University of Puerto Rico

Río Piedras Campus

Faculty of Natural Sciences

Department of Mathematics

Diophantine equations of binomials coefficients and exponential sums of symmetric Boolean functions

By

Luisiany Pomales Negrón

A thesis presented for the degree of

Master in Sciences in Mathematics at the University of Puerto Rico, Río Piedras Campus

July 27, 2022

Approved by the Master Committee in Partial Fullfillment of the Requirements for the Degree of Master in Sciences in Mathematics at the University of Puerto Rico

---

Dr. Luis A. Medina

Department of Mathematics

Thesis Advisor

Dr. M. Reza Emamy-K

Department of Mathematics

Thesis Committee Member

Dr. Heeralal Janwa

Department of Mathematics

Thesis Committee Member

Abstract of M.S. Thesis Presented to the Graduated School of the University of Puerto Rico, Río Piedras Campus in Partial Fulfilments of the Requirements for the Master Degree in Sciences in Mathematics

Diophantine Equations of Binomial Coefficients and exponential sums of symmetric Boolean functions

In this thesis, we study the link that exists between solutions to Diophantine equations that involve binomial coefficient over a bounded set of integers and exponential sums of perturbations of symmetric Boolean functions. This link was established by Castro and Medina in [3]. They extended the concepts of trivially balanced functions and sporadic balanced functions to these perturbations. This problem also is similar to the interesting problem of bisecting binomials which was first studied by Ionascu, Stanica and Martinsen [11], but our study is from the point of view of the theory of exponential sums of symmetric Boolean functions.

Here in this thesis it is presented an identity of two exponential sums of perturbations of two different symmetric Boolean functions. We also study the balancedness of these perturbations of fixed degree when the number of variables grows and we show that these balanced perturbation of fixed degree do not exist when the number of variables grow based on an observation of Canteaut and Videau. Finally, we present some examples of sporadic balanced perturbations and their corresponding Diophantine equation with binomial coefficients.

Dedicated to my former mentors and high school math teachers, Pedro M. Pérez and José E. Cruz, who always believed in my potential and inspired me on studying mathematics and submerging into this beautiful world.

Acknowledgements

I would like to thank to all my family members and friends who always encouraged and believed in me, especially to my loving father who always trusted in my potential.

I would like to give a special thank to my advisor and mentor, Dr. Luis Medina for his patience, advice and guidance throughout this process and to the members of the thesis committee, Dr. Heeralal Janwa and Dr. Reza-Emamy for their review and feedback for the final version of this thesis.

List of Symbols

$\mathbb{N}$	set of natural numbers
$\mathbb{Z}_{\geq 0}$	set of non-negative integers
$\mathbb{Q}$	set of rational numbers
$\mathbb{Q}[x]$	set of polynomials over $\mathbb{Q}$
$\lfloor x \rfloor$	floor function of x
$\mathbb{F}_2$	binary field
$S(F)$	exponential sum of the Boolean function F over $\mathbb{F}_2$
$C(F)$	correlation of the Boolean function F
$\sim$	equivalent to
$\boldsymbol{e}_{n,k}$	symmetric function in n variables of degree k
S_n	symmmetric group on a finite set of n symbols
C_n	cyclic group on a finite set of n symbols
$\binom{n}{k}$	binomial coefficient
$\det(A)$	determinant of a matrix A
$A^\star$	Hermitian conjugate of the matrix A
$\mathrm{wt}(F)$	Hamming weight of a Boolean function F
$\lvert A \rvert$	cardinality of a finite set A
$\Phi_n(x)$	nth cyclotomic polynomial
$\mathrm{Re}(z)$	real part of the complex number z
$\bar{z}$	conjugate of the complex number z
$\arg(z)$	argument of the complex number z

Contents

List of Tables

Chapter 1

Introduction

The study that is presented in this thesis lies on a beautiful branch of combinatorics known as the theory of Boolean functions. This area has many applications in other branches of mathematics like number theory and cryptography as well as in branches outside of these discipline such as electrical engineering and in coding theory.

In this work, we study symmetric Boolean functions. However, the symmetric property may imply that implementations of these functions are vulnerable to attacks. For that reason, we also study of perturbations of symmetric Boolean functions because, in general, these functions happen not to be symmetric.

In this chapter we are going to give a preliminary introduction to the theory of Boolean functions and some of its concepts. Here is where we present a closed formula for the exponential sum of a symmetric Boolean functions proved by Cai et al in [1]. In section 2, we discuss the fact that exponential sums of symmetric Boolean function satisfy homogeneous linear recurrences with integer coefficients that follows directly from the proof of Cai et al [1].

Finally, in section 3, we provide a proof that sequences of exponential sums and its perturbation satisfies the same linear recurrence with integer coefficients using the fact that the linear combination of sequences that satisfy a homogeneous linear recurrence with integer coefficients satisfy the same linear recurrence.

1.1 Preliminaries

Let $\mathbb{F}_2 = \{0, 1\}$ be the binary field and let

$$\mathbb{F}_2^n = \{\boldsymbol{X} = (X_1, X_2, \cdots, X_n) | X_i \in \mathbb{F}_2, i = 1, 2, \cdots, n\}. \tag{1.1}$$

An n-variable *Boolean function* $F(X_1, X_2, \cdots, X_n) = F(\mathbf{X})$ is a function $F : \mathbb{F}_2^n \to \mathbb{F}_2$.

Example 1.1.1. Consider the Boolean function defined as $F(X_1, X_2, X_3) = X_1 + X_2 + X_3$. Computing its values (arithmetic done over $\mathbb{F}_2$), we obtain

$$\begin{aligned} F(0,0,0) &= 0, & F(1,1,0) &= 0, \\ F(1,0,0) &= 1, & F(1,0,1) &= 0, \\ F(0,1,0) &= 1, & F(0,1,1) &= 0, \\ F(0,0,1) &= 1, & F(1,1,1) &= 1. \end{aligned}$$

Example 1.1.2. Consider the Boolean polynomial $F(X_1, X_2, X_3) = X_1X_2X_3 + X_1 + X_3$. Then,

$$\begin{aligned} F(0,0,0) &= 0, & F(1,1,0) &= 1, \\ F(1,0,0) &= 1, & F(1,0,1) &= 0, \\ F(0,1,0) &= 0, & F(0,1,1) &= 1, \\ F(0,0,1) &= 1, & F(1,1,1) &= 1. \end{aligned}$$

It is well known in the theory of Boolean functions that every Boolean function F in n variables can be uniquely expressed in the form

$$F(X_1, X_2, \cdots, X_n) = \bigoplus_{\boldsymbol{a}=(a_1,\ldots,a_n)\in\mathbb{F}_2^n} \lambda_{\boldsymbol{a}} \prod_{j=1}^{n} X_j^{a_j}, \tag{1.2}$$

where $\bigoplus$ represents the fact that the arithmetic is done over $\mathbb{F}_2$ and $\lambda_a \in \mathbb{F}_2$. The

expression (1.2) is called the *algebraic normal form* of the Boolean function F (ANF, for short).

In certain applications, especially the ones related to cryptography, it is important for Boolean functions to be balanced. We say that a Boolean function F is a *balanced* function if F is a function for which the number of 0's and 1's in its truth table (output table) are the same. For example, the function in Example 1.1.1 is a balanced function, whereas the function of Example 1.1.2 is not balanced.

The balancedness of Boolean functions can be studied from the point of view of Hamming weights or from the point of view of exponential sums. The *Hamming weight* of F, which is denoted as $\mathrm{wt}(F)$ is the number of 1's that appear in the truth table of F. Observe that the balancedness of F implies that $\mathrm{wt}(F) = 2^{n-1}$.

Let F be a Boolean function. We define the *exponential sum* associated to F over $\mathbb{F}_2$ as

$$S(F) = \sum_{\mathbf{X} \in \mathbb{F}_2^n} (-1)^{F(\mathbf{X})}. \tag{1.3}$$

Observe that F is balanced if and only if $S(F) = 0$. In this work we study balancedness from the point of view of exponential sums.

Example 1.1.3. Consider the Boolean function of Example 1.1.1. Thus, the exponential sum associated to F over $\mathbb{F}_2$ is given by

$$\begin{aligned} S(F) = \sum_{\mathbf{X} \in \mathbb{F}_2^3} (-1)^{F(\mathbf{X})} &= (-1)^{F(0,0,0)} + (-1)^{F(1,0,0)} + (-1)^{F(0,1,0)} + (-1)^{F(0,0,1)} + \\ &\quad (-1)^{F(1,1,0)} + (-1)^{F(1,0,1)} + (-1)^{F(1,0,1)} + (-1)^{F(1,1,1)} \\ &= 1 + (-1) + (-1) + (-1) + 1 + 1 + 1 + (-1) = 0. \end{aligned}$$

Therefore, F is balanced (that can be easily seen from its truth table).

Example 1.1.4. Consider the Boolean function of Example 1.1.2. The exponential sum

associated to F over $\mathbb{F}_2$ is given by

$$\begin{aligned} S(F) = \sum_{\mathbf{X}\in\mathbb{F}_2^3} (-1)^{F(\mathbf{X})} &= (-1)^{F(0,0,0)} + (-1)^{F(1,0,0)} + (-1)^{F(0,1,0)} + (-1)^{F(0,0,1)} + \\ &\quad (-1)^{F(1,1,0)} + (-1)^{F(1,0,1)} + (-1)^{F(1,0,1)} + (-1)^{F(1,1,1)} \\ &= 1 + (-1) + 1 + (-1) + (-1) + 1 + (-1) + (-1) = -2. \end{aligned}$$

Thus, F is not balanced.

Define the *correlation* of F, denoted as $C(F)$, as the difference between the number of times F gives a value of 0 minus the number of times F gives us the values of 1 in the truth table divided by 2^n, that is

$$C(F) = \frac{1}{2^n} \sum_{\mathbf{X}\in\mathbb{F}_2^n} (-1)^{F(\mathbf{X})}. \tag{1.4}$$

The correlation of a Boolean function F measures how a Boolean function F behave in terms of its outputs. If $C(F) = 1$, then F agree in all its 2^n outputs. If $C(F) = 0$, then the number of values of which F agree and disagree are equal (in such case F is balanced). If $|C(F)| < 1$, but $C(F) \neq 0$, then F almost agree in all its output points. It is easy to see that $S(F) = 2^n C(F)$.

Boolean functions are used vastly because of their cryptographic implementations. However, their applications are challenging due to memory restrictions of current technology. Because of this, symmetric Boolean functions are excellent candidates for efficient implementations (but they might not be robust from the point of view of security). We say that a Boolean function F is *symmetric* if it is invariant under the action of the symmetric group S_n on $\mathbb{F}_2$, that is

$$F(X_1, X_2, \cdots, X_n) = F(X_{\sigma(1)}, X_{\sigma(2)}, \cdots, X_{\sigma(n)}), \tag{1.5}$$

where $\sigma \in S_n$ is any permutation in n symbols. Otherwise, we say that the function is *not symmetric*. For example, the Boolean function on Example 1.1.1 is symmetric. On the

other hand, the function on Example 1.1.2 is not symmetric, because if $(X_1, X_2, X_3) = (1, 1, 0)$, then

$$F(X_1, X_2, X_3) = F(1, 1, 0) = 1 \neq F(X_1, X_3, X_2) = F(1, 0, 1) = 0. \tag{1.6}$$

It is well-known that every symmetric Boolean function can be expressed as a linear combination of elementary symmetric Boolean polynomials, that is, if $F(\mathbf{X})$ is a symmetric Boolean function, then $F(\mathbf{X})$ can be written as

$$F(\boldsymbol{X}) = \boldsymbol{e}_{n,k_1}(\boldsymbol{X}) + \boldsymbol{e}_{n,k_2}(\boldsymbol{X}) + \cdots + \boldsymbol{e}_{n,k_s}(\boldsymbol{X}), \tag{1.7}$$

for $1 < k_1 < \cdots < k_s$ are non-negative integers and where $\mathbf{X} = (X_1, X_2, \cdots, X_n)$, and $\boldsymbol{e}_{n,k}(\mathbf{X})$ is defined as

$$\boldsymbol{e}_{n,k}(X_1, X_2, \cdots, X_n) = \begin{cases} 1, & \text{for } k = 0, \\ \sum_{i_1 < i_2 < i_3 < \cdots < i_k} X_{i_1} X_{i_2} X_{i_3} \cdots X_{i_k}, & \text{for } 0 < k \leq n, \\ 0, & \text{for } k \geq n + 1. \end{cases} \tag{1.8}$$

For example,

$$\boldsymbol{e}_{4,3} = X_1X_2X_3 + X_1X_2X_4 + X_1X_3X_4 + X_2X_3X_4. \tag{1.9}$$

For the simplicity in the writing, we express $\boldsymbol{e}_{n,k}(\boldsymbol{X})$ as $\boldsymbol{e}_{n,k}$ and (1.7) as $\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]}$. For example,

$$\boldsymbol{e}_{3,[1,2]} = \boldsymbol{e}_{3,1} + \boldsymbol{e}_{3,2} = X_1 + X_2 + X_3 + X_1X_2 + X_1X_3 + X_2X_3. \tag{1.10}$$

Consider the symmetric Boolean function $\boldsymbol{e}_{n,k}$. Fix $k \geq 2$ and let n vary. Consider the sequence $\{S(\boldsymbol{e}_{n,k})\}$ of exponential sums. Define

$$A_{n,\ell} = \{(X_1, X_2, X_3, \cdots, X_n) \in \mathbb{F}_2^n \mid w_2(X_1, X_2, X_3, \cdots, X_n) = \ell\}, \tag{1.11}$$

where $w_2(\mathbf{X})$ is the Hamming weight of the tuple $\mathbf{X}$, which is the number of 1's in the entries of $\mathbf{X}$. It is easy to see that $|A_{n,\ell}| = \binom{n}{\ell}$ and $\boldsymbol{e}_{n,k}(\boldsymbol{X}) = \binom{\ell}{k}$ for every $\mathbf{X} \in A_{n,\ell}$. Then we can rewrite our initial expression of the exponential sum 1.3 as

$$S(\boldsymbol{e}_{n,k}) = \sum_{\mathbf{X}\in\mathbb{F}_2^n} (-1)^{\boldsymbol{e}_{n,k}(\mathbf{X})} = \sum_{\ell=0}^{n} \sum_{w_2(\mathbf{X})=\ell} (-1)^{\binom{\ell}{k}} = \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{k}} \binom{n}{\ell}. \tag{1.12}$$

Therefore,

$$S(\boldsymbol{e}_{n,k}) = \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{k}} \binom{n}{\ell}. \tag{1.13}$$

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

$$S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]}) = \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{k_1}+\cdots+\binom{\ell}{k_s}} \binom{n}{\ell}. \tag{1.14}$$

Notice that the left hand side of equation (1.14) does not make sense for values of n less than k_s, while the right hand side does. So, throughout this thesis we will make the convention of defining $S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]})$ be as the sum in the right hand side even for values of n that are less that k_s.

1.2 The Linear Recurrence

Castro and Medina proved in [4] (by induction) that sequences of exponential sums $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]})\}$ satisfy linear recurrences. We are going to present the proof of this result by another approach: by using elementary linear algebra. With the aid of computers, Castro and Medina conjectured and later proved that the sequence of exponential sums $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]})\}$ satisfy the homogeneous linear recurrence with integer coefficients given by

$$x_n = \sum_{\ell=1}^{2^r-1} (-1)^{\ell-1} \binom{2^r}{\ell} x_{n-\ell}, \tag{1.15}$$

where $1 \leq k_1 < k_2 < ... < k_s$ are fixed integers and $r = \lfloor \log_2(k_s) \rfloor + 1$.

For example, the sequence of exponential sums $\{S(\boldsymbol{e}_{n,15})\}_{n\geq 1}$ satisfies the homogeneous

linear recurrence given by

$$\begin{aligned} x_n &= 16x_{n-1} - 120x_{n-2} + 560x_{n-3} - 1820x_{n-4} + 4368x_{n-5} - 8008x_{n-6} \\ &\quad +11440x_{n-7} - 12870x_{n-8} + 11440x_{n-9} - 8008x_{n-10} + 4368x_{n-11} \\ &\quad -1820x_{n-12} + 560x_{n-13} - 120x_{n-14} + 16x_{n-15} \end{aligned}$$

with initial conditions

$$\begin{array}{lll} x_1 = 2, & x_6 = 64, & x_{11} = 2048, \\ x_2 = 4, & x_7 = 128, & x_{12} = 4096, \\ x_3 = 8, & x_8 = 256, & x_{13} = 8192, \\ x_4 = 16, & x_9 = 512, & x_{14} = 16384, \\ x_5 = 32, & x_{10} = 1024, & x_{15} = 32766. \end{array}$$

The fact that the sequence $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]})\}_{n\geq 1}$ satisfies the recurrence (1.15) follows from the following theorem of J. Cai et al [1]. The proof provided is inspired by their proof.

Theorem 1.2.1 (Cai, Green and Thierauf [1])**.** *Let $1 \leq k_1 < k_2 < \cdots < k_s$ be fixed integers and let $r = \lfloor \log_2(k_s) \rfloor + 1$. Then the value of the exponential sum $S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]})$ is given by*

$$S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]}) = \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{k_1}+\cdots+\binom{\ell}{k_s}} \binom{n}{\ell} \tag{1.16}$$

$$= c_0(k_1,\cdots,k_s)2^n + \sum_{j=1}^{2^r-1} c_j(k_1,\cdots,k_s)(1+\zeta_j^{-1})^n, \tag{1.17}$$

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

$$c_j(k_1,\cdots,k_s) = \frac{1}{2^r} \sum_{\ell=0}^{2^r-1} (-1)^{\binom{\ell}{k_1}+\cdots+\binom{\ell}{k_s}} \zeta_j^{\ell}. \tag{1.18}$$

Proof. This proof relies on linear algebra. Let $r = \lfloor \log_2(k_s) \rfloor + 1$. Notice that 2^r is a

period for $\binom{\ell}{k_1} + \cdots + \binom{\ell}{k_s} \pmod 2$ because

$$\binom{2^r j + \ell}{k_i} \equiv \binom{\ell}{k_i} \pmod 2 \tag{1.19}$$

for all non negative integers j, $\ell = 1, 2, 3, \cdots, 2^r - 1$ and $i = 1, 2, 3, \cdots, s$. Thus, we can partition the sum

$$\sum_{\ell=0}^{n} (-1)^{\binom{\ell}{k_1}+\cdots+\binom{\ell}{k_s}} \binom{n}{\ell} \tag{1.20}$$

into 2^r sums, one for each remainder modulo 2^r. Doing so, we can rewrite our initial expression for the exponential sum $S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]})$ as

$$S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]}) = \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{k_1}+\cdots+\binom{\ell}{k_s}} \binom{n}{\ell} = \sum_{\ell=0}^{2^r-1} (-1)^{\binom{\ell}{k_1}+\cdots+\binom{\ell}{k_s}} a_{n,r,\ell}, \tag{1.21}$$

where $a_{n,r,\ell}$ is defined as

$$a_{n,r,\ell} = \sum_{j \equiv \ell \pmod{2^r}} \binom{n}{j}. \tag{1.22}$$

Now, using the elementary binomial identity

$$\binom{n-1}{j} + \binom{n-1}{j-1} = \binom{n}{j} \tag{1.23}$$

we rewrite $a_{n,r,\ell}$ as

$$\begin{aligned} a_{n,r,\ell} = \sum_{j \equiv \ell \pmod{2^r}} \binom{n}{j} &= \sum_{j \equiv \ell \pmod{2^r}} \left[\binom{n-1}{j} + \binom{n-1}{j-1} \right] \\ &= \sum_{j \equiv \ell \pmod{2^r}} \binom{n-1}{j} + \sum_{j-1 \equiv \ell-1 \pmod{2^r}} \binom{n-1}{j-1} \\ &= a_{n-1,r,\ell} + a_{n-1,r,\ell-1}. \end{aligned}$$

Define the vector $\mathbf{a}(n)$ as

$$\mathbf{a}(n) = \begin{pmatrix} a_{n,r,0} \\ a_{n,r,1} \\ a_{n,r,2} \\ \vdots \\ a_{n,r,2^r-1} \end{pmatrix}. \tag{1.24}$$

Then

$$\mathbf{a}(n) = \mathbf{M}\mathbf{a}(n-1) \tag{1.25}$$

where $\mathbf{M}$ is a $2^r \times 2^r$ matrix given by

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots 0 & 1 \\ 1 & 1 & 0 & 0 & \cdots 0 & 0 \\ 0 & 1 & 1 & 0 & \cdots 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots 1 & 1 \end{pmatrix}. \tag{1.26}$$

It is easy to show, inductively, that $\mathbf{a}(n) = \mathbf{M}^n\mathbf{a}(0)$, for all positive integers n, where $a_{0,r,0} = 1$ and $a_{n,r,0} = 0$. We now proceed on computing the matrix $\mathbf{M}^n$. We can compute $\mathbf{M}^n$ by diagonalizing the matrix $\mathbf{M}$. To do this, we compute the characteristic polynomial $x\mathbf{I} - \mathbf{M}$ and equal it to 0. That will give us the eigenvalues of the matrix $\mathbf{M}$. The matrix $x\mathbf{I} - \mathbf{M}$ is given by

$$x\mathbf{I} - \mathbf{M} = \begin{pmatrix} x-1 & 0 & 0 & \cdots & 0 & -1 \\ -1 & x-1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & x-1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & x-1 \end{pmatrix}. \tag{1.27}$$

Expanding the expression $\det(x\mathbf{I}-\mathbf{M})$ by cofactors we get

$$\begin{aligned}
\det(x\mathbf{I}-\mathbf{M}) &= \begin{vmatrix} x-1 & 0 & 0 & \cdots & 0 & -1 \\ -1 & x-1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & x-1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & x-1 \end{vmatrix} \\
&= (x-1)\begin{vmatrix} x-1 & 0 & 0 & \cdots & 0 \\ -1 & x-1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots -1 & x-1 \end{vmatrix} + \begin{vmatrix} -1 & x-1 & 0 & \cdots & 0 \\ 0 & -1 & x-1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{vmatrix} \\
&= (x-1)^{2^r} + (-1)^{2^r-1} \\
&= (x-1)^{2^r} - 1 \\
&= \prod_{i=0}^{r} \Phi_{2^i}(x-1) \\
&= \Phi_1(x-1)\Phi_2(x-1)\Phi_4(x-1)\Phi_8(x-1)\cdots\Phi_{2^r}(x-1).
\end{aligned}$$

