

A New Representation of Binary Sequences by means of Boolean Functions

Miguel Beltrá¹, Sara D. Cardell², Amparo Fúster-Sabater³,
Verónica Requena^{4*}

²Departamento de Matemática, Instituto de Geociências e Ciências
Exatas, Unesp, Campus Rio Claro, Brazil.

³Instituto de Tecnologías Físicas y de la Información, Consejo Superior
de Investigaciones Científicas, Madrid, Spain.

^{1,3*}Departamento de Matemáticas, Universidad de Alicante, Alicante,
Spain.

*Corresponding author(s). E-mail(s): vrequena@ua.es;
Contributing authors: miguel.beltra@es; sd.cardell@unesp.br;
amparo.fuster@csic.es;

Abstract

Boolean functions and binary sequences are main tools used in cryptography. In this work, we introduce a new bijection between the set of Boolean functions and the set of binary sequences with period a power of two. We establish a connection between them which allows us to study some properties of Boolean functions through binary sequences and vice versa. Then, we define a new representation of sequences, based on Boolean functions and derived from the algebraic normal form, named reverse-ANF. Next, we study the relation between such a representation and other representations of Boolean functions as well as between such a representation and the binary sequences. Finally, we analyse the generalized self-shrinking sequences in terms of Boolean functions and some of their properties using the different representations.

Keywords: Binary sequence, binomial sequence, Boolean function, Sierpinski triangle, generalized sequence

1 Introduction

Boolean functions have applications in propositional logic, electrical engineering, game theory, reliability, combinatorics, and linear programming [1]. They are also very important in complexity theory [2, 3], coding theory and cryptography [4, 5], social sciences [6, 7], medicine and biology [8, 9]. Boolean functions are a useful tool for building models inside a large number of processes in nature, logic, engineering, or science. In particular, in *cryptography* they are used intensively in the construction and analysis of cryptosystems. They play an important role in block ciphers, stream ciphers, and hash functions [10, 11]. For example, the implementation of a substitution box (or S-box) needs non-linear Boolean functions to resist cryptanalytic attacks such as differential cryptanalysis [12–14]. In the most common models of stream ciphers, the keystream is produced by means of Boolean functions [15–17]. One of their basic requirements is that they allow to increase the linear complexity [11, 18, 19], provided that such functions have a high algebraic degree [20]. A variety of criteria for the selection of Boolean functions define their abilities to provide security and thus to be used in different applications.

Cryptographic random bit generators are often constructed by combining several maximum-length LFSRs (Linear Feedback Shift Registers [21]) in a nonlinear way. The output bits of different LFSRs are combined using a nonlinear Boolean function to get the keystream sequence [22, 23]. In order to eliminate the predictability of the keystream, the Boolean function should have several properties such as high algebraic degree, high non-linearity and high correlation immunity [11, 24, 25].

Binary sequences with a period power of two can be represented by the binomial representation introduced in [26]. This representation allows to analyse certain cryptographic properties of these sequences such as linear complexity and period [26, 27]. Some studies relate Boolean functions and binary sequences; see, for example, [28]. In this work, authors consider the truth table of a Boolean function as a binary sequence, and the polynomial representation used is the ANF (Algebraic Normal Form). They focus their study in the relation of a pseudorandomness measure (correlation measure of order k) for these sequences and two quality measures (sparsity and combinatorial complexity) for the corresponding Boolean functions. Furthermore, they apply their results to Sidelnikov sequences. In our work, we make use of the binomial representation of a binary sequence and its reverse-ANF to research other cryptographic properties; we do not focus on the pseudorandomness measures of these binary sequences. We examine more general properties such as: algebraic degree, balancedness or linear complexity, among others; moreover, we apply these results to the family of generalized self-shrunk sequences.

In this work, we define a new bijection between the set of Boolean functions (with any number of variables) and the set of binary sequences of period a power of two. For that purpose, we have introduced a new representation, for Boolean functions, based on the Algebraic Normal form, called the reverse-ANF (r-ANF). This new representation facilitates us the link with the binary sequences. There exist different ways to represent a Boolean function (for more details, see [5]). However, we have found it necessary to define this new Boolean representation because using other existing representations in the literature as, for instance, the ANF, the bijection defined fails. The binary

sequence obtained is not unique, it depends on the number of variables of the ANF. This justification will be explained in more detail in the subsection 4. The aim of this connection is to analyse properties of Boolean functions through binary sequences and vice versa. We focus our study on the reverse-ANF of the generalized self-shrunken sequences; since that they have many interesting cryptographic properties[27, 29, 30], and some open problems could be solved using Boolean functions. As far as we know, there are no works that analyse these sequences through Boolean functions.

This paper is organized as follows. In Section 2, we present the main concepts, related to Boolean functions and binary sequences, needed to understand the rest of the paper. Next, in Section 3, we analyse the relation between the binomial representation and the algebraic normal form of a sequence. We observe that it coincides with the known relation between the truth table and the ANF of a Boolean function. After that, in Section 4 we introduce the new representation of a binary sequence, the called reverse-ANF, and study the existing relation between such a representation and the rest of representations previously introduced. Moreover, we expose some results about the Sierpinski triangle in terms of the r-ANF. In Section 5, we define the reverse sequence of a given sequence and analyse its connection with shifted versions of the original sequence. The r-ANF of the family of generalized sequences is examined in Section 6. Finally, the paper closes in Section 7 with some conclusions and future work.

