Linear recurrences and asymptotic behavior of exponential sums of symmetric boolean functions

Francis N. Castro

Department of Mathematics
University of Puerto Rico, San Juan, PR 00931

francis.castro@upr.edu

Luis A. Medina

Department of Mathematics
University of Puerto Rico, San Juan, PR 00931

luis.medina17@upr.edu

Submitted: Jan 28, 2011; Accepted: May 13, 2011; Published: May 25, 2011
Mathematics Subject Classification: 11T23, 05E05

Dedicated to Doron Zeilberger on the occasion of his 60th birthday

Abstract

In this paper we give an improvement of the degree of the homogeneous linear recurrence with integer coefficients that exponential sums of symmetric Boolean functions satisfy. This improvement is tight. We also compute the asymptotic behavior of symmetric Boolean functions and provide a formula that allows us to determine if a symmetric boolean function is asymptotically not balanced. In particular, when the degree of the symmetric function is a power of two, then the exponential sum is much smaller than 2^n.

Keywords: Exponential sums, recurrences, Cusick et al. Conjecture for elementary balanced symmetric boolean functions

1 Introduction

Boolean functions are one of the most studied objects in mathematics. They are important in many applications, for example, in the design of stream ciphers, block and hash functions. These functions also play a vital role in cryptography as they are used as filter and combination generator of stream ciphers based on linear feed-back shift registers. The

case of boolean functions of degree 2 has been intensively studied because of its relation to bent functions (see [11], [1]).

One can find many papers and books discussing the properties of boolean functions (see [5], [9], [2] and [6]). The subject can be studied from the point of view of complexity theory or from the algebraic point of view as we do in this paper, where we compute the asymptotic behavior of exponential sums of symmetric boolean functions.

The correlation between two Boolean functions of n inputs is defined as the number of times the functions agree minus the number of times they disagree all divided by 2^n, i.e.,

$$C(F_1, F_2) = \frac{1}{2^n} \sum_{x_1,\ldots,x_n \in \{0,1\}} (-1)^{F_1(x_1,\ldots,x_n)+F_2(x_1,\ldots,x_n)}. \tag{1.1}$$

In this paper we are interested in the case when F_1 and F_2 are symmetric boolean functions. Without loss of generality, we write $C(F)$ instead of $C(F_1, F_2)$, where F is a symmetric boolean function. In [4], A. Canteaut and M. Videau studied in detail symmetric boolean functions. They established a link between the periodicity of the simplified value vector of a symmetric Boolean function and its degree. They also determined all balanced symmetric functions of degree less than or equal to 7. In [13], J. von zur Gathen and J. Rouche found all the balanced symmetric boolean functions up to 128 variables.

In[3], J. Cai et. al. computed a closed formula for the correlation between any two symmetric Boolean functions. This formula implies that $C(F)$ satisfies a homogeneous linear recurrence with integer coefficients and provides an upper bound for the degree of the minimal recurrence of this type that $C(F)$ satisfies. In this paper we give an improvement to the degree of the minimal homogeneous linear recurrence with integer coefficients satisfying by $C(F)$. In particular, our lower and upper bounds are tight in many cases. Also, in the case of an elementary symmetric function we provide the minimal homogeneous linear recurrence.

We also compute the asymptotic value of $C(F)$. In particular, we give infinite families of boolean functions that are asymptotically not balanced, i.e., $\lim_{n\to\infty} C(F) \neq 0$. In [7], T. Cusick et al. conjectured that there are no nonlinear balanced elementary symmetric polynomials except for the elementary symmetric boolean function of degree $k = 2^r$ in $2^r \cdot l - 1$ variables, where r and l are any positive integers. In this paper, we prove that

$$\lim_{n\to\infty} C(\sigma_{n,k}) = \frac{2^{w_2(k)-1} - 1}{2^{w_2(k)-1}}, \tag{1.2}$$

where $\sigma_{n,k}$ is the elementary symmetric polynomial of degree k in n variables and $w_2(k)$ is the sum of the binary digits of k. Note that this implies that Cusick et al.'s conjecture holds for sufficiently large n. In particular, an elementary symmetric function is asymptotically not balanced if and only if its degree is not a power of 2. In [8], Cusick et al. presented some progress on proving this conjecture. In particular, they presented the following stronger version of their conjecture: If $n \geq 2(k-1)$, where k is fixed and $w_2(k) \geq 6$, then $C(F) > 1/2$. Formula (1.2) implies that this holds for sufficiently large n when $w_2(k) \geq 3$.

When the asymptotic value of $C(F)$ is zero, we compute the asymptotic values of

$$\frac{1}{|\lambda|^n}\sum_{x_1,\ldots,x_n\in\{0,1\}}(-1)^{F(x_1,\ldots,x_n)}, \tag{1.3}$$

where λ and $\bar{\lambda}$ are the roots with the biggest modulus of the characteristic polynomial associated to the exponential sum of F. We prove that the coefficient of λ is not identically zero and obtain information about the spectrum of $F(X_1,\ldots,X_n)$. In particular, its limit is a periodic function in n.

2 Preliminaries

Let $\mathbb{F}$ be the binary field, $\mathbb{F}^n = \{(x_1,\ldots,x_n)\,|\,x_i \in \mathbb{F}, i = 1,\ldots,n\}$, and $F(\mathbf{X}) = F(X_1,\ldots,X_n)$ be a polynomial in n variables over $\mathbb{F}$. The exponential sum associated to F over $\mathbb{F}$ is:

$$S(F) = \sum_{\mathbf{x}\in\mathbb{F}^n}(-1)^{F(\mathbf{x})}. \tag{2.1}$$

Note that $S(F) = 2^n C(F)$. A boolean function $F(\mathbf{X})$ is called balanced if $S(F) = 0$. This property is important for some applications in cryptography. P. Sarkar and S. Maitra [12] found a lower bound for the number of symmetric balanced boolean functions. In particular, in the case that $n \geq 13$ is odd and $n+3$ is a perfect square, this number is bigger than or equal to $2^{(n+1)/2} + 2^{(n+1)/2-3}$.

In this paper we study exponential sums associated to symmetric boolean functions F. Any symmetric function is a linear combination of elementary symmetric polynomials, thus we start with exponential sums of elementary symmetric polynomials.

Let $\sigma_{n,k}$ be the elementary symmetric polynomial in n variables of degree k. For example,

$$\sigma_{4,3} = X_1X_2X_3 + X_1X_4X_3 + X_2X_4X_3 + X_1X_2X_4. \tag{2.2}$$

Fix $k \geq 2$ and let n vary. Consider the sequence of exponential sums $\{S(\sigma_{n,k})\}_{n\in\mathbb{N}}$ where

$$S(\sigma_{n,k}) = \sum_{x_1,\cdots,x_n\in\mathbb{F}}(-1)^{\sigma_{n,k}(x_1,\cdots,x_n)}. \tag{2.3}$$

Define A_j to be the set of all $(x_1,\cdots,x_n) \in \mathbb{F}^n$ with exactly j entries equal to 1. Clearly, $|A_j| = \binom{n}{j}$ and $\sigma_{n,k}(\mathbf{x}) = \binom{j}{k}$ for $\mathbf{x} \in A_j$. Therefore,

$$S(\sigma_{n,k}) = \sum_{j=0}^{n}(-1)^{\binom{j}{k}}\binom{n}{j}. \tag{2.4}$$

In general, if $1 \leq k_1 < k_2 < \cdots < k_s$ are fixed integers, then

$$S(\sigma_{n,k_1} + \sigma_{n,k_2} + \cdots + \sigma_{n,k_s}) = \sum_{j=0}^{n}(-1)^{\binom{j}{k_1}+\binom{j}{k_2}+\cdots+\binom{j}{k_s}}\binom{n}{j}. \tag{2.5}$$

Remark 1 *Note that the sum on the right hand side of (2.5) makes sense for values of n less than k_s, while $S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})$ does not. However, throughout the paper we let $S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})$ to be defined by the sum in (2.5), even for values of n less than k_s.*