Thus, eigenvalues of $\mathbf{M}$ are given by $\lambda_0, \lambda_1, \lambda_2, \cdots, \lambda_{2^r-1}$, where $\lambda_j = 1+\zeta_j^{-1}$ and $\zeta_j = \exp\left(\frac{\pi\sqrt{-1}j}{2^{r-1}}\right)$. Observe that the corresponding eigenvector to λ_j is $(1, \zeta_j^{-1}, \cdots, \zeta_j^{-(2^r-1)})^T$. Such eigenvectors are linearly independent, so we can actually diagonalize the matrix $\mathbf{M}$. The diagonalizing matrix $\mathbf{V}$ is

$$\mathbf{V} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \zeta_1^{-1} & \zeta_2^{-1} & \cdots & \zeta_{2^r-1}^{-1} \\ 1 & \zeta_1^{-2} & \zeta_2^{-2} & \cdots & \zeta_{2^r-1}^{-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \zeta_1^{-(2^r-1)} & \zeta_2^{-(2^r-1)} & \cdots & \zeta_{2^r-1}^{-(2^r-1)} \end{pmatrix} \tag{1.28}$$

and the diagonal matrix Δ is given by

$$\Delta = \begin{pmatrix} \lambda_0 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_1 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{2^r-1} \end{pmatrix}. \tag{1.29}$$

We will verify now that the matrix $\frac{1}{\sqrt{2^r}}\mathbf{V}$ is a unitary matrix (conjugate transpose of its inverse), that is, $\mathbf{V}\mathbf{V}^* = \mathbf{V}^*\mathbf{V} = 2^r\mathbf{I}$, where $\mathbf{V}^*$ denotes the Hermitian conjugate of $\mathbf{V}$ (the matrix obtained by taking the transpose of $\mathbf{V}$ and the complex conjugate of each entry), i.e., $\mathbf{V}^* = \overline{\mathbf{V}}^T$ and $\mathbf{I}$ is the identity matrix. Calculating the Hermitian conjugate of the matrix $\mathbf{V}$, yields the following matrix

$$\mathbf{V}^* = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \zeta_1 & \zeta_1^2 & \cdots & \zeta_1^{(2^r-1)} \\ 1 & \zeta_2 & \zeta_2^2 & \cdots & \zeta_2^{(2^r-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \zeta_{2^r-1} & \zeta_{2^r-1}^2 & \cdots & \zeta_{2^r-1}^{(2^r-1)} \end{pmatrix}. \tag{1.30}$$

Therefore,

$$\mathbf{V}\mathbf{V}^* = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \zeta_1^{-1} & \zeta_2^{-1} & \cdots & \zeta_{2^r-1}^{-1} \\ 1 & \zeta_1^{-2} & \zeta_2^{-2} & \cdots & \zeta_{2^r-1}^{-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \zeta_1^{-(2^r-1)} & \zeta_2^{-(2^r-1)} & \cdots & \zeta_{2^r-1}^{-(2^r-1)} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \zeta_1 & \zeta_1^2 & \cdots & \zeta_1^{(2^r-1)} \\ 1 & \zeta_2 & \zeta_2^2 & \cdots & \zeta_2^{(2^r-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \zeta_{2^r-1} & \zeta_{2^r-1}^2 & \cdots & \zeta_{2^r-1}^{(2^r-1)} \end{pmatrix}$$

$$= \begin{pmatrix} 2^r & 0 & 0 & \cdots & 0 \\ 0 & 2^r & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 2^r \end{pmatrix} = 2^r \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} = 2^r \mathbf{I}.$$

In a similar way, we obtain $\mathbf{V}^*\mathbf{V} = 2^r\mathbf{I}$, so we have shown that the matrix $\frac{1}{\sqrt{2^r}}\mathbf{V}$ is unitary. In the above computation, we used the identity

$$\sum_{j=0}^{2^r-1} \zeta_j^k = \begin{cases} 2^r, & \text{if } k = 0 \\ 0, & \text{otherwise.} \end{cases} \tag{1.31}$$

So we obtain, $\mathbf{M} = \frac{1}{2^r}\mathbf{V}^*\Delta\mathbf{V}$ which implies that $\mathbf{M}^n = \frac{1}{2^r}\mathbf{V}^*\Delta^n\mathbf{V}$. Therefore,

$$\mathbf{a}(n) = \mathbf{M}^n\mathbf{a}(0) = \big(\frac{1}{2^r}\mathbf{V}^*\Delta^n\mathbf{V}\big)\mathbf{a}(0) = \frac{1}{2^r}\mathbf{V}^*\Delta^n\mathbf{V}\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} = \frac{1}{2^r}\mathbf{V}^*\Delta^n\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}$$

$$= \frac{1}{2^r}\mathbf{V}^*\begin{pmatrix}\lambda_0^n\\\lambda_1^n\\\vdots\\\lambda_{2^r-1}^n\end{pmatrix}$$

$$= \begin{pmatrix}\frac{1}{2^r}V_0\\\frac{1}{2^r}V_1\\\vdots\\\frac{1}{2^r}V_{2^r-1}\end{pmatrix},$$

where

$$V_\ell = \sum_{j=0}^{2^r-1}\zeta_j^\ell\lambda_j^n, \tag{1.32}$$

for $0 \le \ell \le 2^r - 1$. Hence,

$$a_{n,r,\ell} = \frac{1}{2^r}\sum_{j=0}^{2^r-1}\zeta_j^\ell\lambda_j^n. \tag{1.33}$$

Therefore, the value of the exponential sum $S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]})$ is

$$\begin{aligned}
S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]}) &= \sum_{\ell=0}^{2^r-1}(-1)^{\binom{\ell}{k_1}+\cdots+\binom{\ell}{k_s}}a_{n,r,\ell} \\
&= \sum_{\ell=0}^{2^r-1}(-1)^{\binom{\ell}{k_1}+\cdots+\binom{\ell}{k_s}}\left(\frac{1}{2^r}\sum_{j=0}^{2^r-1}\zeta_j^{\ell}\lambda_j^n\right) \\
&= \sum_{j=0}^{2^r-1}\left(\frac{1}{2^r}\sum_{i=0}^{2^r-1}(-1)^{\binom{\ell}{k_1}+\cdots+\binom{\ell}{k_s}}\zeta_j^{\ell}\right)(1+\zeta_j^{-1})^n \\
&= \sum_{j=0}^{2^r-1}c_j(k_1,\cdots,k_s)(1+\zeta_j^{-1})^n \\
&= c_0(k_1,\cdots,k_s)2^n + \sum_{j=1}^{2^r-1}c_j(k_1,\cdots,k_s)(1+\zeta_j^{-1})^n
\end{aligned}$$

where $c_j(k_1,\cdots,k_s)$ is defined as

$$c_j(k_1,\cdots,k_s) = \frac{1}{2^r}\sum_{\ell=0}^{2^r-1}(-1)^{\binom{\ell}{k_1}+\cdots+\binom{\ell}{k_s}}\zeta_j^{\ell}. \tag{1.34}$$

This concludes the proof. (Q.E.D.)

From Theorem 1.2.1, it is now evident that the sequence of exponential sums $\{S(\boldsymbol{e}_{n,[k_1.\cdots,k_s]})\}$ satisfies (1.15). Moreover, the roots of the characteristic polynomial associates to the linear recurrence (1.15) are all distinct and the polynomial is given by

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

One should note even though the sequence $\{S(\boldsymbol{e}_{n,[k_1.\cdots,k_s]})\}$ satisfies (1.15), in some instances (1.15) may not necessarily be the minimal linear recurrence with integer coefficients that $\{S(\boldsymbol{e}_{n,[k_1.\cdots,k_s]})\}$ satisfies. For instance, consider the sequence $\{S(\boldsymbol{e}_{n,[2,9]})\}_{n\geq 1}$.

This sequence satisfies the linear recurrence

$$\begin{aligned} x_n &= 16x_{n-1} - 120x_{n-2} + 560x_{n-3} - 1820x_{n-4} + 4368x_{n-5} - 8008x_{n-6} \\ &\quad +11440x_{n-7} - 12870x_{n-8}, 11440x_{n-9} - 8008x_{n-10} + 4368x_{n-11} \\ &\quad -1820x_{n-12} + 560x_{n-13} - 120x_{n-14} + 16x_{n-15}, \end{aligned}$$

but its minimal linear recurrence is given by

$$\begin{aligned} x_n &= 10x_{n-1} - 46x_{n-2} + 128x_{n-3} - 238x_{n-4} + 308x_{n-5} - 280x_{n-6} \\ &\quad +176x_{n-7} - 74x_{n-8} + 20x_{n-9} - 4x_{n-10}, \end{aligned}$$

with initial values

$$\begin{aligned} x_1 &= 2, & x_6 &= -8, \\ x_2 &= 2, & x_7 &= 0, \\ x_3 &= 0, & x_8 &= 16, \\ x_4 &= -4, & x_9 &= 30, \\ x_5 &= -8, & x_{10} &= 12. \end{aligned}$$

The *minimal linear recurrence* is the linear recurrence of *least* degree for which a sequence satisfy. In their study of the homogeneous linear recurrence (1.15), Castro and Medina [4] gave some tight improvements in the degree of the minimal linear recurrence with integer coefficients that $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]})\}$ satisfy. In particular, they provided lower and upper bounds to the degree of the minimal linear recurrence. As part of their study, Castro and Medina studied the asymptotic behaviour of $S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]})$ and they introduced the concepts of *asymptotically balancedness* of Boolean functions. These concepts were used to show that a conjecture of Cusick, Li and Stanica [8] is true asymptotically. This result was later re-established by Guo, Gao and Zhao in [10].

Some of the Castro and Medina's results in [4] were extended to Boolean polynomials of

the form $\boldsymbol{e}_{n,[k_1.\cdots,k_s]}+F(\mathbf{X})$, where F is a Boolean polynomial in j variables with j fixed in [5]. In [5], they also proved that the sequences $\{S(\boldsymbol{e}_{n,[k_1.\cdots,k_s]})\}$ and $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]}+F(\mathbf{X}))\}$ satisfy the same linear recurrence with integers coefficients as we will see in the next section.

1.3 Linear Recurrences of Perturbations

A *perturbation* of a symmetric Boolean function $\boldsymbol{e}_{n,[k_1,\cdots,k_s]}$, is a Boolean polynomial of the form $\boldsymbol{e}_{n,[k_1,\cdots,k_s]}+F(\mathbf{X})$, where F is any Boolean polynomial in the variables $X_1,\cdots,X_j$ (j is fixed) and $j<n$. These perturbations of symmetric Boolean functions are, in principle, not longer symmetric, but the symmetry of the underlying function can be exploited in order to make fast calculations. These perturbations were first studied by Castro and Medina [5]. They showed that exponential sums of symmetric Boolean functions and their perturbations satisfy the same homogeneous linear recurrence with integer coefficients. They also provided a closed formula for the exponential sum of a perturbation of a symmetric Boolean function which eventually helped them in the study of the asymptotic behavior of the sequences $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]}+F(\mathbf{X}))\}$. However, the asymptotic behavior of these sequences is beyond the scope of this writing. For more details about that topic, the reader is invited to read [5].

In this section, we are going to be studying these perturbations. As part of our study, we are going to present some of the results about these perturbations of Boolean symmetric functions obtained by Castro and Medina in [5]. Among those results, we include the closed formula for the exponential sum of these perturbations. In addition, we include a proof that both sequences, $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]})\}$ and $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]}+F(\mathbf{X}))\}$, satisfy the homogeneous linear recurrence given by

$$x_n=\sum_{\ell=1}^{2^r-1}(-1)^{\ell-1}\binom{2^r}{\ell}x_{n-\ell}, \tag{1.35}$$

where $r=\lfloor\log_2(k_s)\rfloor+1$.

Recall that the exponential sum of F over $\mathbb{F}_2$ is given by

$$S(F) = \sum_{\boldsymbol{X} \in \mathbb{F}_2^n} (-1)^{F(\mathbf{X})}. \tag{1.36}$$

In [4], Castro and Medina showed that the sequence of exponential sums $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]})\}_{n\geq 1}$, where $1 \leq k_1 < \cdots < k_s$ are fixed integers, satisfies the homogeneous linear recurrence with integer coefficients given by

$$x_n = \sum_{\ell=1}^{2^r-1} (-1)^{\ell-1} \binom{2^r}{\ell} x_{n-\ell}, \tag{1.37}$$

where $r = \lfloor \log_2(k_s) \rfloor + 1$. Surprisingly, the sequence $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X}))\}_{n\geq 1}$ satisfy the same homogeneous linear recurrence with integer coefficients.

Before we go into the proof of this fact, we will first present the closed formula for the exponential sum $S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X}))$ in terms of exponential sums of symmetric Boolean functions.

Theorem 1.3.1 (Castro and Medina [5])**.** *Suppose that $1 \leq k_1 < \cdots < k_s$ are fixed integers and $F(\boldsymbol{X}) \in F[X_1, X_2, \cdots, X_j]$ (Boolean polynomial in the first j variables), where j is fixed. Define*

$$C_m(F) = \sum_{\boldsymbol{X} \in \mathbb{F}_2^n, \text{ with } w_2(\boldsymbol{X})=m} (-1)^{F(\boldsymbol{X})}, \tag{1.38}$$

for $m = 0, 1, \cdots, j$. Then,

$$S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\boldsymbol{X})) = \sum_{m=0}^{j} C_m(F) S\left(\sum_{i=0}^{m} \binom{m}{i} [\boldsymbol{e}_{n-j,[k_1-i,\cdots,k_s-i]}] \right). \tag{1.39}$$

Proof. This proof is taken from [5], we include it for completion. We will prove the case of one symmetric polynomial, that is, we will show that

$$S(\boldsymbol{e}_{n,k} + F(\mathbf{X})) = \sum_{m=0}^{j} C_m(F) S\left(\sum_{i=0}^{m} \binom{m}{i} \boldsymbol{e}_{n-j,k-i} \right). \tag{1.40}$$

This is done for the simplicity of the writing of the proof. The general case follows the same argument. Proceed first by writing the exponential sum $S(\boldsymbol{e}_{n,k} + F(\mathbf{X}))$, as

$$\begin{aligned} S(\boldsymbol{e}_{n,k} + F(\mathbf{X})) &= \sum_{\mathbf{X}\in\mathbb{F}_2^n} (-1)^{\boldsymbol{e}_{n,k}(\mathbf{X})+F(\mathbf{X})} \\ &= \sum_{\mathbf{X}\in\mathbb{F}_2^n} (-1)^{\boldsymbol{e}_{n,k}(\mathbf{X})}(-1)^{F(\mathbf{X})} \\ &= \sum_{m=0}^{j} \sum_{\mathbf{X}_{(m)}^j\in\mathbb{F}_2^n} (-1)^{\boldsymbol{e}_{n,k}(\mathbf{X})}(-1)^{F(\mathbf{X})}, \end{aligned}$$

where $\mathbf{X}_{(m)}^j \in \mathbb{F}_2^n$ are the tuples which have exactly m ones in the first j entries of $\mathbf{X}$. We will make the assignment to the first j entries while the other entries will vary. Suppose that $\mathbf{X} \in \mathbb{F}_2^n$ has m ones in the first j entries, then $\mathbf{X}$ has the following form:

$$\mathbf{X} = (\delta_1, \delta_2, \cdots, \delta_j, X_{j+1}, \cdots, X_n), \tag{1.41}$$

where $\delta_i \in \{0, 1\}$, for each i, and $\delta_1 + \delta_2 + \cdots + \delta_j = m$. For this particular assignment, one has

$$\boldsymbol{e}_{n,k}(\mathbf{X}) = \sum_{i=0}^{m} \binom{m}{i} \boldsymbol{e}_{n-j,k-i}(\mathbf{X}) \tag{1.42}$$

which is a polynomial in the variables $X_{j+1}, \cdots, X_n$. Then,

$$\begin{aligned} \sum_{(\delta_1,\delta_2,\cdots,\delta_j,X_{j+1},\cdots,X_n)} (-1)^{\boldsymbol{e}_{n,k}(\mathbf{X})}(-1)^{F(\mathbf{X})} &= (-1)^{F(\delta_1,\delta_2,\cdots,\delta_j)} \left(\sum_{(\delta_1,\delta_2,\cdots,\delta_j,X_{j+1},\cdots,X_n)} (-1)^{\boldsymbol{e}_{n,k}(\mathbf{X})} \right) \\ &= (-1)^{F(\delta_1,\delta_2,\cdots,\delta_j)} S(\boldsymbol{e}_{n,k}(\mathbf{X})) \\ &= (-1)^{F(\delta_1,\delta_2,\cdots,\delta_j)} S\left(\sum_{i=0}^{m} \binom{m}{i} \boldsymbol{e}_{n-j,k-i}(\mathbf{X}) \right). \end{aligned}$$

Observe that

$$\begin{aligned} \sum_{\mathbf{X}_{(m)}^j\in\mathbb{F}_2^n} (-1)^{\boldsymbol{e}_{n,k}(\mathbf{X})}(-1)^{F(\mathbf{X})} &= \left(\sum_{\mathbf{X}\in\mathbb{F}_2^n, \text{ with } w_2(\mathbf{X})=m} (-1)^{F(\mathbf{X})} \right) S\left(\sum_{i=0}^{m} \binom{m}{i} \boldsymbol{e}_{n-j.k-i}(\mathbf{X}) \right) \\ &= C_m(F) S\left(\sum_{i=0}^{m} \binom{m}{i} \boldsymbol{e}_{n-j.k-i}(\mathbf{X}) \right). \end{aligned}$$

Hence,

$$\begin{aligned} S(\boldsymbol{e}_{n,k} + F(\mathbf{X})) &= \sum_{\mathbf{X}\in\mathbb{F}_2^n} (-1)^{\boldsymbol{e}_{n,k}(\mathbf{X})+F(\mathbf{X})} \\ &= \sum_{\mathbf{X}\in\mathbb{F}_2^n} (-1)^{\boldsymbol{e}_{n,k}(\mathbf{X})}(-1)^{F(\mathbf{X})} \\ &= \sum_{m=0}^{j} \sum_{\mathbf{X}_{(m)}^j\in\mathbb{F}_2^n} (-1)^{\boldsymbol{e}_{n,k}(\mathbf{X})}(-1)^{F(\mathbf{X})} \\ &= \sum_{m=0}^{j} C_m(F) S\left(\sum_{i=0}^{m} \binom{m}{i} \boldsymbol{e}_{n-j.k-i}(\mathbf{X}) \right). \end{aligned}$$

This proves the theorem for the case of one elementary symmetric polynomial $\boldsymbol{e}_{n,k}$. (Q.E.D.)

Remark: The reader should note that the binomial coefficients inside the exponential sum

$$S\left(\sum_{i=0}^{m} \binom{m}{i} [\boldsymbol{e}_{n-j,[k_1-i,\cdots,k_s-i]}] \right) \tag{1.43}$$

are taken modulo 2 because the parity only matters here. Also, when we are studying the symmetric polynomial $\boldsymbol{e}_{n-j,k-i}$, when $k = i$, then $\boldsymbol{e}_{n-j,k-i}$ is interpreted as 1. If $k < i$, then $\boldsymbol{e}_{n-j,k-i}$ doesn't exist.

The next corollary shows that the sequence of exponential sums $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X}))\}_{n\in\mathbb{N}}$ satisfies the same linear recurrence that the sequence $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]})\}_{n\in\mathbb{N}}$ satisfies.

Corollary 1.3.1 ([5]). *Let $1 \leq k_1 < k_2 < \cdots < k_s$ be fixed integers and $F(\boldsymbol{X})$ be a binary polynomial in the variables $X_1, \cdots, X_j$ where j is fixed. Then the sequence $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\boldsymbol{X}))\}_{n\in\mathbb{N}}$ satisfies the recurrence*

$$x_n = \sum_{\ell=1}^{2^r-1} (-1)^{\ell-1} \binom{2^r}{\ell} x_{n-\ell}, \tag{1.44}$$

where $r = \lfloor \log_2(k_s) \rfloor + 1$.

Proof. Recall that, by Theorem 1.3.1,

$$S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X})) = \sum_{m=0}^{j} C_m(F) S\left(\sum_{i=0}^{m} \binom{m}{i} [\boldsymbol{e}_{n-j,k_1-i} + \cdots + \boldsymbol{e}_{n-j,k_s-i}] \right). \quad (1.45)$$

It is not hard to see that

$$S\left(\sum_{i=0}^{m} \binom{m}{i} [\boldsymbol{e}_{n-j,k_1-i} + \cdots + \boldsymbol{e}_{n-j,k_s-i}] \right) \quad (1.46)$$

is satisfies the recurrence (1.37). Then $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X}))\}$ satisfy recurrence (1.37). This is because $\{S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X}))\}$ is a linear combination of sequences that satisfy the linear recurrence (1.37). This completes the proof. (Q.E.D.)

Cusick et al [8] conjectured that there are non-linear balanced elementary symmetric Boolean functions except for the trivial cases $\boldsymbol{e}_{2^{\ell+1}D-1,2^\ell}$, for ℓ and D positive integers. Canteaout and Videau conjectured an analogue in [2] to what conjectured Cusick et al [8] when the number of variables vary. They conjectured that balanced symmetric functions of fixed degree do not exist when the number of variables grows, except for the trivially balanced cases $\boldsymbol{e}_{2^{\ell+1}D-1,2^\ell}$, for ℓ and D non-negative integers. However, Canteaout and Videau conjecture was proved by Guo et al in [10]. Castro and Medina [4] provided a proof that the conjecture that made Cusick, Li and Stanica is true asymptotically. Particularly, Castro and Medina [4] showed that for the case of an elementary symmetric Boolean function $\boldsymbol{e}_{n,k}$, when $k = 2^\ell$ (k a power of two), then $\boldsymbol{e}_{n,k}$ is asymptotically balanced. Castro and Medina [4] also provided some families of symmetric Boolean functions that are asymptotically balanced and other families that are not.

The concept of asymptotically balanced were extended to perturbations as well by Castro and Medina [5]. The same authors gave necessary and sufficient conditions for a perturbation to be asymptotically balanced. In this same article, they presented the relation that exists between asymptotic coefficients and symmetric Boolean functions. However, this discussion is beyond the scope of this thesis.

In Chapter 2, we provide an identity which was unexpectedly discovered and proved by Castro et al [3]. This identity will later be useful for obtaining one trivially balanced

perturbation from another different trivially balanced perturbation (the same works for sporadic balanced perturbations). Next, we study the link that exists between balanced perturbations and Diophantine equations with binomial coefficients over a bounded set of integers. We also provide a proof that if a perturbation is trivially balanced for some natural number n, then it is trivially balanced for infinitely many n. Finally, we provide some families of trivially balanced and give some particular examples of these perturbations.