2 Preliminaries

In this section, we introduce the necessary concepts to understand the rest of the paper. In the following subsections, we present some important results concerning binary sequences and Boolean functions which help us to define a new connection between them.

2.1 Sequences

Let $\mathbb{F}_2 = \{0, 1\}$ be the Galois field of two elements. Consider $\{u_\tau\}_{\tau \geq 0} = \{u_0, u_1, u_2, \dots\}$ a binary sequence with $u_\tau \in \mathbb{F}_2$, for $\tau = 0, 1, 2, \dots$. We say that the sequence $\{u_\tau\}_{\tau \geq 0}$, or simply $\{u_\tau\}$, is periodic if there exists an integer T , called period, such that $u_{\tau+T} = u_\tau$, for all $\tau \geq 0$. In the sequel, all the sequences considered will be binary sequences and the symbol $+$ will denote the Exclusive-OR (XOR) logic operation.

Let r be a positive integer, and let $a_0, a_1, a_2, \dots, a_{r-1}$ be constant coefficients with $a_j \in \mathbb{F}_2$, for $j = 0, 1, \dots, r-1$. A binary sequence $\{u_\tau\}$ satisfying the relation

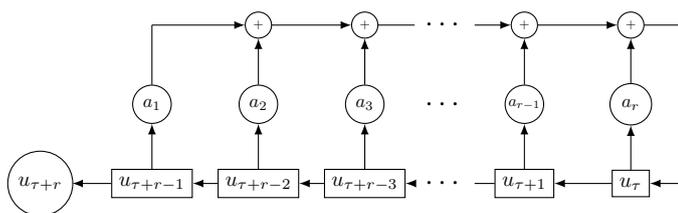
$$u_{\tau+r} = a_{r-1}u_{\tau+r-1} + a_{r-2}u_{\tau+r-2} + a_{r-3}u_{\tau+r-3} + \dots + a_1u_{\tau+1} + a_0u_\tau, \quad \tau \geq 0,$$

is called a (r -th order) linear recurring sequence in \mathbb{F}_2 . The terms $\{u_0, u_1, \dots, u_{r-1}\}$ are referred to as the initial terms of the sequence and define uniquely such a sequence.

The monic polynomial

$$p(x) = x^r + a_{r-1}x^{r-1} + a_{r-2}x^{r-2} + a_{r-3}x^{r-3} + \dots + a_1x + a_0 \in \mathbb{F}_2[x]$$

Fig. 1: LFSR of length r



is called the characteristic polynomial of the linear recurring sequence and $\{u_\tau\}$ is said to be generated by $p(x)$.

We can generate linear recurring sequences using LFSRs [21]. In fact, an LFSR can be defined as an electronic device with r interconnected memory cells (stages) with binary content. At every clock pulse, the binary element of each stage is shifted to the adjacent stage as well as a new element is computed through the linear feedback to fill the empty stage (see Figure 1). We say that the LFSR has maximal-length if its characteristic polynomial is primitive. Then, its output sequence is called PN-sequence (Pseudo-Noise sequence) and its period is $T = 2^r - 1$ (see [21]).

The linear complexity of a sequence $\{u_\tau\}$, denoted by LC , is defined as the length of the shortest LFSR that generates such a sequence or, equivalently, as the lowest order linear recurrence relationship that generates such a sequence. In cryptographic applications, the linear complexity must be as large as possible. The expected value should be at least half the period (see [18]).

2.2 Boolean functions

A *Boolean function* of n variables is a map of the form $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$. We denote by \mathcal{B}_n the set of all Boolean functions of n variables; it is well known that \mathcal{B}_n , with the usual addition of functions, is a linear space of dimension 2^n over \mathbb{F}_2 .

If we denote by \mathbf{i} the coefficient vector of the binary expansion of the integer i in n digits, for $i = 0, 1, 2, \dots, 2^n - 1$, then we can represent $\mathbb{F}_2^n = \{\mathbf{i} \mid i \in \mathbb{Z}_{2^n}\}$. The *truth table (TT)* of f (see, for example [31, 32]) is the binary sequence of length 2^n given by

$$\boldsymbol{\xi}_f = (f(\mathbf{0}), f(\mathbf{1}), \dots, f(\mathbf{2}^n - \mathbf{1})). \quad (1)$$

Next, we show that the truth table of a Boolean function is perfectly determined from its minterms, fact that will be important for our results. A *minterm* of n variables, defined by the vector $\mathbf{u} = (u_1, u_2, \dots, u_n) \in \mathbb{F}_2^n$, is the function $m_{\mathbf{u}} \in \mathcal{B}_n$ given by

$$m_{\mathbf{u}}(\mathbf{x}) = (1 \oplus u_1 \oplus x_1)(1 \oplus u_2 \oplus x_2) \cdots (1 \oplus u_n \oplus x_n),$$

for all $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{F}_2^n$. We use $m_{\mathbf{u}}(\mathbf{x})$ or $m_u(\mathbf{x})$, indistinctly because, as it was previously commented, $\mathbf{u} \in \mathbb{F}_2^n$ is the binary expansion of $u \in \mathbb{Z}_{2^n}$. For $i = 0, 1, 2, \dots, 2^n - 1$, it is evident that $m_i(\mathbf{x}) = 1$ if and only if $\mathbf{x} = \mathbf{i}$. Therefore, the truth table

$$(m_i(\mathbf{0}), m_i(\mathbf{1}), \dots, m_i(\mathbf{2}^n - \mathbf{1}))$$

of the function $m_i(\mathbf{x})$ has 1 in the i -th position and 0 in the remaining. As a consequence, we have that

$$\bigoplus_{i=0}^{2^n-1} m_i(\mathbf{x}) = 1, \quad \text{for all } \mathbf{x} \in \mathbb{F}_2^n.$$