3 The recurrence

Computer experimentation suggests that for fix $1 \leq k_1 < \cdots < k_s$, the sequence $\{S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})\}_{n\in\mathbb{N}}$ satisfies a homogeneous linear recurrence with integer coefficients. For example, if we consider $\{S(\sigma_{n,7})\}_{n\in\mathbb{N}}$ and type

```
FindLinearRecurrence[Table[Sum[((-1)^Binomial[m,7])*
                    Binomial[n,m],{m,0,n}],{n,1,30}]]
```

into *Mathematica* 7, then it returns

```
{8,-28,56,-70,56,-28,8}.
```

This suggests that $\{S(\sigma_{n,7})\}_{n\in\mathbb{N}}$ satisfies the recurrence

$$x_n = 8x_{n-1} - 28x_{n-2} + 56x_{n-3} - 70x_{n-4} + 56x_{n-5} - 28x_{n-6} + 8x_{n-7}. \tag{3.1}$$

If we continue with these experiments, we arrive to the observation that if $r = \lfloor \log_2(k_s) \rfloor + 1$, then $\{S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})\}_{n\in\mathbb{N}}$ seems to satisfy the recurrence

$$x_n = \sum_{m=1}^{2^r-1} (-1)^{m-1} \binom{2^r}{m} x_{n-m}. \tag{3.2}$$

This result can be proved using elementary machinery. The idea is to use the fact that if $r = \lfloor \log_2(k_s) \rfloor + 1$, then

$$\binom{j+i2^r}{k_m} \equiv \binom{j}{k_m} (\text{mod } 2) \tag{3.3}$$

for all non-negative integers i and $m = 1, 2, \cdots, s,$ to show inductively that the family of sequences

$$a_{n,r,i} = \sum_j \binom{n}{2^r j + i} = \sum_{j \equiv i \ (\text{mod } 2^r)} \binom{n}{j}, \tag{3.4}$$

$i = 0, 1, \cdots 2^r - 1$ satisfies the same recurrence (3.2). However, we should point out the fact that $\{S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})\}_{n\in\mathbb{N}}$ satisfies (3.2) is a consequence of the following theorem of J. Cai et al. [3].

Theorem 3.1 *Fix $1 \leq k_1 < \cdots < k_s$ and let $r = \lfloor \log_2(k_s) \rfloor + 1$. The value of the exponential sum $S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})$ is given by*

$$\begin{aligned} S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s}) &= \sum_{i=0}^{n} (-1)^{\binom{i}{k_1}+\cdots+\binom{i}{k_s}} \binom{n}{i} \\ &= c_0(k_1,\cdots,k_s)2^n + \sum_{j=1}^{2^r-1} c_j(k_1,\cdots,k_s)(1+\zeta_j)^n, \end{aligned} \tag{3.5}$$

where $\zeta_j = \exp\left(\frac{\pi\sqrt{-1}\,j}{2^{r-1}}\right)$ *and*

$$c_j(k_1,\cdots,k_s) = \frac{1}{2^r}\sum_{i=0}^{2^r-1}(-1)^{\binom{i}{k_1}+\cdots+\binom{i}{k_s}}\zeta_j^{-i}. \tag{3.6}$$

The proofs of Cai et al. rely on linear algebra. They wrote

$$S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s}) = \sum_{i=0}^{2^r-1}(-1)^{\binom{i}{k_1}+\cdots+\binom{i}{k_s}}a_{n,r,i}, \tag{3.7}$$

and used the elementary identity

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}, \tag{3.8}$$

to find a recurrence for

$$\mathbf{a}_{n,r} = \begin{pmatrix} a_{n,r,1} \\ a_{n,r,2} \\ \vdots \\ a_{n,r,2^r-1} \end{pmatrix} \tag{3.9}$$

of the form $\mathbf{a}_{n,r} = \mathbf{M}\mathbf{a}_{n-1,r}$, for some matrix $\mathbf{M}$. Finding the eigenvalues and corresponding eigenvectors of $\mathbf{M}$, they were able to solve this recurrence and prove Theorem 3.1. See [3] for more details.

From Theorem 3.1 it is now evident that $\{S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})\}_{n\in\mathbb{N}}$ satisfies (3.2). Moreover, the roots of the characteristic polynomial associated to the linear recurrence (3.2) are all different and the polynomial is given by

$$\begin{aligned} P_r(x) &= \sum_{m=0}^{2^r-1}(-1)^m\binom{2^r}{m}x^{2^r-1-m} \\ &= (x-2)\Phi_4(x-1)\Phi_8(x-1)\cdots\Phi_{2^r}(x-1), \end{aligned} \tag{3.10}$$

where $\Phi_m(x)$ represents the m-th cyclotomic polynomial

$$\Phi_m(x) = \prod_{\zeta^m=1 \text{ primitive}}(x-\zeta). \tag{3.11}$$

Even though $\{S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})\}_{n\in\mathbb{N}}$ satisfies (3.2), in many instances (3.2) is not the minimal homogeneous linear recurrence with integer coefficients that $\{S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})\}_{n\in\mathbb{N}}$ satisfies. For example, $\{S(\sigma_{n,3}+\sigma_{n,5})\}_{n\in\mathbb{N}}$ satisfies (3.1), but its minimal recurrence is

$$x_n = 6x_{n-1} - 14x_{n-2} + 16x_{n-3} - 10x_{n-4} + 4x_{n-5}. \tag{3.12}$$

In the next section we use Theorem 3.1 to give some improvements on the degree of the minimal linear recurrence associated to $\{S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})\}_{n\in\mathbb{N}}$.

4 On the degree of the recurrence relation

Now that we are equipped with equation (3.6), we move to the problem of reducing the degree of the recurrence relation that our sequences of exponential sums satisfy. The idea behind our approach is very simple. Consider all roots $1+\zeta$'s of $\Phi_{2^{t+1}}(x-1)$ where $1 \leq t \leq r-1$. We know that $(1+\zeta)^n$ appears in (3.5). If we show that the coefficient that corresponds to $(1+\zeta)^n$ is zero for each $1+\zeta$, then we reduce the degree of the characteristic polynomial, and therefore the degree of the recurrence, by 2^t.

However, note that $\Phi_{2^{t+1}}(x-1)$ is irreducible over $\mathbb{Q}$ (according to Eisenstein's criterion on $\Phi_{2^{t+1}}(x-1)$ with $\Phi_{2^{t+1}}(x) = x^{2^t}+1$, see [10]). Therefore, the coefficients related to the roots of $\Phi_{2^{t+1}}(x-1)$ are either all zeros or all non-zeros. In view of (3.6), this can be determined by checking whether or not the sum

$$\sum_{m=0}^{2^r-1} (-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}} \exp\left(\frac{\pi\sqrt{-1}m}{2^t}\right) \tag{4.1}$$

is zero.

First, we discuss the case of the exponential sum of one elementary symmetric polynomial, i.e. $\{S(\sigma_{n,k}\}_{n\in\mathbb{N}}$. We start with the following elementary result.

Lemma 4.1 (Lucas' theorem) *Let n be a natural number with 2-adic expansion $n = 2^{a_1}+2^{a_2}+\cdots+2^{a_l}$. The binomial coefficient $\binom{n}{k}$ is odd if and only if k is either 0 or a sum of some of the 2^{a_i}'s.*

Proof: Recall that $(1+x)^{2^m} \equiv 1+x^{2^m} \pmod 2$ for all non-negative integer m, thus $(1+x)^n \equiv (1+x^{2^{a_1}})(1+x^{2^{a_2}})\cdots(1+x^{2^{a_l}}) \pmod 2$. Note that the coefficient of x^k in $(1+x^{2^{a_1}})(1+x^{2^{a_2}})\cdots(1+x^{2^{a_l}})$ is 1 if and only if $k=0$ or a sum of some of the 2^{a_i}'s. □

The next result is an immediate consequence of the above lemma.