Chapter 2

Diophantine Equations of Binomial Coefficients

In this chapter, we are going to study the link that exists between the balancedness of perturbations of symmetric Boolean functions $\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X})$ and solutions to the Diophantine equation with binomial coefficients

$$\sum_{\ell=0}^{n} \delta_\ell \binom{n}{\ell} = 0 \tag{2.1}$$

over a bounded subset Γ of the integers.

Recall that the exponential sum of the perturbation $\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X})$ over $\mathbb{F}_2$ is

$$S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X})) = \sum_{m=0}^{j} C_m(F) S\left(\sum_{i=0}^{m} \binom{m}{i} (\boldsymbol{e}_{n-j,[k_1-i,\cdots,k_s-i]}) \right), \tag{2.2}$$

where $C_m(F)$ is defined as

$$C_m(F) = \sum_{\mathbf{x}\in\mathbb{F}_2^n, \text{ with } w_2(\mathbf{X})=m} (-1)^{F(\mathbf{X})}. \tag{2.3}$$

Observe that a perturbation $\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X})$ is balanced if and only if

$$S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X})) = 0,$$

so finding a balanced perturbation, will allow us to find a solution to (2.1) over a bounded set of integers Γ as we will discuss in this chapter.

This problem of finding solutions to (2.1) is similar to the problem of bisecting binomial coefficients. That problem (bisecting binomial coefficients) was studied by Ionascu, Stanica and Martinsen in [11]. In their study, they provided an integral representation for the formula of the number of binomial bisections of (2.1) as well as some upper and lower bounds for the number of bisections for some particular cases of n.

In this chapter, we provide some perturbation identities. We define the concepts of trivially and sporadic balanced Boolean functions and their extensions to perturbations. As said at the beginning, we study the relation that exists between balanced perturbations and solutions to Diophantine equations that involve binomial coefficients. We provide some families of trivially balanced perturbations and some examples. We also present a proof that if a perturbation is trivially balanced of some natural number n, then it is trivially balanced for infinitely many n.

2.1 Some Perturbation Identities

In this section, we will establish a quite interesting identity between perturbations of two different Boolean functions. We will start by providing a particular example in order to gain an intuition of what is behind this identity. The general idea will be studied later.

Consider the symmetric Boolean polynomials $\boldsymbol{e}_{n,8}$ and $\boldsymbol{e}_{n,9}$ and their corresponding exponential sums

$$S(\boldsymbol{e}_{n,8}) = \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{8}} \binom{n}{\ell} \text{ and } S(\boldsymbol{e}_{n,9}) = \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{9}} \binom{n}{\ell}. \tag{2.4}$$

Of course, these sums, in principle, are different. Consider the expressions

$$(-1)^{\binom{\ell}{8}} \text{ and } (-1)^{\binom{\ell}{9}} \tag{2.5}$$

which are the coefficients of the binomial numbers in the sums (2.4). Both expressions,

when considered as sequences in ℓ, have period length of 16. Within the range of $0 \leq \ell \leq 15$, we can construct a table of values of these expressions as shown below:

ℓ	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$(-1)^{\binom{\ell}{8}}$	1	1	1	1	1	1	1	1	−1	−1	−1	−1	−1	−1	−1	−1
$(-1)^{\binom{\ell}{9}}$	1	1	1	1	1	1	1	1	1	−1	1	−1	1	−1	1	−1

Table 2.1: Values of $(-1)^{\binom{\ell}{8}}$ and $(-1)^{\binom{\ell}{9}}$, where $0 \leq \ell \leq 15$.

Thus, within a period, the values of the sequences differ on 4 positions and that can be used to explain why the sums in (2.4) have different behaviors. Moreover, observe that $S(\boldsymbol{e}_{n,4}) = 0$, that is $\boldsymbol{e}_{n,8}$ is balanced whenever $n = 16k + 15$, for some $k \in \mathbb{N}$ (see [8, Thm. 3, p. 4]). On the other hand $\boldsymbol{e}_{n,9}$ is not balanced for any n. The following table shows the values of $S(\boldsymbol{e}_{n,8})$ and $S(\boldsymbol{e}_{n,9})$ for $1 \leq n \leq 14$.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$S(\boldsymbol{e}_{n,8})$	1	2	4	8	16	32	64	128	254	492	912	1584	2508	3432
$S(\boldsymbol{e}_{n,9})$	1	2	4	8	16	32	64	128	256	510	1004	1936	3632	6604

Table 2.2: Values of $S(\boldsymbol{e}_{n,8})$ and $S(\boldsymbol{e}_{n,9})$, for $1 \leq n \leq 14$.

Thus, it is evident that the sequences $\{S(\boldsymbol{e}_{n,8})\}$ and $\{S(\boldsymbol{e}_{n,9})\}$ are different (see Table 2.2). Remarkably, it turns out that these sums can be altered to make them equal up to a shift in the number of variables.

The trick is quite simple: just add the linear polynomial X_1 to both symmetric polynomials $\boldsymbol{e}_{n,8}$ and $\boldsymbol{e}_{n,9}$. The next table contains the values of $S(\boldsymbol{e}_{n,8} + X_1)$ and $S(\boldsymbol{e}_{n+1,9} + X_1)$.

n	2	3	4	5	6	7	8	9	10	11	12	13	14
$S(\boldsymbol{e}_{n,8} + X_1)$	0	0	0	0	0	0	2	16	72	240	660	1584	3432
$S(\boldsymbol{e}_{n,9} + X_1)$	0	0	0	0	0	0	0	2	16	72	240	660	1584

Table 2.3: Values of $S(\boldsymbol{e}_{n,8} + X_1)$ and $S(\boldsymbol{e}_{n,9} + X_1)$, for $2 \leq n \leq 14$.

This suggest that

$$S(\boldsymbol{e}_{n,8} + X_1) = S(\boldsymbol{e}_{n+1,9} + X_1). \tag{2.6}$$

This trick not only works for $S(\boldsymbol{e}_{n,8} + X_1)$ and $S(\boldsymbol{e}_{n+1,9} + X_1)$, but also works for $\boldsymbol{e}_{n,2k}$

and $\boldsymbol{e}_{n+1,2k+1}$, for every positive integer k, that is, it appears that

$$S(\boldsymbol{e}_{n,2k} + X_1) = S(\boldsymbol{e}_{n+1,2k+1} + X_1). \tag{2.7}$$

Let's see another example with the same symmetric polynomials $\boldsymbol{e}_{n,8}$ and $\boldsymbol{e}_{n,9}$, but instead of adding the linear polynomial X_1 to both symmetric functions in their respective exponential sums, we will just add the polynomial $X_1 + X_2$ (which is in fact balanced):

n	3	4	5	6	7	8	9	10	11	12	13	14
$S(\boldsymbol{e}_{n,8} + X_1 + X_2)$	0	0	0	0	0	-2	-12	-40	-96	-180	-264	-264
$S(\boldsymbol{e}_{n,9} + X_1 + X_2)$	0	0	0	0	0	0	-2	-12	-40	-96	-180	-264

Table 2.4: Values of $S(\boldsymbol{e}_{n,8} + X_1 + X_2)$ and $S(\boldsymbol{e}_{n,9} + X_1 + X_2)$, for $3 \leq n \leq 14$.

In general, it appears that for any positive integer k we have

$$S(\boldsymbol{e}_{n,2k} + X_1 + X_2) = S(\boldsymbol{e}_{n+1,2k+1} + X_1 + X_2). \tag{2.8}$$

We will provide a proof of a general formula that includes (2.7) and (2.8), but first, we need to establish some results. We should mention that the proof of (2.7) relies on the fact that sequences of the form $\{S(\boldsymbol{e}_{n,k})\}$ and $\{S(\boldsymbol{e}_{n,k} + F(\mathbf{X}))\}$ satisfy linear recurrences with integers coefficients. That is, the idea behind the proof of (2.7) is to show that both sequences $\{S(\boldsymbol{e}_{n,k})\}$ and $\{S(\boldsymbol{e}_{n,k} + F(\mathbf{X}))\}$ satisfy the same linear recurrence, and then, once this is done, it will suffice to show that their first initial conditions are equal.

Lemma 2.1.1. *Let $k > 1$ and $m \geq 1$. Then the sequence*

$$\left\{ S\left(\sum_{j=0}^{m} \binom{m}{j} \boldsymbol{e}_{n,k-j} \right) \right\}_{n \geq 1} \tag{2.9}$$

satisfy the same homogeneous linear recurrence as $\{S(\boldsymbol{e}_{n,k})\}_{n \geq 1}$.

Proof. The proof is by induction on m. Let $m = 1$ in the theorem (base case). Define

$\mathbf{X}_0 = (X_1, X_2, \ldots, X_{n-1}, 0)$, and $\mathbf{X}_1 = (X_1, X_2, \ldots, X_{n-1}, 1)$. Then we obtain the identity

$$\begin{aligned} S(\boldsymbol{e}_{n,k}) &= \sum_{(X_1,X_2,\ldots,X_{n-1},X_n)\in\mathbb{F}_2^n} (-1)^{\boldsymbol{e}_{n,k}(X_1,X_2,\ldots,X_{n-1},X_n)} \\ &= \sum_{(X_1,X_2,\ldots,X_{n-1})\in\mathbb{F}_2^{n-1}} (-1)^{\boldsymbol{e}_{n,k}(\mathbf{X}_0)} + \sum_{(X_1,X_2,\ldots,X_{n-1})\in\mathbb{F}_2^{n-1}} (-1)^{\boldsymbol{e}_{n,k}(\mathbf{X}_1)} \\ &= S(\boldsymbol{e}_{n-1,k}) + S(\boldsymbol{e}_{n-1,k} + \boldsymbol{e}_{n-1,k-1}) \end{aligned}$$

which is written equivalently as

$$S(\boldsymbol{e}_{n-1,k} + \boldsymbol{e}_{n-1,k-1}) = S(\boldsymbol{e}_{n,k}) - S(\boldsymbol{e}_{n-1,k}) \tag{2.10}$$

Making the relabel $n \to n+1$, one gets

$$S(\boldsymbol{e}_{n,k} + \boldsymbol{e}_{n,k-1}) = S(\boldsymbol{e}_{n+1,k}) - S(\boldsymbol{e}_{n,k}). \tag{2.11}$$

Now, since the sequences $\{S(\boldsymbol{e}_{n+1,k})\}_{n\geq 1}$ and $\{S(\boldsymbol{e}_{n,k})\}_{n\geq 1}$ satisfies the same linear recurrence, so it does its linear combination $\{S(\boldsymbol{e}_{n+1,k}) - S(\boldsymbol{e}_{n,k})\}_{n\geq 1}$, so the sequence $\{S(\boldsymbol{e}_{n,k} + \boldsymbol{e}_{n,k-1})\}_{n\geq 1}$ satisfies the same linear recurrence. Hence, the result holds for $m = 1$.

Suppose now that the statement is true for some natural number m' with $1 < m' < m$. Then observe that

$$\begin{aligned} S(\boldsymbol{e}_{n,k}) &= \sum_{\ell=0}^{m} \binom{m}{\ell} S\left(\sum_{i=0}^{\ell} \binom{\ell}{i} \boldsymbol{e}_{n-m,k-i} \right) \\ &= \sum_{\ell=0}^{m-1} \binom{m}{\ell} S\left(\sum_{i=0}^{\ell} \binom{\ell}{i} \boldsymbol{e}_{n-m,k-i} \right) + S\left(\sum_{i=0}^{m} \binom{m}{i} \boldsymbol{e}_{n-m,k-i} \right). \end{aligned}$$

The latter equality can be written in an equivalent way as

$$S\left(\sum_{i=0}^{m} \binom{m}{i} \boldsymbol{e}_{n-m,k-i} \right) = S(\boldsymbol{e}_{n,k}) - \sum_{\ell=0}^{m-1} \binom{m}{\ell} S\left(\sum_{i=0}^{\ell} \binom{\ell}{i} \boldsymbol{e}_{n-m,k-i} \right). \tag{2.12}$$

Making the relabel $n \to n+m$, equation (2.12) is transformed into

$$S\left(\sum_{i=0}^{m}\binom{m}{i}\boldsymbol{e}_{n,k-i}\right) = S(\boldsymbol{e}_{n+m,k}) - \sum_{\ell=0}^{m-1}\binom{m}{\ell}S\left(\sum_{i=0}^{\ell}\binom{\ell}{i}\boldsymbol{e}_{n,k-i}\right). \tag{2.13}$$

By the induction hypothesis, each term in the right hand side of equation (2.13) satisfy the same linear recurrence as $\{S(\boldsymbol{e}_{n,k})\}_{n\geq 1}$. Therefore, the theorem is true for all natural numbers m. (Q.E.D.)

The next step for our proof of identity (2.7) is to show that both sequences $\{S(\boldsymbol{e}_{n,k})\}$ and $\{S(\boldsymbol{e}_{n,k}+F(\mathbf{X}))\}$ satisfy the same linear recurrence with integer coefficients. The following result states that fact.

Theorem 2.1.1. *Let $k>1$ and j be fixed integers and let $F(\boldsymbol{X})$ be a binary polynomial in the variables $X_1, X_2, \ldots, X_j$. Let $\bar{k} = 2\lfloor k/2\rfloor + 1 = 2^{a_1}+\cdots+2^{a_s}+1$ and define $\epsilon(k)$ as*

$$\epsilon(k) = \begin{cases} 0, \text{ if } k \text{ is a power of 2,} \\ 1, \text{ otherwise.} \end{cases} \tag{2.14}$$

Then the sequence $\{S(\boldsymbol{e}_{n,k}+F(\boldsymbol{X}))\}_{n\geq 1}$ satisfies the homogeneous linear recurrence whose characteristic polynomial is given by

$$f(X) = (X-2)^{\epsilon(k)}\prod_{\ell=1}^{s}\Phi_{2^{a_\ell+1}}(X-1). \tag{2.15}$$

Moreover, if $F(\boldsymbol{X})$ is balanced, then the sequence satisfies the homogeneous linear recurrence with characteristic polynomial

$$\bar{f}(X) = \prod_{\ell=1}^{s}\Phi_{2^{a_\ell+1}}(X-1), \tag{2.16}$$

and $\deg(\bar{f}(X)) = \bar{k}-1 = 2^{a_1}+\cdots+2^{a_s}$.

Proof. Recall that

$$S(\boldsymbol{e}_{n,k}+F(\mathbf{X})) = \sum_{m=0}^{j}C_m(F)S\left(\sum_{i=0}^{m}\binom{m}{i}\boldsymbol{e}_{n-j,k-i}\right), \tag{2.17}$$

where

$$C_m(F) = \sum_{\mathbf{X}\in\mathbb{F}_2^j,\text{ with } w_2(\mathbf{X})=m} (-1)^{F(\mathbf{X})}. \tag{2.18}$$

Castro and Medina showed that the sequence $\{S(\boldsymbol{e}_{n,k})\}$ satisfies the homogeneous linear recurrence whose characteristic polynomial is given by (2.15) (see [4]). Since the sequence $\left\{S\left(\sum_{j=0}^{m}\binom{m}{j}\boldsymbol{e}_{n,k-j}\right)\right\}_{n\geq 1}$ satisfy the same recurrence as $\{S(\boldsymbol{e}_{n,k})\}_{n\geq 1}$, then Lemma 2.1.1 implies that the sequence $\{S(\boldsymbol{e}_{n,k}+F(\mathbf{X}))\}_{n\geq 1}$ also satisfy the the linear recurrence whose characteristic polynomial is given by (2.15). This implies that $\{S(\boldsymbol{e}_{n,k}+F(\mathbf{X}))\}_{n\geq 1}$ satisfies the homogeneous linear recurrence whose characteristic polynomial is given by

$$f(X) = (X-2)^{\epsilon(k)}\prod_{\ell=1}^{s}\Phi_{2^{a_\ell+1}}(X-1). \tag{2.19}$$

To prove the last statement, observe that Castro and Medina proved (see [5]) that,

$$S(\boldsymbol{e}_{n,k}+F(\mathbf{X})) = d_0\cdot 2^n + \sum_{f(\lambda)=0,\lambda\neq 2} d_\lambda\cdot\lambda^n, \tag{2.20}$$

where d_λ's are the coefficients associated to the roots $\lambda\neq 0$ of the characteristic polynomial $f(X)$, and

$$d_0 = c_0(k)\cdot\frac{S(F)}{2^j}. \tag{2.21}$$

If F is balanced, then $S(F)=0$, so $d_0=0$, which means that the characteristic polynomials associated to the homogeneous linear recurrence that the sequence satisfies is given by

$$\bar{f}(X) = \prod_{\ell=1}^{s}\Phi_{2^{a_\ell+1}}(X-1), \tag{2.22}$$

which is of degree equal to $\bar{k}-1 = 2^{a_1}+\cdots+2^{a_s}$. This concludes the proof. (Q.E.D.)

With the above results at hand, we are now ready to show that $S(\boldsymbol{e}_{n,2k}+X_1) = S(\boldsymbol{e}_{n+1,2k+1}+X_1)$. First note that

$$\begin{aligned}\overline{2k} &= 2\lfloor(2k)/2\rfloor+1 = 2k+1\\ \overline{2k+1} &= 2\lfloor(2k+1)/2\rfloor+1 = 2k+1.\end{aligned}$$

Since $F(\mathbf{X}) = X_1$ is balanced, then the previous theorem tells us that $\{S(\boldsymbol{e}_{n,2k} + X_1)\}_{n\geq 1}$ and $\{S(\boldsymbol{e}_{n+1,2k+1} + X_1)\}_{n\geq 1}$ satisfy the same linear recurrence of order $2k$. To prove that both sequences are equal, it is sufficient to show that their first $2k$ initial values coincide. Define

$$f(n,k) = S(\boldsymbol{e}_{n,2k} + X_1) \text{ and } g(n,k) = S(\boldsymbol{e}_{n+1,2k+1} + X_1). \tag{2.23}$$

Write $f(n,k)$ as

$$\begin{aligned} f(n,k) &= S(\boldsymbol{e}_{n,2k} + X_1) \\ &= S(\boldsymbol{e}_{n-1,2k}) + S(\boldsymbol{e}_{n-1,2k} + \boldsymbol{e}_{n-1,2k-1}) \\ &= \sum_{\ell=0}^{n-1} (-1)^{\binom{\ell}{2k}} \binom{n-1}{\ell} - \sum_{\ell=0}^{n-1} (-1)^{\binom{\ell}{2k}+\binom{\ell}{2k-1}} \binom{n-1}{\ell} \\ &= \sum_{\ell=0}^{n-1} (-1)^{\binom{\ell}{2k}} \left[1 - (-1)^{\binom{\ell}{2k-1}}\right] \binom{n-1}{\ell} \end{aligned}$$

and $g(n,k)$ as

$$\begin{aligned} g(n,k) &= S(\boldsymbol{e}_{n+1,2k+1} + X_1) \\ &= S(\boldsymbol{e}_{n,2k+1}) + S(\boldsymbol{e}_{n,2k+1} + \boldsymbol{e}_{n,2k}) \\ &= \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{2k+1}} \binom{n}{\ell} - \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{2k+1}+\binom{\ell}{2k}} \binom{n}{\ell} \\ &= \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{2k+1}} \left[1 - (-1)^{\binom{\ell}{2k}}\right] \binom{n}{\ell}. \end{aligned}$$

Note that

$$\begin{aligned} f(1,k) &= 0 = g(1,k) \\ f(2,k) &= 0 = g(2,k) \\ &\vdots \\ f(2k-1,k) &= 0 = g(2k-1,k) \\ f(2k,k) &= 2 = g(2k,k). \end{aligned}$$

So the sequences $\{f(n,k)\}_{n\geq 1}$ and $\{g(n,k)\}_{n\geq 1}$ satisfy the same recurrence of order $2k$ with the same initial conditions. Therefore,

$$f(n,k) = S(\boldsymbol{e}_{n,2k} + X_1) = S(\boldsymbol{e}_{n+1,2k+1} + X_1) = g(n,k), \tag{2.24}$$

for every positive integers n and k. The above discussion, leads to the following consequence involving sums of binomial coefficients.

Corollary 2.1.1. *Let n and k be positive integers. Then,*

$$\sum_{\ell=0}^{n-1}(-1)^{\binom{\ell}{2k}}\left[1-(-1)^{\binom{\ell}{2k-1}}\right]\binom{n-1}{\ell} = \sum_{\ell=0}^{n}(-1)^{\binom{\ell}{2k+1}}\left[1-(-1)^{\binom{\ell}{2k}}\right]\binom{n}{\ell}.$$

Therefore, even though the sequences $\{S(\boldsymbol{e}_{n,2k})\}$ and $\{S(\boldsymbol{e}_{n+1,2k+1})\}$ are different, they can be altered in such way they are equal, up to a shift in the number of variables, i.e.,

$$S(\boldsymbol{e}_{n,2k} + X_1) = S(\boldsymbol{e}_{n+1,2k+1} + X_1). \tag{2.25}$$

Equation (2.24) leads to the following question: for which Boolean polynomials $F(\mathbf{X})$ the identity

$$S(\boldsymbol{e}_{n,2k} + F(\mathbf{X})) = S(\boldsymbol{e}_{n+1,2k+1} + F(\mathbf{X})). \tag{2.26}$$

holds for every positive integer n and k? The answer to this question was provided by Castro and Medina. They show that it holds for all balanced Boolean polynomials $F(\mathbf{X})$. Their proof of this fact depends on the following classical result from number theory, whose proof is included for completeness.

Theorem 2.1.2 (**Lucas' Theorem**). *Let $n \in \mathbb{N}$ with 2-adic expansion $m = 2^{a_1} + 2^{a_2} + \cdots + 2^{a_\ell}$. Then the binomial coefficient $\binom{m}{n}$ is odd if and only if either $n = 0$ or n is a sum of some of the a_i's.*

Proof. Recall that

$$(x+1)^{2^k} = \sum_{i=0}^{2^k} \binom{2^k}{i} x^{2^k-i} = x^{2^k} + \binom{2^k}{1} x^{2^k-1} + \binom{2^k}{2} x^{2^k-2} + \cdots + \binom{2^k}{2^k-1} x + 1$$
$$\equiv x^{2^k} + 1 \pmod 2.$$

Now,

$$\sum_{n=0}^{m} \binom{m}{n} x^n = (x+1)^m = (x+1)^{2^{a_1}+2^{a_2}+\cdots+2^{a_\ell}}$$
$$= (x+1)^{2^{a_1}} (x+1)^{2^{a_2}} \cdots (x+1)^{2^{a_\ell}}$$
$$\equiv (x^{2^{a_1}}+1)(x^{2^{a_2}}+1)\cdots(x^{2^{a_\ell}}+1) \pmod 2.$$

Therefore, it is clear that $\binom{m}{n} \equiv 1 \pmod 2$ if and only if $n = 0$ or n is the sum of some of the 2^{a_i}'s. (Q.E.D.)