Furthermore, $m_i(\mathbf{x}) = m_j(\mathbf{x})$ if and only if $i = j$, so we can identify the minterm $m_i(\mathbf{x})$ with the integer i (or the vector \mathbf{i} as required).

Now, for all $f \in \mathcal{B}_n$ is easy to check that

$$f(\mathbf{x}) = \bigoplus_{i=0}^{2^n-1} f(\mathbf{i}) m_i(\mathbf{x}) \quad (2)$$

and, as

$$\bigoplus_{i=0}^{2^n-1} a_i m_i(\mathbf{x}) = 0 \quad \text{implies that} \quad a_i = 0 \quad \text{for } i = 0, 1, 2, \dots, 2^n - 1,$$

then, we can state that the set $\{m_0, m_1, \dots, m_{2^n-1}\}$ is a basis of \mathcal{B}_n . Consequently, $\dim \mathcal{B}_n = 2^n$ and, therefore, $|\mathcal{B}_n| = 2^{2^n}$.

The *support* of f , denoted by $\text{Supp}(f)$, is the set of vectors of \mathbb{F}_2^n whose image by f is 1, that is,

$$\text{Supp}(f) = \{\mathbf{a} \in \mathbb{F}_2^n \mid f(\mathbf{a}) = 1\}. \quad (3)$$

Therefore, $\text{Supp}(f)$ is composed by the vectors of \mathbb{F}_2^n corresponding to the components of $\boldsymbol{\xi}_f$ equal to 1. Due to the identification given between the elements of \mathbb{F}_2^n and \mathbb{Z}_{2^n} , through the binary representation, we can rewrite:

$$\text{Supp}(f) = \{i \in \mathbb{Z}_{2^n} \mid f(\mathbf{i}) = 1\}.$$

Thus, according to expression (2), we can express $f(\mathbf{x})$ as follows:

$$f(\mathbf{x}) = \bigoplus_{i \in \text{Supp}(f)} m_i(\mathbf{x}) \quad (4)$$

which allows us to identify $\text{Supp}(f)$ with the set of indices of the minterms of $f(\mathbf{x})$.

If $f \in \mathcal{B}_n$, then we call *weight* of f , denoted by $w(f)$, to the number of 1s in its truth table. Therefore, $w(f) = |\text{Supp}(f)|$ and, clearly

$$w(f) = \sum_{\mathbf{a} \in \mathbb{Z}_2^n} f(\mathbf{a}).$$

We said that f is *balanced* if $w(f) = 2^{n-1}$.

There exist several ways to represent a Boolean function (see [5]). For example, we can write $f(\mathbf{x})$ uniquely (see [31, 32]) as

$$f(\mathbf{x}) = f(x_1, x_2, \dots, x_n) = \sum_{\mathbf{u} \in \mathbb{F}_2^n} \mu_f(\mathbf{u}) \mathbf{x}^{\mathbf{u}}, \quad (5)$$

where $\mu_f(\mathbf{u}) \in \mathbb{F}_2$ is also a Boolean function, called the Möbius transform of f (see, [33]). If $\mathbf{u} = (u_1, u_2, \dots, u_n)$, then

$$\mathbf{x}^{\mathbf{u}} = x_1^{u_1} x_2^{u_2} \cdots x_n^{u_n} \quad \text{with} \quad x_j^{u_j} = \begin{cases} x_j, & \text{if } u_j = 1, \\ 1, & \text{if } u_j = 0. \end{cases}$$

Observe that in equation (5) each one of the terms $\mathbf{x}^{\mathbf{u}}$ is a monomial whose degree is $w(\mathbf{u})$, corresponding to the number of variables appearing in the product. Note that $w(\mathbf{u})$ is the number of nonzero components of the vector \mathbf{u} . This representation of f is known as the *algebraic normal form* (ANF) of $f(\mathbf{x})$ [1]. Denote by \mathbf{m}_{μ_f} the vector composed by the coefficients of the ANF of $f(\mathbf{x})$ given in expression (5), that is,

$$\mathbf{m}_{\mu_f} = (\mu_f(\mathbf{0}), \mu_f(\mathbf{1}), \dots, \mu_f(\mathbf{2}^n - \mathbf{1})). \quad (6)$$

We call *degree* of $f(\mathbf{x})$, denoted by $\deg f$, to the maximum of degrees of the monomials of its ANF, that is

$$\deg f = \max\{w(\mathbf{u}) \mid \mu_f(\mathbf{u}) = 1\}.$$

2.3 Relation between the TT and the ANF

In this section, we summarise the well-known relation between the truth table and the ANF of f using Hadamard matrices (see [34]). This connection will be crucial in order to understand the new Boolean representation for binary sequences named reverse-ANF and introduced later in Section 4.