Corollary 4.2 *Fix a natural number k. Suppose its 2-adic expansion is $k = 2^{a_1}+2^{a_2}+\cdots+2^{a_l}$. A natural number m is such that $\binom{m}{k}$ is odd if and only if m has a 2-adic expansion of the form*

$$m = k + \sum_{2^i \notin \{2^{a_1},2^{a_2},\cdots,2^{a_l}\}} \delta_i 2^i \tag{4.2}$$

where $\delta_i \in \{0,1\}$.

Remark 2 *Let $k \geq 1$ be an integer with 2-adic expansion $k = 2^{a_1}+\cdots+2^{a_l}$. Suppose $m \in \{0,1,2,3,\cdots,2^r-1\}$ is such that $\binom{m}{k}$ is odd. Note that Corollary 4.2 implies*

$$m = k + \delta_1 2^{b_1} + \delta_2 2^{b_2} + \cdots + \delta_t 2^{b_f}, \tag{4.3}$$

where $\{2^{b_1},2^{b_2},\cdots,2^{b_f}\} = \{1,2,2^2,\cdots,2^{r-1}\}\backslash\{2^{a_1},2^{a_2},\cdots,2^{a_l}\}$.

We now proceed to show which coefficients $c_j(k)$ are zero. We start with $c_0(k)$.

Lemma 4.3 *Suppose $k \geq 2$ is an integer. Then,*

$$c_0(k) = \frac{2^{w_2(k)-1} - 1}{2^{w_2(k)-1}}, \tag{4.4}$$

where $w_2(k)$ is the sum of the binary digits of k. In particular, $c_0(k) = 0$ if and only if k is a power of two.

Proof: Recall from Theorem 3.1 that

$$c_0(k) = \frac{1}{2^r} \sum_{m=0}^{2^r-1} (-1)^{\binom{m}{k}},$$

where $r = \lfloor \log_2(k) \rfloor + 1$. Re-write $c_0(k)$ as

$$c_0(k) = \frac{1}{2^r} \left(2^r - 2 \sum_{m \in N} 1 \right), \tag{4.5}$$

where

$$N = \left\{ m \in \{0, 1, 2, 3, \cdots, 2^r - 1\} : \binom{m}{k} \text{ is odd.} \right\} \tag{4.6}$$

Note that (4.3) implies that the cardinality of N is $2^{r-w_2(k)}$. A simple calculation yields the result. □

Remark 3 *In [7], T. Cusick et al. conjectured that there are no nonlinear balanced elementary symmetric polynomials except for the elementary symmetric boolean function of degree $k = 2^r$ in $2^r \cdot l - 1$ variables, where r and l are any positive integers. Note that Lemma 4.3 implies that Cusick et al. conjecture holds for sufficiently large n. In particular, Lemma 4.3 shows that if k is not a power of two, then $\sigma_{n,k}$ is not balanced for sufficiently large n. We say that $\sigma_{n,k_1} + \cdots + \sigma_{n,k_s}$ is asymptotically not balanced if*

$$\lim_{n \to \infty} \frac{S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})}{2^n} \neq 0. \tag{4.7}$$

In the case that

$$\lim_{n \to \infty} \frac{S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})}{2^n} = 0, \tag{4.8}$$

we cannot conclude anything about whether or not $S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})$ is balanced for some values of n. For instance,

$$\lim_{n \to \infty} \frac{S(\sigma_{n,2} + \sigma_{n,5})}{2^n} = 0, \tag{4.9}$$

but $S(\sigma_{n,2} + \sigma_{n,5}) \neq 0$.

Consider now the coefficients

$$c_j(k) = \frac{1}{2^r} \sum_{m=0}^{2^r-1} (-1)^{\binom{m}{k}} \exp\left(\frac{-\pi\sqrt{-1}mj}{2^{r-1}}\right)$$

with $j > 0$. From Theorem 3.1 we know each $c_j(k)$ is the coefficient of $(1+\zeta_j)^n$ where $1+\zeta_j$ is a root of $\Phi_{2^{t+1}}(x-1)$ for some $t = 1, 2, \cdots, 2^{r-1}$.

Lemma 4.4 *Let $k \geq 2$ be an integer with 2-adic expansion $k = 2^{a_1} + \cdots + 2^{a_l}$. Then $c_j(k) = 0$ if and only if it is the coefficient of $(1+\zeta)^n$, where $1+\zeta$ is a root of $\Phi_{2^{b+1}}(x-1)$ and $b \neq a_i$ for all $i = 1, \cdots, l$, i.e. 2^b does not appear in the 2-adic expansion of k.*

Proof: Recall that to show that the coefficients of the roots of $\Phi_{2^{t+1}}(x-1)$ are zero is equivalent to show that

$$\sum_{m=0}^{2^r-1} (-1)^{\binom{m}{k}} \exp\left(\frac{\pi\sqrt{-1}m}{2^t}\right) = 0. \tag{4.10}$$

If $\{2^{b_1}, \cdots, 2^{b_f}\} = \{1, 2, 2^2, \cdots, 2^{r-1}\} \backslash \{2^{a_1}, \cdots, 2^{a_l}\}$, then equation (4.3) implies,

$$\begin{aligned}\sum_{m=0}^{2^r-1} (-1)^{\binom{m}{k}} \exp\left(\frac{\pi\sqrt{-1}m}{2^t}\right) &= -2 \sum_{(\delta_1,\cdots,\delta_f)\in\mathbb{F}_2^f} \exp\left(\frac{\pi\sqrt{-1}}{2^t}(k + \delta_1 2^{b_1} + \cdots + \delta_t 2^{b_f})\right) \\ &= -2\exp\left(\frac{\pi\sqrt{-1}k}{2^t}\right) \sum_{(\delta_1,\cdots,\delta_f)\in\mathbb{F}_2^f} \exp\left(\frac{\pi\sqrt{-1}}{2^t}(\delta_1 2^{b_1} + \cdots + \delta_t 2^{b_f})\right).\end{aligned} \tag{4.11}$$

Thus, (4.10) holds if and only if $\exp\left(\frac{\pi\sqrt{-1}}{2^t}\right)$ is a root of

$$\sum_{(\delta_1,\cdots,\delta_f)\in\mathbb{F}_2^f} x^{\delta_1 2^{b_1} + \cdots + \delta_t 2^{b_f}}. \tag{4.12}$$

Consider first $t = b_1$. If we set $\delta_1 = 0$ in the last sum of (4.11), then we have

$$\sum_{(\delta_2,\cdots,\delta_f)\in\mathbb{F}_2^{f-1}} \exp\left(\frac{\pi\sqrt{-1}}{2^{b_1}}(\delta_2 2^{b_2} + \cdots + \delta_t 2^{b_f})\right). \tag{4.13}$$

However, if we set $\delta_1 = 1$, then we have

$$-\sum_{(\delta_2,\cdots,\delta_f)\in\mathbb{F}_2^{f-1}} \exp\left(\frac{\pi\sqrt{-1}}{2^{b_1}}(\delta_2 2^{b_2} + \cdots + \delta_t 2^{b_f})\right). \tag{4.14}$$

We conclude that (4.10) holds for $t = b_1$, i.e. the 2^{b_1} coefficients related to the roots of $\Phi_{2^{b_1+1}}(x-1)$ are zero. Repeat this argument with $t = b_2, \cdots, b_f$ to conclude that the

coefficients related to the roots of $\Phi_{2^{b_i+1}}(x-1)$, $i=1,\cdots,f$ are zero. Since (4.12) is of degree $d=2^{b_1}+\cdots+2^{b_f}$, then only d of the coefficients $c_j(k)$ can be zero. Since we already found d coefficients that are zero, then we conclude that these are all of them. The claim follows. □

Lemmas 4.3 and 4.4 are put together in the following theorem. The function $\epsilon(n)$ used in the theorem is defined as

$$\epsilon(n)=\begin{cases}0, & \text{if } n \text{ is a power of 2,}\\ 1, & \text{otherwise.}\end{cases} \tag{4.15}$$

Theorem 4.5 *Let k be a natural number and $P_k(x)$ be the characteristic polynomial associated to the minimal linear recurrence with integer coefficients that $\{S(\sigma_{n,k})\}_{n\in\mathbb{N}}$ satisfies. Let $\bar{k}=2\lfloor k/2\rfloor+1$. We know $\bar{k}$ has a 2-adic expansion of the form*

$$\bar{k}=1+2^{a_1}+2^{a_2}+\cdots+2^{a_l}, \tag{4.16}$$

where the last exponent is given by $a_l=\lfloor\log_2(\bar{k})\rfloor$. Then $P_k(x)$ equals

$$(x-2)^{\epsilon(k)}\prod_{j=1}^{l}\Phi_{2^{a_j+1}}(x-1). \tag{4.17}$$

In particular, the degree of the minimal linear recurrence that $\{S(\sigma_{n,k})\}_{n\in\mathbb{N}}$ satisfies is equal to $2\lfloor k/2\rfloor+\epsilon(k)$.