We are ready to provide a proof of Castro and Medina's assertion.

Theorem 2.1.3 ([5])**.** *Suppose that $k \geq 1$ is an integer. Let $F(\boldsymbol{X})$ be a Boolean polynomial in j variables for j fixed. Then $S(\boldsymbol{e}_{n+j,2k} + F(\boldsymbol{X})) = S(\boldsymbol{e}_{n+1+j,2k+1}, +F(\boldsymbol{X}))$ if and only if $F(\boldsymbol{X})$ is balanced.*

Proof. This proof is inspired by the one presented in [5]. Suppose that $F(\mathbf{X})$ is not balanced, that is, $S(F) \neq 0$. Recall that

$$S(\boldsymbol{e}_{n+j,2k} + F(\mathbf{X})) = \sum_{m=0}^{j} C_m(F) S\left(\sum_{i=0}^{m} \binom{m}{i} \boldsymbol{e}_{n,2k-i} \right)$$
$$= \sum_{m=0}^{j} C_m(F) \left[\sum_{\ell=0}^{n} (-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{\ell}{2k-i}} \binom{n}{\ell} \right]$$
$$= \sum_{\ell=0}^{n} \left[\sum_{m=0}^{j} C_m(F) \cdot (-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{\ell}{2k-i}} \right] \binom{n}{\ell}.$$

In a similar way, we have,

$$\begin{aligned} S(\boldsymbol{e}_{n+1+j,2k+1} + F(\mathbf{X})) &= \sum_{m=0}^{j} C_m(F) S\left(\sum_{i=0}^{m} \binom{m}{i} \boldsymbol{e}_{n+1,2k+1-i} \right) \\ &= \sum_{m=0}^{j} C_m(F) \left[\sum_{\ell=0}^{n+1} (-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{\ell}{2k+1-i}} \binom{n+1}{\ell} \right] \\ &= \sum_{\ell=0}^{n+1} \left[\sum_{m=0}^{j} C_m(F) \cdot (-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{\ell}{2k+1-i}} \right] \binom{n+1}{\ell}. \end{aligned}$$

Theorem 2.1.1 implies that both $\{S(\boldsymbol{e}_{n+j,2k} + F(\mathbf{X}))\}$ and $\{S(\boldsymbol{e}_{n+1+j,2k+1}, +F(\mathbf{X}))\}$ satisfy the same linear recurrence of order $2k+1$. Observe that the first initial condition of $\{S(\boldsymbol{e}_{n,2k} + F(\mathbf{X}))\}$ is $S(F)$(when $n = j$), because

$$\begin{aligned} S(\boldsymbol{e}_{j,2k} + F(\mathbf{X})) &= \sum_{\ell=0}^{0} \left[\sum_{m=0}^{j} C_m(F) \cdot (-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{\ell}{2k-i}} \right] \binom{0}{\ell} \\ &= \sum_{m=0}^{j} C_m(F) \cdot 1 \\ &= S(F), \end{aligned}$$

whereas the first initial condition of $\{S(\boldsymbol{e}_{n+1,2k+1}, +\mathbf{X}))\}$ is $2S(F)$, so similarly,

$$\begin{aligned} S(\boldsymbol{e}_{j+1,2k+1} + F(\mathbf{X})) &= \sum_{\ell=0}^{1} \left[\sum_{m=0}^{j} C_m(F) \cdot (-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{\ell}{2k+1-i}} \right] \binom{1}{\ell} \\ &= 2S(F). \end{aligned}$$

Since $S(F) \neq 0$, then $S(F) \neq 2S(F)$, so the sequences $\{S(\boldsymbol{e}_{n+j,2k}+F(\mathbf{X}))\}$ and $\{S(\boldsymbol{e}_{n+1+j,2k+1},$ $+ F(\mathbf{X}))\}$ are different.

Conversely, suppose that $F(\mathbf{X})$ is balanced. Then $S(F) = 0$. For simplicity in the writing, let $C_m = C_m(F)$. Recall that $S(F) = C_0 + C_1 + C_2 + \cdots + C_j$. Then $C_0 + C_1 + C_2 + \cdots + C_j = 0$. Then use the following identity

$$\sum_{i=0}^{m} \binom{m}{i} \binom{\ell}{k-i} = \binom{\ell+m}{k} \tag{2.27}$$

to simplify the formulas

$$S(\boldsymbol{e}_{n+j,2k} + F(\mathbf{X})) = \sum_{\ell=0}^{n} \left[\sum_{m=0}^{j} C_m \cdot (-1)^{\binom{\ell+m}{2k}} \right] \binom{n}{\ell}, \tag{2.28}$$

and

$$S(\boldsymbol{e}_{n+1+j,2k+1} + F(\mathbf{X})) = \sum_{\ell=0}^{n+1} \left[\sum_{m=0}^{j} C_m \cdot (-1)^{\binom{\ell+m}{2k+1}} \right] \binom{n+1}{\ell}. \tag{2.29}$$

Observe that $S(\boldsymbol{e}_{n+1+j,2k+1} + F(\mathbf{X}))$ can be rewritten as

$$\begin{aligned}
S(\boldsymbol{e}_{n+1+j,2k+1} + F(\mathbf{X})) &= \sum_{\ell=0}^{n+1} \left[\sum_{m=0}^{j} C_m \cdot (-1)^{\binom{\ell+m}{2k+1}} \right] \binom{n+1}{\ell} \\
&= \sum_{\ell=0}^{n+1} \left[\sum_{m=0}^{j} C_m \cdot (-1)^{\binom{\ell+m}{2k+1}} \right] \left[\binom{n}{\ell} + \binom{n}{\ell-1} \right] \\
&= \sum_{\ell=0}^{n+1} \left[\sum_{m=0}^{j} C_m \cdot (-1)^{\binom{\ell+m}{2k+1}} \right] \binom{n}{\ell} + \sum_{\ell=0}^{n+1} \left[\sum_{m=0}^{j} C_m \cdot (-1)^{\binom{\ell+m}{2k+1}} \right] \binom{n}{\ell-1} \\
&= \sum_{\ell=0}^{n} \left[\sum_{m=0}^{j} C_m \cdot (-1)^{\binom{\ell+m}{2k+1}} \right] \binom{n}{\ell} + \sum_{\ell=0}^{n} \left[\sum_{m=0}^{j} C_m \cdot (-1)^{\binom{\ell+1+m}{2k+1}} \right] \binom{n}{\ell} \\
&= \sum_{\ell=0}^{n} \left[\sum_{m=0}^{j} C_m \left((-1)^{\binom{\ell+m}{2k+1}} + (-1)^{\binom{\ell+1+m}{2k+1}} \right) \right] \binom{n}{\ell}.
\end{aligned}$$

Since $C_0 = -C_1 - C_2 - \cdots - C_j$, then we can rewrite $S(\boldsymbol{e}_{n+j,2k} + F(\mathbf{X}))$ as

$$\begin{aligned}
S(\boldsymbol{e}_{n+j,2k} + F(\mathbf{X})) &= \sum_{\ell=0}^{n} \left[\sum_{m=0}^{j} C_m \cdot (-1)^{\binom{\ell+m}{2k}} \right] \binom{n}{\ell} \\
&= C_0 \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{2k}} \binom{n}{\ell} + \sum_{m=1}^{j} C_m \sum_{\ell=0}^{n} (-1)^{\binom{\ell+m}{2k}} \binom{n}{\ell} \\
&= (-C_1 - C_2 - \cdots - C_j) \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{2k}} \binom{n}{\ell} + \sum_{m=1}^{j} C_m \sum_{\ell=0}^{n} (-1)^{\binom{\ell+m}{2k}} \binom{n}{\ell} \\
&= \sum_{m=1}^{j} C_m \sum_{\ell=0}^{n} \left[(-1)^{\binom{\ell+m}{2k}} - (-1)^{\binom{\ell}{2k}} \right] \binom{n}{\ell}.
\end{aligned}$$

We can write $S(\boldsymbol{e}_{n+1+j,2k+1} + F(\boldsymbol{X}))$ as

$$S(\boldsymbol{e}_{n+1+j,2k+1} + F(\mathbf{X})) = \sum_{m=1}^{j} C_m \sum_{\ell=0}^{n} \left[(-1)^{\binom{\ell+m}{2k+1}} + (-1)^{\binom{\ell+1+m}{2k+1}} - (-1)^{\binom{\ell}{2k+1}} - (-1)^{\binom{\ell+1}{2k+1}}\right] \binom{n}{\ell}.$$

Lucas' theorem implies that

$$(-1)^{\binom{\ell+m}{2k}} - (-1)^{\binom{\ell}{2k}} = (-1)^{\binom{\ell+m}{2k+1}} + (-1)^{\binom{\ell+1+m}{2k+1}} - (-1)^{\binom{\ell}{2k+1}} - (-1)^{\binom{\ell+1}{2k+1}}, \tag{2.30}$$

for all $k \in \mathbb{N}$ and $\ell, m \in \mathbb{Z}_{\geq 0}$. Therefore, $S(\boldsymbol{e}_{n+j,2k} + F(\mathbf{X})) = S(\boldsymbol{e}_{n+1+j,2k+1} + F(\mathbf{X}))$, for all positive integers n and k. This concludes the proof. (Q.E.D.)

Before we go in discussing some examples, we will introduce a new definition. A *rotation symmetric Boolean function* $R(\mathbf{X})$ in n variables is a functions which is invariant under the action of the cyclic group C_n, that is, $R(\mathbf{X})$ is rotation symmetric if

$$R(X_{\sigma(1)}, X_{\sigma(2)}, \cdots, X_{\sigma(n)}) = R(X_1, X_2, \cdots, X_n) \tag{2.31}$$

for every $\sigma \in C_n$ (here we are seeing C_n as a subgroup of S_n). An example of such function is

$$R(\mathbf{X}) = X_1X_2X_3 + X_2X_3X_4 + X_3X_4X_1 + X_4X_1X_2, \tag{2.32}$$

which is a 4-variable function invariant under the action of C_4.

Rotation symmetric functions can be studied from the point of view of recurrences of sequences of their exponential sums over binary field $\mathbb{F}_2$ as well as in the general setting of Galois fields $\mathbb{F}_q$, where $q = p^r$, p a prime number and $r \geq 1$. These functions have very interesting properties on its own right, however, we will not go deeply into discussing them. For more details about the recursive behavior of these functions, please refer to [7].

Example 2.1.1. Consider the rotation

$$R(\mathbf{X}) = X_1X_2 + X_2X_3 + X_3X_4 + X_4X_1. \tag{2.33}$$

This rotation is balanced, thus both sequences $\{S(\boldsymbol{e}_{n,4}+R(\mathbf{X}))\}$ and $\{S(\boldsymbol{e}_{n+1,5}+R(\mathbf{X}))\}$ satisfy the same linear recurrence of order 4. Moreover, we have that

$$\begin{aligned} S(\boldsymbol{e}_{5,4}+R(\mathbf{X})) &= S(\boldsymbol{e}_{6,5}+R(\mathbf{X})) = 4 \\ S(\boldsymbol{e}_{6,4}+R(\mathbf{X})) &= S(\boldsymbol{e}_{7,5}+R(\mathbf{X})) = 20 \\ S(\boldsymbol{e}_{7,4}+R(\mathbf{X})) &= S(\boldsymbol{e}_{8,5}+R(\mathbf{X})) = 48 \\ S(\boldsymbol{e}_{8,4}+R(\mathbf{X})) &= S(\boldsymbol{e}_{9,5}+R(\mathbf{X})) = 92. \end{aligned}$$

Then it follows that both sequences $\{S(\boldsymbol{e}_{n,4}+R(\mathbf{X}))\}$ and $\{S(\boldsymbol{e}_{n+1,5}+R(\mathbf{X}))\}$ are equal for all positive integers n.

Example 2.1.2. Consider the function

$$F(\mathbf{X}) = X_1X_2 + X_1X_3 + X_2X_3. \tag{2.34}$$

One can easily check that $F(\mathbf{X})$ is balanced. By Theorem 2.1.3, we have that

$$S(\boldsymbol{e}_{n,2k}+F(\mathbf{X})) = S(\boldsymbol{e}_{n+1,2k+1}+F(\mathbf{X})). \tag{2.35}$$

Let us see that with explicit numbers. Set $2k=6$. The first few values of $\{S(\boldsymbol{e}_{n,6}+F(\mathbf{X}))\}$ (starting from $n=6$), are

$$2, 16, 60, 152, 292, 432, 432, 0, -1384, -4544, -10608, \ldots \tag{2.36}$$

while the first few values of the sequence $\{S(\boldsymbol{e}_{n,7}+F(\mathbf{X}))\}$ (starting from $n=6$), are

$$0, 2, 16, 60, 152, 292, 432, 432, 0, -1384, -4544, \ldots \tag{2.37}$$

suggesting that both sequences $\{S(\boldsymbol{e}_{n,6}+F(\mathbf{X}))\}$ and $\{S(\boldsymbol{e}_{n+1,7}+F(\mathbf{X}))\}$ are equal up to a shift in the number of variables.

The result of the previous theorem can be further generalized to perturbations of the

form

$$\sum_{i=0}^{t}\binom{t}{i}\boldsymbol{e}_{n,[k_1-i,k_2-i,\cdots,k_s-i]}+F(\mathbf{X}), \tag{2.38}$$

where $1\leq k_1<\cdots<k_s$ are fixed integers and F is a Boolean function in the variables $X_1,\cdots,X_j$. To do this, observe that in the proof of Theorem 2.1.3, one has the identity

$$(-1)^{\binom{\ell+m}{2k}}-(-1)^{\binom{\ell}{2k}}=(-1)^{\binom{\ell+m}{2k+1}}+(-1)^{\binom{\ell+1+m}{2k+1}}-(-1)^{\binom{\ell}{2k+1}}-(-1)^{\binom{\ell+1}{2k+1}}, \tag{2.39}$$

for all positive integers k and non-negative integers ℓ and m. Using the identity

$$\binom{\ell+m}{2k}=\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+m-t}{2k-i} \tag{2.40}$$

and other similar ones for

$$\binom{\ell}{2k},\ \binom{\ell+m}{2k+1},\ \binom{\ell+m+1}{2k+1},\ \binom{\ell}{2k+1},\ \text{and}\ \binom{\ell+1}{2k+1}$$

we obtain

$$\begin{aligned}&(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+m-t}{2k-i}}-(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell-t}{2k-i}}=\\&(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+m-t}{2k+1-i}}+(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+m+1-t}{2k+1-i}}-(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell-t}{2k+1-i}}-(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+1-t}{2k+1-i}}.\end{aligned} \tag{2.41}$$

for all k and non-negative ℓ and m.

Before we go into the generalized result of Theorem 2.1.3, we will state the following two results, which are extensions of Lemma 2.1.1 and Theorem 2.1.1, respectively and whose proof of Lemma 2.1.2 will be provided just for completeness and the proof of Theorem 2.1.4 will be omitted since it follows a similar argument as in the proof of Theorem 2.1.1.

Lemma 2.1.2. *Let $1\leq k_1<k_2<\cdots<k_s$ and $m\geq 1$ be fixed integers. Then the sequence*

$$\left\{S\left(\sum_{j=0}^{m}\binom{m}{j}\boldsymbol{e}_{n,[k_1-j,k_2-j,\cdots,k_s-j]}\right)\right\} \tag{2.42}$$

satisfies the same homogeneous linear recurrence as $\{S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]})\}$

Proof. The proof is similar to the one of Lemma 2.1.1. We will proceed by induction on m. Let $m = 1$ and write $S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]})$ as

$$S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]}) = S(\boldsymbol{e}_{n-1,[k_1,k_2,\cdots,k_s]}) + S(\boldsymbol{e}_{n-1,[k_1,k_2,\cdots,k_s]} + \boldsymbol{e}_{n-1,[k_1-1,k_2-1,\cdots,k_s-1]}).$$

After making the relabel $n \to n+1$, we obtain the equivalent expression

$$S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]} + \boldsymbol{e}_{n,[k_1-1,k_2-1,\cdots,k_s-1]}) = S(\boldsymbol{e}_{n+1,[k_1,k_2,\cdots,k_s]}) - S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]}), \tag{2.43}$$

and from this, it is clear that $\{S(\boldsymbol{e}_{n\,[k_1,k_2,\cdots,k_s]} + \boldsymbol{e}_{n,[k_1-1,k_2-1,\cdots,k_s-1]})\}$ satisfies the same recurrence as $\{S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]})\}$.

Now, suppose that the statement holds for all values of m' that are less than some $m > 1$. Then write $S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]})$ as

$$\begin{aligned} S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]}) &= \sum_{\ell=0}^{m} \binom{m}{\ell} S\left(\sum_{i=0}^{\ell} \binom{\ell}{i} \boldsymbol{e}_{n-m,[k_1-m,k_2-m,\cdots,k_s-m]} \right) \\ &= \sum_{\ell=0}^{m-1} \binom{m}{\ell} S\left(\sum_{i=0}^{\ell} \binom{\ell}{i} \boldsymbol{e}_{n-m,[k_1-m,k_2-m,\cdots,k_s-m]} \right) \\ &\quad + S\left(\sum_{i=0}^{m} \binom{m}{i} \boldsymbol{e}_{n-m,[k_1-m,k_2-m,\cdots,k_s-m]} \right). \end{aligned}$$

After making the relabel $n \to n+m$, this implies that

$$\begin{aligned} S\left(\sum_{i=0}^{m} \binom{m}{i} \boldsymbol{e}_{n,[k_1-m,k_2-m,\cdots,k_s-m]} \right) &= S(\boldsymbol{e}_{n+m,[k_1,k_2,\cdots,k_s]}) \\ &\quad - \sum_{\ell=0}^{m-1} \binom{m}{\ell} S\left(\sum_{i=0}^{\ell} \binom{\ell}{i} \boldsymbol{e}_{n,[k_1-m,k_2-m,\cdots,k_s-m]} \right). \end{aligned} \tag{2.44}$$

Then by the induction hypothesis, each term in the right hand side satisfies the same recurrence as $\{S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]})\}$. The theorem is proved. (Q.E.D.)

We define the "OR" operator $\vee$ operator on $\mathbb{F}_2$ as

$$0 \vee 0 = 0$$

$$0 \vee 1 = 1$$

$$1 \vee 0 = 1$$

$$1 \vee 1 = 1.$$

We can extend this definition to the natural numbers by letting $a \vee b$ be the natural number obtained by applying $\vee$ coordinatewise to the binary digits of a and b, that is, if

$$a = \delta_t \cdot 2^t + \delta_{t-1} \cdot 2^{t-1} + \cdots + \delta_1 \cdot 2 + \delta_0$$
$$b = \delta'_t \cdot 2^t + \delta'_{t-1} \cdot 2^{t-1} + \cdots + \delta'_1 \cdot 2 + \delta'_0,$$

for $\delta_i, \delta'_i \in \{0, 1\}$, then

$$a \vee b = (\delta_t \vee \delta'_t) \cdot 2^t + (\delta_{t-1} \vee \delta'_{t-1}) \cdot 2^{t-1} + \cdots + (\delta_1 \vee \delta'_1) \cdot 2 + (\delta_0 \vee \delta'_0). \tag{2.45}$$

For example, let $a = 7$, and $b = 11$. Then

$$\begin{aligned} 7 \vee 11 &= (0 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2 + 1 \cdot 1) \vee (1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2 + 1 \cdot 1) \\ &= (0 \vee 1) \cdot 2^3 + (1 \vee 0) \cdot 2^2 + (1 \vee 1) \cdot 2 + (1 \vee 1) \cdot 1 \\ &= 1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2 + 1 \cdot 1 \\ &= 8 + 4 + 2 + 1 \\ &= 15. \end{aligned}$$

Next is a result that shows that the sequences $\{S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]})\}$ and $\{S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]} + F(\mathbf{X}))\}$ satisfy the same linear recurrence whose proof will be omitted.

Theorem 2.1.4. *Let $1 \leq k_1 < k_2 < \cdots < k_s$ and j and let $F(\boldsymbol{X})$ be a Boolean polynomial in the variables $X_1, X_2, \cdots, X_j$ (j is fixed). Let $\bar{k} = 2\lfloor (k_1 \vee k_2 \vee \cdots \vee k_s)/2 \rfloor + 1$ where $\bar{k}$*

has a 2-adic expansion of the form

$$\bar{k} = 1 + 2^{a_1} + 2^{a_2} + \cdots + 2^{a_s},$$

where the last exponent a_s is given by $a_s = \lfloor \log_2(\bar{k}) \rfloor$. Then the sequences $\{S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]} + F(\boldsymbol{X}))\}$ and $\{S(\boldsymbol{e}_{n,[k_1,k_2,\cdots,k_s]})\}$ satisfy the homogeneous linear recurrence whose characteristic polynomial $P_{k_1,k_2,\cdots,k_s}(X)$ divides

$$(X-2)\prod_{\ell=1}^{s} \Phi_{2^{a_\ell}}(X-1).$$

The next theorem is just a generalization of Theorem 2.1.3.

Theorem 2.1.5. *Suppose that k and t are integers with k positive and t non-negative. Let $F(\boldsymbol{X})$ be a Boolean function in j variables (j is fixed). Then,*

$$S\left(\left[\sum_{i=0}^{t} \binom{t}{i} \boldsymbol{e}_{n+j,2k-i}\right] + F(\boldsymbol{X})\right) = S\left(\left[\sum_{i=0}^{t} \binom{t}{i} \boldsymbol{e}_{n+1+j,2k+1-i}\right] + F(\boldsymbol{X})\right)$$

for each positive integer n if and only if $F(\boldsymbol{X})$ is balanced.

Proof. Suppose that $F(\mathbf{X})$ is not balanced. For simplicity, let $C_m = C_m(F)$. Write

$$S\left(\left[\sum_{i=0}^{t} \binom{t}{i} \boldsymbol{e}_{n+j,2k-i}\right] + F(\mathbf{X})\right) = \sum_{\ell=0}^{n} \left(\sum_{m=0}^{j} C_m (-1)^{\sum_{i=0}^{t} \binom{t}{i}\binom{\ell+m-t}{2k-i}}\right) \binom{n}{\ell}.$$

Similarly,

$$S\left(\left[\sum_{i=0}^{t} \binom{t}{i} \boldsymbol{e}_{n+1+j,2k+1-i}\right] + F(\mathbf{X})\right) = \sum_{\ell=0}^{n+1} \left(\sum_{m=0}^{j} C_m (-1)^{\sum_{i=0}^{t} \binom{t}{i}\binom{\ell+m-t}{2k+1-i}}\right) \binom{n+1}{\ell}.$$

It it not hard to prove that the first initial condition of $\left\{S\left(\left[\sum_{i=0}^{t}\binom{t}{i}\boldsymbol{e}_{n+j,2k-i}\right] + F(\mathbf{X})\right)\right\}$ is $S(F)$ and the first initial condition of $\left\{S\left(\left[\sum_{i=0}^{t}\binom{t}{i}\boldsymbol{e}_{n+1+j,2k+1-i}\right] + F(\mathbf{X})\right)\right\}$ is $2S(F)$ (letting $n = 0$). Since $S(F) \neq 0$, we have that $S(F) \neq 2S(F)$, so the given sequences are different.