The transformation of f into its ANF can be performed by using its corresponding truth table (see [5, 34]). We call *binomial matrix* to the binary Hadamard matrix of size $2^t \times 2^t$ defined by

$$H_t = \begin{bmatrix} H_{t-1} & H_{t-1} \\ 0_{t-1} & H_{t-1} \end{bmatrix} \quad \text{and} \quad H_0 = [1],$$

with 0_{t-1} the null matrix of size $2^{t-1} \times 2^{t-1}$ and $t > 0$ a positive integer. It is well known that the binomial matrix H_t provides a relation between the ANF of a Boolean function and its truth table through the following equation

$$\boldsymbol{\xi}_f = \mathbf{m}_{\mu_f} \cdot H_t \pmod{2}, \quad (7)$$

where the vector $\boldsymbol{\xi}_f$ corresponds to the truth table of $f(\mathbf{x})$, given in equation (1), and \mathbf{m}_{μ_f} is the vector associated with the ANF of $f(\mathbf{x})$, defined in expression(6). As

$$= x_3x_4 + x_2 + x_2x_4 + x_1 + x_1x_3x_4 + x_1x_2 + x_1x_2x_3 + x_1x_2x_3x_4$$

Next example proceeds in a reverse way.

Example 2: Consider the Boolean function of 4 variables given by

$$f(\mathbf{x}) = f(x_1, x_2, x_3, x_4) = 1 + x_4 + x_2x_3x_4 + x_1x_2.$$

We have that its ANF is $f(\mathbf{x}) = x^0 + x^1 + x^7 + x^{12}$ and, therefore, $\mathbf{m}_{\mu_f} = (1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0)$. Applying equation (7), we have that the truth table of $f(\mathbf{x})$ is

$$\xi_f = \mathbf{m}_{\mu_f} \cdot H_4 \bmod 2 = (1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0, 0, 1, 0, 0),$$

and, therefore, we can represent $f(\mathbf{x})$ through its minterms as follows

$$f(\mathbf{x}) = m_0(\mathbf{x}) + m_2(\mathbf{x}) + m_4(\mathbf{x}) + m_6(\mathbf{x}) + m_7(\mathbf{x}) + m_8(\mathbf{x}) + m_{10}(\mathbf{x}) + m_{13}(\mathbf{x}).$$

2.4 B-representation

From now on, we always consider sequences $\{s_\tau\}$ of period a power of 2.

The binomial representation (*B-representation*) of binary sequences was first introduced in [26, 27]. In fact, every binary sequence $\{s_\tau\}$ with period 2^t can be expressed as a linear combination of binomial sequences as

$$b(\{s_\tau\}) = \sum_{i=0}^{2^t-1} c_i \left\{ \binom{n}{i} \right\}, \quad (9)$$

where $c_i \in \mathbb{F}_2$ for $i = 0, 1, \dots, 2^t - 1$ and $\left\{ \binom{n}{i} \right\}$ is the i -th binomial sequence [26]. This sequence can be computed as

$$\left\{ \binom{n}{i} \right\} = \left\{ \binom{0}{i}, \binom{1}{i}, \binom{2}{i}, \dots \right\},$$

where the coefficients $\binom{j}{i}$, for $j = 0, 1, \dots$, are reduced modulo 2 and $\binom{j}{i} = 0$ if $i > j$. If there is no ambiguity, we can make an abuse of notation and simply denote the binomial sequence by $\binom{n}{i}$. In Table 1, one can find the first 16 binomial sequences.

The expression (9) is called the *B-representation* of a sequence $\{s_\tau\}$, denoted by $b(\{s_\tau\})$. Sometimes, for simplicity, we will denote the B-representation by $b(\{s_\tau\}) = \left\{ \sum_{i=0}^{2^t-1} c_i \binom{n}{i} \right\}$ or simply $b(\{s_\tau\}) = \sum_{i=0}^{2^t-1} c_i \binom{n}{i}$.

Consider \mathbf{c}_{μ_f} the vector of length 2^t composed by the coefficients of the B-representation of $\{s_\tau\}$, that is, we have that

$$c_i = \begin{cases} 1 & \text{if } \binom{n}{i} \in b(\{s_\tau\}), \\ 0 & \text{if } \binom{n}{i} \notin b(\{s_\tau\}), \end{cases} \quad (10)$$

for $i = 0, 1, \dots, 2^t - 1$. For convenience, we can also denote the B-representation by the vector (i_1, i_2, \dots, i_k) , where $c_{i_j} \neq 0$ in (10), with $(i_1 < i_2 < \dots < i_k)$ for $j = 1, \dots, k$.

Next, we present an example to clarify these two representations.

Example 3: Consider the sequence of period $T = 8$, $\{s_\tau\} = (1, 1, 1, 0, 0, 0, 0, 1)$ whose B-representation is given by $b(\{s_\tau\}) = \binom{n}{0} + \binom{n}{3} + \binom{n}{4}$. We have that $\mathbf{c}_{\mu_f} = (1, 0, 0, 1, 1, 0, 0, 0)$. Moreover, this sequence can also be represented by the vector $b(\{s_\tau\}) = (0, 3, 4)$.

