} g_a t^a = \sum_{a \in \mathbb{N}^n} h_a t^a \tag{2}$$

if there exist integers $\{d_a : a \in \mathbb{N}^n\}$ so that $d_a = 0$ if any component of a is zero, and for all a

$$\sum_{b+c=a} f_b g_c + d_a = h_a + 2d_{a+\overline{1}}.$$

(This is not simply multiplication along the diagonals.) With respect to this structure, the number $-1 = -1t^{\overline{0}}$ is represented by the multi-power-series

$$-1 = t^{\overline{0}} + t^{\overline{1}} + t^{\overline{2}} + \cdots$$

as may be seen by adding 1 to both sides of this equation. So the multi 2-adic number represented in Figure 1 is $-t_1^2 t_2^2(1+t_2+t_2^2)$.

Theorem 1.1 *If $\mathbb{Z}[[t_1,\cdots,t_n]]$ is the power series ring in n variables over the integers, then R_n is isomorphic to the quotient ring*

$$S_n = \mathbb{Z}[[t_1,\cdots,t_n]]/(t_1 t_2 \cdots t_n - 2).$$

Proof: As described above, to each lattice point $a=(a_1,\cdots,a_n)\in\mathbb{N}^n$ we associate the monomial $t^a = t_1^{a_1}t_2^{a_2}\cdots t_n^{a_n}$ so that an integer valued function $\mathbf{f}:\mathbb{N}^n\to\mathbb{Z}$ becomes identified with the multi power series

$$\sum_{a\in\mathbb{N}^n}\mathbf{f}(a)t^a = \sum_{(a_1,\cdots,a_n)\in\mathbb{N}^n}\mathbf{f}(a_1,\cdots,a_n)t_1^{a_1}\cdots t_n^{a_n}\in\mathbb{Z}[[t_1,\cdots,t_n]].$$

Then R_n is the set of multi power series with coefficients in $\{0,1\}$. Let us check that the resulting mapping $\phi: R_n\to S_n=\mathbb{Z}[[t_1,\cdots,t_n]]/(t_1\cdots t_n-2)$ is a homomorphism. From equation (1) we have:

$$\begin{aligned}\sum_{a\geq 0}f_a t^a+\sum_{a\geq 0}g_a t^a-\sum_{a\geq 0}h_a t^a &= \sum_{a\geq 0}2d_{a+\bar{1}}t^a-\sum_{a\geq\bar{1}}d_a t^a\\ &= \sum_{a\geq 0}2d_{a+\bar{1}}t^a-\sum_{b\geq\bar{0}}d_{b+\bar{1}}t^{b+\bar{1}}\\ &= \left(2-t^{\bar{1}}\right)\sum_{a\geq 0}d_{a+\bar{1}}t^a\end{aligned}$$

which is in the ideal $\left(t^{\bar{1}}-2\right)$. A similar calculation applies to the multiplication law.

The mapping $\phi: R_n\to\mathbb{Z}[[t_1,\cdots,t_n]]/(t_1\cdots t_n-2)$ is surjective, for the following reason. First observe in S_n that

$$-1=\sum_{i=0}^{\infty}t^{\bar{i}}\in S_n,$$

so $\phi(-1)=-1$. If k is a nonnegative integer, write $k=k_0+k_1 2+\cdots+k_r 2^r$ for its binary expansion, with $k_i\in\{0,1\}$. Then, in the ring S_n we have an equality

$$k=k_0+k_1 t^{\bar{1}}+\cdots+k_r t^{\bar{r}}\in R_n.$$

We conclude that $\mathbb{Z}$ is in the image of ϕ. Each monomial is also in the image of ϕ, and since ϕ is a homomorphism it follows that it is surjective.

Now let us prove that ϕ is injective. By a *minimal degree nonzero term* of $\mathbf{f}$ we mean an element $a \in \mathbb{N}^n$ with $\mathbf{f}(a) \neq 0$ and $\mathbf{f}(b) = 0$ whenever $b \neq a$ and $b_i \leq a_i$ for all i. It suffices to prove the following: if $\alpha = \sum \mathbf{f}(a)t^a \in (t_1 \cdots t_n - 2)$ and if $a \in \mathbb{N}^n$ is a minimal degree nonzero term of $\mathbf{f}$ then $\mathbf{f}(a)$ is even. For, suppose $\alpha = \beta.(t^{\overline{1}} - 2)$ for some $\beta = \sum_a \mathbf{g}(a)t^a$. Then

$$\begin{aligned} \alpha &= \sum_a \mathbf{g}(a)t^{a+\overline{1}} - 2\sum_a \mathbf{g}(a)t^a \\ &= \sum_a t^a \left(\mathbf{g}(a-\overline{1}) - 2\mathbf{g}(a)\right) \end{aligned}$$

where we have written $\mathbf{g}(a_1, \cdots, a_n) = 0$ if $a_i < 0$ for any index i. In other words, $\mathbf{f}(a) = \mathbf{g}(a - \overline{1}) - 2\mathbf{g}(a)$. It follows that

$$\mathbf{g}(a-\overline{1}) = \frac{1}{2}\mathbf{g}(a-\overline{2}) = \frac{1}{4}\mathbf{g}(a-\overline{3}) = \cdots$$

which is zero. Therefore $\mathbf{f}(a) = -2\mathbf{g}(a)$ is even, and is non-zero by assumption. □

Corollary 1.2 *The addition and multiplication operations defined above make R_n into a commutative ring. The zero (additive identity) is the element $z \in R_n$ with $z_a = 0$ for all a, and the one (multiplicative identity) is the element $e \in R_n$ with $e_{0^n} = 1$ and $e_a = 0$ if $a \neq 0^n$.*

It follows from the proof of Theorem 1.1 that an element $\mathbf{f} \in R_n$ can be represented either as a function $\mathbf{f}$ from $\mathbb{N}$ to $\{0, 1\}$ (i.e., as an element of $\mathbb{Z}[[t_1, \cdots, t_n]]$ with coefficients in $\{0, 1\}$), or as a function $\bar{f}$ from $\{a = (a_1, \cdots, a_n) : a_1, \cdots, a_n \in \mathbb{N}$ and at least one $a_i = 0\}$ to $\mathbb{Z}_2$. These representations are connected by the formula

$$\bar{f}(a) = \sum_{i=0}^{\infty} \mathbf{f}(a + i \cdot 1^n)2^i. \tag{3}$$

We refer to $\bar{f}(a)$ as the *restriction* of f to the diagonal D_a. The same notation and terminology will be used even if a does not have a zero component.

It is important to note that the set P_n of elements of R_n that have period 2 in all directions is not a subring of R_n. In fact the sum and difference of elements of P_n may not be in P_n. However, since addition is just 2-adic addition on each diagonal, and the sum of two periodic 2-adic integers is eventually periodic (i.e., periodic beyond some point), the sum and difference of two elements of P_n are ultimately periodic along each

diagonal. Moreover, the set restrictions to diagonals are periodic (that is, the restriction of an element $\mathbf{f} \in P_n$ to a diagonal D_a is the same as the restriction of $\mathbf{f}$ to D_{a+2b} for any $b \in V_n$). Thus if $\mathbf{f}, \mathbf{g} \in P_n$, then $\mathbf{f}+\mathbf{g}$ and $\mathbf{f}-\mathbf{g}$ (where the sum and difference of $\mathbf{f}$ and $\mathbf{g}$ are in the ring R_n) are eventually 2-periodic in the following sense.

Definition 1.3 *The element $\mathbf{f} \in R_n$ is eventually p-periodic if there is an integer k so that if $a = (a_1, \cdots, a_n) \in V_n$, and $a_i \geq k$ for $i = 1, \cdots, n$, then for every $b \in V_n$, $\mathbf{f}(a + pb) = \mathbf{f}(a)$. If $a = (a_1, \cdots, a_n) \in V_n$, and $a_i \geq k$ for $i = 1, \cdots, n$, then the restriction of $\mathbf{f}$ to the set $\{a + b : b = (b_1, \cdots, b_n), 0 \leq b_i < p, i = 1, \cdots, n\}$ is called a* complete period *of $\mathbf{f}$.*

It is possible to be more explicit. Let $\mathbf{f} = \sum_{i=0}^{\infty} f_i 2^i$ and $\mathbf{g} = \sum_{i=0}^{\infty} g_i 2^i$ be a pair of 2-adic integers whose coefficient sequences have period 2. Then the coefficient sequences of $-\mathbf{f}$, $\mathbf{f}+\mathbf{g}$, and $\mathbf{f}-\mathbf{g}$ are periodic from the coefficients with index 2 on. Tables 1, 2, and 3 show the first four coefficients of the negation of $\mathbf{f}$ and the sum and difference $\mathbf{f}+\mathbf{g}$ and $\mathbf{f}-\mathbf{g}$ for all possible combinations of periodic 2-adic integers with period 2.

$\mathbf{f}$	$-\mathbf{f}$
00	0000
01	0111
10	1111
11	1011

Table 1: Negation of a 2-periodic 2-adic integer.

$\mathbf{f}+\mathbf{g}$	00	01	10	11
00	0000	0101	1010	1111
01	0101	0010	1111	1001
10	1010	1111	0101	0010
11	1111	1001	0010	0111

Table 2: Sum of 2-periodic 2-adic integers.