Now we will proceed with the sufficient part. Suppose that $S(F) = 0$. Then $C_0 + C_1 + \cdots + C_m = 0$. Rewrite the expression $S\left(\left[\sum_{i=0}^{t}\binom{t}{i}\boldsymbol{e}_{n+j,2k+1-i}\right] + F(\mathbf{X})\right)$

$$
\begin{aligned}
&S\left(\left[\sum_{i=0}^{t}\binom{t}{i}\boldsymbol{e}_{n+j,2k+1-i}\right] + F(\mathbf{X})\right)\\
&\quad= \sum_{\ell=0}^{n}\left(\sum_{m=0}^{j} C_m(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+m-t}{2k-i}}\right)\binom{n}{\ell}\\
&\quad= \sum_{m=0}^{j} C_m\left(\sum_{\ell=0}^{n}(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+m-t}{2k-i}}\right)\binom{n}{\ell}\\
&\quad= (-C_1 - C_2 - \cdots - C_m)\left(\sum_{\ell=0}^{n}(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell-t}{2k-i}}\right)\binom{n}{\ell} + \sum_{m=1}^{j} C_m\left(\sum_{l=0}^{n}(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+m-t}{2k-i}}\right)\binom{n}{\ell}\\
&\quad= \sum_{m=1}^{j} C_m \sum_{\ell=0}^{n}\left[(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+m-t}{2k-i}} - (-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell-t}{2k-i}}\right]\binom{n}{\ell}.
\end{aligned}
$$

In a similar way, rewrite $S\left(\left[\sum_{i=0}^{t}\binom{t}{i}\boldsymbol{e}_{n+1+j,2k-i}\right] + F(\mathbf{X})\right)$ as

$$
\begin{aligned}
&S\left(\left[\sum_{i=0}^{t}\binom{t}{i}\boldsymbol{e}_{n+1+j,2k+1-i}\right] + F(\mathbf{X})\right)\\
&= \sum_{m=1}^{j} C_m \sum_{\ell=0}^{n}\left[(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+m-t}{2k+1-t}} + (-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+m+1-t}{2k+1}} - (-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell-t}{2k+1}} - (-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+1-t}{2k+1}}\right]\binom{n}{\ell}.
\end{aligned}
\tag{2.46}
$$

An implication of equation (2.30) in the proof of Theorem 2.1.3 is that

$$
\begin{aligned}
&(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+m-t}{2k-i}} - (-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell-t}{2k-i}} =\\
&(-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+m-t}{2k+1-i}} + (-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+m+1-t}{2k+1-i}} - (-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell-t}{2k+1-i}} - (-1)^{\sum_{i=0}^{t}\binom{t}{i}\binom{\ell+1-t}{2k+1-i}},
\end{aligned}
\tag{2.47}
$$

for all k and non-negative ℓ and m. Therefore,

$$
S\left(\left[\sum_{i=0}^{t}\binom{t}{i}\boldsymbol{e}_{n+j,2k-i}\right] + F(\mathbf{X})\right) = S\left(\left[\sum_{i=0}^{t}\binom{t}{i}\boldsymbol{e}_{n+1+j,2k+1-i}\right] + F(\mathbf{X})\right), \tag{2.48}
$$

for all positive integers n. This concludes the proof. (Q.E.D.)

Example 2.1.3. Let $F(\mathbf{X})$ be a balanced Boolean function in the variables $X_1, X_2, \cdots, X_j$. Then by Theorem 2.1.5, we have the following

$$S\left(\left[\sum_{i=0}^{0}\binom{0}{i}\boldsymbol{e}_{n+j,10-i}\right]+F(\mathbf{X})\right)=S(\boldsymbol{e}_{n+j,10})=S(\boldsymbol{e}_{n+j+1,11}+F(\mathbf{X}))$$
$$S\left(\left[\sum_{i=0}^{1}\binom{1}{i}\boldsymbol{e}_{n+j,10-i}\right]+F(\mathbf{X})\right)=S(\boldsymbol{e}_{n+j,[10,9]}+F(\mathbf{X}))=S(\boldsymbol{e}_{n+j+1,[11,10]}+F(\mathbf{X}))$$
$$S\left(\left[\sum_{i=0}^{2}\binom{2}{i}\boldsymbol{e}_{n+j,10-i}\right]+F(\mathbf{X})\right)=S(\boldsymbol{e}_{n+j,[10,8]}+F(\mathbf{X}))=S(\boldsymbol{e}_{n+j+1,[11,9]}+F(\mathbf{X}))$$
$$S\left(\left[\sum_{i=0}^{3}\binom{3}{i}\boldsymbol{e}_{n+j,10-i}\right]+F(\mathbf{X})\right)=S(\boldsymbol{e}_{n+j,[10,9,8,7]}+F(\mathbf{X}))=S(\boldsymbol{e}_{n+j+1,[11,10,9,8]}+F(\mathbf{X}))$$
$$S\left(\left[\sum_{i=0}^{4}\binom{4}{i}\boldsymbol{e}_{n+j,10-i}\right]+F(\mathbf{X})\right)=S(\boldsymbol{e}_{n+j,[10,6]}+F(\mathbf{X}))=S(\boldsymbol{e}_{n+j+1,[11,7]}+F(\mathbf{X}))$$
$$S\left(\left[\sum_{i=0}^{5}\binom{5}{i}\boldsymbol{e}_{n+j,10-i}\right]+F(\mathbf{X})\right)=S(\boldsymbol{e}_{n+j,[10,9,6,5]}+F(\mathbf{X}))=S(\boldsymbol{e}_{n+j+1,[11,10,7,6]}+F(\mathbf{X}))$$
$$S\left(\left[\sum_{i=0}^{6}\binom{6}{i}\boldsymbol{e}_{n+j,10-i}\right]+F(\mathbf{X})\right)=S(\boldsymbol{e}_{n+j,[10,8,6,4]}+F(\mathbf{X}))=S(\boldsymbol{e}_{n+j+1,[11,9,7,5]}+F(\mathbf{X}))$$
$$S\left(\left[\sum_{i=0}^{7}\binom{7}{i}\boldsymbol{e}_{n+j,10-i}\right]+F(\mathbf{X})\right)=S(\boldsymbol{e}_{n+j,[10,9,8,7,6,5,4,3]}+F(\mathbf{X}))=S(\boldsymbol{e}_{n+j+1,[11,10,9,8,7,6,5,4]}+F(\mathbf{X}))$$
$$S\left(\left[\sum_{i=0}^{8}\binom{8}{i}\boldsymbol{e}_{n+j,10-i}\right]+F(\mathbf{X})\right)=S(\boldsymbol{e}_{n+j,[10,2]}+F(\mathbf{X}))=S(\boldsymbol{e}_{n+j+1,[11,3]}+F(\mathbf{X}))$$
$$S\left(\left[\sum_{i=0}^{9}\binom{9}{i}\boldsymbol{e}_{n+j,10-i}\right]+F(\mathbf{X})\right)=S(\boldsymbol{e}_{n+j,[10,9,2,1]}+F(\mathbf{X}))=S(\boldsymbol{e}_{n+j+1,[11,10,3,2]}+F(\mathbf{X}))$$
$$S\left(\left[\sum_{i=0}^{10}\binom{10}{i}\boldsymbol{e}_{n+j,10-i}\right]+F(\mathbf{X})\right)=S(\boldsymbol{e}_{n+j,[10,8,2,0]}+F(\mathbf{X}))=S(\boldsymbol{e}_{n+j+1,[11,9,3,1]}+F(\mathbf{X})).$$

2.2 Diophantine Equations of Binomial Coefficients

In this section, we are interested in equations of the form

$$\sum_{\ell=0}^{n}\delta_\ell\binom{n}{\ell}=0, \tag{2.49}$$

where $\delta_\ell \in \Gamma$ and Γ is a bounded subset of $\mathbb{Z}$. This equation was studied by Castro, González and Medina [3]. They connected this equation to exponential sums of symmetric Boolean functions.

Recall that,

$$S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]}) = \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{k_1}+\cdots+\binom{\ell}{k_s}} \binom{n}{\ell}, \tag{2.50}$$

where $1 \leq k_1 < k_2 < \cdots < k_s$ are fixed integers. It is clear that each time we find a balanced symmetric Boolean function, we also find a solution to the above Diophantine equation where $\delta_\ell \in \{\pm 1\}$. Conversely, if we find a solution to (2.49), where $\delta_\ell \in \{\pm 1\}$, then we also find a balanced symmetric Boolean function whose exponential sums corresponds to (2.49).

To see this claim, let $\boldsymbol{e}_{n,[k_1,\cdots,k_s]}$ be a balanced symmetric function. Then

$$S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]}) = \sum_{\ell=0}^{n} (-1)^{\binom{\ell}{k_1}+\cdots+\binom{\ell}{k_s}} \binom{n}{\ell} = 0,$$

so we obviously have a solution $\delta = (\delta_0, \delta_1, \cdots, \delta_n)$ to (2.49) over $\Gamma = \{\pm 1\}$ (such solution is given by $\delta = (\delta_0, \delta_1, \cdots, \delta_n)$ with $\delta_\ell = (-1)^{\binom{\ell}{k_1}+\cdots+\binom{\ell}{k_s}}$).

Now, suppose that $\delta = (\delta_0, \delta_1, \cdots, \delta_n)$ is a solution to (2.49) over $\Gamma = \{\pm 1\}$. Then we have

$$\sum_{\ell=0}^{n} \delta_\ell \binom{n}{\ell} = 0, \text{ where } \delta_\ell \in \{\pm 1\}. \tag{2.51}$$

Let $\mathbf{x}_{(\ell)} \in \mathbb{F}_2^n$ and associate δ_ℓ to a vector $\mathbf{x}_{(\ell)}$ of weight ℓ, for $0 \leq \ell \leq n$, where $\delta_\ell = (-1)^{F(\mathbf{x}_{(\ell)})}$, and

$$(-1)^{F(\mathbf{x}_{(\ell)})} = \begin{cases} 1, & F(\mathbf{x}_{(\ell)}) \equiv 0 \pmod 2, \\ -1, & F(\mathbf{x}_{(\ell)}) \equiv 1 \pmod 2. \end{cases} \tag{2.52}$$

There are $\binom{n}{\ell}$ vectors of weight equal to ℓ. Since F is symmetric, the value of $F(\mathbf{x})$ is the same for each of these vectors (such value is given by $\binom{j}{\ell} \pmod 2$). From here, it is not hard to determine the unique truth table and unique symmetric function associated to a

solution δ to (2.49) over $\Gamma = \{\pm 1\}$. For example, consider the equation

$$\binom{5}{0} + \binom{5}{1} - \binom{5}{2} + \binom{5}{3} - \binom{5}{4} - \binom{5}{5} = 0. \tag{2.53}$$

Its corresponding balanced symmetric Boolean polynomial is $\boldsymbol{e}_{5,[2,3,4]}$. Another example is the equation

$$\begin{aligned}\binom{11}{0} + \binom{11}{1} - \binom{11}{2} - \binom{11}{3} - \binom{11}{4} - \binom{11}{5} + \\ \binom{11}{6} + \binom{11}{7} + \binom{11}{8} + \binom{11}{9} - \binom{11}{10} - \binom{11}{11} = 0.\end{aligned} \tag{2.54}$$

Its corresponding symmetric Boolean function is $\boldsymbol{e}_{11,[4,2]}$. When the set considered is $\Gamma = \{\pm 1\}$, then any solution to (2.49) is said to give a *bisection of the binomial coefficients* $\binom{n}{\ell}$, for $0 \leq \ell \leq n$. Such solution provides us with two sets A and B with $A \cap B = \varnothing$ such that $A \cup B = \{0, 1, 2, \ldots, n\}$ and

$$\sum_{\ell \in A} \binom{n}{\ell} = \sum_{\ell \in B} \binom{n}{\ell} = 2^{n-1}. \tag{2.55}$$

To see this last equation, observe that if

$$\sum_{\ell \in A} \binom{n}{\ell} - \sum_{\ell \in B} \binom{n}{\ell} = 0 \tag{2.56}$$

and

$$\sum_{\ell \in A} \binom{n}{\ell} + \sum_{\ell \in B} \binom{n}{\ell} = \sum_{\ell=0}^{n} \binom{n}{\ell} = 2^n, \tag{2.57}$$

then we have the system of equations

$$\begin{aligned}\sum_{\ell \in A} \binom{n}{\ell} - \sum_{\ell \in B} \binom{n}{\ell} = 0 \\ \sum_{\ell \in A} \binom{n}{\ell} + \sum_{\ell \in B} \binom{n}{\ell} = 2^n.\end{aligned}$$

Adding both equations we obtain

$$\sum_{\ell \in A} \binom{n}{\ell} = 2^{n-1}. \tag{2.58}$$

Thus, (2.55) holds.

Consider (2.49) with $\Gamma = \{-1, 1\}$. Some of its solutions are easy to get. Observe that if n is even, then the Binomial Theorem implies that $\delta_\ell = \pm(-1)^\ell$ is a solution to (2.49) for $0 \leq \ell \leq n$, because $\pm \sum_\ell (-1)^\ell \binom{n}{\ell} = \pm(1-1)^n = 0$. For instance, the equation

$$\delta_0 \binom{4}{0} + \delta_1 \binom{4}{1} + \delta_2 \binom{4}{2} + \delta_3 \binom{4}{3} + \delta_4 \binom{4}{4} = 0 \tag{2.59}$$

has $(\delta_0, \delta_1, \delta_2, \delta_3, \delta_4) = (1, -1, 1, -1, 1)$ and $(-\delta_0, -\delta_1, -\delta_2, -\delta_3, -\delta_4) = (-1, 1, -1, 1, -1)$ as solutions. When n is odd, then the symmetry of the binomial coefficients implies that $(\delta_0, \delta_1, \ldots, \delta_{\frac{n-1}{2}}, -\delta_{\frac{n-1}{2}}, \ldots, -\delta_1, -\delta_0)$ is a solution to (2.49).

Consider, for instance, the equation

$$\delta_0 \binom{3}{0} + \delta_1 \binom{3}{1} + \delta_2 \binom{3}{2} + \delta_3 \binom{3}{3} = 0 \tag{2.60}$$

over $\Gamma = \{-1, 1\}$. One can easily check that the tuples $(1, 1, -1, -1)$, $(1, -1, 1, -1)$, $(-1, 1, -1, 1)$ and $(-1, -1, 1, 1)$ are all solutions to the above Diophantine equation.

The solutions discussed above are called *trivial solutions*. It is not hard to see that when n is even, there are two trivial solutions. On the other hand, when n is odd, there are $2^{\frac{n+1}{2}}$ trivial solutions. A balanced symmetric Boolean function in n variables which corresponds to one of the trivial solutions over $\Gamma = \{-1, 1\}$ is called a *trivially balanced function*. For example, symmetric Boolean function $\boldsymbol{e}_{5,[2,3,4]}$, which corresponds to the equation (2.53) is trivially balanced. The same holds true for $\boldsymbol{e}_{11,[4,2]}$, the symmetric Boolean function that corresponds to (2.54). Balanced symmetric Boolean functions that are not trivially balanced are called *sporadic balanced functions*. Computational experiments suggest that trivially balanced functions are quite common, thus it is of great interest to find those functions that are sporadic balanced. In [13, Thm. 1, p.

2354], Sarkar and Maitra showed that there are an infinite number of sporadic balanced symmetric functions.

We will focus our attention to the balancedness of perturbations of the form $\boldsymbol{e}_{n,[k_1,\cdots,k_s]}+F(\mathbf{X})$, where $1 \leq k_1 < \cdots < k_s$ are fixed integers and $F(\mathbf{X})$ is a Boolean polynomial in j variables (j fixed) and its connection to (2.49). For purposes of simplicity, we will consider the case $\boldsymbol{e}_{n,k} + F(\mathbf{X})$. Recall that,

$$S(\boldsymbol{e}_{n,k} + F(\mathbf{X})) = \sum_{m=0}^{j} C_m(F) S\left(\sum_{i=0}^{m} \binom{m}{i} \boldsymbol{e}_{n-j,k-i} \right), \tag{2.61}$$

where

$$C_m(F) = \sum_{\mathbf{X}\in\mathbb{F}_2^n, w_2(\mathbf{X})=m} (-1)^{F(\mathbf{X})}. \tag{2.62}$$

We can re-write (2.61) as,

$$S(\boldsymbol{e}_{n+j,k} + F(\mathbf{X})) = \sum_{\ell=0}^{n} \Big(\sum_{m=0}^{j} C_m(F) \cdot (-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{\ell}{k-i}} \Big) \binom{n}{\ell}. \tag{2.63}$$

Castro, González and Medina noticed that

$$\begin{aligned}
\left| \sum_{m=0}^{j} C_m(F) \cdot (-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{\ell}{k-i}} \right| &\leq \sum_{m=0}^{j} \left| C_m(F) \cdot (-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{\ell}{k-i}} \right| \\
&= \sum_{m=0}^{j} |C_m(F)| \cdot \left| (-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{\ell}{k-i}} \right| \\
&= \sum_{m=0}^{j} |C_m(F)| \\
&\leq \sum_{m=0}^{j} \binom{j}{m} = 2^j.
\end{aligned}$$

Hence,

$$\sum_{m=0}^{j} C_m(F) \cdot (-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{\ell}{k-i}} \equiv \sum_{m=0}^{j} C_m(F) = S(F) \equiv 0 \pmod 2. \tag{2.64}$$

Therefore, balancedness of a perturbation of the form $\boldsymbol{e}_{n+j,k} + F(\mathbf{X})$ is connected to

solutions of (2.49) over the set

$$\Gamma_j^{(e)} = \{x \in 2\mathbb{Z} : |x| \leq 2^j\} = \{0, \pm 2, \pm 4, \pm 6, \ldots, \pm 2^j\}. \tag{2.65}$$

Any solution to (2.49) over the set $\Gamma_j^{(e)}$ divided by 2 produces a solution of (2.49) over the set

$$\Gamma_j = \{x \in \mathbb{Z} : |x| \leq 2^{j-1}\} = \{0, \pm 1, \pm 2, \pm 3, \ldots, \pm 2^{j-1}\}. \tag{2.66}$$

We can also do the converse way, that is, any solution to (2.49) over Γ_j multiplied by 2 produces a solution over $\Gamma_j^{(e)}$.

We can reach to the same conclusion from a perturbation of the form $\boldsymbol{e}_{n+j,[k_1,\cdots,k_s]} + F(\mathbf{X})$. If $F(\mathbf{X}) = 0$, then we are back to the initial problem of bisecting binomial coefficients with $\Gamma = \{-1, 1\}$ as its corresponding set.

Castro, González and Medina [3] defined *trivial solutions* to (2.49) over Γ_j for bisections of binomial coefficients. If n is odd, then $(\delta_0, \ldots, \delta_{\frac{n-1}{2}}, -\delta_{\frac{n-1}{2}}, \ldots, -\delta_0)$, where $\delta_\ell \in \Gamma_j$, by the symmetry of the binomial coefficients, are solutions to (2.49). There are $(2^j + 1)^{\frac{n+1}{2}}$ such solutions.

Otherwise, if n is even, then $\delta_\ell = (-1)^\ell m$, for $0 \leq \ell \leq n$ and $m \in \Gamma_j$ are the $2^j + 1$ solutions to (2.49) over Γ_j. By the symmetry of the binomial coefficients, this implies that $(\delta_0, \ldots, \delta_{\frac{n}{2}-1}, 0, -\delta_{\frac{n}{2}-1}, \ldots, -\delta_0)$ are the $(2^j + 1)^{\frac{n}{2}}$ solutions to (2.49) over Γ_j. Note that the trivial solution $(0, 0, 0, \ldots, 0)$ is of the form $(\delta_0, \ldots, \delta_{\frac{n-1}{2}}, -\delta_{\frac{n-1}{2}}, \ldots, -\delta_0)$ (for n odd) and $(\delta_0, \ldots, \delta_{\frac{n}{2}-1}, 0, -\delta_{\frac{n}{2}-1}, \ldots, -\delta_0)$ (for n even) and are called *trivial solutions* to (2.49) over Γ_j.

Example 2.2.1. Consider the Diophantine equation

$$\delta_0\binom{5}{0} + \delta_1\binom{5}{1} + \delta_2\binom{5}{2} + \delta_3\binom{5}{3} + \delta_4\binom{5}{4} + \delta_5\binom{5}{5} = 0 \tag{2.67}$$

over the set $\Gamma_5 = \{0 \pm 1, \pm 2, \pm 3, \pm 4, \ldots, \pm 16\}$. Then $(5, -2, -7, 7, 2, -5)$ is one of the $(2^5 - 1)^{\frac{5+1}{2}} = (31)^3$ solutions to (2.67) over Γ_5. Observe that $(0, 0, 0, 0, 0, 0)$ is the trivial solution to (2.67) over Γ_5.

Example 2.2.2. Consider the Diophantine equation

$$\delta_0\binom{8}{0}+\delta_1\binom{8}{1}+\delta_2\binom{8}{2}+\delta_3\binom{8}{3}+\delta_4\binom{8}{4}+\delta_5\binom{8}{5}+\delta_6\binom{8}{6}+\delta_7\binom{8}{7}+\delta_8\binom{8}{8}=0 \quad (2.68)$$

over the set $\Gamma_3 = \{0, \pm1, \pm2, \pm3, \pm4\}$. Then $(-3, 1, -4, 0, 4, -1, 3)$ is a trivial solution over Γ_3. Also, $(0, -1, 2, 4, 0, -4, -2, 1, 0)$ is another trivial solution over Γ_4. There are a total of $(2^3+1)^{\frac{8}{2}} = 9^4$ solutions to (2.68) over Γ_4.

There may be solutions, of course, that may not look trivial at first glance. For example, consider the equation,

$$-2\binom{10}{3}+2\binom{10}{4}-\binom{10}{5}+2\binom{10}{8}-2\binom{10}{9}+2\binom{10}{10}=0 \quad (2.69)$$

which is equivalent to (by the symmetry in the binomial coefficients)

$$\binom{10}{0}-\binom{10}{1}+\binom{10}{2}-\binom{10}{3}+\binom{10}{4}-\binom{10}{5}+\binom{10}{6}-\binom{10}{7}+\binom{10}{8}-\binom{10}{9}+\binom{10}{10}=0. \quad (2.70)$$

So $(0, 0, 0, -2, 2, -1, 0, 0, 2, -2, 2)$ and $(1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1)$ are solutions. Both solutions are said to be *equivalent*.