Table 1: The first 16 binomial coefficients, their binomial sequences $\binom{n}{i}$, periods and complexities

Binomial sequence	Binary representation	Period	LC
$\binom{n}{0}$	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 ...	1	1
$\binom{n}{1}$	0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 ...	2	2
$\binom{n}{2}$	0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1 ...	4	3
$\binom{n}{3}$	0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 1 ...	4	4
$\binom{n}{4}$	0 0 0 0 1 1 1 1 0 0 0 0 1 1 1 1 ...	8	5
$\binom{n}{5}$	0 0 0 0 0 1 0 1 0 0 0 0 0 1 0 1 ...	8	6
$\binom{n}{6}$	0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 1 ...	8	7
$\binom{n}{7}$	0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 ...	8	8
$\binom{n}{8}$	0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 ...	16	9
$\binom{n}{9}$	0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 1 ...	16	10
$\binom{n}{10}$	0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 1 ...	16	11
$\binom{n}{11}$	0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 ...	16	12
$\binom{n}{12}$	0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 ...	16	13
$\binom{n}{13}$	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 ...	16	14
$\binom{n}{14}$	0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 ...	16	15
$\binom{n}{15}$	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 ...	16	16

The coefficient c_{i_k} and the B-representation provide us some information about two fundamental parameters of the sequence: period and linear complexity (see [26] for more details). Indeed, note that:

- The period of the sequence $\{s_\tau\}$ is the same as that of the binomial sequence $\binom{n}{i_k}$.
- The linear complexity of the sequence $\{s_\tau\}$ is the same as that of the binomial sequence $\binom{n}{i_k}$, that is $LC = i_k + 1$.

Fixed a linear complexity LC , the period of the sequence is uniquely determined. Nevertheless, fixed a period there exist distinct sequences with such a period but different linear complexities (see for example Table 1).

In [35], authors prove that there exists a relation between a binary sequence of period a power of 2 and its B-representation. This connection is similar to the one between ANF of a Boolean function and its truth table given in expression (7). Due to the particular structure of the binomial sequences, the B-representation of a sequence $\{s_\tau\}$ can be turned into a matrix equation, as it is showed in the following theorem.

Theorem 1 (Theorem 2 of [35]): Consider the B-representation of a sequence $\{s_\tau\}$ of period $T = 2^t$, with t a non-negative integer, and let H_t be the binomial matrix

of size $2^t \times 2^t$. Then,

$$(s_0, s_1, \dots, s_{2^t-1}) = (c_0, c_1, \dots, c_{2^t-1}) \cdot H_t \text{ mod } 2, \quad (11)$$

where the vector $(s_0, s_1, \dots, s_{2^t-1})$ corresponds to the first 2^t terms of the sequence $\{s_\tau\}$, and $(c_0, c_1, \dots, c_{2^t-1})$ is the vector \mathbf{c}_{μ_f} given in expression (10).

As the matrix H_t is an idempotent matrix, we can write equation (11) as

$$(s_0, s_1, \dots, s_{2^t-1}) \cdot H_t \text{ mod } 2 = (c_0, c_1, \dots, c_{2^t-1}). \quad (12)$$

Therefore, the B-representation of a sequence $\{s_\tau\}$ can be computed in terms of the binomial matrix H_t and of the elements of the sequence $\{s_\tau\}$, and vice versa.

The following example clarifies this construction.

Example 4: Consider the sequence $\{s_\tau\}$ with period $T = 16$ given by the vector

$$(s_0, s_1, \dots, s_{15}) = (0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1).$$

From equation (12), we have that

$$(c_0, c_1, \dots, c_{15}) = (s_0, s_1, \dots, s_{15}) \cdot H_4 \text{ mod } 2 = (0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0),$$

and, therefore, the B-representation of $\{s_\tau\}$ is

$$b(\{s_\tau\}) = \binom{n}{3} + \binom{n}{4} + \binom{n}{6} + \binom{n}{7} + \binom{n}{8} + \binom{n}{9} + \binom{n}{10} + \binom{n}{11} + \binom{n}{12}.$$

3 Relation between B-representation and ANF

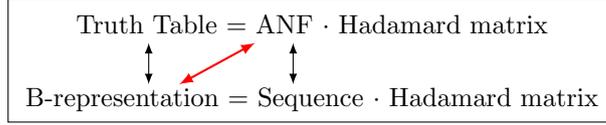
Due to the close relation between expressions (8) and (11), we will provide a connection between binary sequences $\{s_\tau\}$ and Boolean functions. This link will allow us to analyse some parameters of the Boolean functions from the study of the binary sequences, and vice versa. In particular, in this section, we analyse the relation between the B-representation of a sequence $\{s_\tau\}$ and the ANF of a Boolean function.

An important fact which helps us to understand this relation is that the rows of the binomial matrix H_t are the first 2^t binomial sequences (see, [35, 36]), that is, we can express

$$H_t = \left[\left\{ \binom{n}{0} \right\} \left\{ \binom{n}{1} \right\} \cdots \left\{ \binom{n}{2^t-2} \right\} \left\{ \binom{n}{2^t-1} \right\} \right]^T.$$

Therefore, from expression (9), we have that every binary sequence $\{s_\tau\}$ with period 2^t can be expressed as a linear combination of the rows of H_t .