The possibilities for $\mathbf{f}$ and $\mathbf{g}$ are given by the first two coefficients of each. In the second and third tables, the various $\mathbf{f}$s are listed down the left hand side and the various $\mathbf{g}$s are listed across the top.

$\mathbf{f}-\mathbf{g}$	00	01	10	11
00	0000	0111	1111	1011
01	0101	0000	1010	1101
10	1010	1101	0000	0110
11	1111	1010	0101	0000

Table 3: Difference of 2-periodic 2-adic integers.

Let us return to the case of R_n, and suppose that $\mathbf{f} : \mathbb{N} \to \{0, 1\}$ is strictly 2-periodic. Then in the representation in equation (3) we have

$$\begin{aligned} \bar{f}(a) &= \sum_{i=0}^{\infty} \mathbf{f}(a + i \cdot 1^n) 2^i \\ &= f(a) + f(a + 1^n)2 + f(a)2^2 + f(a + 1^n)2^3 + \cdots \\ &= -\frac{f(a) + 2f(a + 1^n)}{3}. \end{aligned} \tag{4}$$

2 Arithmetic Correlations and Walsh Transforms

Now we can define the arithmetic correlations and Walsh transforms. First note that when defining classical correlation functions and Walsh transforms, we can do just as well to use differences of functions rather than sums of functions. The result is the same in the Boolean case, but when these concepts are generalized to N-ary functions, it becomes apparent that differences are needed. This is the point of view we use here. First we extend the notion of imbalance to eventually 2-periodic elements.

Definition 2.1 *Let $\mathbf{f} \in R_n$ be eventually p-periodic. Then the* imbalance *of $\mathbf{f}$ is*

$$Z(\mathbf{f}) = \sum_a (-1)^{\mathbf{f}(a)},$$

where the sum is extended over one complete period of $\mathbf{f}$.

Note that $Z(\mathbf{f})$ is independent of the choice of complete period. Also, this definition is consistent with the definition of the imbalance of Boolean functions in the sense that the imbalance of a Boolean function equals its the imbalance of its periodic extension to $\mathbb{N}^n$.

Definition 2.2 *The* arithmetic cross-correlation *of two eventually periodic functions* $\mathbf{f}$ *and* $\mathbf{g}$ *in* R_n *is the real valued function* $C^a_{\mathbf{f},\mathbf{g}} : V_n \to \mathbb{R}$ *defined by*

$$C^a_{\mathbf{f},\mathbf{g}}(a) = Z(\mathbf{f} - \mathbf{g}_a).$$

If f *and* g *are two Boolean functions on* V_n*, then the* arithmetic cross-correlation *of* f *and* g *is the arithmetic cross-correlation of their extensions* $\mathbf{f}$ *and* $\mathbf{g}$*,*

$$C^a_{f,g}(a) = C^a_{\mathbf{f},\mathbf{g}}(a).$$

The arithmetic autocorrelation *of* f *is*

$$A^a_f(b) = C^a_{f,g}(b).$$

In defining $\mathbf{g}_a$ it doesn't matter whether we translate by a and then extend to $\mathbb{N}^n$ or extend to $\mathbb{N}^n$ and then translate by a. A linear function is a Boolean function T_a, $a \in V_n$, where $T_a(b) = [a \cdot b]_2$. Thus the extension $\mathbf{T}_a$ is also defined by $\mathbf{T}_a(b) = [a \cdot b]_2$ for $b \in \mathbb{N}^n$.

Definition 2.3 *The* arithmetic Walsh Transform *of an eventually periodic* $\mathbf{f} \in R_n$ *is the real valued function* $\widetilde{\mathbf{f}} : V_n \to \mathbb{R}$ *defined by*

$$\widetilde{\mathbf{f}}(a) = Z(\mathbf{f} - \mathbf{T}_a).$$

If f *is a Boolean function on* V_n*, then the* arithmetic Walsh Transform *of* f *is the arithmetic Walsh Transform of the extension* $\mathbf{f}$ *of* f*,* $\widetilde{f}(a) = \widetilde{\mathbf{f}}(a)$*.*

We want to use the representation in equations (3) and (4) to compute correlations. Let

$$U_n = \{a = (a_1, \cdots, a_n) : a_i \in \{0, 1\} \text{ and } a_1 = 0\}.$$

The restriction of an eventually periodic function $\mathbf{f} \in R_n$ to a diagonal D_a with $a \in U_n$ is eventually periodic. If we select one full period from each of these diagonals, altogether we will have one complete period of $\mathbf{f}$. It follows that the imbalance of $\mathbf{f}$ is the sum of the imbalances of the restrictions of $\mathbf{f}$ to the diagonals. The imbalance of the restriction of $\mathbf{f}$ to diagonal D_a in turn is the imbalance of the 2-adic number $\bar{f}(a)$ (defined in equation (3)). This then is the imbalance of the 2-adic representation of the rational number in equation (4). Thus

$$Z(f) = \sum_{a \in U_n} Z(\bar{f}(a)). \tag{5}$$

Theorem 2.4 *Let $f : V_n \to \mathbb{F}_2$ be a Boolean function. If $b \cdot 1^n = 0$, then*

$$\begin{aligned}
\widetilde{f}(b) &= \sum_{a \in U_n} 2(1 - f(a) - f(a + 1^n) + 2f(a)f(a + 1^n)[a \cdot b]_2) \\
&= 2^n - 2\sum_{a \in V_n} f(a) + 4\sum_{a \in U_n} f(a)f(a + 1^n)[a \cdot b]_2 \qquad (6) \\
&= 2^n - 2\sum_{a \in V_n} f(a) + 2\sum_{a \in V_n} f(a)f(a + 1^n)[a \cdot b]_2 \qquad (7)
\end{aligned}$$

If $b \cdot 1^n = 1$, then

$$\begin{aligned}
\widetilde{f}(b) &= 2\sum_{a \in U_n} (f(a + 1^n) - f(a)f(a + 1^n) + (f(a) - f(a + 1^n))[a \cdot b]_2) \qquad (8) \\
&= \sum_{a \in V_n} (f(a + 1^n) - f(a)f(a + 1^n) + (f(a) - f(a + 1^n))[a \cdot b]_2). \qquad (9)
\end{aligned}$$

Proof: If $b \cdot 1^n = 1 \pmod 2$, then

$$[(a + 1^n) \cdot b]_2 = [a \cdot b]_2 + 1 \pmod 2 = 1 - [a \cdot b]_2.$$

It then follows from the discussion above that that

$$\begin{aligned}
\widetilde{f}(b) &= \sum_{a \in U_n} Z((\bar{f} - \bar{T}_b)(a)) \\
&= \sum_{a \in U_n} Z\left(-\frac{f(a) + 2f(a + 1^n) - [a \cdot b]_2 - 2[(a + 1^n) \cdot b]_2}{3}\right) \\
&= \begin{cases} \displaystyle\sum_{a \in U_n} Z\left(-\frac{f(a) + 2f(a + 1^n)}{3} + [a \cdot b]_2\right) & \text{if } b \cdot 1^n = 0 \\ \displaystyle\sum_{a \in U_n} Z\left(-\frac{f(a) + 2f(a + 1^n) + [a \cdot b]_2 - 2}{3}\right) & \text{if } b \cdot 1^n = 1. \end{cases}
\end{aligned}$$

The 2-adic expansion of $u/3$ is eventually periodic with period 2 and each period equals 10 or 01 unless u is a multiple of 3. In these cases the imbalance is always 0. If u is a multiple of 3, then the eventual period is 1 and each period is either 1 (if u is negative) or 0 (if u is nonnegative). The imbalance is thus -2 if u is negative and is 2 if u is nonnegative. Let $Z_a = Z((\bar{f} - \bar{T}_b)(a))$.

For $b \cdot 1^n = 0$, we have the following table of values:

$f(a)$	$f(a+1^n)$	$[a \cdot b]_2$	Z_a
0	0	0	2
1	0	0	0
0	1	0	0
1	1	0	-2
0	0	1	2
1	0	1	0
0	1	1	0
1	1	1	2

Using Lagrange interpolation we find that

$$Z_a = 2(1 - f(a) - f(a+1^n) + 2f(a)f(a+1^n)[a \cdot b]_2).$$

For $b \cdot 1^n = 1$, we have the following table of values:

$f(a)$	$f(a+1^n)$	$[a \cdot b]_2$	Z_a
0	0	0	0
1	0	0	0
0	1	0	2
1	1	0	0
0	0	1	0
1	0	1	2
0	1	1	0
1	1	1	0

Using Lagrange interpolation we find that

$$Z_a = 2(f(a+1^n) - f(a)f(a+1^n) + (f(a) - f(a+1^n))[a \cdot b]_2).$$

This definition makes sense for $a \notin U_n$ as well. It can be checked that $Z_a = Z_{a+1^n}$. Thus the last equality holds. This proves the theorem. □

Corollary 2.5 *If f is a Boolean function on V_n, and $b \cdot 1^n = 0$, then $\widetilde{f}(b)$ is even.*

Corollary 2.6 *Let $L : V_n \to V_n$ be a nonsingular linear transformation such that $L(1^n) = 1^n$. Let f be a Boolean function on V_n. Then the set of arithmetic Walsh coefficients of f is invariant under composition with L. That is,*

$$\{\widetilde{f}(b) : b \in V_n\} = \{\widetilde{(f \circ L)}(b) : b \in V_n\}.$$

Proof: Note that if $a \in V_n$, then $L(a + 1^n) = L(a) + L(1^n) = L(a) + 1^n$. Let M denote the representation of L as a matrix using the standard basis of V_n. Thus $L(a) = aM$. For any matrix N let N^t denote the transpose of N. Then for any $a, b \in V_n$, we have $a \cdot b = ab^t$. We claim that for any $b \in V_n$, $\widetilde{(f \circ L)}(b) = \widetilde{f}(b(M^{-1})^t)$.