That led Castro, González and Medina [3] to define equivalence of solutions. We say that $(\delta_0^{(1)}, \delta_1^{(1)}, \ldots, \delta_n^{(1)})$ and $(\delta_0^{(2)}, \delta_1^{(2)}, \ldots, \delta_n^{(2)})$ are *equivalent* and we write $(\delta_0^{(1)}, \delta_1^{(1)}, \ldots, \delta_n^{(1)}) \sim (\delta_0^{(2)}, \delta_1^{(2)}, \ldots, \delta_n^{(2)})$ if,

1. both solutions are non-zero and,

$$\frac{1}{g_1}(\delta_0^{(1)}, \delta_1^{(1)}, \ldots, \delta_n^{(1)}) = \pm\frac{1}{g_2}(\delta_0^{(2)}, \delta_1^{(2)}, \ldots, \delta_n^{(2)}), \quad (2.71)$$

   where $g_i = \gcd(\delta_0^{(i)}, \delta_1^{(i)}, \ldots, \delta_n^{(i)})$.

2. one can obtain one solution from the other by using the symmetry of binomial coefficients.

3. one solution can be obtained from another by combining the previous two cases.

Example 2.2.3. For example,

$$(0,-1,2,4,0,-4,-2,1,0) \sim (0,-9,18,36,0,-36,-18,9,0). \tag{2.72}$$

In general,

$$(0,-1,2,4,0,-4,-2,1,0) \sim (0,-k,2k,4k,0,-4k,-2k,k,0), \tag{2.73}$$

for every integer k.

Example 2.2.4. Observe that

$$(0,0,0,-2,2,-1,0,0,2,-2,2) \sim (1,-1,1,-1,1,-1,1,-1,1,-1,1). \tag{2.74}$$

Also,

$$(2,-2,2,-2,2,-2,2,-2,2,-2,2) \sim (4,-4,4,-4,4,-4,4,-4,4,-4,4). \tag{2.75}$$

Because of the above three cases, now we say that the solution of the previous three forms are written in *trivial form* or just they are called *trivial form solutions*. From this, we now extend the definition of trivial solutions that is equivalent to one of the trivial form solutions.

Let $\delta = (\delta_0, \delta_1, \ldots, \delta_n) \in \Gamma_j$ and define

$$[\delta_0, \delta_1, \ldots, \delta_n] = \{(\delta_0', \delta_1', \ldots, \delta_n') \in \Gamma_j | (\delta_0', \delta_1', \ldots, \delta_n') \sim (\delta_0, \delta_1, \ldots, \delta_n)\}. \tag{2.76}$$

The above set is just the equivalence class of δ under the equivalence relation $\sim$. For example, if n is odd, then every trivial form solution is equivalent to $(0,0,0,\ldots,0)$, so hence, they belong to the equivalence class $[0,0,0,\ldots,0]$. If, on the other hand, n is even, then every trivial form solution is either equivalent to $(1,-1,1,-1,\ldots,-1,1)$ or $(0,0,0,\ldots,0)$, so they belong to either one of the equivalence classes $[1,-1,1,-1,\ldots,-1,1]$ or $[0,0,0,\ldots,0]$.

Castro, González and Medina [3] defined $\Omega(n,j) = \{\delta \in \Gamma_j^n | \delta \text{ is a solution to (2.49)}\}$

and $\gamma_j(n) := |\Omega(n,j)|$, that is, the number of solutions to (2.49) over Γ_j. They also defined $\gamma_j^*(n)$ to be the number of trivial form solutions. Then

$$\gamma_j^*(n) = \begin{cases} (2^j+1)^{\frac{n}{2}} + 2^j, & \text{if } n \text{ is even,} \\ (2^j+1)^{\frac{n+1}{2}}, & \text{if } n \text{ is odd.} \end{cases} \tag{2.77}$$

So from this, we can see that the number of solutions to Γ_j grows exponentially as n grows, which can already be seen in the number of trivial form solutions. Note that in the set $\Omega(n,j)$ it includes both the trivial and the non-trivial solutions to (2.49) over Γ_j, so it is expected that $\gamma_j(n) \geq \gamma_j^*(n)$, for each n and j.

The following theorem gives an integral representation for $\gamma_j(n)$, whose proof uses similar techniques as the ones used in [11]. This formula was provided in [3], but without proof.

Theorem 2.2.1. *Let $\mathbb{V}_j = [0, 2^{j-1}] \cap \mathbb{Z}$, and $w(\boldsymbol{x})$ represent the number of non-zero entries of $\boldsymbol{x} = (x_0, x_1, \ldots, x_n) \in \mathbb{V}_j^{n+1}$. Then,*

$$\gamma_j(n) = \sum_{\boldsymbol{x} \in \mathbb{V}_j^{n+1}} 2^{w(\boldsymbol{x})} \int_0^1 \prod_{\ell=0}^{n} \cos\left(\pi x_\ell \binom{n}{\ell} s\right) ds.$$

Proof. We will consider the number of solutions (all the possible choices of the signs $+$ and $-$) of the equation

$$\pm x_0 \binom{n}{0} \pm x_1 \binom{n}{1} \pm \cdots \pm x_n \binom{n}{n} = 0, \tag{2.78}$$

over the set $\Gamma_j^* = \{0, 1, 2, 3, \ldots, 2^{j-1}\}$. Use the following identity from trigonometry

$$\prod_{\ell=0}^{n} \cos\left(x_\ell \binom{n}{\ell} t\right) = \frac{1}{2^{w(\mathbf{x})}} \sum \cos\left(\left(\pm x_0 \binom{n}{0} \pm x_1 \binom{n}{1} \pm \cdots \pm x_n \binom{n}{n}\right) t\right). \tag{2.79}$$

and integrate both sides with respect to t in the interval $[-\pi, \pi]$ to obtain

$$\int_{-\pi}^{\pi} \prod_{\ell=0}^{n} \cos\left(x_\ell \binom{n}{\ell} t\right) dt = \frac{1}{2^{w(\mathbf{x})}} \sum \int_{-\pi}^{\pi} \cos\left(\left(\pm x_0 \binom{n}{0} \pm x_1 \binom{n}{1} \pm \cdots \pm x_n \binom{n}{n}\right) t\right) dt.$$

Remember that in the expression

$$\pm x_0\binom{n}{0} \pm x_1\binom{n}{1} \pm \cdots \pm x_n\binom{n}{n} \tag{2.80}$$

we are considering all the possible choices of the the + and − signs, so there may be some choices that may not constitute solutions to (2.78). But the terms in which the choices does not constitute possible solutions vanishes because of the identity

$$\int_{-\pi}^{\pi} \cos(mt)dt = \begin{cases} 2\pi, & \text{if } m = 0, \\ 0, & \text{otherwise.} \end{cases} \tag{2.81}$$

Hence, if N represent the number of solutions of (2.78), then we have

$$\frac{1}{2^{w(\mathbf{x})}} \sum \int_{-\pi}^{\pi} \cos\left(\left(\pm x_0\binom{n}{0} \pm x_1\binom{n}{1} \pm \cdots \pm x_n\binom{n}{n}\right)t\right)dt = \frac{2\pi N}{2^{w(\mathbf{x})}}. \tag{2.82}$$

So we re-write this last formula as

$$\frac{2\pi N}{2^{w(\mathbf{x})}} = \int_{-\pi}^{\pi} \prod_{\ell=0}^{n} \cos\left(x_\ell\binom{n}{\ell}t\right)dt, \tag{2.83}$$

and solving for N, we get

$$N = \frac{2^{w(\mathbf{x})}}{2\pi} \int_{-\pi}^{\pi} \prod_{\ell=0}^{n} \cos\left(x_\ell\binom{n}{\ell}t\right)dt = \frac{2^{w(\mathbf{x})}}{\pi} \int_{0}^{\pi} \prod_{\ell=0}^{n} \cos\left(x_\ell\binom{n}{\ell}t\right)dt. \tag{2.84}$$

Making the substitution $t = \pi s$, we transform (2.84) into

$$N = 2^{w(\mathbf{x})} \int_{0}^{1} \prod_{\ell=0}^{n} \cos\left(\pi x_\ell\binom{n}{\ell}s\right)ds.$$

Therefore, if we add over all possible tuples $\boldsymbol{x} = (x_0, x_1, \cdots, x_n) \in \mathbb{V}_j^{n+1}$ of (2.78), we get

$$\gamma_j(n) = \sum_{\mathbf{x}\in\mathbb{V}_j^{n+1}} 2^{w(\mathbf{x})} \int_{0}^{1} \prod_{\ell=0}^{n} \cos\left(\pi x_\ell\binom{n}{\ell}s\right)ds, \tag{2.85}$$

which is our desired result. (Q.E.D.)

The number of solutions $\gamma_j(n)$ over Γ_j grows exponentially in n, which can be seen from the number of trivial form solutions. Table 2.5 shows the values of $\gamma_j(n)$ for various n's and j's. Note that the solutions considered in Table 2.5 include the trivial and the non-trivial solutions which are bigger than the numbers that consider only the trivial form solutions which are shown in Table 2.6.

n	1	2	3	4	5	6	7	8	9	10
$\gamma_1(n)$	3	5	9	15	39	45	129	149	243	369
$\gamma_2(n)$	5	13	41	103	275	685	2525	5221	13897	32717
$\gamma_3(n)$	9	41	219	1033	5181	23035	121921	*	*	*
$\gamma_4(n)$	17	145	1469	12969	120521	*	*	*	*	*
$\gamma_5(n)$	33	545	10659	183477	*	*	*	*	*	*
$\gamma_6(n)$	65	2113	81421	*	*	*	*	*	*	*
$\gamma_7(n)$	129	8321	636099	*	*	*	*	*	*	*
$\gamma_8(n)$	257	33025	*	*	*	*	*	*	*	*
$\gamma_9(n)$	513	131585	*	*	*	*	*	*	*	*
$\gamma_{10}(n)$	1025	*	*	*	*	*	*	*	*	*

Table 2.5: Number of solutions to (2.49) that lies in Γ_j, for $1 \leq n, j \leq 10$.

One should note that many of the solutions that are counted in Table 2.5 are equivalent to some others, so the amount of "meaningful" solutions should be expected to be significantly smaller than the numbers presented in Table 2.5.

Define $\omega_j(n)$ to be the number of different equivalent classes on $\Omega(n, j)$ under the equivalent relation $\sim$, that is, the cardinality of the quotient set $\Omega(n, j)/\sim$. For example,

n	1	2	3	4	5	6	7	8	9	10
$\gamma_1^*(n)$	3	5	9	11	27	29	81	83	243	245
$\gamma_2^*(n)$	5	9	25	29	125	129	625	629	3125	3129
$\gamma_3^*(n)$	9	17	81	89	729	737	6561	6569	59049	59057
$\gamma_4^*(n)$	17	33	289	305	4913	4929	83521	83537	*	*
$\gamma_5^*(n)$	33	65	1089	1121	35937	35969	*	*	*	*
$\gamma_6^*(n)$	65	129	4225	4289	274625	274689	*	*	*	*
$\gamma_7^*(n)$	129	257	16641	16769	*	*	*	*	*	*
$\gamma_8^*(n)$	257	513	*	*	*	*	*	*	*	*
$\gamma_9^*(n)$	513	1025	*	*	*	*	*	*	*	*
$\gamma_{10}^*(n)$	1025	2049	*	*	*	*	*	*	*	*

Table 2.6: Number of trivial form solutions to (2.49) that lie in Γ_j, for $1 \leq n, j \leq 10$.

if $n = 7$, and $j = 2$, then

$$\Omega(7,2)/\sim = \{[0,0,0,0,0,0,0,0], [0,-2,1,0,0,0,-1,0], [0,-1,2,-1,0,0,0,0],$$
$$[0,1,2,-1,-1,1,0,0], [0,2,1,-1,0,0,0,0], [0,2,2,-2,-1,2,1,0], [0,2,-2,1,0,-1,2,0]\}. \tag{2.86}$$

Thus, $\omega_2(7) = 7$.

For $n = 5$ and $j = 3$, we have

$$\Omega(5,3)/\sim = \{[0,0,0,0,0,0], [0,2,-1,0,0,0], [4,-4,3,0,-3,1], [4,-4,2,0,-1,1],$$
$$[4,-3,1,0,0,1], [4,-1,0,0,0,1], [4,1,-1,0,0,1], [4,3,-2,0,0,1],$$
$$[4,4,-3,0,1,1], [4,4,-4,0,3,1]\}, \tag{2.87}$$

and so, $\omega_3(5) = 10$. Table 2.7 shows the values of $\omega_j(n)$ for various n's and j's. These numbers are quite smaller in comparison with the values of $\gamma_j(n)$ in Table 2.5, as we expected. Of course, there are many "meaningful" solutions that are not trivial. We don't known if an explicit formula exist for the number $\omega_j(n)$ of equivalence classes on $\Omega(n,j)$.

n	1	2	3	4	5	6	7	8	9	10
$\omega_1(n)$	1	2	1	3	2	3	3	7	1	5
$\omega_2(n)$	1	2	2	5	2	13	7	36	26	71
$\omega_3(n)$	1	2	2	13	10	72	77	389	274	1681
$\omega_4(n)$	1	2	2	45	37	504	443	5076	4336	*
$\omega_5(n)$	1	2	2	161	127	3811	3119	*	*	*
$\omega_6(n)$	1	2	2	649	481	29742	*	*	*	*
$\omega_7(n)$	1	2	2	2521	2005	*	*	*	*	*
$\omega_8(n)$	7	36	389	5076	*	*	*	*	*	*
$\omega_9(n)$	1	26	274	*	*	*	*	*	*	*
$\omega_{10}(n)$	5	71	1681	*	*	*	*	*	*	*

Table 2.7: Values of $\omega_j(n)$, for $1 \leq n, j \leq 10$.

Castro, González and Medina extended the concepts of trivially and sporadic balanced functions to perturbations of symmetric Boolean functions as well. We say that a perturbation $\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X})$ is *trivially balanced* if it corresponds to one of the trivial

solutions of (2.49) over Γ_j. Otherwise, we say that it is *sporadic balanced*. Unlike for the case of sporadic balanced symmetric functions, it is (currently) unknown if there is an infinite number of sporadic balanced perturbations (assuming, of course, that $F \neq 0$).

As for the case of balanced symmetric functions, it seems that most perturbations are trivially balanced. As a particular case, consider the simplest perturbation $\boldsymbol{e}_{n,k} + X_1$. It appears that this perturbation is trivially balanced when the number of variables is $n = 2^r m + k - 1$, where $r = \lfloor \log_2(k) \rfloor + 1$. Before we go into that result, we need an auxiliary lemma.

Lemma 2.2.1. *Let $n = 2^r m + k - 1$, where $r = \lfloor \log_2(k) \rfloor + 1$. Define $N(k) = \{\ell \in \mathbb{Z} | \binom{\ell}{k}$ is odd$\}$. Then for $0 \leq \ell \leq n$, $\ell \in N(k) \Leftrightarrow n - \ell \in N(k)$.*

Proof. Suppose that $\ell \in N(k)$. Then by the definition of $N(k)$, the binomial coefficient $\binom{\ell}{k}$ is odd, so it follows that, by Lucas' theorem, the binomial coefficient $\binom{\ell}{k}$ is odd if and only if

$$\ell = k + 2^{\delta_1} + 2^{\delta_2} + \cdots + 2^{\delta_t}, \tag{2.88}$$

where $k = 2^{\alpha_1} + 2^{\alpha_2} + \cdots + 2^{\alpha_s}$, and $2^{\delta_i} \notin \{2^{\alpha_1}, 2^{\alpha_2}, \ldots, 2^{\alpha_s}\}$. Now,

$$n - \ell = 2^r m + k - \ell - (k + 2^{\delta_1} + 2^{\delta_2} + \cdots + 2^{\delta_t}) = 2^r m - (2^{\delta_1} + 2^{\delta_2} + \cdots + 2^{\delta_t}). \tag{2.89}$$

Consider two cases:

(a) (k is even): If m is even, then $m = 2c$, for some $c \in \mathbb{Z}$.

$$\begin{aligned}
n - \ell &= 2^r m - \ell - (2^{\delta_1} + 2^{\delta_2} + \cdots + 2^{\delta_t}) \\
&= 2^{r+1} c - 1 - (2^{\delta_1} + 2^{\delta_2} + \cdots + 2^{\delta_t}) \\
&= 2^{\beta_1} + \cdots + 2^{\beta_w} - 1 - (2^{\delta_1} + 2^{\delta_2} + \cdots + 2^{\delta_t}) \\
&= (2^{\beta_1} + \cdots + 2^{\beta_w} - 2^r) + [2^r - 1 - (2^{\delta_1} + 2^{\delta_2} + \cdots + 2^{\delta_t})] \\
&= (2^{\gamma_1} + \cdots + 2^{\gamma_r}) + (k + 2^{\epsilon_1} + \cdots + 2^{\epsilon_q}) \\
&= k + (2^{\epsilon_1} + \cdots + 2^{\epsilon_q} + 2^{\gamma_1} + \cdots + 2^{\gamma_r}),
\end{aligned}$$

where $2^{\epsilon_i}, 2^{\gamma_j} \notin \{2^{\alpha_1}, \ldots, 2^{\alpha_s}\}$. This implies that $\binom{n-\ell}{k}$ is odd, by Lucas' theorem, so $n - \ell \in N(k)$.

(b) (k is odd): If m is odd, then $m = 2c + 1$, for some $c \in \mathbb{Z}$. Then we write,

$$\begin{aligned} n - \ell &= 2^r m - \ell - (2^{\delta_1} + 2^{\delta_2} + \cdots + 2^{\delta_t}) \\ &= 2^{r+1}c + 2^r - 1 - (2^{\delta_1} + 2^{\delta_2} + \cdots + 2^{\delta_t}) \\ &= 2^{r+1}c + [2^r - 1 - (2^{\delta_1} + 2^{\delta_2} + \cdots + 2^{\delta_t})] \\ &= (2^{r+1}c) + k + (2^{\epsilon_1} + \cdots + 2^{\epsilon_q}) \\ &= k + (2^{\beta_1} + \cdots + 2^{\beta_w}) + (2^{\epsilon_1} + \cdots + 2^{\epsilon_q}), \end{aligned}$$

where $2^{\beta_i}, 2^{\epsilon_j} \notin \{2^{\alpha_1}, \ldots, 2^{\alpha_s}\}$. Hence, $\binom{n-\ell}{k}$ is odd, by Lucas' theorem. Therefore, $n - \ell \in N(k)$.

Now, if $n - \ell \in N(k)$, then $\binom{n-\ell}{k}$ is odd. So we write (by Lucas' theorem)

$$2^r m + k - 1 - \ell = n - \ell = k + 2^{\alpha_1} + \cdots + 2^{\alpha_s}, \tag{2.90}$$

Solving for ℓ, we get

$$\ell = 2^r m - 1 - (2^{\alpha_1} + \cdots + 2^{\alpha_s}). \tag{2.91}$$

Dividing into cases as above (according to the parity of m) and following the same steps as in each of the cases above, we reach to the conclusion that $\binom{\ell}{k}$ is odd, so $\ell \in N(k)$. (Q.E.D.)

Theorem 2.2.2. *Let k be a natural number and $r = \lfloor \log_2(k) \rfloor + 1$. Then the perturbation $\boldsymbol{e}_{n,k} + X_1$ is trivially balanced where $n = 2^r m + k - 1$, for $m \in \mathbb{N}$.*

Proof. Recall the identity

$$\begin{aligned} S(\boldsymbol{e}_{n,k} + X_1) &= S(\boldsymbol{e}_{n-1,k}) - S(\boldsymbol{e}_{n-1,[k,k-1]}) \\ &= \sum_{\ell=0}^{n-1} \left[(-1)^{\binom{\ell}{k}} - (-1)^{\binom{\ell}{k} + \binom{\ell}{k-1}} \right] \binom{n-1}{\ell}. \end{aligned}$$

Define $N(k) = \{\ell \in \mathbb{Z} | \binom{\ell}{k} \text{ is odd}\}$ and $N(k, k-1) = \{\ell \in \mathbb{Z} | \binom{\ell}{k} + \binom{\ell}{k-1} \text{ is odd}\}$. Then $\ell \in N(k) \Leftrightarrow \ell - 1 \in N(k, k-1)$, because $\ell \in N(k) \Leftrightarrow \binom{\ell}{k} = \binom{\ell-1}{k} + \binom{\ell-1}{k-1}$ is odd $\Leftrightarrow \ell - 1 \in N(k, k-1)$. Now, since $n = 2^r m + k - 1$, then for $0 \leq \ell \leq n$, $\ell \in N(k) \Leftrightarrow n - \ell \in N(k)$ (by Lemma 2.2.1).

Suppose now that $\ell_0 \in N(k)$. Then $\binom{\ell_0}{k}$ is odd and the coefficient of $\binom{n-1}{\ell_0}$ in the first sum is -1. Then $n - \ell_0 \in N(k)$ and so $n - \ell_0 - 1 \in N(k, k-1)$, so the coefficient $\binom{n-1}{n-\ell_0-1} = \binom{n-1}{\ell_0}$ in the second sum is also -1, that is, $\binom{n-\ell_0}{k} + \binom{n-\ell_0}{k-1}$ is odd. Conversely, if $\ell_0' \in N(k, k-1)$, then $\binom{\ell_0'}{k} + \binom{\ell_0'}{k-1} = \binom{\ell_0'+1}{k}$ is odd and hence, the coefficient of $\binom{n-1}{\ell_0'}$ in the second sum is -1, so $\ell_0' + 1 \in N(k)$. This implies that $n - (\ell_0' + 1) = n - \ell_0' - 1 \in N(k)$, so the coefficient $\binom{n-1}{n-\ell_0'-1} = \binom{n-1}{\ell_0'}$ in the first sum is also -1. Hence,

$$\sum_{\ell=0}^{n-1} \left[(-1)^{\binom{\ell}{k}} - (-1)^{\binom{\ell}{k}+\binom{\ell}{k-1}}\right] \binom{n-1}{\ell} = 0.$$

so $S(\boldsymbol{e}_{n,k} + X_1) = 0$ and the perturbation $\boldsymbol{e}_{n,k} + X_1$ is trivially balanced. The proof is complete. (Q.E.D.)