On the one hand, we have that equation (7), $\xi_f = \mathbf{m}_{\mu_f} \cdot H_t$, means that the vector ξ_f is obtained as the linear combination of the rows of H_t which is determined by the elements of \mathbf{m}_{μ_f} . Therefore, the vector ξ_f is obtained summing up binomial sequences. Observe that we can define a relation between minterms and binomial sequences. Moreover, as the support of a Boolean function is the set of its minterms, and the set



$\xi_f = m_{\mu_f} \cdot H_t$	
↕	↕
$c_{\mu_f} = \{s_\tau\} \cdot H_t$	

Fig. 2: Relation between B-representation and ANF

of the binomial sequences is the B-representation, we can relate both representations. Later, we will go into detail about this idea.

On the other hand, we have that \mathbf{c}_{μ_f} , the vector of coefficients of the B-representation of $\{s_\tau\}$, is determined by the coefficients of the linear combination of the binomial sequences of $b(\{s_\tau\})$. That is, from equation (12) we have that, $\mathbf{c}_{\mu_f} = \{s_\tau\} \cdot H_t$, and we can deduce that the vector \mathbf{c}_{μ_f} is obtained as the linear combination of the rows of H_t which is determined by the elements of $\{s_\tau\}$. Moreover, the vector $b(\{s_\tau\}) = (i_1, i_2, \dots, i_k)$ indicates the positions of the 1s in \mathbf{c}_{μ_f} , i.e., the binomial sequences which form the B-representation. Note that this vector is closely related with the concept of the support of a Boolean function.

Therefore, we can deduce that there exists a connection between the vectors ξ_f and \mathbf{c}_{μ_f} , since both provide the B-representation associated to certain sequence. It implies that the ANF of a function f , with vector \mathbf{m}_{μ_f} , and the sequence $\{s_\tau\}$ are related too. In Figure 2, we show the different connections.

From now on, we work indistinctly with the sequence $\{s_\tau\}$ of length 2^t (which can be expressed by a vector \mathbf{s}) and f the Boolean function of t variables whose ANF is determined by the vector \mathbf{m}_{μ_f} , which coincides with the sequence \mathbf{s} . In this case, we say that the sequence $\{s_\tau\}$ is associated with f .

Through these relations, we can state that the set of minterms of a Boolean function f coincides with the vector $b(\{s_\tau\})$ of the B-representation of the sequence $\{s_\tau\}$ associated with f .

The following example shows this relation.

Example 5: Consider again Example 4. We have that $b(\{s_\tau\}) = (3, 4, 6, 7, 8, 9, 10, 11, 12)$. We can identify this set of indices with the set of minterms of a Boolean function $f(\mathbf{x})$ of 4 variables, that is, we have that

$$\xi_f = (0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0).$$

On the one hand, from equation (11), we obtain the sequence

$$(s_0, s_1, \dots, s_{15}) = (0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1).$$

On the other hand, from equation (8), we have that

$$\mathbf{m}_{\mu_f} = (0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1) = \{s_\tau\}.$$

Therefore, the ANF of f is given by $f(\mathbf{x}) = x^3 + x^4 + x^5 + x^8 + x^{11} + x^{12} + x^{14} + x^{15}$. We have related a sequence with a Boolean function.

4 New representation of sequences

In this section, we introduce a new representation of sequences, named reverse-ANF, based on the ANF of a Boolean function and the B-representation of such sequences. We study this representation and analyse its relation with the rest of representations studied in this paper.

4.1 Reverse-ANF of a sequence

In order to introduce this new representation of a sequence, we need to consider \mathbf{x}^u in a reverse way (see, for example Table 2), that is, we denote by

$$\overline{\mathbf{x}^u} = x_n^{u_1} x_{n-1}^{u_2} \cdots x_1^{u_n} \quad \text{with} \quad x_{n-j+1}^{u_j} = \begin{cases} x_{n-j+1}, & \text{if } u_j = 1, \\ 1, & \text{if } u_j = 0. \end{cases} \quad (13)$$

Table 2: First 8 powers of $\overline{\mathbf{x}^u}$

$\overline{\mathbf{x}^u}$	\mathbf{u}	$x_3^{u_1} x_2^{u_2} x_1^{u_3}$
$\overline{\mathbf{x}^0}$	000	1
$\overline{\mathbf{x}^1}$	001	x_1
$\overline{\mathbf{x}^2}$	010	x_2
$\overline{\mathbf{x}^3}$	011	$x_2 x_1$
$\overline{\mathbf{x}^4}$	100	x_3
$\overline{\mathbf{x}^5}$	101	$x_3 x_1$
$\overline{\mathbf{x}^6}$	110	$x_3 x_2$
$\overline{\mathbf{x}^7}$	111	$x_3 x_2 x_1$

Next theorem is the most important result of our paper since that it establishes a bijection between the B-representation of a sequence and a Boolean function.

Theorem 2: Let \mathcal{S} be the set of all binary sequences of period power of two and \mathcal{B} the set of all Boolean functions. Consider the map given by

$$\phi : \begin{array}{ccc} \mathcal{S} & \longrightarrow & \mathcal{B} \\ \sum_{i=0}^{2^t-1} c_i \binom{n}{i} & \longrightarrow & c_0 \overline{\mathbf{x}^0} + c_1 \overline{\mathbf{x}^1} + \cdots + c_{2^t-1} \overline{\mathbf{x}^{2^t-1}} \end{array} \quad (14)$$

where $\overline{\mathbf{x}^i}$ is as in equation (13), \mathbf{i} is the shortest binary representation of i and t represents any positive integer. Then, ϕ is a bijection.