Suppose that $[b \cdot 1^n]_2 = 0$. Then by equation (7),

$$\begin{aligned}
\widetilde{(f \circ L)}(b) &= 2^n - 2\sum_{a \in V_n} f(L(a)) + 2\sum_{a \in V_n} f(L(a))f(L(a + 1^n))[a \cdot b]_2 \\
&= 2^n - 2\sum_{a \in V_n} f(L(a)) + 2\sum_{a \in V_n} f(L(a))f(L(a) + 1^n)[ab^t]_2 \\
&= 2^n - 2\sum_{a \in V_n} f(L(a)) + 2\sum_{a \in V_n} f(L(a))f(L(a) + 1^n)[aMM^{-1}b^t]_2 \\
&= 2^n - 2\sum_{a \in V_n} f(L(a)) + 2\sum_{a \in V_n} f(L(a))f(L(a) + 1^n)[L(a) \cdot b(M^{-1})^t]_2 \\
&= 2^n - 2\sum_{a \in V_n} f(a) + 2\sum_{a \in V_n} f(a)f(a + 1^n)[a \cdot b(M^{-1})^t]_2 \\
&= \widetilde{f}(b(M^{-1})^t),
\end{aligned}$$

where the penultimate line holds because L is a permutation. Moreover,

$$b(M^{-1})^t \cdot 1^n = b(M^{-1})^t(1^n)^t = b(1^n M)^t = b(1^n)^t = b \cdot 1^n,$$

so this is the correct expression for $\widetilde{f}(b(M^{-1})^t)$. This proves the claim in this case. A similar argument works when $[b \cdot 1^n]_2 = 1$. □

Let us return for a moment to the classical case of Boolean functions and Walsh transforms. If f and g are Boolean functions, then the distance between f and g is

$$\delta(f, g) = |\{a \in V_n : f(a) \neq g(a)\}|.$$

This is a true distance measure. It is well-known that $Z(f - g) = 2^n - \delta(f, g)$. In particular, $Z(f - g) = 2^n$ if and only if $\delta(f, g) = 0$ if and only if $f = g$. Also, $Z(f - g) = -2^n$ if and only if f is the complement of g. Thus f has a Walsh coefficient equal to 2^n if and only if f is linear, and has a Walsh coefficient equal to -2^n if and only if f is affine and nonlinear.

Now we return to the arithmetic case. Let f and g be Boolean functions and let $\mathbf{f}$ and $\mathbf{g}$ be their extensions. Suppose that $Z(\mathbf{f} - \mathbf{g}) = 2^n$. From equation (5) it follows

that for every $a \in U_n$, $Z(\bar{f}(a) - \bar{g}(a)) = 2$. That is,

$$Z\left(\frac{g(a) - f(a) + 2(g(a+1^n) - f(a+1^n))}{3}\right) = 2.$$

This holds if and only if either (1) $f(a) = g(a)$ and $f(a+1^n) = g(a+1^n)$ or (2) $g(a) = g(a+1^n) = 1$ and $f(a) = f(a+1^n) = 0$. Thus g is obtained from f by choosing some elements $X \subseteq U_n$ so that f is 0 on the diagonal determined by each $a \in X$ and changing the value on these diagonals to 1. Alternatively, f is obtained from g by choosing some elements $Y \subseteq U_n$ so that g is 1 on the diagonal determined by each $a \in Y$ and changing the value on these diagonals to 0.

Now suppose g is a linear function, say $g(a) = [a \cdot b]_2$, $b \neq 0^n$. The function g is constant on some diagonal if and only if $g(1^n) = 0$. In this case g is 1 on exactly 2^{n-2} diagonals, so there are $2^{2^{n-2}} - 1$ nonlinear functions f so that $\widetilde{f}(b) = 2^n$.

3 Computing Arithmetic Correlations

Let f be a Boolean function. In this section we use equations (3) and (4) to compute the arithmetic correlations of f. Surprisingly, we see that all arithmetic autocorrelations are nonnegative.

As before, we let $U_n = \{a = (a_1, a_2, \cdots, a_n) \in V_n : a_1 = 0\}$.

Suppose first that $b \in U_n$. Then $a + b \in U_n$ if and only if $a \in U_n$. If $b = 0^n$, then $A_f^a(b) = 2^n$. Now assume that $b \neq 0^n$. Using arguments similar to those in Section 2, the arithmetic autocorrelation of f with shift $a \in V_n$ is

$$A_f^a(b) = \sum_{a \in U_n} Z\left(\frac{f(a+b) - f(a) + (f(a+b+1^n) - f(a+1^n))2}{3}\right). \tag{10}$$

Then for any $a \in U_n$, both terms in

$$\begin{aligned} Z_a \;=\; & Z\left(\frac{f(a+b) - f(a) + (f(a+b+1^n) - f(a+1^n))2}{3}\right) \\ & + Z\left(\frac{f(a) - f(a+b) + (f(a+1^n) - f(a+b+1^n))2}{3}\right) \end{aligned} \tag{11}$$

appear in equation (10). The sum depends on $f(a)$, $f(a+1^n)$, $f(a+b)$, and $f(a+b+1^n)$, and no other terms in equation (10) depend on these values. We want to determine Z_a in terms of these four values.

The numerators of the two terms are negatives of each other, so one numerator is divisible by three if and only if the other is. If neither is a multiple of three, then both imbalances are zero. If either numerator is positive, then the other is negative so the imbalances are negatives of each other. Thus the only nonzero contribution to $A_f^a(b)$ is from those as for which $f(a+b) - f(a) + (f(a+b+1^n) - f(a+1^n))2 = 0$. This happens exactly when $f(a) = f(a+b)$ and $f(a+1^n) - f(a+b+1^n)$, and the two imbalances add to 4. We account for each term once if we sum just over all $a < a+b$ (say in lexicographic order). Thus

$$\begin{aligned} A_f^a(b) &= 4|\{a \in U_n : a < a+b, f(a) = f(a+b), \text{ and } f(a+1^n) = f(a+b+1^n)\}| \\ &= 2|\{a \in U_n : f(a) = f(a+b), \text{ and } f(a+1^n) = f(a+b+1^n)\}| \\ &= |\{a \in V_n : f(a) = f(a+b), \text{ and } f(a+1^n) = f(a+b+1^n)\}|. \end{aligned} \tag{12}$$

This expression is also correct when $b = 0^n$.

Now suppose that $b \in V_n - U_n$. If $b = 1^n$, then $a + b = a + 1^n$ and $a + b + 1^n = a$. Thus the contribution from the term corresponding to any $a \in U_n$ is

$$\begin{aligned} Z\left(\frac{f(a+1^n) - f(a) + (f(a) - f(a+1^n))2}{3}\right) &= Z\left(\frac{f(a) - f(a+1^n)}{3}\right) \\ &= \begin{cases} 2 & \text{if } f(a) = f(a+1^n) \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Thus

$$\begin{aligned} A_f^a(b) &= 2|\{a \in U_n : f(a) = f(a+1^n)\}| \\ &= |\{a \in V_n : f(a) = f(a+1^n)\}|. \end{aligned}$$

Equation (12) agrees with this value when $b = 1^n$.

Lastly, let $b \in V_n - U_n$ and $b \neq 1^n$. Then for any $a \in U_n$, both terms in

$$\begin{aligned} Z_a &= Z\left(\frac{f(a+b) - f(a) + (f(a+b+1^n) - f(a+1^n))2}{3}\right) \\ &\quad + Z\left(\frac{f(a+1^n) - f(a+b+1^n) + (f(a) - f(a+b))2}{3}\right) \end{aligned} \tag{13}$$

appear in equation (10). The sum depends on $f(a)$, $f(a+1^n)$, $f(a+b)$, and $f(a+b+1^n)$, and no other terms depend on these values. Let u and v denote the numerators of the two terms. Then $-2v = u + 3(f(a) - f(a+b))$. Thus v is divisible by 3 if and only if u is divisible by 3. If u and v are not divisible by 3 then both terms contribute 0 to the

autocorrelation. Suppose u and v are divisible by 3. Not that $f(a)-f(a+b) \in \{-1,0,1\}$. We have $u=0$ if and only if $f(a)=f(a+b)$ and $f(a+1^n)=f(a+b+1^n)$, and this holds if and only if $v=0$. If $u>0$ then $u \geq 3$ and so $-2v = u+3(f(a)-f(a+1^n)) \geq 0$. Thus $v \leq 0$. But $u>0$ implies $v \neq 0$, so $v<0$. Conversely, if $v<0$ then $v \leq -3$ so $6 \leq -2v = u+3(f(a)-f(a+b))$ which implies $u \geq 3 > 0$. Thus u and v have opposite signs. It follows that the two terms cancel unless $f(a)=f(a+b)$ and $f(a+1^n)=f(a+b+1^n)$. We obtain the same expression for $A_f^a(b)$. We summarize this analysis in the following theorem

Theorem 3.1 *If f is any Boolean function on n variables and $b \in V_n$, then*

$$A_f^a(b) = |\{a \in V_n : f(a)=f(a+b), \text{ and } f(a+1^n)=f(a+b+1^n)\}|.$$

In particular, $A_f^a(b) \geq 0$.