Example 2.2.5. Let $k = 7$, then $r = \lfloor \log_2(7) \rfloor + 1 = 3$, so if we let $m = 3$, then $n = 2^r m + k - 1 = 2^3(3) + 7 - 1 = 8(3) + 7 - 1 = 30$, so by Theorem 2.2.2, $\boldsymbol{e}_{30,7} + X_1$ is trivially balanced. In particular,

$$\begin{aligned} S(\boldsymbol{e}_{30,7} + X_1) &= \sum_{\ell=0}^{29} \left[(-1)^{\binom{\ell}{7}} - (-1)^{\binom{\ell}{7}+\binom{\ell}{6}}\right] \binom{29}{\ell} \\ &= 2\binom{29}{6} - 2\binom{29}{7} + 2\binom{29}{14} - 2\binom{29}{15} + 2\binom{29}{22} - 2\binom{29}{23} \\ &= 0 \text{ (trivially balanced).} \end{aligned}$$

Example 2.2.6. Let $k = 12$, then $r = \lfloor \log_2(12) \rfloor + 1 = 4$. Let $m = 4$, then $n = 2^r m + k - 1 = 2^4(4) + 12 - 1 = 43$. Then Theorem 2.2.2 says that $\boldsymbol{e}_{43,12} + X_1$ is balanced,

so

$$\begin{aligned} S(\boldsymbol{e}_{43,12}+X_1) &= \sum_{\ell=0}^{42}\left[(-1)^{\binom{\ell}{12}}-(-1)^{\binom{\ell}{12}+\binom{\ell}{11}}\right]\binom{42}{\ell} \\ &= 2\binom{42}{11}-2\binom{42}{15}+2\binom{42}{27}-2\binom{42}{31} \\ &= 0 \text{ (trivially balanced).} \end{aligned}$$

The authors in [3] tried to find sporadic balanced function of the form $\boldsymbol{e}_{n,k}+X_1$, but their attempts failed, and that led them to believe that a conjecture similar to the one presented by Cusick et al in [6] for the case of elementary symmetric Boolean functions holds. So from this, Castro, González and Medina [3] conjectured the following: "*No perturbation of the form* $\boldsymbol{e}_{n,k}+X_1$ *is balanced expect for the trivial cases, that is, when* $n=2^r m+k-1$, *where* $r=\lfloor \log_2(k)\rfloor+1$, *and* $m\in\mathbb{N}$".

The next result shows a particular family of a trivially balanced perturbation whose proof will be ommited.

Theorem 2.2.3. *The perturbation*

$$\boldsymbol{e}_{2^{\ell+1}D-1,2^\ell}+X_1+X_2+\cdots+X_{2m}, \tag{2.92}$$

where D,ℓ *and* m *are positive integers, is trivially balanced. In view of Theorem 2.1.3, the perturbation*

$$\boldsymbol{e}_{2^{\ell+1}D,2^\ell+1}+X_1+X_2+\cdots+X_{2m}, \tag{2.93}$$

is also trivially balanced.

The two families of perturbations in Theorems 2.2.2 and 2.2.3 are the only ones that are known so far. As part of future work, we will try searching for other families of trivially balanced perturbations.

As we mentioned before, it appears that most of perturbations that are balanced are actually trivially balanced. Observe that, in the case of Theorem 2.2.2, if the perturbation $\boldsymbol{e}_{n,k}+X_1$ is balanced at one point, that it, $n=2^r+k-1$, then $\boldsymbol{e}_{n,k}+X_1$ it is trivially

balanced for infinitely many n, because of the choice of $n = 2^r m + k - 1$, for every positive integer m. This allows us to have infinite families of trivially balanced functions of the form $\boldsymbol{e}_{n,k} + X_1$, where $n = 2^r m + k - 1$, for all $m \in \mathbb{N}$. This it turns out to be true, not only for this particular perturbation, but for any perturbation as the next theorem shows.

Theorem 2.2.4. *[3] Let $1 \leq k_1 < \cdots < k_s$ be integers and $F(\boldsymbol{X})$ be a Boolean polynomial in the variables $X_1, \cdots, X_j$. Let $r = \lfloor \log_2(k_s) \rfloor + 1$. Suppose that n_0 is a positive integer such that $\boldsymbol{e}_{n_0+j,[k_1,\cdots,k_s]} + F(\boldsymbol{X})$ is trivially balanced. Then, $\boldsymbol{e}_{n_0+t\cdot 2^r+j,[k_1,\cdots,k_s]} + F(\boldsymbol{X})$ is trivially balanced for all non-negative integers t.*

Proof. This proof is taken from [3]. We include it for the reader to see the argument.

Suppose that n_0 is such that the perturbation $\boldsymbol{e}_{n_0+j,[k_1,\cdots,k_s]}+F(\mathbf{X})$ is trivially balanced, where $F(\mathbf{X})$ is a Boolean polynomial in the variables $X_1, \cdots, X_j$. For simplicity in our writing, suppose that $(\delta_0, \delta_1, \cdots, \delta_{n_0})$ is a trivial solution of the form $\delta_\ell = -\delta_{n_0-\ell}$ to the equation

$$\sum_{\ell=0}^{n_0} \delta_\ell \binom{n_0}{\ell} = 0. \tag{2.94}$$

Since $\boldsymbol{e}_{n_0+j,[k_1,\cdots,k_s]} + F(\mathbf{X})$ is trivially balanced, then it corresponds to a trivial solution $(\delta_0, \delta_1, \cdots, \delta_{n_0})$ to the equation above. Now, δ_ℓ is of the form

$$\begin{aligned}\delta_\ell &= \sum_{m=0}^{j} C_m(F)(-1)^{\sum_{i=0}^{m}\binom{m}{i}\left[\binom{\ell}{k_1-i}+\cdots+\binom{\ell}{k_s-i}\right]} \\ &= \sum_{m=0}^{j} C_m(F)(-1)^{\binom{\ell+m}{k_1}+\cdots+\binom{\ell+m}{k_s}}.\end{aligned}$$

We know that the binomial coefficients

$$\binom{\ell+m}{k_1}, \cdots, \binom{\ell+m}{k_s} \tag{2.95}$$

are all periodic modulo 2 with a period length of 2^r, that is

$$\binom{\ell+m+t\cdot 2^r}{k_i} \equiv \binom{\ell+m}{k_i} \pmod 2, \tag{2.96}$$

for each $i = 1, 2, \cdots, s$, and each non-negative integer t. From this fact, we get

$$\delta_\ell = \delta_{\ell+t\cdot 2^r} \text{ and } \delta_{n_0-\ell} = \delta_{n_0-\ell+t\cdot 2^r}. \tag{2.97}$$

But $\delta_\ell = -\delta_{n_0-\ell} = -\delta_{n_0-\ell+t\cdot 2^r}$. Hence the tuple $(\delta_0, \delta_1, \cdots, \delta_{n_0+t\cdot 2^r})$ is a trivial solution to (2.49) over Γ_j, so this tuple corresponds to $S(\boldsymbol{e}_{n_0+t\cdot 2^r+j,[k_1,\cdots,k_s]} + F(\mathbf{X}))$, that is

$$S(\boldsymbol{e}_{n_0+t\cdot 2^r+j,[k_1,\cdots,k_s]} + F(\mathbf{X})) = \sum_{\ell=0}^{n_0+t\cdot 2^r} (2\delta_\ell)\binom{n+t\cdot 2^r}{\ell} = 0. \tag{2.98}$$

Therefore, $\boldsymbol{e}_{n_0+t\cdot 2^r+j,[k_1,\cdots,k_s]} + F(\mathbf{X})$ is trivially balanced, so we are done. (Q.E.D.)

Example 2.2.7. Consider the perturbation $\boldsymbol{e}_{11,2} + X_1 + X_2 + X_3 + X_4$. Then by Theorem 2.2.3 (taking $\ell = 1$, $D = 3$), this perturbation is trivially balanced. Write

$$\begin{aligned} S(\boldsymbol{e}_{11,2} + X_1 + X_2 + X_3 + X_4) &= \sum_{m=0}^{4} C_m(F) S\left(\sum_{i=0}^{m}\binom{m}{i}\boldsymbol{e}_{7,2-i}\right) \\ &= \sum_{\ell=0}^{7}\left(\sum_{m=0}^{4} C_m(F)(-1)^{\binom{m+\ell}{2}}\right)\binom{7}{\ell}. \end{aligned}$$

Then the corresponding equation to this perturbation is

$$\sum_{\ell=0}^{7}\delta_\ell\binom{7}{\ell} = 0, \tag{2.99}$$

whose corresponding solution is

$$\begin{aligned} \delta &= (\delta_0, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7) \\ &= (-4, -4, 4, 4, -4, -4, 4, 4). \end{aligned}$$

Then Theorem 2.2.4 implies that the perturbation $S(\boldsymbol{e}_{11+16t,2} + X_1 + X_2 + X_3 + X_4)$ is trivially balanced for every non-negative integer t. Since $X_1 + X_2 + X_3 + X_4$ is balanced, then Theorem 2.2.3 implies that $\boldsymbol{e}_{12,3} + X_1 + X_2 + X_3 + X_4$ is also balanced. Now, we

write

$$S(\boldsymbol{e}_{12,3} + X_1 + X_2 + X_3 + X_4) = \sum_{\ell=0}^{8} \left(\sum_{m=0}^{4} C_m(F)(-1)^{\binom{m+\ell}{3}} \right) \binom{8}{\ell} \tag{2.100}$$

This perturbation corresponds to the Diophantine equation

$$\sum_{\ell=0}^{8} \delta_\ell \binom{8}{\ell} = 0, \tag{2.101}$$

whose solution is given by

$$\begin{aligned} \delta &= (\delta_0, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8) \\ &= (8, -12, 8, -4, 8, -12, 8, -4, 8) \\ &\sim (8, -8, 8, -8, 8, -8, 8 - 8, 8) \\ &\sim (1, -1, 1, -1, 1, -1, 1, -1, 1). \end{aligned}$$

Example 2.2.8. Consider the perturbation $\boldsymbol{e}_{23,4} + X_1 + X_2 + X_3 + X_4 + X_5 + X_6$. This perturbation is trivially balanced (take $\ell = 2$, $D = 3$), and write

$$S(\boldsymbol{e}_{23,4} + X_1 + X_2 + X_3 + X_4 + X_5 + X_6) = \sum_{\ell=0}^{17} \left(\sum_{m=0}^{6} C_m(F)(-1)^{\binom{\ell+m}{4}} \right) \binom{17}{\ell}$$

whose corresponding Diophantine equation is

$$\sum_{\ell=0}^{17} \delta_\ell \binom{17}{\ell} = 0,$$

where

$$\begin{aligned} \delta &= (\delta_0, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8, \delta_9, \delta_{10}, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14}, \delta_{15}, \delta_{16}, \delta_{17}) \\ &= (-20, 20, -8, -8, 20, -20, 8, 8, -20, 20, -8, -8, 20, -20, 8, 8, -20, 20). \end{aligned}$$

By Theorem 2.1.3, the perturbation $\boldsymbol{e}_{24,5} + X_1 + X_2 + X_3 + X_4 + X_5 + X_6$ is also trivially

balanced. Write

$$S(\boldsymbol{e}_{24,5} + X_1 + X_2 + X_3 + X_4 + X_5 + X_6) = \sum_{\ell=0}^{18} \left(\sum_{m=0}^{6} C_m(F)(-1)^{\binom{m+\ell}{5}} \right) \binom{18}{\ell},$$

with corresponding Diophantine equation

$$\sum_{\ell=0}^{18} \delta_\ell \binom{18}{\ell} = 0,$$

and solution given by

$$\begin{aligned} \delta &= (\delta_0, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8, \delta_9, \delta_{10}, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14}, \delta_{15}, \delta_{16}, \delta_{17}, \delta_{18}) \\ &= (12, -32, 52, -60, 52, -32, 12, -4, 12, -32, 52, -60, 52, -32, 12, -4, 12, -32, 52) \\ &\sim (1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1). \end{aligned}$$

Example 2.2.9. Consider the perturbation $\boldsymbol{e}_{7,4} + X_1 + X_2$. This perturbation is trivially balanced. Its corresponding Diophantine equation is

$$-2\binom{5}{2} + 2\binom{5}{3} = 0.$$

By Theorem 2.1.3, the perturbation $\boldsymbol{e}_{8,5} + X_1 + X_2$ is trivially balanced.

Chapter 3

Balancedness of Perturbations as n Grows

In this chapter, we study the characterization of balanced perturbations of fixed degree when the number of variables is big enough, which is the central topic of the first section of this chapter. Canteaout and Videau in [2] conjectured this situation, but for the case of symmetric Boolean functions. They conjectured that balanced symmetric functions of fixed degree do not exist when the number of variables grows, except for the trivial cases. However, Canteaout and Videau's conjecture was proved by Guo et al in [10]. Particularly, they showed that for n big enough, balanced elementary Boolean functions of fixed degree do not exist, except for the trivial cases $\boldsymbol{e}_{2^{\ell+1}D-1,2^\ell}$, for $\ell, D \in \mathbb{Z}_{\geq 0}$. At the end of this chapter, we provide some examples of sporadic balanced perturbations which are linked with some special Diophantine binomial equations.

3.1 Balancedness of Perturbations as the Number of Variables Grows

Consider a perturbation of the form $\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X})$, where $1 \leq k_1 < \cdots < k_s$ are fixed integers and $F(\mathbf{X})$ is a Boolean polynomial in the variables $X_1, X_2, \cdots, X_j$ (j fixed). The problem is to characterize $S(\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + F(\mathbf{X})) = 0$ for n big enough. We already know that this exponential sum, when seen as a sequence (let n vary), it satisfies the linear

recurrence with integer coefficients

$$x_n = \sum_{\ell=1}^{2^r-1} (-1)^{\ell-1} \binom{2^r}{\ell} x_{n-\ell}, \tag{3.1}$$

where $r = \lfloor \log_2(k_s) \rfloor + 1$. The sequence of exponential sums of our perturbation is a real solution to the linear recurrence (3.1). Recall that the characteristic polynomial associated to (3.1) is

$$(X-2)\Phi_4(X-1)\Phi_8(X-1)\cdots\Phi_{2^r}(X-1). \tag{3.2}$$

Any solution $\{a_n\}$ to (3.1) is of the form

$$a_n = d_0 \cdot 2^n + \sum_{\ell=1}^{2^r-1} d_\ell \cdot \lambda_\ell^n, \tag{3.3}$$

where $\lambda_\ell = 1 + \xi_\ell^{-1}$, with $\xi_\ell = \exp(\frac{i\pi\ell}{2^{r-1}})$, with $i = \sqrt{-1}$. Observe that $\xi_{2^r-\ell} = \bar{\xi}_\ell$, $\lambda_{2^r-\ell} = \bar{\lambda}_\ell$, and $\lambda_{2^r-1} = 0$. If $\{a_n\}$ is a real solution to (3.1), then following the similar techniques used by Guo et al [10], we express $\{a_n\}$ as

$$\begin{aligned}
a_n &= d_0 \cdot 2^n + \sum_{\ell=1}^{2^r-1} d_\ell \cdot \lambda_\ell^n \\
&= d_0 \cdot 2^n + \sum_{\ell=1}^{2^{r-1}-1} d_\ell \cdot \lambda_\ell^n + \sum_{\ell=2^{r-1}}^{2^r-1} d_\ell \cdot \lambda_\ell^n \\
&= d_0 \cdot 2^n + \sum_{\ell=1}^{2^{r-1}-1} d_\ell \cdot \lambda_\ell^n + \sum_{l=1}^{2^{r-1}-1} \overline{d_\ell \cdot \lambda_\ell^n} \\
&= d_0 \cdot 2^n + \sum_{\ell=1}^{2^{r-1}-1} \left(d_\ell \cdot \lambda_\ell^n + \overline{d_\ell \cdot \lambda_\ell^n} \right) \\
&= d_0 \cdot 2^n + 2 \sum_{\ell=1}^{2^{r-1}-1} \mathrm{Re}(d_\ell \cdot \lambda_\ell^n).
\end{aligned}$$

In the last equality we used the fact that $z + \overline{z} = 2\,\mathrm{Re}(z)$, where $\mathrm{Re}(z)$ denotes the real part of the complex number z. Now, define $t_\ell(n) = \mathrm{Re}(d_\ell \cdot \lambda_\ell^n)$, for $0 \le \ell \le 2^{r-1} - 1$.

Then we can rewrite our latter sum in terms of $t_\ell(n)$ as

$$a_n = t_0(n) + 2\sum_{\ell=0}^{2^{r-1}-1} t_\ell(n). \tag{3.4}$$

The next lemma gives a characterization of when $a_n = 0$, for some n big enough.

Lemma 3.1.1. *Suppose that $\{a_n\}$ is a real solution to (3.1). Then there exists an integer n_0 such that for any $n > n_0$, $a_n = 0$ if and only if $t_\ell(n) = 0$, for $0 \leq \ell \leq 2^{r-1} - 1$.*

Proof. The proof is inspired by the one presented in [1]. Suppose that $a_n = 0$. The number $t_\ell(n)$ can be expressed in the following form

$$t_\ell(n) = |d_\ell|\left|2\cos\left(\frac{\pi\ell}{2^r}\right)\right|^n \cos\left(\arg(d_\ell) - \frac{\pi n\ell}{2^r}\right). \tag{3.5}$$

Since $\left|\cos\left(\arg(d_\ell) - \dfrac{\pi n\ell}{2^r}\right)\right| \leq 1$, then we have

$$|t_\ell(n)| \leq |d_\ell|\left|2\cos\left(\frac{\pi\ell}{2^r}\right)\right|^n. \tag{3.6}$$

Now, if $t_\ell(n) \neq 0$, then since the cosine is periodic in n, then there is a positive constant c_ℓ, which does not depend on n, such that

$$|t_\ell(n)| \geq c_\ell\left|2\cos\left(\frac{\pi\ell}{2^r}\right)\right|^n. \tag{3.7}$$

Then each $|t_\ell(n)|$ is either zero or in a constant range of $\left|2\cos\left(\frac{\pi\ell}{2^r}\right)\right|^n$.

When n is big enough the expression $\left|2\cos\left(\frac{\pi\ell}{2^r}\right)\right|^n$ dominates $\left|2\cos\left(\frac{\pi(\ell+1)}{2^r}\right)\right|^n$, for $\ell < 2^r - 1$, so any $t_\ell(n) \neq 0$ dominate all the $t_{\ell'}(n)$, for $\ell < \ell' < 2^{r-1}$. Let ℓ_0 be the least ℓ such that $t_\ell(n) \neq 0$. Then the subsequent terms cannot cancel $t_{\ell_0}(n)$. Hence, $a_n \neq 0$, a contradiction. So we must have $t_\ell(n) = 0$, for $0 \leq \ell \leq 2^{r-1} - 1$. The converse is clear. The proof is now complete. (Q.E.D.)

Lemma 3.1.2. *We have $t_\ell(n) = 0$ if and only if $d_\ell = -\xi_\ell^n\overline{d_\ell}$, for any $0 \leq \ell \leq 2^{r-1} - 1$.*

Proof. The proof of this lemma is also inspired by the proof presented in [1]. If $d_\ell =$

0, then the lemma is trivial. Otherwise, suppose that $d_\ell \neq 0$. Then, $t_\ell(n) = 0 \Leftrightarrow \cos\left(\arg(d_\ell) - \frac{\pi n \ell}{2^r}\right) = 0 \Leftrightarrow \arg(d_\ell) - \frac{\pi n \ell}{2^r} = \frac{\pi}{2} + k\pi$, for some $k \in \mathbb{Z} \Leftrightarrow \exp\left(2i \arg(d_\ell)\right) = \exp\left(2i\left(\frac{\pi n \ell}{2^r} + \frac{\pi}{2} + k\pi\right)\right) \Leftrightarrow |d_\ell| e^{i \arg(d_\ell)} = -|d_\ell| e^{-i \arg(d_\ell)} \exp\left(\frac{\pi n \ell}{2^{r-1}}\right) \Leftrightarrow d_\ell = -\xi_\ell^n \overline{d_\ell}$, for $0 \leq \ell \leq 2^{r-1} - 1$. The proof is complete. (Q.E.D.)

This characterizes the cases when a real solution a_n to the recurrence is zero for n big enough. The above two lemmas have the following implication:

Lemma 3.1.3. *Suppose that $\{a_n\}$ is a real solution to (3.1). Then there exists an integer n_0 such that for any $n > n_0$, we have $a_n = 0$ if and only if $d_\ell = -\xi_\ell^n \overline{d_\ell}$, for any $0 \leq \ell \leq 2^{r-1} - 1$.*

Consider now the perturbation $\boldsymbol{e}_{n,k} + F(\mathbf{X})$. Then recall that

$$S(\boldsymbol{e}_{n,k} + F(\mathbf{X})) = \sum_{\ell=0}^{n} \left(\sum_{m=0}^{j} C_m(F) (-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{\ell}{k-i}} \right) \binom{n}{\ell}. \tag{3.8}$$

Define $\delta_\ell^{(F)}(k)$ as

$$\delta_\ell^{(F)}(k) = \sum_{m=0}^{j} C_m(F) (-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{\ell}{k-i}}. \tag{3.9}$$

Then (3.8) can be rewritten as

$$S(\boldsymbol{e}_{n,k} + F(\mathbf{X})) = \sum_{\ell=0}^{n} \delta_\ell^{(F)}(k) \binom{n}{\ell}. \tag{3.10}$$

If n is such that $\boldsymbol{e}_{n,k} + F(\mathbf{X})$ is balanced, then we find a solution to the Diophantine equation

$$\sum_{\ell=0}^{n} \delta_\ell \binom{n}{\ell} = 0, \tag{3.11}$$

over $\Gamma_j^{(e)}$ and the solution would be given by $(\delta_0^{(F)}(k), \delta_1^{(F)}(k), \cdots, \delta_n^{(F)}(k))$.

Now, if we have

$$a_n = d_0 \cdot 2^n + \sum_{\ell=1}^{2^r-1} d_\ell \cdot \lambda_\ell^n, \tag{3.12}$$

then,

$$\begin{aligned} d_\ell &= \sum_{m=0}^{j} C_m(F)\left(\frac{1}{2^r}\sum_{a=0}^{2^r-1}(-1)^{\sum_{i=0}^{m}\binom{m}{i}\binom{a}{k-i}}\xi_\ell^a\right) \\ &= \frac{1}{2^r}\sum_{a=0}^{2^r-1}\left(\sum_{m=0}^{j} C_m(F)(-1)^{\sum_{i=0}^{m}\binom{m}{i}\binom{a}{k-i}}\right)\xi_\ell^a \\ &= \frac{1}{2^r}\sum_{a=0}^{2^r-1}\delta_a^{(F)}(k)\cdot\xi_\ell^a. \end{aligned}$$

The next result is a generalization of Canteaut and Videau's observation for symmetric Boolean functions of fixed degree. It shows, excluding the trivial cases, that balanced functions of fixed degree do not exist when the number of variables grows.