Theorem 4.5 can be generalized to the case $\{S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})\}_{n\in\mathbb{N}}$. Define the "OR" operator $\vee$ on $\mathbb{F}_2$ as

$$\begin{aligned}0\vee 0 &= 0\\ 0\vee 1 &= 1\\ 1\vee 0 &= 1\\ 1\vee 1 &= 1.\end{aligned} \tag{4.18}$$

Extend $\vee$ to $\mathbb{N}$ by letting $m\vee n$ be the natural number obtained by applying $\vee$ coordinatewise to the binary digits of n and m. For example,

$$\begin{aligned}4\vee 6 &= (0\cdot 1+0\cdot 2+1\cdot 2^2)\vee(0\cdot 1+1\cdot 2+1\cdot 2^2)\\ &= (0\vee 0)\cdot 1+(0\vee 1)\cdot 2+(1\vee 1)\cdot 2^2=6.\end{aligned} \tag{4.19}$$

and

$$\begin{aligned}3\vee 8 &= (1\cdot 1+1\cdot 2+0\cdot 2^2+0\cdot 2^3)\vee(0\cdot 1+0\cdot 2+0\cdot 2^2+1\cdot 2^3)\\ &= (1\vee 0)\cdot 1+(0\vee 1)\cdot 2+(0\vee 0)\cdot 2^2+(0\vee 1)\cdot 2^3=11.\end{aligned} \tag{4.20}$$

Next is a generalization of Theorem 4.5.

Theorem 4.6 *Let $1 \leq k_1 < k_2 < \cdots < k_s$ be fixed integers and $P_{k_1,\cdots,k_s}(x)$ be the characteristic polynomial associated to the minimal linear recurrence with integer coefficients that $\{S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})\}_{n\in\mathbb{N}}$ satisfies. Let $\bar{k} = 2\lfloor (k_1 \vee \cdots \vee k_s)/2 \rfloor + 1$. We know $\bar{k}$ has a 2-adic expansion of the form*

$$\bar{k} = 1 + 2^{a_1} + 2^{a_2} + \cdots + 2^{a_l}, \tag{4.21}$$

where the last exponent is given by $a_l = \lfloor \log_2(\bar{k}) \rfloor$. Then $P_{k_1,\cdots,k_s}(x)$ divides the polynomial

$$(x-2)\prod_{j=1}^{l} \Phi_{2^{a_j+1}}(x-1). \tag{4.22}$$

Proof: The proof is similar to the one of Lemma 4.4. Let $\bar{k} = 2\lfloor (k_1 \vee \cdots \vee k_s)/2 \rfloor + 1$ and $r = \lfloor \log_2(\bar{k}) \rfloor + 1$. Define

$$N = \left\{ m \in \{1, 2, 3, \cdots, 2^r - 1\} : \binom{m}{k_1} + \cdots + \binom{m}{k_s} \text{ is odd} \right\}. \tag{4.23}$$

Suppose $2^b \in \{2, 2^2, \cdots, 2^{r-1}\}$ is such that 2^b does not appear in the 2-adic expansion of $\bar{k}$. We will show that

$$\sum_{m=0}^{2^r-1} (-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}} \exp\left(\frac{\pi\sqrt{-1}}{2^b}\right) = 0, \tag{4.24}$$

which implies that the coefficients related to the roots of $x^{2^b} + 1$ are all zero.

Observe that

$$\sum_{m=0}^{2^r-1} (-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}} \exp\left(\frac{\pi\sqrt{-1}}{2^b}\right) = -2\sum_{m\in N} \exp\left(\frac{\pi\sqrt{-1}m}{2^b}\right). \tag{4.25}$$

Suppose $m \in N$ is such that 2^b does not appear in the 2-adic expansion of m. Note that equation (4.3) implies $m + 2^b \in N$. Thus, the same argument as in (4.13) and (4.14) imply that (4.24) is true. Hence, the claim follows. □

The following example presents a case when Theorem 4.6 is tight.

Example 4.7 *Consider $k_1 = 6$ and $k_2 = 17$. Note that $2\lfloor (6 \vee 17)/2 \rfloor + 1 = 23 = 1+2+4+16$. In this case, the characteristic polynomial associated to $\{S(\sigma_{n,6}+\sigma_{n,17})\}_{n\in\mathbb{N}}$ is $P_{6,17}(x) = (x-2)\Phi_4(x-1)\Phi_8(x-1)\Phi_{32}(x-1)$. This is the best case scenario of Theorem 4.6, i.e we have equality rather than just divisibility. Also, note that in this case the recurrence given by Theorem 3.1 is of degree* 31*, while the minimal linear recurrence is of degree* 23*.*

The next example presents a case in which Theorem 4.6 improves the degree of the homogeneous linear recurrence provided by Theorem 3.1. However it did not provide the minimal degree of the recurrence.

Example 4.8 *Consider $k_1 = 3$, $k_2 = 5$, and $k_3 = 17$. We have $2\lfloor (3 \vee 5 \vee 17)/2 \rfloor + 1 = 23$. In this case, the characteristic polynomial of the minimal recurrence is $P_{3,5,17}(x) = (x - 2)\Phi_{32}(x-1)$. It divides $(x-2)\Phi_4(x-1)\Phi_8(x-1)\Phi_{32}(x-1)$ as Theorem 4.6 predicted, but are clearly not equal. The factors $\Phi_4(x-1)$ and $\Phi_8(x-1)$ do not appear in $P_{3,5,17}(x)$. This means that the coefficients $c_j(3,5,17)$ related to the roots of $\Phi_4(x) = x^2+1$ and $\Phi_8(x) = x^4+1$ are zero. However, since 2 and 4 appear in the 2-adic expansion of 23, then Theorem 4.6 cannot detect this.*

We now provide bounds on the degree of the minimal linear recurrence that $\{S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})\}_{n\in\mathbb{N}}$ satisfies. We start with the following theorem.