Now let g be a second Boolean function. By similar reasoning, the arithmetic cross-correlation of f with shift $a \in V_n$ is

$$C_{f,g}^a(b) = \sum_{a \in U_n} Z\left(\frac{g(a+b)-f(a)+2(g(a+b+1^n)-f(a+1^n))}{3}\right).$$

Then for any $a \in U_n$, the term

$$Z_a = Z\left(\frac{g(a+b)-f(a)+2(g(a+b+1^n)-f(a+1^n))}{3}\right) \tag{14}$$

depends on $f(a)$, $f(a+1^n)$, $g(a+b)$, and $g(a+b+1^n)$, and no other terms depend on these values. We want to determine equation (14 in terms of these four values.

The fraction in equation (14) has an eventually balanced 2-adic expansion if and only if the numerator is not a multiple of 3. If the numerator is an odd multiple of 3, then the expansion is eventually all 1s, so the imbalance is -2. If the numerator is 0 or a positive multiple of 3, then the expansion is eventually all 0s, so the imbalance is 2. This gives the following theorem.

Theorem 3.2 *Let f and g be Boolean functions on n variables and let $b \in V_n$. Then*

$$\begin{aligned} C_{f,g}^a(b) \;=\; & |\{a \in V_n : g(a+b)=f(a) \text{ and } g(a+b+1^n)=f(a+1^n)\}| \\ & +|\{a \in V_n : g(a+b)=g(a+b+1^n)=1 \text{ and } f(a)=f(a+1^n)=0\}| \\ & -|\{a \in V_n : g(a+b)=g(a+b+1^n)=0 \text{ and } f(a)=f(a+1^n)=1\}|. \end{aligned} \tag{15}$$

This implies that if $f(a) = 0$ for all a and $g(a) = 1$ for all a, then $C^a_{f,g}(b) = 2^n$ for all b and that $C^a_{g,f}(b) = -2^n$ for all b.

If f is a Boolean function, let f' denote the complement of f. That is, $f'(a) = 1$ if and only if $f(a) = 0$, so $f'(a) = 1 - f(a)$ as integers.

Corollary 3.3 *Let f and g be Boolean functions on n variables and let $b \in V_n$. Then for every $b \in V_n$*

$$C^a_{f,g}(b) = C^a_{g',f'}(b).$$

3.1 Arithmetic Correlations of Linear and Affine Functions

In this section we use Theorems 3.1 and 3.2 to compute the arithmetic auto- and cross-correlations of linear and affine functions.

First consider autocorrelations. If f is constant (identically 0 or identically 1), then $A^a_f(b) = 2^n$ for all b. If f is nonzero and linear, then $f(a) = f(a + b)$ and $f(a + 1^n) = f(a + b + 1^n)$ if and only if $f(b) = 0$. Similarly, if f is affine but not linear, then $f(x) = 1 - h(x)$ with h linear, and these equations hold if and only if $h(b) = 0$. Thus in either case

$$A^a_f(b) = \begin{cases} 2^n & \text{if } f(b) = 0 \\ 0 & \text{otherwise.} \end{cases}$$

We can compare this to the classical autocorrelations, where if f is affine, then

$$A_f(b) = \begin{cases} 2^n & \text{if } f(b) = 0 \\ -2^n & \text{otherwise.} \end{cases}$$

Now let us consider the cross-correlation. Let f and g be linear or affine. Then the sets in the last two terms of equation (15) are solutions to inhomogeneous systems of degree 1 equations. In the two sets the homogeneous parts of the equations are the same. It is only the constant terms that differ. It follows that the numbers of solutions is the same for both systems, depending only on the rank of the homogeneous part. Thus for linear and affine functions $C^a_{f,g}(b)$ is the number of $a \in V_n$ such that

$$g(a + b) = f(a) \tag{16}$$

and

$$g(a + b + 1^n) = f(a + 1^n). \tag{17}$$

Let $f(a) = f_1(a) + c \pmod 2$ and $g(a) = g_1(a) + d \pmod 2$, where f_1 and g_1 are linear and $c, d \in \{0, 1\}$. Then equations (16) and (17) hold if and only equation (16) and

$$g_1(1^n) = f_1(1^n) \tag{18}$$

hold. Thus if equation (18) does not hold, then $C^a_{f,g}(b) = 0$. Otherwise $C^a_{f,g}(b)$ is the number of $a \in V_n$ such that equation (16) holds.

Suppose equation (18) holds. If $f_1 \neq g_1$, then equation (16) is a rank one affine equation, so it holds for 2^{n-1} values of a. If f and g are the same constant, then $C^a_{f,g}(b) = 2^n$. If they are different constants, then $C^a_{f,g}(b) = 0$. If $f_1 = g_1 \neq 0$, then equation (16) holds if and only if $g(b) = f(0^n)$. That is, if and only if $g_1(b) = c - d \pmod 2$. This occurs for 2^{n-1} values of b.

Theorem 3.4 *Let $f(a) = f_1(a) + c \pmod 2$ and $g(a) = g_1(a) + d \pmod 2$, where f_1 and g_1 are linear and $c, d \in \{0, 1\}$. If $g_1(1^n) \neq f_1(1^n)$ then $C^a_{f,g}(b) = 0$ for all b. If $f_1 = g_1 = 0$ and $c = d$, then $C^a_{f,g}(b) = 2^n$ for all b. If $f_1 = g_1 = 0$ and $c \neq d$, then $C^a_{f,g}(b) = 0$. If $f_1 = g_1 \neq 0$, then $C^a_{f,g}(b) = 2^n$ for 2^{n-1} values of b and is 0 for 2^{n-1} values of b. If $f_1 \neq g_1$, then $C^a_{f,g}(b) = 2^{n-1}$ for all b.*

By contrast, the classical cross-correlation is 0 if $f_1 \neq g_1$. If $f_1 = g_1 = 0$, then it is 2^n for all b or -2^n for all b. If $f_1 = g_1 \neq 0$, then it is 2^n for 2^{n-1} values of b and is -2^n for 2^{n-1} values of b.

4 Uniqueness of Arithmetic Walsh Spectra

The arithmetic Walsh spectrum of a Boolean function is the set of its arithmetic Walsh coefficients. In this section we show that the mapping from Boolean functions to their arithmetic Walsh spectra is one to one. That is, we show that a Boolean function is uniquely determined by its arithmetic Walsh spectrum. We do not, however, know a simple expression for the inverse arithmetic Walsh transform, or even an efficient way to compute it.

It follows from equation (6) that if $b \neq 0^n$ and $\mathrm{wt}(b)$ is even, then

$$\sum_{a \in U_n} f(a) f(a + 1^n)[a \cdot b]_2 = \frac{\widetilde{f}(b) + \widetilde{f}(0^n)}{4}. \tag{19}$$

Let M_n be the $(2^{n-1} - 1) \times (2^{n-1} - 1)$ rational matrix indexed by $U_n - \{0^n\}$ and $W_n = \{b \in V_n : \mathrm{wt}(b) \text{ even}, b \neq 0^n\}$ whose entry with index (a, b) is $[a \cdot b]_2$ treated as a rational number. Similarly, let N_n be the $(2^{n-1} - 1) \times (2^{n-1} - 1)$ rational matrix indexed by $U_n - \{0^n\}$ and $T_n = \{b \in V_n : \mathrm{wt}(b) \text{ odd}, b \neq 10^{n-1}\}$ whose entry with index (a, b) is $[a \cdot b]_2$ treated as a rational number.

Let $v(a) = f(a)f(a + 1^n)$ and let v be the vector indexed by $U_n - \{0^n\}$ whose entries are the $v(a)$. Let $z(b) = (\widetilde{f}(b) + \widetilde{f}(0^n))/4$ and let z be the vector indexed by W_n whose

entries are the $z(b)$. Then equation (19) implies that $vM_n = z$. Thus if M_n is invertible, then the $v(a)$ with $a \neq 0^n$ or 1^n can be determined uniquely from the $\widetilde{f}(b)$.

Similarly, it follows from equation (8) that if $b \neq 10^{n-1}$ and $\mathrm{wt}(b)$ is odd, then

$$\sum_{a \in U_n} (f(a) - f(a+1^n))[a \cdot b]_2 = \frac{\widetilde{f}(b) - \widetilde{f}(10^{n-1})}{2}. \tag{20}$$

Let $u(a) = f(a) - f(a+1^n)$ and let u be the vector indexed by $U_n - \{0^n\}$ whose entries are the $u(a)$. Let $w(b) = (\widetilde{f}(b) - \widetilde{f}(10^{n-1}))/2$ and let w be the vector indexed by T_n whose entries are the $w(b)$. Then equation (20) implies that $uN_n = w$. Thus if N_n is invertible, then the $u(a)$ with $a \neq 10^{n-1}$ can be determined uniquely from the $\widetilde{f}(b)$.