PROOF: On the one hand, two different sequences in \mathcal{S} have different B-representations, therefore, the vector $(c_0, c_1, \dots, c_{2^t-1})$ is unique for each sequence. Thus, two different sequences in \mathcal{S} correspond to two different Boolean functions in \mathcal{B} .

On the other hand, consider a Boolean function $f \in \mathcal{B}$ of t variables, whose Boolean representation is

$$f(\mathbf{x}) = \sum_{i=0}^{2^t-1} c_i \overline{\mathbf{x}^i}.$$

This function represents the sequence

$$\sum_{i=0}^{2^t-1} c_i \left\{ \binom{n}{i} \right\}.$$

As a consequence, the map ϕ is bijective. \square

Note that, if we had considered the terms \mathbf{x}^j , instead of $\overline{\mathbf{x}^j}$ in the bijection given in expression (14), then the Boolean representation of a binary sequence would not be unique; since that this would depend on the number of digits considered in the binary representation of j , and, therefore, of the number of variables. To illustrate this idea, consider the map:

$$\varphi : \begin{array}{ccc} \mathcal{S} & \longrightarrow & \mathcal{B} \\ \sum_{i=0}^{2^t-1} c_i \binom{n}{i} & \longrightarrow & c_0 \mathbf{x}^0 + c_1 \mathbf{x}^1 + \dots + c_{2^t-1} \mathbf{x}^{2^t-1} \end{array} \quad (15)$$

In (14) we have that the r-ANF of $\{s_\tau\} = \left\{ \binom{n}{1} \right\}$ is $\mathcal{B}(\{s_\tau\}) = \phi(\{s_\tau\}) = \overline{\mathbf{x}^1}$. However, the form of the representation in (15) depends on the number of variables. If we consider Boolean functions of 4 variables, the representation in (15) is

$$\varphi(\{s_\tau\}) = \mathbf{x}^1 = x_1^0 x_2^0 x_3^0 x_4^1 = x_4.$$

Now, if we consider 5 variables, then we have that

$$\varphi(\{s_\tau\}) = \mathbf{x}^1 = x_1^0 x_2^0 x_3^0 x_4^1 x_5^1 = x_5.$$

Therefore, considering the map given in (15), the same sequence has two different representations, which is not practical for our purposes. For this reason, we consider the bijection ϕ given in expression (14) instead of φ in expression (15). On that account, we found necessary the introduction of the r-ANF of a sequence s_τ in order to obtain a unique Boolean function representing the sequence.

According to Theorem 2, we can introduce a new representation, in terms of Boolean functions, for binary sequences with period a power of two.

Definition 1: Given a binary sequence $\{s_\tau\}$ with period 2^t , and its B-representation $b(\{s_\tau\}) = \sum_{i=0}^{2^t-1} c_i \binom{n}{i}$, the Boolean representation of this sequence, called reverse-ANF of $\{s_\tau\}$, and denoted by $\mathcal{B}(\{s_\tau\})$, is given by $\mathcal{B}(\{s_\tau\}) = \phi(\{s_\tau\})$, where ϕ is as

Table 3: Boolean representation of the first 16 binomial sequences

Binomial sequence	Boolean representation
$\binom{n}{0}$	1
$\binom{n}{1}$	x_1
$\binom{n}{2}$	x_2
$\binom{n}{3}$	x_2x_1
$\binom{n}{4}$	x_3
$\binom{n}{5}$	x_3x_1
$\binom{n}{6}$	x_3x_2
$\binom{n}{7}$	$x_3x_2x_1$
$\binom{n}{8}$	x_4
$\binom{n}{9}$	x_4x_1
$\binom{n}{10}$	x_4x_2
$\binom{n}{11}$	$x_4x_2x_1$
$\binom{n}{12}$	x_4x_3
$\binom{n}{13}$	$x_4x_3x_1$
$\binom{n}{14}$	$x_4x_3x_2$
$\binom{n}{15}$	$x_4x_3x_2x_1$

in expression (14); that is,

$$\mathcal{B}(\{s_\tau\}) = \sum_{i=0}^{2^t-1} c_i \overline{\mathbf{x}^i}. \quad (16)$$

Observe that the vector of the coefficients of the reverse-ANF of $\{s_\tau\}$, given in expression (16), coincides with \mathbf{c}_{μ_f} , the vector of the coefficients of the B-representation of $\{s_\tau\}$. Therefore, we use the vector \mathbf{c}_{μ_f} indistinctly for the B-representation and the r-ANF of a sequence $\{s_\tau\}$.

In Table 3 we can find the r-ANF of the first 16 binomial sequences. Note that these r-ANF can be considered as Boolean functions of 4 variables. Moreover, if we compute their truth tables, we obtain the binomial sequences of length 16 given in Table 1 (evaluating the elements $\mathbf{x} \in \mathbb{F}_2^n$ in reverse order. See Table 4 for the case $\binom{n}{1}$). Next example illustrates this idea.