Theorem 4.9 *Suppose $1 \leq k_1 < \cdots < k_s$ are integers. Let $r = \lfloor \log_2(k_s) \rfloor + 1$. Then $\Phi_{2^r}(x-1)$ divides $P_{k_1,\cdots,k_s}(x)$, the characteristic polynomial associated to $\{S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})\}_{n\in\mathbb{N}}$.*

Proof: Note that the theorem will follow if we show that

$$\sum_{m=0}^{2^r-1} (-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}} \exp\left(\frac{\pi\sqrt{-1}m}{2^{r-1}}\right) \neq 0. \tag{4.26}$$

This is equivalent to showing that $x^{2^{r-1}} + 1$ does not divide

$$\sum_{m=0}^{2^r-1} (-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}} x^m. \tag{4.27}$$

We present the core of our proof with a particular example. The general case will follow in a similar manner. Consider the case $k_1 = 3$, $k_2 = 5$, and $k_3 = 10$. Then (4.27) equals,

$$1 + x + x^2 - x^3 + x^4 - x^5 + x^6 + x^7 + x^8 + x^9 - x^{10} + x^{11} + x^{12} - x^{13} - x^{14} - x^{15}. \tag{4.28}$$

Look at the sign of x^j for $j = 8, 9, \cdots, 15$. If x^j and x^{j-8} have the same sign, then leave the sign of x^j as it is. Otherwise, change the sign of x^j. After doing this, we get

$$1 + x + x^2 - x^3 + x^4 - x^5 + x^6 + x^7 + x^8 + x^9 + x^{10} - x^{11} + x^{12} - x^{13} + x^{14} + x^{15}, \tag{4.29}$$

which equals

$$(1 + x^8)(1 + x + x^2 - x^3 + x^4 - x^5 + x^6 + x^7). \tag{4.30}$$

Of course, in order to get (4.28) back, we need to add to (4.30) two times the terms for which we changed their signs:

$$(1 + x^8)(1 + x + x^2 - x^3 + x^4 - x^5 + x^6 + x^7) - 2x^{10} + 2x^{11} - 2x^{14} - 2x^{15}. \tag{4.31}$$

This last polynomial equals

$$(1 + x^8)(1 + x + x^2 - x^3 + x^4 - x^5 + x^6 + x^7) + 2x^8(-x^2 + x^3 - x^6 - x^7). \tag{4.32}$$

In general,

$$\sum_{m=0}^{2^r-1}(-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}}x^m = (x^{2^{r-1}}+1)\left(\sum_{m=0}^{2^{r-1}-1}(-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}}x^m\right) + 2x^{2^{r-1}}q(x), \tag{4.33}$$

where $q(x)$ is a polynomial of degree at most $2^{r-1}-1$. We conclude that $x^{2^{r-1}}+1$ does not divide (4.27) and so the claim follows. □

Corollary 4.10 *Let $1 \leq k_1 < \cdots < k_s$ be integers. Let D be the degree of the minimal homogeneous linear recurrence with integer coefficients that $\{S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})\}_{n\in\mathbb{N}}$ satisfies. Then $2^{\lfloor \log_2(k_s)\rfloor} \leq D \leq 2\lfloor (k_1 \vee \cdots \vee k_s)/2\rfloor + 1$.*

Proof: Note that the upper bound follows from Theorem 4.6 while the lower bound is a consequence of Theorem 4.9. □

Note that Corollary 4.10 is an improvement of Theorem 3.1 with respect to the degree D of the minimal homogeneous linear recurrence with integer coefficients that $\{S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})\}_{n\in\mathbb{N}}$ satisfies. From Theorem 3.1 we can only infer that $D \leq 2^r-1$, where $r = \lfloor \log_2(k_s)\rfloor + 1$. However, now we know that $2^{\lfloor \log_2(k_s)\rfloor} \leq D \leq 2\lfloor (k_1 \vee \cdots \vee k_s)/2\rfloor + 1$ and $2\lfloor (k_1 \vee \cdots \vee k_s)/2\rfloor + 1 \leq 2^r - 1$. Also, example 4.7 shows that the upper bound of Corollary 4.10 can be attained. In the next theorem, we show that when k_s (the highest degree) is a power of two, then the lower bound is tight.

Theorem 4.11 *Suppose $1 \leq k_1 < k_2 < \cdots < k_s$ are fixed integers with $k_s = 2^{r-1}$ a power of two. Let $P_{k_1,k_2,\cdots,2^{r-1}}(x)$ be the characteristic polynomial associated to the minimal linear recurrence that $\{S(\sigma_{n,k_1}+\sigma_{n,k_2}+\cdots+\sigma_{n,2^{r-1}})\}_{n\in\mathbb{N}}$ satisfies. Then*

$$P_{k_1,k_2,\cdots,2^{r-1}}(x) = \Phi_{2^r}(x-1) = 2 + \sum_{m=1}^{2^{r-1}}(-1)^m\binom{2^{r-1}}{m}x^m. \tag{4.34}$$

In particular, $\deg(P_{k_1,k_2,\cdots,2^{r-1}}(x)) = 2^{r-1} = 2^{\lfloor \log_2(k_s)\rfloor}$.

Proof: The theorem will follow if we show that $c_0(k_1,k_2,\cdots,2^{r-1}) = 0$, and

$$\sum_{m=0}^{2^r-1}(-1)^{\binom{m}{k_1}+\binom{m}{k_2}+\cdots+\binom{m}{2^{r-1}}}\exp\left(\frac{\pi\sqrt{-1}m}{2^j}\right) = 0, \tag{4.35}$$

for each $j = 1,2,\cdots,r-2$, and

$$\sum_{m=0}^{2^r-1}(-1)^{\binom{m}{k_1}+\binom{m}{k_2}+\cdots+\binom{m}{2^{r-1}}}\exp\left(\frac{\pi\sqrt{-1}m}{2^{r-1}}\right) \neq 0. \tag{4.36}$$

From Theorem 4.9 we know that (4.36) holds true. Now, the coefficient $c_0(k_1, \cdots, k_{s-1}, 2^{r-1})$ is zero if and only if

$$\sum_{m=0}^{2^r-1} (-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_{s-1}}+\binom{m}{2^{r-1}}} = 0. \tag{4.37}$$

From (3.3) we see that the period of $(-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_{s-1}}}$ is a proper divisor of 2^r, so

$$\sum_{m=0}^{2^{r-1}-1} (-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_{s-1}}} = \sum_{m=2^{r-1}}^{2^r-1} (-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_{s-1}}}. \tag{4.38}$$

However,

$$(-1)^{\binom{m}{2^{r-1}}} = \begin{cases} 1, & \text{if } m \leq 2^{r-1}-1 \\ -1, & \text{if } m \geq 2^{r-1}. \end{cases} \tag{4.39}$$

Thus, (4.37) holds and therefore $c_0(k_1, \cdots, k_{s-1}, 2^{r-1}) = 0$. Similarly, the periodicity of $(-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_{s-1}}}$ and $\exp\left(\frac{\pi\sqrt{-1}m}{2^j}\right)$ implies

$$\sum_{m=0}^{2^{r-1}-1} (-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_{s-1}}} \exp\left(\frac{\pi\sqrt{-1}m}{2^j}\right) = \sum_{m=2^{r-1}}^{2^r-1} (-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_{s-1}}} \exp\left(\frac{\pi\sqrt{-1}m}{2^j}\right). \tag{4.40}$$

So, (4.39) and (4.40) imply (4.35). This concludes the proof. □

We conclude this section with the following result, which shows that when k_s is a power of two, then, as n increases, $|S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})|$ is much smaller than 2^n.

Corollary 4.12 *Suppose $1 \leq k_1 < k_2 < \cdots < k_s$ are fixed integers with $k_s = 2^{r-1}$ a power of two. Then, for $0 \leq j \leq 2^r - 1$, $c_j(k_1, \cdots, k_{s-1}, 2^{r-1}) \neq 0$ if and only if j is odd. In particular, $c_0(k_1, \cdots, k_{s-1}, 2^{r-1}) = 0$.*

5 Asymptotic behavior

In this section we discuss the asymptotic behavior of $\{S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})\}_{n\in\mathbb{N}}$. Note that Theorem 3.1 implies that

$$\lim_{n\to\infty} \frac{S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})}{2^n} = c_0(k_1, \cdots, k_s) = \frac{1}{2^r} \sum_{m=0}^{2^r-1} (-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}}. \tag{5.1}$$

Thus, we study $c_0(k_1, \cdots, k_s)$ first.

We already discussed the case of one elementary symmetric polynomial $\{S(\sigma_{n,k})\}_{n\in\mathbb{N}}$, see (4.4):

$$c_0(k) = \frac{2^{w_2(k)-1} - 1}{2^{w_2(k)-1}}. \tag{5.2}$$

For instance, we know that $c_0(k) \geq 0$, and the equality holds if and only if k is a power of two.