Theorem 4.1 *The matrices M_n and N_n have nonzero determinants.*

Proof: We order the indices in both dimensions lexicographically, with most significant position on the right. For both types of matrices, we think of the rows (the as) as being divided into three segments:

1. The rows indexed by $a = 0a'0$ with $a' \neq 0^{n-2}$;
2. The row indexed by $a = 0^{n-1}1$; and
3. The rows indexed by $a = 0a'1$ with $a' \neq 0^{n-2}$.

For M_n, we think of the columns (the bs) as being divided into three segments:

1. The columns indexed by $b = b'0$ with $\mathrm{wt}(b')$ even and $b' \neq 0^{n-1}$;
2. The column indexed by $b = 10^{n-2}1$; and
3. The columns indexed by $b = b'1$ with $\mathrm{wt}(b')$ odd and $b' \neq 10^{n-2}$.

Similarly, for N_n, we think of the columns as being divided into three segments:

1. The columns indexed by $b = b'0$ with $\mathrm{wt}(b')$ odd;
2. The column indexed by $b = 0^{n-1}1$; and
3. The columns indexed by $b = b'1$ with $\mathrm{wt}(b')$ even and $b' \neq 0^{n-1}$.

Let $\langle 1 \rangle_n$ denote the $2^n \times 2^n$ matrix all of whose entries are 1. Following these decompositions of the indices, we can decompose M_n into blocks as follows.

1. If $a = 0a'0$ with $a' \neq 0^{n-2}$ and $b = b'0$ with $b' \neq 0^{n-1}$ and $\mathrm{wt}(b')$ even, then $[0a' \cdot b']_2 = [a \cdot b]_2$. Thus the upper left hand block of M_n equals M_{n-1}.

2. If $a = 0a'1$ with $a' \neq 0^{n-2}$ and $b = b'0$ with $b' \neq 0^{n-1}$ and $\mathrm{wt}(b)$ even, then $[0a' \cdot b']_2 = [a \cdot b]_2$. Thus the lower left hand block of M_n equals M_{n-1}.

3. If $a = 0a'0$ with $a' \neq 0^{n-2}$ and $b = b'1$ with $b' \neq 10^{n-2}$ and $\mathrm{wt}(b')$ odd, then $[0a' \cdot b']_2 = [a \cdot b]_2$. Thus the upper right hand block of M_n equals N_{n-1}.

4. If $a = 0a'1$ with $a' \neq 0^{n-2}$ and $b = b'1$ with $b' \neq 10^{n-2}$ and $\mathrm{wt}(b')$ odd, then $[0a' \cdot b']_2 = 1 - [a \cdot b]_2$. Thus the lower right hand block of M_n equals $\langle 1 \rangle_{n-1} - N_{n-1}$.

5. If $a = 0^{n-1}1$, then the row indexed by a is $0^{2^{n-1}-1}1^{2^{n-1}}$.

6. If $b = 10^{n-2}1$, then the column indexed by b is $0^{2^{n-1}-1}1^{2^{n-1}}$.

We can summarize this by saying

$$M_n = \begin{bmatrix} & & & 0 & & & \\ & M_{n-1} & & \vdots & & N_{n-1} & \\ & & & 0 & & & \\ 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\ & & & 1 & & & \\ & M_{n-1} & & \vdots & & \langle 1 \rangle_{n-1} - N_{n-1} & \\ & & & 1 & & & \end{bmatrix}$$

By subtracting the first block of rows from the last block of rows, then subtracting the

row indexed by $a = 0^{n-1}1$ from each of the last block of rows, we have

$$\begin{aligned}
\det(M_n) &= \det \begin{bmatrix} & & & 0 & & & \\ & M_{n-1} & & \vdots & & N_{n-1} & \\ & & & 0 & & & \\ 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\ & & & 1 & & & \\ & 0 & & \vdots & & \langle 1 \rangle_{n-1} - 2N_{n-1} & \\ & & & 1 & & & \end{bmatrix} \\
&= \det \begin{bmatrix} & & & 0 & & & \\ & M_{n-1} & & \vdots & & N_{n-1} & \\ & & & 0 & & & \\ 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\ & & & 0 & & & \\ & 0 & & \vdots & & -2N_{n-1} & \\ & & & 0 & & & \end{bmatrix} \\
&= 2^{2^{n-1}-1} \det(M_{n-1}) \det(N_{n-1}). \qquad (21)
\end{aligned}$$

Similarly, we can decompose N_n into blocks as follows.

1. If $a = 0a'0$ with $a' \neq 0^{n-2}$ and $b = b'0$ with $b' \neq 0^{n-1}$ and $\mathrm{wt}(b')$ odd, then $[0a' \cdot b']_2 = [a \cdot b]_2$. Thus the upper left hand block of N_n equals N_{n-1}.

2. If $a = 0a'1$ with $a' \neq 0^{n-2}$ and $b = b'0$ with $b' \neq 0^{n-1}$ and $\mathrm{wt}(b)$ odd, then $[0a' \cdot b']_2 = [a \cdot b]_2$. Thus the lower left hand block of N_n equals N_{n-1}.

3. If $a = 0a'0$ with $a' \neq 0^{n-2}$ and $b = b'1$ with $b' \neq 10^{n-2}$ and $\mathrm{wt}(b')$ even, then $[0a' \cdot b']_2 = [a \cdot b]_2$. Thus the upper right hand block of N_n equals M_{n-1}.

4. If $a = 0a'1$ with $a' \neq 0^{n-2}$ and $b = b'1$ with $b' \neq 10^{n-2}$ and $\mathrm{wt}(b')$ even, then $[0a' \cdot b']_2 = 1 - [a \cdot b]_2$. Thus the lower right hand block of N_n equals $\langle 1 \rangle_{n-1} - M_{n-1}$.

5. If $a = 0^{n-1}1$, then the row indexed by a is $0^{2^{n-1}-1}1^{2^{n-1}}$.

6. If $b = 0^{n-1}1$, then the column indexed by b is $0^{2^{n-1}-1}1^{2^{n-1}}$.

As for M_n, this means that

$$N_n = \begin{bmatrix} & & 0 & & \\ & N_{n-1} & \vdots & M_{n-1} & \\ & & 0 & & \\ 0 \quad \cdots \quad 0 & & 1 & 1 \quad \cdots \quad 1 & \\ & & 1 & & \\ & N_{n-1} & \vdots & \langle 1 \rangle_{n-1} - M_{n-1} & \\ & & 1 & & \end{bmatrix},$$

so that we also have

$$\det(N_n) = 2^{2^{n-1}-1} \det(M_{n-1}) \det(N_{n-1}). \tag{22}$$

Finally, we see that M_2 and N_2 are 1×1 matrices whose single entries are 1, hence whose determinants are 1. It follows that the determinant of M_n and N_n are nonzero as claimed. □

In fact it follows from equations (21) and (22) that

$$\det(M_n) = \det(N_n) = 2^{(n-2)2^{n-1}+1},$$

but we shall not use this fact.

Corollary 4.2 *The values of $u(a)$ and $v(a)$ for $a \in U_n - \{0^n\}$ are uniquely determined by the $\widetilde{f}(b)$. This in turn uniquely determines the values of the $f(a)$ for $a \neq 0^n, 1^n$.*

Proof: The first statement follows from Theorem 4.1. Since $f(a), f(a+1^n) \in \{0,1\}$, the following table gives the possible values of $f(a)$, $f(a+1^n)$, $u(a)$, and $v(a)$.

$f(a)$	$f(a+1^n)$	$u(a)$	$v(a)$
0	0	0	0
1	0	1	0
0	1	-1	0
1	1	0	1

It follows that $f(a)$ and $f(a+1^n)$ are uniquely determined by any valid value of $u(a)$ and $v(a)$. □

Having determined the $f(a)$ with $a \neq 0^n, 1^n$, we are left with two equations in the unknowns $f(0^n)$ and $f(1^n)$. From equation (7) with $b = 0^n$ we have

$$f(0^n) + f(1^n) = x$$

for some $x \in \mathbb{Q}$, and from equation (9) with $b = 10^{n-1}$ we have

$$f(1^n) - f(0^n)f(1^n) = y$$

for some $y \in \mathbb{Q}$. The values x and y are uniquely determined by the $\widetilde{f}(b)$.

Again, we can make a table of possible values

$f(0^n)$	$f(1^n)$	x	y
0	0	0	0
1	0	1	0
0	1	1	1
1	1	2	0

It follows that $f(0^n)$ and $f(1^n)$ are uniquely determined by any valid value of x and y. We have proved the following theorem.

Theorem 4.3 *Every Boolean function on V_n is uniquely determined by its arithmetic Walsh transform.*

Embedded in this proof is a method of computing the function f from its arithmetic Walsh transform. It is, however, more complicated than the situation for classical Walsh transforms where one simply computes essentially the Walsh transform of the Walsh transform. We do not know of such an idempotency law for the arithmetic Walsh transform.