Theorem 3.1.1. *[3] Suppose that $1 \le k_1 < \cdots < k_s$ are integers and $F(\boldsymbol{X})$ is a Boolean polynomial in the variables $X_1, \cdots, X_j$ (j is fixed). Then there is an n_0 such that for all $n > n_0$, the perturbation $\boldsymbol{e}_{n+j,[k_1,\cdots,k_s]} + F(\boldsymbol{X})$ is balanced if and only if it is trivially balanced.*

Proof. For the simplicity of the proof, we will present the proof for the case of a perturbation of the form $\boldsymbol{e}_{n+j,k} + F(\mathbf{X})$. The general case follows by the same argument. The sufficient part is clear since any trivially balanced perturbation is balanced by definition, so we will only prove the necessary part.

Recall that $\{S(\boldsymbol{e}_{n+j,k} + F(\mathbf{X}))\}$ is a real solution to (3.1). So by Lemma 3.1.3, there is an n_0 such that for every $n > n_0$, we have $S(\boldsymbol{e}_{n+j,k} + F(\mathbf{X})) = 0$ if and only if $d_\ell = -\xi_\ell^n\overline{d_\ell}$, for any $0 \le \ell \le 2^{r-1} - 1$, where

$$d_\ell = \frac{1}{2^r}\sum_{a=0}^{2^r-1}\delta_a^{(F)}(k)\cdot\xi_\ell^a, \tag{3.13}$$

and $r = \lfloor \log_2(k) \rfloor + 1$. Suppose that $n > n_0$ and that $d_\ell = -\xi_\ell^n\overline{d_\ell}$, for any $0 \le \ell \le 2^{r-1} - 1$.

Then

$$\begin{aligned} d_\ell &= -\xi_\ell^n \overline{d_\ell} \\ &= -\frac{\xi_\ell^n}{2^r} \sum_{a=0}^{2^r-1} \left(\sum_{m=0}^{j} C_m(F)(-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{a}{k-i}} \right) \xi_\ell^{-a} \\ &= -\frac{1}{2^r} \sum_{a=0}^{2^r-1} \left(\sum_{m=0}^{j} C_m(F)(-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{a}{k-i}} \right) \xi_\ell^{n-a}. \end{aligned}$$

Make the substitution $t = n - a$, then the latter sum becomes

$$\begin{aligned} -\frac{1}{2^r} \sum_{a=0}^{2^r-1} \left(\sum_{m=0}^{j} C_m(F)(-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{a}{k-i}} \right) \xi_\ell^{n-a} &= -\frac{1}{2^r} \sum_{t=n-2^r-1}^{n} \left(\sum_{m=0}^{j} C_m(F)(-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{n-t}{k-i}} \right) \xi_\ell^{t} \\ &= -\frac{1}{2^r} \sum_{a=0}^{2^r-1} \left(\sum_{m=0}^{j} C_m(F)(-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{n-a}{k-i}} \right) \xi_\ell^{a} \\ &= -\frac{1}{2^r} \sum_{a=0}^{2^r-1} \delta_{n-a}^{(F)}(k) \cdot \xi_\ell^a. \end{aligned}$$

The previous identity holds because the sum

$$\sum_{m=0}^{j} C_m(F)(-1)^{\sum_{i=0}^{m} \binom{m}{i}\binom{n-a}{k-i}} \tag{3.14}$$

has a period of 2^r. Therefore, $d_\ell = -\xi_\ell^n \overline{d_\ell}$, for any $0 \leq \ell \leq 2^{r-1} - 1$ holds if and only if

$$\frac{1}{2^r} \sum_{a=0}^{2^r-1} \delta_a^{(F)}(k) \cdot \xi_\ell^a = -\frac{1}{2^r} \sum_{a=0}^{2^r-1} \delta_{n-a}^{(F)}(k) \cdot \xi_\ell^a \tag{3.15}$$

which is equivalent to

$$\sum_{a=0}^{2^r-1} (\delta_a^{(F)}(k) + \delta_{n-a}^{(F)}(k)) \cdot \xi_\ell^a = 0. \tag{3.16}$$

Define $\Delta(X)$ as

$$\Delta(X) = \sum_{a=0}^{2^r-1} (\delta_a^{(F)}(k) + \delta_{n-a}^{(F)}(k)) X^a \tag{3.17}$$

Since $\xi_\ell = \exp(\frac{\sqrt{-1}\pi\ell}{2^r})$ are all roots of $\Delta(X)$, for all $0 \leq \ell \leq 2^{r-1} - 1$, this implies that all the polynomials in the list

$$X-1,\, X^2+1,\, X^4+1,\, \cdots,\, X^{2^{r-1}}+1$$

divide $\Delta(X)$. But these polynomials are irreducible in $\mathbb{Q}[X]$, so

$$(X-1)\prod_{t=1}^{r-1}(X^{2^t}+1) \tag{3.18}$$

divides $\Delta(X)$. However, the degree of $\Delta(X)$ and (3.18) is 2^r-1. Since $\mathbb{Q}[X]$ is a UFD (Unique Factorization Domain), then there exists a constant z, such that

$$\Delta(X) = z\cdot(X-1)\prod_{t=1}^{r-1}(X^{2^t}+1). \tag{3.19}$$

It is not hard to see that such constant z is in fact an integer. By comparing coefficients, one gets

$$\delta_a^{(F)}(k)+\delta_{n-a}^{(F)}(k) = (-1)^{a-1}z, \tag{3.20}$$

for $0\le a\le 2^r-1$. This equation holds beyond the range $0\le a\le 2^r-1$ because $\delta_a^{(F)}(k)$ has period of 2^r. Then, when n is big enough, (3.20) characterizes all solutions

$$(\delta_0^{(F)}(k), \delta_1^{(F)}(k), \cdots, \delta_n^{(F)}(k))$$

to the Diophantine equation

$$\sum_{\ell=0}^{n}\delta_\ell\binom{n}{\ell} = 0, \tag{3.21}$$

over $\Gamma_j^{(e)}$, that comes from the perturbation $\boldsymbol{e}_{n+j,k}+F(\mathbf{X})$. We now wish to prove that all of these solutions are trivial.

Suppose first that n is odd, that is, $n=2m+1$. Then equation (3.20) becomes

$$\delta_a^{(F)}(k)+\delta_{2m+1-a}^{(F)}(k) = (-1)^{a-1}z. \tag{3.22}$$

Let $a=m$, then (3.22) becomes

$$\delta_m^{(F)}(k)+\delta_{m+1}^{(F)}(k) = (-1)^{m-1}z. \tag{3.23}$$

Now, let $a = m + 1$, then (3.23) becomes

$$\delta^{(F)}_{m+1}(k) + \delta^{(F)}_{m}(k) = (-1)^m z. \tag{3.24}$$

Subtracting both equations (3.23) and (3.24) one gets $((-1)^{m-1} - (-1)^m)z = 0$, so $z = 0$. Then equation (3.22) becomes

$$\delta^{(F)}_{2m+1-a}(k) = -\delta^{(F)}_{a}(k). \tag{3.25}$$

Therefore, the perturbation is trivially balanced for n odd.

Suppose now that n is even, that is, $n = 2m$. Then

$$\delta^{(F)}_{a}(k) + \delta^{(F)}_{2m-a}(k) = (-1)^{a-1} z. \tag{3.26}$$

If $z = 0$, then the perturbation is trivially balanced and we are done. Otherwise, suppose that $z \neq 0$. Then let $a = m$, so (3.26) becomes

$$2\delta^{(F)}_{m}(k) = (-1)^{m-1} z. \tag{3.27}$$

It follows from this that z is even, say $z = 2z_0$, where $z_0 \neq 0$. Then $\delta^{(F)}_{m}(k) = (-1)^{m-1} z_0$, and

$$\begin{aligned}
&(\delta^{(F)}_0(k), \delta^{(F)}_1(k), \cdots, \delta^{(F)}_{2m}(k)) \\
&\quad \sim (\delta^{(F)}_0(k) + \delta^{(F)}_{2m}(k), \delta^{(F)}_1(k) + \delta^{(F)}_{2m-1}(k)), \cdots, \delta^{(F)}_{m-1}(k) + \delta^{(F)}_{m+1}(k), \delta^{(F)}_{m}(k), 0, 0, \cdots, 0) \\
&\quad \sim (2z_0. - 2z_0, \cdots, (-1)^m 2z_0, (-1)^{m-1} z_0, 0, 0, \cdots, 0) \\
&\quad \sim (2, -2, \cdots, (-1)^m 2, (-1)^{m-1}, 0, 0, \cdots, 0) \\
&\quad \sim (1, -1, 1, -1, \cdots, -1, 1).
\end{aligned}$$

Therefore, the perturbation $\boldsymbol{e}_{n,k} + F(\mathbf{X})$ is trivially balanced when n is even. The theorem is now proved. (Q.E.D.)

3.2 Some Examples of Sporadic Balanced Perturbations

In the previous section we provided a proof that balanced perturbations of fixed degree k, except for the trivial cases, do not exist when the number of variables n is big enough. This implies that up to a certain natural number n, sporadic balanced function of fixed degree k do not exist. In this section, we are going to provide some examples of these types of balanced functions which are linked with some special Diophantine binomial equations. These functions happen to be non-trivially balanced and they are less common in comparison to the ones that are trivially balanced, so it is of great interest to find these types of functions.

Example 3.2.1. Consider the perturbation $\boldsymbol{e}_{22,9} + X_1 + X_2 + X_3$. This perturbation is sporadic balanced. Moreover, this perturbation corresponds to the Diophantine equation

$$\binom{19}{6} - 3\binom{19}{7} + 4\binom{19}{8} - 4\binom{19}{9} + 4\binom{19}{10} - 4\binom{19}{11} + 4\binom{19}{12} - 4\binom{19}{13} + 3\binom{19}{14} - \binom{19}{15} = 0 \tag{3.28}$$

and its corresponding solution is given by

$$\begin{aligned}\delta &= (\delta_0, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8, \delta_9, \delta_{10}, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14}, \delta_{15}, \delta_{16}, \delta_{17}, \delta_{18}, \delta_{19}) \\ &= (0, 0, 0, 0, 0, 0, 1, -3, 4, -4, 4, -4, 4, -4, 3, -1, 0, 0, 0, 0).\end{aligned}$$

Example 3.2.2. The perturbation $\boldsymbol{e}_{9,4} + X_1 + X_2 + X_3 + X_4 + X_5$ is sporadic balanced and it corresponds to the Diophantine equation

$$-2\binom{4}{0} + 3\binom{4}{1} - 2\binom{4}{2} + 2\binom{4}{4} = 0. \tag{3.29}$$

By making use of the identity

$$\binom{n}{k} = \binom{n}{n-k} \tag{3.30}$$

we can see that (3.29) is equivalent to the Diophantine equation

$$3\binom{4}{1} - 2\binom{4}{2} = 0. \tag{3.31}$$

The corresponding solution to (3.29) is

$$\begin{aligned}\delta &= (\delta_0, \delta_1, \delta_2, \delta_3, \delta_4) \\ &= (-2, 3, -2, 0, 2),\end{aligned}$$

while the corresponding solution to (3.31) is

$$\begin{aligned}\delta &= (\delta_0, \delta_1, \delta_2, \delta_3, \delta_4) \\ &= (0, 3, -2, 0, 0).\end{aligned}$$

Therefore, $(-2, 3, -2, 0, 2) \sim (0, 3, -2, 0, 0)$.

Theorem 2.1.3 implies that the perturbation $\boldsymbol{e}_{10,5} + X_1 + X_2 + X_3 + X_4 + X_5$ is also sporadic balanced, with corresponding solution

$$\begin{aligned}\delta &= (\delta_0, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5) \\ &= (1, -5, 11, -15, 15, 11)\end{aligned}$$

and corresponding Diophantine equation

$$\binom{5}{0} - 5\binom{5}{1} + 11\binom{5}{2} - 15\binom{5}{3} + 15\binom{5}{4} + 11\binom{5}{5} = 0. \tag{3.32}$$

Example 3.2.3. Consider the perturbation $\boldsymbol{e}_{9,7} + X_1 + X_2 + X_3$. This perturbation is sporadic balanced. Its corresponding Diophantine equation is given by

$$\binom{6}{4} - 3\binom{6}{5} + 3\binom{6}{6} = 0. \tag{3.33}$$

Equation (3.33) corresponds to equations of the form

$$A\binom{n}{k} + B\binom{n}{k+1} + C\binom{n}{k+2} = 0, \tag{3.34}$$

where A, B, C are integers, with $A > 0$, $C \neq 0$ and $\gcd(A, B, C) = 1$. It was shown by Luca and Szalay (see [12]) that for suitable integers A, B and C, the above equation has infinitely many solutions. Particularly, they found out that the Diophantine equation

$$\binom{n}{k} - 2\binom{n}{k+1} + \binom{n}{k+2} = 0 \tag{3.35}$$

has infinitely many solutions given by $n = \frac{1}{2}(t^2 - 2)$ and $k = \frac{1}{2}(t^2 + t - 4)$, for any integer t satisfying $|t| \geq 3$.

Example 3.2.4. Singmaster showed (see [14]) that the Diophantine equation

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n}{k+2} \tag{3.36}$$

has infinitely many solutions given by $n = f_{2i+2}f_{2i+3} - 1$ and $k = f_{2i}f_{2i+3} - 1$, where f_n denotes the nth Fibonacci number. The smallest of these solutions (letting $i = 1$) is given by

$$\binom{f_4f_5 - 1}{f_2f_5 - 1} + \binom{f_4f_5 - 1}{f_2f_5} - \binom{f_4f_5 - 1}{f_2f_5 + 1} = \binom{14}{4} + \binom{14}{5} - \binom{14}{6} = 0 \tag{3.37}$$

which is clearly a non-trivial solution to (2.49) over Γ_1. Castro, González and Medina found [3] only 4 perturbations of the form $\boldsymbol{e}_{15,[k_1,k_2,\ldots,k_s]} + X_1$ of degree less than or equal to 14 that corresponds to (3.37). One of the four perturbations they found was $\boldsymbol{e}_{15,[6,7,11,12,14]} + X_1$, which corresponds to the equivalent solution

$$\binom{14}{5} - \binom{14}{6} + \binom{14}{10} = 0, \tag{3.38}$$

which is obtained from (3.37) by using the identity

$$\binom{n}{k} = \binom{n}{n-k}. \tag{3.39}$$

Consider equation (3.34) with $A = C = 1$ and $B = -2$, that is

$$\binom{n}{k} - 2\binom{n}{k+1} + \binom{n}{k+2} = 0. \tag{3.40}$$

The smallest solution to (3.40) with the n and k as defined before (letting $t = 3$) is given by

$$\binom{7}{4} - 2\binom{7}{5} + \binom{7}{6} = 0. \tag{3.41}$$

Castro, González and Medina found all sporadic balanced perturbations of the form $\boldsymbol{e}_{8,[k_1,k_2,\ldots,k_s]} + X_1$, for $k_s \leq 7$ for which their corresponding solution are equivalent to (3.41) with the aid of a *Mathematica* implementation. One of these perturbations is $\boldsymbol{e}_{8,[2,3,4,7]} + X_1$ which its corresponding solution is $(0, 1, -1, 1, 0, -1, 0, 0)$. Indeed we have $(0, 0, 0, 0, 1, -2, 1, 0) \sim (0, 1, -1, 1, 0, -1, 0, 0)$.

Castro, González and Medina also found all sporadic perturbations of the form $\boldsymbol{e}_{9,[k_1,k_2,\ldots,k_s]} + X_1 + X_2$, for $k_s \leq 8$ for which their corresponding solutions are equivalent to (3.41). By the same *Mathematica* implementation, they found that there are 265 sporadic balanced perturbations of the form $\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + X_1$ with $n, k_s \leq 17$. Also, they found 606 sporadic balanced perturbations of the form $\boldsymbol{e}_{n,[k_1,\cdots,k_s]} + X_1 + X_2$ with $n, k_s \leq 17$. Tables 3.1 and 3.2 shows some examples of these perturbations with their corresponding solutions that we were able to find using the same *Mathematica* implementation.

Perturbation	Corresponding solution
$\boldsymbol{e}_{8,[1,6]} + X_1 + X_2$	$(2, -2, 2, -2, 1, 1, -1)$
$\boldsymbol{e}_{8,[1,2,3,6]} + X_1 + X_2$	$(1, 0, 1, -2, 2, -1, 0)$
$\boldsymbol{e}_{8,[2,5]} + X_1 + X_2$	$(-1, 1, 1, -2, 1, 1, -1)$
$\boldsymbol{e}_{8,[2,5,6]} + X_1 + X_2$	$(-1, 1, 1, -2, 2, -2, 2)$
$\boldsymbol{e}_{8,[2,6,9]} + X_1 + X_2$	$(-1, 1, 1, -1, 0, 0, 0)$
$\boldsymbol{e}_{8,[2,3,5,6]} + X_1 + X_2$	$(-1, 2, -1 - 1, 2 - 1, 0)$
$\boldsymbol{e}_{8,[2,5,9,10]} + X_1 + X_2$	$(-1, 1, 1, -2, 1, 1, -1)$
$\boldsymbol{e}_{8,[3,5,6]} + X_1 + X_2$	$(0, -1, 2, -2, 1, 0, 1)$
$\boldsymbol{e}_{8,[3,6,9]} + X_1 + X_2$	$(0, -1, 2, -1, -1, 2, -1)$
$\boldsymbol{e}_{8,[5,6,9,10]} + X_1 + X_2$	$(0, 0, 0, -1, 1, 1, -1)$

Table 3.1: Some examples of perturbations of the form $\boldsymbol{e}_{8,[k_1,\cdots,k_s]} + X_1 + X_2$ and their corresponding solutions to (2.49).

Perturbation	Corresponding solution
$\boldsymbol{e}_{7,[1,2,4,5]} + X_1$	$(1, 0, -1, 1, 0, -1, 0)$
$\boldsymbol{e}_{7,[2,5,8]} + X_1$	$(0, 1, 0, -1, 1, 0, -1)$
$\boldsymbol{e}_{7,[2,5,10]} + X_1$	$(1, -1, -1, 1, 0, 0, 0)$
$\boldsymbol{e}_{7,[2,5,17]} + X_1$	$(-1, 0, 0, 1, -1, 1, 0)$
$\boldsymbol{e}_{7,[1,3,4,6]} + X_1$	$(1, -1, 0, 1, -1, 0, 0)$
$\boldsymbol{e}_{7,[3,6,8]} + X_1$	$(0, 0, 1, -1, 0, 1, -1)$
$\boldsymbol{e}_{7,[3,6,9,11,13]} + X_1$	$(0, 0, 0, -1, 1, 1, -1)$

Table 3.2: Some examples of perturbations of the form $\boldsymbol{e}_{7,[k_1,\cdots,k_s]} + X_1$, and their corresponding solutions to (2.49).

3.3 Conclusion

It is established a remarkable yet beautiful identity between two perturbations of two different symmetric Boolean functions. For this, we provided a proof that the sequences of symmetric functions and its perturbations satisfy the same recurrence relation with integer coefficients. We discussed that symmetric Boolean functions have connections with Diophantine equations with binomial coefficients with solutions over a bounded set of integers $\Gamma = \{\pm 1\}$ whose solutions are either trivial or not. From this, it is defined the concepts of trivially and sporadic balanced functions. Perturbations of symmetric Boolean function also have a connection with Diophantine equations with binomial coefficients over a bounded set Γ_j of integers and the definition of trivially and sporadic balanced function are extended to these perturbations. A similar conjecture to the one that Cusick conjectured for elementary symmetric polynomials, is presented for the simplest type of perturbations considered in this thesis. It has been shown, based on an observation made by Canteaut and Videau, that balanced perturbations of symmetric functions of fixed degree, excluding the trivial cases, do not exist when the number of variables is big enough. This implies that sporadic functions are less common and hard to find in comparison to the ones that are trivially balanced. We hope to extend some of these results in a future to perturbations of symmetric functions over a arbitrary finite field $\mathbb{F}_q$, where $q = p^r$, with p prime and $r \geq 1$.

Bibliography

[1] J. Cai, F. Green and T. Thierauf. On the correlation of symmetric functions. *Math. Systems Theory* **29** (1996) 245-258.

[2] A. Canteaut and M. Videau. Symmetric Boolean Functions. *IEEE Trans. Inf. Theory* **51**, no. 8 (2005)

[3] F. N. Castro, O.E. González and L. A. Medina. Diophantine equations with binomial coefficients and perturbation of symmetric Boolean functions. *IEEE Trans. Inf. Theory* **64(2)** (2018) 1347-1360.

[4] F. N. Castro and L.A. Medina. Linear Recurrences and Asymptotic Behavior of Exponential Sums of Symmetric Boolean Functions. *Elec. J. Combinatorics*, 18:#P8, 2011.

[5] F. N. Castro and L. A. Medina. Asymptotic Behavior of Perturbation of Symmetric Functions. *Annals of Combinatorics*, 18:397-417, 2014.

[6] F. N. Castro and L. A. Medina. Modular periodicity of exponential sums of symmetric Boolean functions. *Discrete Appl. Math.* **217**, 455-473, 2017.

[7] F.N. Castro, R. Chapman, L. A. Medina, L. B. Sepúlveda. Recursions associated to trapezoid, symmetric and rotation symmetric functions over Galois fields. *Discrete Mathematics*, **341(7)** (2018) 1915-1931.

[8] T. W. Cusick, Y. Li, and P. Stănică. Balanced Symmetric Functions over $GF(p)$. *IEEE Trnas. Inf. Theory* **54 (3)** (2008) 1304-1307.

[9] T. W. Cusick, Y. Li, and P. Stănică. On a conjecture for balanced symmetric Boolean functions. *J. Math, Cript.* **3** (2009), 1-18.

[10] Y. Guo, G. Gao, Y. Zhao. Recent Results on Balanced Symmetric Boolean Functions. *IEEE Trans. Inf. Theory* (62), no. 9 (2016), 5199–5203.

[11] E. J. Ionascu, Thor Martinsen, Partelimon Stănică. Bisecting binomial coefficients. *Discret. Appl. Math.* 227 (2017) 70–83.

[12] F. Luca and Szalay. Linear diophantine equations with three consecutive binomial coefficients. Acta Academiae Paedagogicae Agriensis, Sectio Mathematicae **31** (2004), 53-60.

[13] P. Sakar and S. Maitra. Balancedness and correlation immunity of symmetric Boolean functions. *Discrete Math.* **307** (2007), 2351-2358.

[14] D. Singmaster. Repeated binomial coefficients and Fibonacci numbers. *Fibonacci Quarterly* **13**(4) (1975), 295-298.