The method of inclusion-exclusion can be used to get a formula in the case that we have more than one symmetric polynomial. For example, in the case of two elementary symmetric polynomials $\{S(\sigma_{n,k_1} + \sigma_{n,k_2})\}_{n\in\mathbb{N}}$, we have

$$c_0(k_1, k_2) \;=\; 1 - 2^{1-w_2(k_1)} - 2^{1-w_2(k_2)} + 2^{2-w_2(k_1\vee k_2)} \tag{5.3}$$

The reader can check that this formula implies $c_0(k_1, k_2) \geq 0$, with equality if and only if $w_2(k_1 \vee k_2) = w(k_1) + w_2(k_2)$ and $w_2(k_i) = 1$, where $i = 1$ or $i = 2$. We start the general case with the following lemma.

Lemma 5.1 *Suppose that $1 \leq k_1 < k_2 < \cdots < k_s$ are integers. Define*

$$N(k_{i_1}, \cdots k_{i_j}) = \left\{ m \in \{1, 2, 3, \cdots, 2^r - 1\} : \binom{m}{k_{i_1}} + \cdots + \binom{m}{k_{i_j}} \text{ is odd} \right\}. \tag{5.4}$$

Then,

$$\begin{aligned} \#N(k_1, k_2, \cdots, k_s) \;=\; & \sum_{i=1}^{s} \#N(k_i) - 2\sum_{i_1<i_2} \#\left(N(k_{i_1}) \cap N(k_{i_2})\right) \\ & +4 \sum_{i_1<i_2<i_3} \#\left(N(k_{i_1}) \cap N(k_{i_2}) \cap N(k_{i_2})\right) - \cdots \\ & +(-1)^{s-1}2^{s-1}\#\left(N(k_1) \cap N(k_2) \cap \cdots \cap N(k_s)\right). \end{aligned} \tag{5.5}$$

Proof: Note that

$$\binom{m}{k_1} + \cdots + \binom{m}{k_s} \tag{5.6}$$

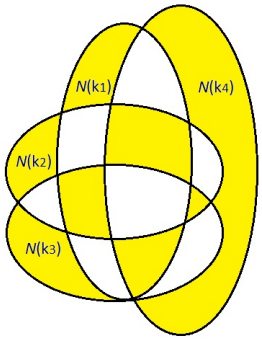
is odd exactly when the amount of odd summands is itself an odd number (trivial). In terms of our sets $N(k_i)$, this implies that $N(k_1, \cdots, k_s)$ is obtained by including all the intersections of an odd amount of the sets $N(k_i)$, while excluding all the intersections of an even amount of them. For example, the case of four k's can be represented by the Venn diagram in Figure 1. In this case, we want to include the shaded regions and exclude the white ones.

We start by adding $\#N(k_1) + \#N(k_2) + \cdots + \#N(k_s)$. Then, we proceed to take out all the intersections of two sets $N(k_i) \cap N(k_j)$, $i \neq j$. In this case, each of them has been added twice in our previous sum. Therefore, to take them out, we should add $-2\#(N(k_1) \cap N(k_2)) - 2\#(N(k_1) \cap N(k_3)) - \cdots - 2\#(N(k_{s-1}) \cap N(k_s))$ to the previous sum. So, now we have

$$\sum_{i=1}^{s} \#N(k_i) - 2\sum_{i_1<i_2} \#(N(k_{i_1}) \cap N(k_{i_2})). \tag{5.7}$$

This takes care of all intersections of two sets. Now, we need to add all intersections of three sets $\#(N(k_{i_1}) \cap N(k_{i_2}) \cap N(k_{i_3}))$. Each of them have been added three times by the

Figure 1: Representation of the case of four k's.



first sum and subtracted six times by the second sum. Thus, in order to add them into the equation, we have to add each of them four times to (5.7). By doing this we have

$$\begin{aligned}\sum_{i=1}^{s}\#N(k_i) \quad &- \quad 2\sum_{i_1<i_2}\#(N(k_{i_1})\cap N(k_{i_2})) \\ &+ \quad 4\sum_{i_1<i_2<i_3}\#(N(k_{i_1})\cap N(k_{i_2})\cap N(k_{i_3})).\end{aligned} \tag{5.8}$$

Continue in this manner and use the identity

$$\sum_{i=1}^{j-1}(-1)^{i-1}2^{i-1}\binom{j}{i} = \begin{cases} 2^{j-1}, & \text{if } j \text{ is even} \\ -2^{j-1}+1, & \text{if } j \text{ is odd,}\end{cases} \tag{5.9}$$

to get the result. □

Now that we have the above lemma, we are ready to state our general formula for $c_0(k_1,\cdots,k_s)$.

Theorem 5.2 *Suppose that $1 \le k_1 < k_2 < \cdots < k_s$ are integers. Then*

$$\begin{aligned}c_0(k_1,\cdots,k_s) \quad &= \quad 1-\sum_{i=1}^{s}2^{1-w_2(k_i)}+\sum_{i_1<i_2}2^{2-w_2(k_{i_1}\vee k_{i_2})} \\ &\quad -\sum_{i_1<i_2<i_3}2^{3-w_2(k_{i_1}\vee k_{i_2}\vee k_{i_3})}+\cdots+(-1)^s 2^{s-w_2(k_1\vee k_2\vee\cdots\vee k_s)}.\end{aligned} \tag{5.10}$$

Proof: Let $r = \lfloor \log_2(k_s) \rfloor + 1$. Note that

$$c_0(k_1,\cdots,k_s) = \frac{2^r - \#N(k_1,\cdots,k_s)}{2^r}. \tag{5.11}$$

Consider the integer k_i. Suppose its 2-adic expansion is $k_i = 2^{a_1}+2^{a_2}+\cdots+2^{a_l}$. Then, by Corollary 4.2 we know that $0 \le m \le 2^r-1$ is such that $\binom{m}{k_i}$ is odd precisely when

$$m = k + \delta_1 2^{b_1} + \delta_2 2^{b_2} + \cdots + \delta_t 2^{b_t}, \tag{5.12}$$

where $\{2^{b_1}, 2^{b_2}, \cdots, 2^{b_t}\} = \{1, 2, 2^2, \cdots, 2^{r-1}\} \backslash \{2^{a_1}, 2^{a_2}, \cdots, 2^{a_l}\}$ and δ_j is either 0 or 1. This implies that

$$\#N(k_i) = 2^{r-w_2(k_i)}. \tag{5.13}$$

Also, (5.12) implies that if $i \neq j$, then

$$\#(N(k_i) \cap N(k_j)) = 2^{r-w_2(k_i \vee k_j)}, \tag{5.14}$$

and in general, if $1 \leq i_1 < i_2 < \cdots i_t \leq s$, then

$$\#(N(k_{i_1}) \cap N(k_{i_2}) \cap \cdots \cap N(k_t)) = 2^{r-w_2(k_{i_1} \vee k_{i_2} \vee \cdots \vee k_{i_t})}. \tag{5.15}$$

Equations (5.11) and (5.15) and Lemma 5.1 imply the theorem. □

Example 5.3 *Suppose* $k_1 = 7$, $k_2 = 9$, $k_3 = 2^{10^5} + 2^{10^4}$, *and* $k_4 = 2^{10^6} + 5$, *then* $c_0(k_1, k_2, k_3, k_4) = 1/4$. *In other words,*

$$\lim_{n \to \infty} \frac{S(\sigma_{n,k_1} + \sigma_{n,k_2} + \sigma_{n,k_3} + \sigma_{n,k_4})}{2^n} = \frac{1}{4}. \tag{5.16}$$

Example 5.4 *Suppose* $k_1 = 31$, $k_2 = 2^{10^4}+64$ *and* $k_3 = 2^{10^4}+32+128$, *then* $c_0(k_1, k_2, k_3) = 45/128$.