5 Statistics of the Arithmetic Walsh Transform

Recall that for a Boolean function f, the mean of the classical Walsh coefficients of f is $(-1)^{f(0)}$ and the second moment is 2^n (independent of f). The picture is quite different in the arithmetic case.

Let

$$H(f) = \sum_{a \in V_n} f(a),$$

the Hamming weight of f, and let

$$Q(f) = \sum_{a \in U_n} f(a)f(a+1^n) = \frac{1}{2}\sum_{a \in V_n} f(a)f(a+1^n),$$

the number of diagonals on which f is a constant 1.

Lemma 5.1 *If f is a Boolean function on n variables, then*

$$\sum_{a \in V_n} f(a)f(a+1^n) \sum_{b \cdot 1^n = 0} [a \cdot b]_2 = 2^{n-1}Q(f) - 2^{n-1}f(0^n)f(1^n),$$

$$\sum_{a \in V_n} (f(a) - f(a+1^n))) \sum_{b \cdot 1^n = 1} [a \cdot b]_2 = 2^{n-1}(f(1^n) - f(0^n)),$$

$$\sum_{a,c \in V_n} f(a)f(a+1^n)f(c)f(c+1^n) \sum_{b \cdot 1^n = 0} [a \cdot b]_2[c \cdot b]_2$$

$$\begin{aligned} &= 2^{n-1}Q(f)^2 + 2^{n-1}Q(f) - 2^n f(0^n)f(1^n)Q(f) \\ &= 2^{n-1}Q(f)(Q(f) + 1 - 2f(0^n)f(1^n)). \end{aligned}$$

and

$$\sum_{a,c \in V_n} (f(a) - f(a+1^n))(f(c) - f(c+1^n)) \sum_{b \cdot 1^n = 1} [a \cdot b]_2[c \cdot b]_2 = 2^{n-1}(H(f) - 2Q(f)).$$

Proof: For any $a \in V_n$, let

$$S_a = \sum_{b \cdot 1^n = 0} [a \cdot b]_2.$$

If $a = 0^n$ or $a = 1^n$, then $S_a = 0$. Otherwise 1^n and a are linearly independent modulo 2. Thus

$$\begin{aligned} \sum_{b \cdot 1^n = 0} [a \cdot b]_2 &= |\{b : [a \cdot b]_2 = 1 \text{ and } [1^n \cdot b]_2 = 0\}| \\ &= |\{b : [1^n \cdot b]_2 = 0\}| - |\{b : [a \cdot b]_2 = 0 \text{ and } [1^n \cdot b]_2 = 0\}| \\ &= 2^{n-1} - 2^{n-2} \\ &= 2^{n-2}. \end{aligned}$$

Thus

$$\sum_{a \in V_n} f(a)f(a+1^n) \sum_{b \cdot 1^n = 0} [a \cdot b]_2 = 2^{n-1}Q(f) - 2^{n-1}f(0^n)f(1^n).$$

Let

$$T_a = \sum_{b \cdot 1^n = 1} [a \cdot b]_2.$$

If $a = 0^n$, then $T_a = 0$. If $a = 1^n$, then $T_a = 2^{n-1}$. Otherwise, as in the previous case, $T_a = 2^{n-2}$. Since $\sum_{a \in V_n} (f(a) - f(a + 1^n))) = 0$, we have

$$\sum_{a \in V_n} (f(a) - f(a + 1^n)) \sum_{b \cdot 1^n = 0} [a \cdot b]_2 = 2^{n-1}(f(1^n) - f(0^n)).$$

Let

$$R_{a,c} = \sum_{b \cdot 1^n = 0} [a \cdot b]_2 [c \cdot b]_2.$$

Then $R_{a,c} = |\{b : [b \cdot 1^n]_2 = 0, [a \cdot b]_2 = 1, \text{ and } [c \cdot b]_2 = 1\}|$. There are several possibilities.

1. If a or c is 0^n or 1^n, then $R_{a,c} = 0$.
2. If $a = c$ and $a, c \notin \{0^n, 1^n\}$, then $R_{a,c} = 2^{n-2}$.
3. If $a = c + 1^n$ and $a, c \notin \{0^n, 1^n\}$, then $R_{a,c} = 2^{n-2}$.
4. Otherwise a, c, and 1^n are linearly independent modulo 2, so $R_{a,c} = 2^{n-3}$.

Thus

$$\begin{aligned}
&\sum_{a,c\in V_n} f(a)f(a+1^n)f(c)f(c+1^n)\sum_{b\cdot 1^n=0}[a\cdot b]_2[c\cdot b]_2\\
&= 2^{n-3}\sum_{a,c\in V_n} f(a)f(a+1^n)f(c)f(c+1^n)\\
&\quad +(2^{n-2}-2^{n-3})\sum_{a\in V_n} f(a)f(a+1^n)f(a+1^n)f(a)\\
&\quad +(2^{n-2}-2^{n-3})\sum_{a\in V_n} f(a)f(a+1^n)f(a)f(a+1^n)\\
&\quad +(-2^{n-3})\sum_{a\in V_n} f(0^n)f(1^n)f(c)f(c+1^n)\\
&\quad +(-2^{n-3})\sum_{a\in V_n} f(1^n)f(0^n)f(c)f(c+1^n)\\
&\quad +(-2^{n-3})\sum_{a\in V_n} f(a)f(a+1^n)f(0^n)f(1^n)\\
&\quad +(-2^{n-3})\sum_{a\in V_n} f(a)f(a+1^n)f(1^n)f(0^n)\\
&= 2^{n-1}Q(f)^2+2^{n-1}Q(f)-2^n f(0^n)f(1^n)Q(f)\\
&= 2^{n-1}Q(f)(Q(f)+1-2f(0^n)f(1^n)).
\end{aligned}$$

Note that in the final two lines one would expect a term $f(0^n)f(1^n)$ with some coefficient, accounting for all the appearances of this term in the various sums. In fact for each of the four choices of $a, c \in \{0^n, 1^n\}$ we have $[a\cdot b]_2[c\cdot b]_2 = 0$. Thus the coefficient of $f(0^n)f(1^n)$ is zero. Let

$$P_{a,c} = \sum_{b\cdot 1^n=1}[a\cdot b]_2[c\cdot b]_2.$$

Then $P_{a,c} = |\{b : [b\cdot 1^n]_2 = 1, [a\cdot b]_2 = 1, \text{ and } [c\cdot b]_2 = 1\}|$. There are several possibilities.

1. If a or c is 0^n or if $a = c + 1^n$, then $P_{a,c} = 0$.
2. If $a = c = 1^n$, then $P_{a,c} = 2^{n-1}$.
3. If exactly two of $a, c, 1^n$ are equal and $a, c \neq 0^n$, then $P_{a,c} = 2^{n-2}$.
4. Otherwise a, c, and 1^n are linearly independent, so $P_{a,c} = 2^{n-3}$.

Thus

$$\begin{aligned}
&\sum_{a,c\in V_n}(f(a)-f(a+1^n))(f(c)-f(c+1^n))\sum_{b\cdot 1^n=1}[a\cdot b]_2[c\cdot b]_2\\
&= 2^{n-3}\sum_{a,c\in V_n}(f(a)-f(a+1^n))(f(c)-f(c+1^n))\\
&\quad +(2^{n-2}-2^{n-3})\sum_{a\in V_n}(f(a)-f(a+1^n))^2\\
&\quad +(2^{n-2}-2^{n-3})\sum_{a\in V_n}(f(a)-f(a+1^n))(f(1^n)-f(0^n))\\
&\quad +(2^{n-2}-2^{n-3})\sum_{c\in V_n}(f(1^n)-f(0^n))(f(c)-f(c+1^n))\\
&\quad +(-2^{n-3})\sum_{c\in V_n}(f(0^n)-f(1^n))(f(c)-f(c+1^n)))\\
&\quad +(-2^{n-3})\sum_{a\in V_n}(f(a)-f(a+1^n))(f(0^n)-f(1^n))\\
&\quad +(-2^{n-3})\sum_{a\in V_n}(f(a)-f(a+1^n))(f(a+1^n)-f(a))\\
&\quad 2^{n-3}(f(0^n)-f(1^n))(f(0^n)-f(1^n))-3\cdot 2^{n-3}(f(1^n)-f(0^n))(f(1^n)-f(0^n))\\
&\qquad -2^{n-3}(f(0^n)-f(1^n))(f(1^n)-f(0^n))-2^{n-3}(f(1^n)-f(0^n))(f(0^n)-f(1^n))\\
&= 2^{n-2}\sum_{a\in V_n}(f(a)-f(a+1^n))^2\\
&= 2^{n-1}(H(f)-2Q(f)).
\end{aligned}$$

□

Theorem 5.2 *Let f be a Boolean function on n variables. The mean arithmetic Walsh transform of f is*

$$E[\widetilde{f}]=2^{n-1}-\frac{H(f)+f(0^n)-f(1^n)}{2}-f(0^n)f(1^n).$$

Proof: We have

$$E[\widetilde{f}]=\frac{1}{2^n}\sum_{b\in V_n}\widetilde{f}(b)=\frac{1}{2^n}\left(\sum_{b\cdot 1^n=0}\widetilde{f}(b)+\sum_{b\cdot 1^n=1}\widetilde{f}(b)\right).$$

We use equations (7) and (9) and Lemma 5.1 to compute these two sums separately. For the first sum we have

$$\begin{aligned}\sum_{b\cdot 1^n=0}\widetilde{f}(b) &= \sum_{b\cdot 1^n=0}\left(2^n - 2\sum_{a\in V_n} f(a) + 2\sum_{a\in V_n} f(a)f(a+1^n)[a\cdot b]_2\right)\\ &= 2^{2n-1} - 2^n H(f) + 2\sum_{a\in V_n} f(a)f(a+1^n)\sum_{b\cdot 1^n=0}[a\cdot b]_2\\ &= 2^{2n-1} - 2^n H(f) + 2^n Q(f) - 2^n f(0^n)f(1^n)\end{aligned}$$

by Lemma 5.1.