We now use the above theorem to provide a family of symmetric polynomials that are not balanced for sufficiently large n. Suppose that k_1 and k_2 are two positive integers. We say that $k_1 \preceq k_2$ if each power of two appearing in the 2-adic expansion of k_1 also appears in the 2-adic expansion of k_2. For example, $10 \preceq 14$ because $10 = 2 + 8$ and $14 = 2 + 4 + 8$.

Corollary 5.5 *Suppose that* $k_1 \preceq k_2 \preceq \cdots \preceq k_s$ *are positive integers. Then,*

$$c_0(k_1, \cdots, k_s) = 1 - \Delta(s) 2^{1-w_2(k_s)} - \sum_{j=1}^{\lfloor s/2 \rfloor} (2^{w_2(k_{2j} - k_{2j-1})} - 1) 2^{1-w_2(k_{2j})}. \tag{5.17}$$

Here $\Delta(s)$ *equals* 0 *if* s *is even and* 1 *otherwise; in other words,* $\Delta(s) = s \mod 2$.

Proof: This follows directly from Theorem 5.2 and the equality $w_2(k_i) = w_2(k_i - k_{i-1}) + w_2(k_{i-1})$. □

Theorem 5.6 *Suppose that* $k_1 \preceq k_2 \preceq \cdots \preceq k_s$ *are positive integers. Then,*

$$c_0(k_1, \cdots, k_s) > 0. \tag{5.18}$$

In particular, $\{S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s}\}_{n \in \mathbb{N}}$ *is asymptotically not balanced.*

Proof: Corollary 5.5 implies that $c_0(k_1, \cdots, k_s) > 0$ if and only if

$$\frac{\Delta(s)}{2^{w_2(k_s)}} + \sum_{j=1}^{\lfloor s/2 \rfloor} \frac{2^{w_2(k_{2j}-k_{2j-1})} - 1}{2^{w_2(k_{2j})}} < \frac{1}{2}. \tag{5.19}$$

Since $k_1 \preceq k_2 \preceq \cdots \preceq k_s$, then we have the inequality

$$w_2(k_i) \geq 1 + w_2(k_{i-1}) \tag{5.20}$$

and the equality

$$w_2(k_i) = w_2(k_i - k_{i-1}) + w_2(k_{i-1}). \tag{5.21}$$

Note that (5.20) and (5.21) imply

$$\begin{aligned}
\frac{\Delta(s)}{2^{w_2(k_s)}} + \sum_{j=1}^{\lfloor s/2 \rfloor} \frac{2^{w_2(k_{2j}-k_{2j-1})} - 1}{2^{w_2(k_{2j})}} &= \frac{\Delta(s)}{2^{w_2(k_s)}} + \frac{1}{2^{w_2(k_1)}} - \frac{1}{2^{w_2(k_2)}} + \sum_{j=2}^{\lfloor s/2 \rfloor} \frac{2^{w_2(k_{2j}-k_{2j-1})} - 1}{2^{w_2(k_{2j})}} \\
&\leq \frac{\Delta(s)}{2^{w_2(k_s)}} + \frac{1}{2^{w_2(k_1)}} - \frac{1}{2^{w_2(k_2)}} + \frac{1}{2^{w_2(k_2)}} \sum_{j=2}^{\lfloor s/2 \rfloor} \frac{1}{2^{2j-3}} \\
&< \frac{1}{2^{w_2(k_1)}} - \frac{1}{2^{w_2(k_2)}} + \frac{1}{2^{w_2(k_2)}} \sum_{j=2}^{\infty} \frac{1}{2^{2j-3}} \\
&= \frac{1}{2^{w_2(k_1)}} + \left(\frac{2}{3} - 1\right) \frac{1}{2^{w_2(k_2)}} < \frac{1}{2}.
\end{aligned} \tag{5.22}$$

This finishes the proof. □

We now turn our attention to the case when $c_0(k_1, \cdots, k_s) = 0$, which happens if and only if 2 is not a root of the characteristic polynomial associated to $\{S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s}\}_{n\in\mathbb{N}}$. In this case

$$\lim_{n\to\infty} \frac{S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})}{2^n} = 0, \tag{5.23}$$

because $|S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})|$ is exponentially smaller than 2^n. However, aside from the size of $S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})$ with respect to 2^n, knowing that $c_0(k_1, \cdots, k_s) = 0$ does not give a real sense of the behavior of $S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})$ as n increases.

Now, when $c_0(k_1, \cdots, k_s) = 0$, the biggest modulus of the roots of the characteristic polynomial associated to $\{S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s}\}_{n\in\mathbb{N}}$ is $2\cos(\pi/2^r)$. This modulus is obtained at the roots $1 + \exp\left(\pi\sqrt{-1}/(2^{r-1})\right)$ and $1 + \exp\left(-\pi\sqrt{-1}/(2^{r-1})\right)$. Thus, as n increases,

$$\frac{S(\sigma_{n,k_1} + \cdots + \sigma_{n,k_s})}{(2\cos(\pi/2^r))^n} \tag{5.24}$$

approaches

$$\frac{c_1(k_1, \cdots, k_s)\left((1 + \exp\left(\frac{\pi\sqrt{-1}}{2^{r-1}}\right)\right)^n + c_{2^r-1}(k_1, \cdots, k_s)\left(1 + \exp\left(-\frac{\pi\sqrt{-1}}{2^{r-1}}\right)\right)^n}{(2\cos(\pi/2^r))^n}. \tag{5.25}$$

Let $c_i = c_i(k_1, \cdots, k_s)$ and $\xi = 1 + \exp\left(\frac{\pi\sqrt{-1}}{2^{r-1}}\right)$. Note that $c_{2^r-1} = \bar{c_1}$. Since

$$1 + \exp\left(\pm\frac{\pi\sqrt{-1}}{2^{r-1}}\right) = 2\cos\left(\frac{\pi}{2^r}\right)\exp\left(\pm\frac{\pi\sqrt{-1}}{2^r}\right), \tag{5.26}$$

then

$$\begin{aligned} c_1\xi^n + \bar{c_1}\bar{\xi}^n &= \frac{(2\cos(\pi/2^r))^n}{2^r}\sum_{m=0}^{2^r-1}(-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}}\times \\ &\quad \left(\exp\left(\frac{(n-2m)\pi\sqrt{-1}}{2^r}\right) + \exp\left(-\frac{(n-2m)\pi\sqrt{-1}}{2^r}\right)\right) \\ &= \frac{(2\cos(\pi/2^r))^n}{2^{r-1}}\sum_{m=0}^{2^r-1}(-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}}\cos\left(\frac{(n-2m)\pi}{2^r}\right). \end{aligned} \tag{5.27}$$

Note that (5.27) is not the zero function, because

$$\sum_{m=0}^{2^r-1}(-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}}\cos\left(\frac{(n-2m)\pi}{2^r}\right) \tag{5.28}$$

is the real part of

$$\exp\left(\frac{\pi\sqrt{-1}n}{2^r}\right)\sum_{m=0}^{2^r-1}(-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}}\exp\left(-\frac{m\pi\sqrt{-1}}{2^{r-1}}\right) \tag{5.29}$$

and from the proof of Theorem 4.10 we know that the sum in (5.29) is a non-zero constant. We conclude that

$$\begin{aligned} S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s}) &= \frac{(2\cos(\pi/2^r))^n}{2^{r-1}}\sum_{m=0}^{2^r-1}(-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}}\cos\left(\frac{(n-2m)\pi}{2^r}\right) \\ &\quad + O\left(\left(2\cos\left(\frac{\pi}{2^{r-1}}\right)\right)^n\right). \end{aligned} \tag{5.30}$$