Similarly, for the second sum we have

$$\begin{aligned}\sum_{b\cdot 1^n=1}\widetilde{f}(b) &= \sum_{b\cdot 1^n=1}\sum_{a\in V_n}(f(a+1^n) - f(a)f(a+1^n) + (f(a) - f(a+1^n))[a\cdot b]_2)\\ &= 2^{n-1}H(f) - 2^n Q(f) + \sum_{a\in V_n}(f(a) - f(a+1^n)))\sum_{b\cdot 1^n=1}[a\cdot b]_2\\ &= 2^{n-1}H(f) - 2^n Q(f) + 2^{n-1}(f(1^n) - f(0^n)),\end{aligned}$$

again by Lemma 5.1. It follows that

$$E[\widetilde{f}] = 2^{n-1} - \frac{H(f) + f(0^n) - f(1^n)}{2} - f(0^n)f(1^n),$$

as claimed. □

Parseval's identity says the the sum of the squares of the Walsh coefficients of a Boolean function on n variables is 2^{2n}. This important fact leads, for example, to the notion of bent functions. Again the picture is more complicated in the arithmetic case.

Theorem 5.3 *Let f be a Boolean function on n variables. The second moment of the arithmetic Walsh transform of f is*

$$\begin{aligned}E[\widetilde{f}^2] &= 2^{2n-1} + \frac{5}{2}H(f)^2 - 6H(f)Q(f) + 4Q(f)^2\\ &\quad -(2^{n+1} - \frac{1}{2} + f(0^n) - f(1^n) - 4f(0^n)f(1^n))H(f)\\ &\quad +(2^{n+1} + 1 + 2f(0^n) - 2f(1^n) - 4f(0^n)f(1^n))Q(f) - 2^{n+1}f(0^n)f(1^n).\end{aligned}$$

Proof: We have

$$E[\widetilde{f}^2] = \frac{1}{2^n}\sum_{b\in V_n}\widetilde{f}(b)^2 = \frac{1}{2^n}\left(\sum_{b\cdot 1^n=0}\widetilde{f}(b)^2 + \sum_{b\cdot 1^n=1}\widetilde{f}(b)^2\right).$$

We again use equations (7) and (9) to compute these two sums separately. For the first sum we have

$$\begin{aligned}
\sum_{b\cdot 1^n=0}\widetilde{f}(b)^2 &= \sum_{b\cdot 1^n=0}\left(2^n - 2\sum_{a\in V_n} f(a) + 2\sum_{a\in V_n} f(a)f(a+1^n)[a\cdot b]_2\right)^2 \\
&= \sum_{b\cdot 1^n=0}\left(2^n - 2H(f) + 2\sum_{a\in V_n} f(a)f(a+1^n)[a\cdot b]_2\right)^2 \\
&= \sum_{b\cdot 1^n=0}(2^n - 2H(f))^2 + 4(2^n - 2H(f))\sum_{a\in V_n} f(a)f(a+1^n)[a\cdot b]_2 \\
&\quad +4\left(\sum_{a\in V_n} f(a)f(a+1^n)[a\cdot b]_2\right)^2 \\
&= 2^{n-1}(2^n - 2H(f))^2 + 4(2^n - 2H(f))\sum_{a\in V_n} f(a)f(a+1^n)\sum_{b\cdot 1^n=0}[a\cdot b]_2 \\
&\quad +4\sum_{a,c\in V_n} f(a)f(a+1^n)f(c)f(c+1^n)\sum_{b\cdot 1^n=0}[a\cdot b]_2[c\cdot b]_2 \\
&= 2^{n-1}(2^n - 2H(f))^2 + 4(2^n - 2H(f))(2^{n-1}Q(f) - 2^{n-1}f(0^n)f(1^n)) \\
&\quad +4(2^{n-1}Q(f)^2 + 2^{n-1}Q(f) - 2^n f(0^n)f(1^n)Q(f)) \\
&= 2^{3n-1} - 2^{2n+1}H(f) + 2^{n+1}H(f)^2 - 2^{n+2}H(f)Q(f) \\
&\quad +2^{n+2}H(f)f(0^n)f(1^n) + (2^{2n+1} + 2^{n+1})Q(f) - 2^{2n+1}f(0^n)f(1^n) \\
&\quad +2^{n+1}Q(f)^2 - 2^{n+2}f(0^n)f(1^n)Q(f).
\end{aligned}$$

Similarly, for the second sum we have

$$\begin{aligned}
\sum_{b\cdot 1^n=1} \widetilde{f}(b)^2 &= \sum_{b\cdot 1^n=1} \left(\sum_{a\in V_n} f(a+1^n) - f(a)f(a+1^n) + (f(a)-f(a+1^n))[a\cdot b]_2 \right)^2 \\
&= \sum_{b\cdot 1^n=1} \left(H(f) - 2Q(f) + \sum_{a\in V_n} (f(a)-f(a+1^n))[a\cdot b]_2 \right)^2 \\
&= 2^{n-1}(H(f)-2Q(f))^2 \\
&\quad +2(H(f)-2Q(f)) \sum_{a\in V_n} (f(a)-f(a+1^n)) \sum_{b\cdot 1^n=1} [a\cdot b]_2 \\
&\quad + \sum_{a,c\in V_n} (f(a)-f(a+1^n))(f(c)-f(c+1^n)) \sum_{b\cdot 1^n=1} [a\cdot b]_2[c\cdot b]_2 \\
&= 2^{n-1}(H(f)-2Q(f))^2 + 2^n(H(f)-2Q(f))(f(1^n)-f(0^n)) \\
&\quad +2^{n-1}(H(f)-2Q(f)).
\end{aligned}$$

It follows that

$$\begin{aligned}
E[\widetilde{f}^2] &= 2^{2n-1} - 2^{n+1}H(f) + 2H(f)^2 - 4H(f)Q(f) + 4H(f)f(0^n)f(1^n) \\
&\quad +(2^{n+1}+2)Q(f) - 2^{n+1}f(0^n)f(1^n) + 2Q(f)^2 - 4f(0^n)f(1^n)Q(f) \\
&\quad +2^{-1}(H(f)-2Q(f))^2 + (H(f)-2Q(f))(f(1^n)-f(0^n)) \\
&\quad +2^{-1}(H(f)-2Q(f)) \\
&= 2^{2n-1} + \frac{5}{2}H(f)^2 - 6H(f)Q(f) + 4Q(f)^2 \\
&\quad -(2^{n+1} - \frac{1}{2} + f(0^n) - f(1^n) - 4f(0^n)f(1^n))H(f) \\
&\quad +(2^{n+1}+1+2f(0^n)-2f(1^n)-4f(0^n)f(1^n))Q(f) - 2^{n+1}f(0^n)f(1^n),
\end{aligned}$$

as claimed. □

6 Arithmetic Walsh Transforms of Linear Functions

In this section we make use of the analysis in Section 2 to completely describe the arithmetic correlations of linear functions. That is, of Boolean functions $f(a) = \mathbf{T}_c(a) = [a\cdot c]_2$, $a, c \in V_n$.

If $c = 0^n$, then f is identically zero. By Theorem 2.4,

$$\widetilde{\mathbf{T}}_{0^n}(b) = \begin{cases} 2^n & \text{if } b\cdot 1^n = 0 \\ 0 & \text{if } b\cdot 1^n = 1. \end{cases}$$

For the remainder of the section we assume that $c \neq 0^n$. By equation (7), if $b \cdot 1^n = 0$, then

$$\begin{aligned}\widetilde{\mathbf{T}}_c(b) &= 2^n - 2\sum_{a\in V_n}[a\cdot c]_2 + 2\sum_{a\in V_n}[a\cdot c]_2[(a+1^n)\cdot c]_2[a\cdot b]_2 \\ &= 2\sum_{a\in V_n}[a\cdot c]_2[(a+1^n)\cdot c]_2[a\cdot b]_2. \end{aligned} \tag{23}$$

By equation (9), if $b \cdot 1^n = 1$, then

$$\widetilde{\mathbf{T}}_c(b) = \sum_{a\in V_n}[(a+1^n)\cdot c]_2(1-[a\cdot c]_2) + ([a\cdot c]_2 - [(a+1^n)\cdot c]_2)[a\cdot b]_2). \tag{24}$$

We treat these equations separately. First suppose that $b \cdot 1^n = 0$. If $b = 0^n$, then $\widetilde{\mathbf{T}}_c(b) = 0$. If $b \neq 0^n$ and $c \cdot 1^n = 0$, then

$$\begin{aligned}\widetilde{\mathbf{T}}_c(b) &= 2\sum_{a\in V_n}[a\cdot c]_2[a\cdot c]_2[a\cdot b]_2 \\ &= 2\sum_{a\in V_n}[a\cdot c]_2[a\cdot b]_2 \\ &= \begin{cases} 2\sum_{a\in V_n}[a\cdot c]_2 = 2^n & \text{if } b = c \\ 2\cdot 2^{n-2} = 2^{n-1} & \text{if } b \neq c. \end{cases}\end{aligned}$$

(The last line holds because $[a\cdot c]_2[a\cdot b]_2 = 1$ on the intersection of two hyperplanes and is 0 everywhere else.) The last case occurs for $2^{n-1} - 2$ values of b for each such c. If $c \cdot 1^n = 1$, then

$$\widetilde{\mathbf{T}}_c(b) = 2\sum_{a\in V_n}[a\cdot c]_2(1-[a\cdot c]_2)[a\cdot b]_2 = 0,$$

since if $x \in \{0,1\}$, then $x(1-x) = 0$. This occurs for 2^{n-1} values of b for each such c.

Now suppose that $b \cdot 1^n = 1$. If $c \cdot 1^n = 0$, then

$$\widetilde{\mathbf{T}}_c(b) = \sum_{a\in V_n}[a\cdot c]_2(1-[a\cdot c]_2) + ([a\cdot c]_2 - [a\cdot c]_2)[a\cdot b]_2) = 0.$$

This occurs for 2^{n-1} values of b for each such c. If $c \cdot 1^n = 1$, then

$$\begin{aligned}
\widetilde{\mathbf{T}}_c(b) &= \sum_{a \in V_n} (1 - [a \cdot c]_2)^2 + (2[a \cdot c]_2 - 1)[a \cdot b]_2) \\
&= \sum_{a \in V_n} (1 - [a \cdot c]_2) + (2[a \cdot c]_2 - 1)[a \cdot b]_2) \\
&= 2^{n-1} + \sum_{a \in V_n} (2[a \cdot c]_2 - 1)[a \cdot b]_2) \\
&= \begin{cases} 2^{n-1} + \sum_{a \in V_n} 2[a \cdot c]_2^2 - [a \cdot c]_2 = 2^n & \text{if } b = c \\ 2^{n-1} + \sum_{a \in V_n} 2[a \cdot c]_2[a \cdot b]_2 - [a \cdot b]_2 = 2^{n-1} & \text{if } b \neq c. \end{cases}
\end{aligned}$$

The second case occurs for $2^{n-1} - 1$ values of b for each such c. Now we fix c and describe the distribution of values of $\widetilde{\mathbf{T}}_c(b)$.

Theorem 6.1 *Let $c \in V_n$. If $c = 0^n$, then the arithmetic Walsh transform of $\mathbf{T}_c$ has values 0, which occurs 2^{n-1} times, and 2^n, which occurs 2^{n-1} times. If $c \cdot 1^n = 0$ and $c \neq 0^n$, then the arithmetic Walsh transform of $\mathbf{T}_c$ has values 0, which occurs $2^{n-1} + 1$ times, 2^{n-1}, which occurs $2^{n-1} - 2$ times, and 2^n, which occurs once. If $c \cdot 1^n = 1$, then the arithmetic Walsh transform of $\mathbf{T}_c$ has values 0, which occurs 2^{n-1} times, 2^{n-1}, which occurs $2^{n-1} - 1$ times, and 2^n, which occurs once.*

7 Arithmetic Walsh Transforms of Affine Functions

In this section we make use of the analysis in Section 2 to completely describe the arithmetic correlations of affine nonlinear functions. That is, of Boolean functions $f(a) = \mathbf{S}_c(a) = 1 - [a \cdot c]_2$, $a, c \in V_n$.

If $c = 0^n$, then f is identically one. By Theorem 2.4,

$$\widetilde{\mathbf{S}}_{0^n}(b) = \begin{cases} -2^n & \text{if } b = 0 \\ 0 & \text{if } b \neq 0. \end{cases}$$

For the remainder of the section we assume that $c \neq 0^n$. Theorem 2.4 implies that if $b \cdot 1^n = 0$, then

$$\begin{aligned}
\widetilde{\mathbf{S}}_c(b) &= 2^n - 2 \sum_{a \in V_n} (1 - [a \cdot c]_2) + 2 \sum_{a \in V_n} (1 - [a \cdot c]_2)(1 - [(a + 1^n) \cdot c]_2)[a \cdot b]_2 \\
&= 2 \sum_{a \in V_n} (1 - [a \cdot c]_2)(1 - [(a + 1^n) \cdot c]_2)[a \cdot b]_2. \qquad (25)
\end{aligned}$$

If $b \cdot 1^n = 1$, then

$$\widetilde{\mathbf{S}}_c(b) = \sum_{a \in V_n} (1 - [(a + 1^n) \cdot c]_2)[a \cdot c]_2 + ([(a + 1^n) \cdot c]_2 - [a \cdot c]_2)[a \cdot b]_2. \qquad (26)$$

We treat these equations separately. First suppose that $b \cdot 1^n = 0$. If $b = 0^n$, then $\widetilde{\mathbf{S}}_c(b) = 0$. If $b \neq 0^n$ and $c \cdot 1^n = 0$, then

$$\begin{aligned} \widetilde{\mathbf{S}}_c(b) &= 2 \sum_{a \in V_n} (1 - [a \cdot c]_2)[a \cdot b]_2 \\ &= \begin{cases} 0 & \text{if } b = c \\ 2 \cdot 2^{n-2} = 2^{n-1} & \text{if } b \neq c. \end{cases} \end{aligned}$$

The last case occurs for $2^{n-1} - 2$ values of b for each such c. If $c \cdot 1^n = 1$, then

$$\widetilde{\mathbf{S}}_c(b) = 2 \sum_{a \in V_n} (1 - [a \cdot c]_2)[a \cdot c]_2[a \cdot b]_2 = 0.$$

This occurs for 2^{n-1} values of b for each such c.

Now suppose that $b \cdot 1^n = 1$. If $c \cdot 1^n = 0$, then

$$\widetilde{\mathbf{S}}_c(b) = \sum_{a \in V_n} (1 - [a \cdot c]_2)[a \cdot c]_2 + ([a \cdot c]_2 - [a \cdot c]_2)[a \cdot b]_2 = 0.$$

This occurs for 2^{n-1} values of b for each such c. If $c \cdot 1^n = 1$, then

$$\begin{aligned} \widetilde{\mathbf{S}}_c(b) &= \sum_{a \in V_n} [a \cdot c]_2^2 + (1 - 2[a \cdot c]_2)[a \cdot b]_2 \\ &= \sum_{a \in V_n} [a \cdot c]_2 + (1 - 2[a \cdot c]_2)[a \cdot b]_2 \\ &= 2^{n-1} + \sum_{a \in V_n} (1 - 2[a \cdot c]_2)[a \cdot b]_2 \\ &= \begin{cases} 2^{n-1} + \sum_{a \in V_n} [a \cdot c]_2 - 2[a \cdot c]_2^2 = 0 & \text{if } b = c \\ 2^{n-1} + \sum_{a \in V_n} [a \cdot b]_2 - 2[a \cdot c]_2[a \cdot b]_2 = 2^{n-1} & \text{if } b \neq c. \end{cases} \end{aligned}$$

The second case occurs for $2^{n-1} - 1$ values of b for each such c. Now we fix c and describe the distribution of values of $\widetilde{\mathbf{S}}_c(b)$.

Theorem 7.1 *Let $c \in V_n$. If $c = 0^n$, then the arithmetic Walsh transform of $1 - \mathbf{S}_c$ has values 0, which occurs 2^{n-1} times, and -2^n, which occurs 2^{n-1} times. If $c \cdot 1^n = 0$ and $c \neq 0^n$, then the arithmetic Walsh transform of $1 - \mathbf{S}_c$ has values 0, which occurs $2^{n-1} + 2$ times and 2^{n-1}, which occurs $2^{n-1} - 2$ times. If $c \cdot 1^n = 1$, then the arithmetic Walsh transform of $1 - \mathbf{S}_c$ has values 0, which occurs $2^{n-1} + 1$ times and 2^{n-1}, which occurs $2^{n-1} - 1$ times.*

References

[1] C. F. Gauss, *Disquisitiones Arithmeticae*, 1801; reprinted in English translation by Yale Univ. Press, New Haven, CT. 1966.

[2] M. Goresky and A. Klapper, Arithmetic Cross-Correlations of FCSR Sequences, *IEEE Trans. Info. Theory.* **43** (1997) pp. 1342-1346.

[3] A. Klapper and M. Goresky, Feedback Shift Registers, Combiners with Memory, and 2-Adic Span, *Journal of Cryptology* **10** (1997) pp. 111-147.

[4] N. Koblitz, *p-Adic Numbers, p-Adic Analysis, and Zeta Functions,* Springer-Verlag: New York, 1984.