Remark 4 *Note that*

$$\sum_{m=0}^{2^r-1}(-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}}\cos\left(\frac{(n-2m)\pi}{2^r}\right), \tag{5.31}$$

is not identically zero, regardless if $c_0(k_1, \cdots, k_s)$ is zero. Hence, the asymptotic expansion of $S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})$ is

$$\begin{aligned} S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s}) &= c_0 2^n + \frac{(2\cos(\pi/2^r))^n}{2^{r-1}}\sum_{m=0}^{2^r-1}(-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}}\cos\left(\frac{(n-2m)\pi}{2^r}\right) \\ &\quad + O\left(\left(2\cos\left(\frac{\pi}{2^{r-1}}\right)\right)^n\right). \end{aligned} \tag{5.32}$$

Remark 5 *If $c_0(k_1,\cdots,k_s)=0$, then we define the function $Error_n(k_1,\cdots,k_s)$ as*

$$\begin{aligned} Error_n(k_1,\cdots,k_s) &= \frac{S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})}{(2\cos(\pi/2^r))^n} \\ &\quad -\frac{1}{2^{r-1}}\sum_{m=0}^{2^r-1}(-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}}\cos\left(\frac{(n-2m)\pi}{2^r}\right). \end{aligned} \tag{5.33}$$

We point out that (5.32) is a consequence of Theorem 3.1. Moreover, it is not hard to continue refining this formula and re-write $S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})$ completely in terms of cosine (which will be a restatement of Theorem 3.1). However, we stress that we are proving that

$$\sum_{m=0}^{2^r-1}(-1)^{\binom{m}{k_1}+\cdots+\binom{m}{k_s}}\cos\left(\frac{(n-2m)\pi}{2^r}\right) \tag{5.34}$$

is not identically zero, so the second term of the asymptotic expansion (5.32) is always present. By the discussion presented in section 4 we know that in many cases some of the c_i's, $i\neq 1, 2^r-1$, are zero. Because of this, we write the asymptotic expansion of $S(\sigma_{n,k_1}+\cdots+\sigma_{n,k_s})$ as in (5.32) and decide not to continue with a refinement of it.

Example 5.7 *Consider $k_1=5$, $k_2=9$, and $k_3=12$. The reader can check that in this case $c_0(5,9,12)=0$. By (5.30) we know that as n increases,*

$$\frac{S(\sigma_{n,5}+\sigma_{n,9}+\sigma_{n,12})}{(2\cos(\pi/16))^n} \tag{5.35}$$

approaches

$$\begin{aligned} &\frac{1}{8}\sum_{m=0}^{15}(-1)^{\binom{m}{5}+\binom{m}{9}+\binom{m}{12}}\cos\left(\frac{(n-2m)\pi}{16}\right)= \\ &\frac{1}{8}\left(\sqrt{2}\left(\sqrt{2+\sqrt{2}}-1\right)\cos\left(\frac{n\pi}{16}\right)+\left(2+\sqrt{2}+\sqrt{2\left(2+\sqrt{2}\right)}\right)\sin\left(\frac{n\pi}{16}\right)\right). \end{aligned} \tag{5.36}$$

Here, we simplified the expression using Mathematica. In Figure 2 you can see a graphical representation of this. The dots represents (5.35) and the curve is the one given by (5.36). In Table 1 you can see the error term.

Example 5.8 *Similarly, consider $k_1=2$, $k_2=4$, $k_3=11$, and $k_4=35$. In this case we also have $c_0(2,4,11,35)=0$, so by (5.30) we know that as n increases,*

$$\frac{S(\sigma_{n,2}+\sigma_{n,4}+\sigma_{n,11}+\sigma_{n,35})}{(2\cos(\pi/64))^n} \tag{5.37}$$

approaches

$$\frac{1}{32}\sum_{m=0}^{63}(-1)^{\binom{m}{2}+\binom{m}{4}+\binom{m}{11}+\binom{m}{35}}\cos\left(\frac{(n-2m)\pi}{16}\right). \tag{5.38}$$

In Figure 3 you can see a graphical representation of this, where the dots represents (5.37) and the curve is (5.38). In Table 2 you can see the error term.

Figure 2: Graphical representation of the asymptotic behavior when $k_1 = 5$, $k_2 = 9$, and $k_3 = 12$.

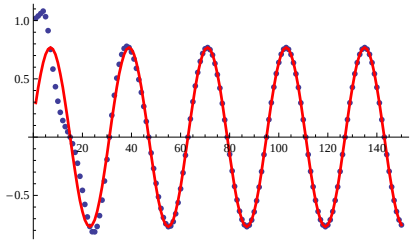


Table 1: The error term for $k_1 = 5$, $k_2 = 9$, and $k_3 = 12$.

n	$\text{Error}_n(5, 9, 12)$
100	0.001530582098
200	$-1.60776038707 \times 10^{-6}$
300	$-9.843230768196 \times 10^{-9}$
400	$1.033957384537 \times 10^{-11}$
500	$6.330222602868 \times 10^{-14}$

Figure 3: Graphical representation of the asymptotic behavior when $k_1 = 2$, $k_2 = 4$, $k_3 = 11$, and $k_4 = 35$.

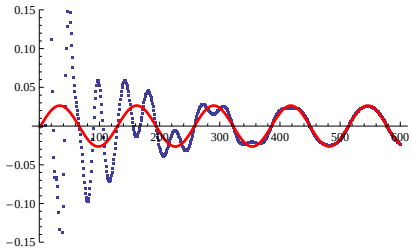


Table 2: The error term for $k_1 = 2$, $k_2 = 4$, $k_3 = 11$, and $k_3 = 35$.

n	$\text{Error}_n(2, 4, 11, 35)$
250	-0.014750
500	-0.0012673
750	-0.000024944
1000	$7.21779483609288 \times 10^{-6}$
1250	$1.01240694303367 \times 10^{-6}$

Acknowledgments. We would like to thank Claude Carlet and Doron Zeilberger for a careful reading of the manuscript and for all their helpful suggestions. A special thank goes to the referee for pointing out relevant references to this work and for improving the presentation of this manuscript.

References

[1] C. Bey and G. M. Kyureghyan, *On Boolean functions with the sum of every two of them being bent*, Des. Codes Cryptogr., **49**, pp. 341–346, 2008.

[2] R. E. Canfield, Z. Gao, C. Greenhill, B. McKay and R. W. Robinson, Asymptotic enumeration of correlation-immune Boolean functions, *Cryptogr. Commun.*, **2**, pp. 111-126, 2010

[3] J. Y. Cai, F. Greeen, and T. Thierauf. On the Correlation of Symmetric Functions. *Theory of Computing Systems*, **29**, pp. 245–258, 1996.

[4] A. Canteaut and M. Videau, *Symmetric Boolean Functions, IEEE Transactions on Information Theory*, **51**, pp. 2791–2807, 2005.

[5] C. Carlet, Boolean Functions for Cryptography and Error Correcting Codes, *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, Eds. Yves Crama and Peter Hammer, Cambridge University Press, 2010.

[6] C. Carlet, X. Zeng, C. Li, and L. Hu, Further properties of several classes of Boolean functions with optimum algebraic immunity, *Des. Codes Cryptogr.*, **52**,pp. 303–338, 2009.

[7] T. Cusick, Y. Li, and P. Stănică, Balanced Symmetric Functions over $GF(p)$. *IEEE Trans. on Information Theory*, **54**, pp. 1304-1307, 2008.

[8] T. Cusick, Y. Li, and P. Stănică, On a conjecture for balanced symmetric Boolean functions, *J. Math. Crypt.*, **3**, pp. 1-18, 2009.

[9] T. Cusick and P. Stănică, *Cryptographic Boolean Functions and Applications*, Academic Press, 2009.

[10] D.J.H. Garling, *A Course in Galois Theory*, Cambridge University Press, 1986.

[11] O. S. Rothaus, On Bent functions. *J. Combin. Theory Ser. A*, **20**, pp. 300-305, 1976.

[12] P. Sarkar and S. Maitra, Balancedness and correlation immunity of symmetric Boolean functions. *Discrete Mathematics*, **307**, pp. 3251-2358, 2007.

[13] J. von zur Gathen and J. Roche, Polynomial with Two Values, *Combinatorica*, **17**, pp. 345-362, 1997.