Example 6: Consider the r-ANF of a sequence $\{s_\tau\}$ of length 2^4 given by

$$\mathcal{B}(\{s_\tau\}) = \overline{\mathbf{x}^1} = x_1.$$

The truth table of this Boolean function, obtained in Table 4, is

$$(0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1),$$

which coincides with the binomial sequence $\binom{n}{1}$ given in Table 1. If we want to know the sequence $\{s_\tau\}$ associated to this r-ANF, we have to apply the equation (11), considering

$$\mathbf{c}_{\mu_f} = (0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),$$

Table 4:
Truth table of
 $\mathcal{B}(\{s_\tau\}) = \mathbf{x}^{\mathbf{1}}$

(x_4, x_3, x_2, x_1)	x_1
(0,0,0,0)	0
(0,0,0,1)	1
(0,0,1,0)	0
(0,0,1,1)	1
(0,1,0,0)	0
(0,1,0,1)	1
(0,1,1,0)	0
(0,1,1,1)	1
(1,0,0,0)	0
(1,0,0,1)	1
(1,0,1,0)	0
(1,0,1,1)	1
(1,1,0,0)	0
(1,1,0,1)	1
(1,1,1,0)	0
(1,1,1,1)	1

with $\mathcal{B}(\{s_\tau\}) = (1)$. Therefore, we have that

$$\{s_\tau\} = (0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1),$$

that is, the truth table of the r-ANF. Moreover, the ANF associated to the sequence $\{s_\tau\}$ is

$$\begin{aligned} f(x_1, x_2, x_3, x_4) &= x^{\mathbf{1}} + x^{\mathbf{3}} + x^{\mathbf{5}} + x^{\mathbf{7}} + x^{\mathbf{9}} + x^{\mathbf{11}} + x^{\mathbf{13}} + x^{\mathbf{15}} \\ &= x_4 + x_3x_4 + x_2x_4 + x_2x_3x_4 + x_1x_4 + x_1x_3x_4 + x_1x_2x_4 + x_1x_2x_3x_4. \end{aligned}$$

Observe that in the previous example the ANF associated to the sequence $\{s_\tau\}$ and the reverse-ANF are completely different and perform distinct functions. Through the ANF, we obtain the sequence $\{s_\tau\}$ which is the truth table of the reverse-ANF; and, from the reverse-ANF, we compute the B-representation and, therefore, the minterms representation of $\{s_\tau\}$.

Next, we introduce some results where we study two important properties about Boolean functions such as balancedness and the maximum degree, both obtained from the r-ANF (or equivalently, from the B-representation) and the ANF of the sequence associated with the function.

The following two theorems allow us to know if the r-ANF of a sequence $\{s_\tau\}$ (and the ANF of the function f associated with it) has maximum degree. These results are based on Theorem 4 of [37], even though we apply this result to sequences and give an alternative proof.

Theorem 3: Consider $\{s_\tau\}$ a sequence of period 2^t , with t a positive integer. Then, the ANF associated to $\{s_\tau\}$ has maximum degree if and only if $w(\mathbf{c}_{\mu_f})$ is odd.

PROOF: Suppose that we have a sequence $\{s_\tau\}$ of period 2^t whose ANF has maximum degree, that is, it has the term $x_1x_2 \cdots x_t$ and, therefore, the vector \mathbf{m}_{μ_f} has 1 in the last position. Remember that all the binomial sequences have 1 in the last position. Due to the fact that the rows of the matrix H_t are composed of the first $2^t - 1$ binomial sequences, and from expression (11), we obtain \mathbf{m}_{μ_f} as a linear combination of the rows of H_t . Thus, $\{s_\tau\}$ is formed by the sum of an odd number of binomial sequences, that is, $w(\mathbf{c}_{\mu_f})$ is odd.

Now, suppose that $w(\mathbf{c}_{\mu_f})$ is odd. In a similar way as in the previous implication, from expression (11), we obtain \mathbf{m}_{μ_f} as a linear combination of an odd number of rows of H_t . Therefore, \mathbf{m}_{μ_f} has a 1 in the last position and the ANF has maximum degree. \square

We obtain a similar result for the r-ANF of a sequence.

Theorem 4: *Consider $\{s_\tau\}$ a sequence of period 2^t , with t a positive integer. Then, the r-ANF of $\{s_\tau\}$ has maximum degree if, and only if, $\binom{n}{2^t-1}$ is a member of the r-ANF of $\{s_\tau\}$.*

PROOF: Consequence of Theorem 3 and the isomorphism properties of (14). \square

Next theorem provides a necessary condition for a sequence to be balanced.

Theorem 5: *Let $\{s_\tau\}$ be a balanced sequence of period 2^t . Then, the term $\binom{n}{2^t-1}$ is not part of the B-representation. Moreover, the term $\overline{x^{2^t-1}}$ is not part of the r-ANF.*

PROOF: If $\{s_\tau\}$, with period 2^t , is balanced, then its weight is 2^{t-1} . From equation (12), we have that the vector \mathbf{c}_{μ_f} is the result of the sum of 2^{t-1} binomial sequences, determined by the components of the vector $(s_0, s_1, \dots, s_{2^t-1})$. We know that all the binomial sequences has 1 in the last position. Therefore, the last component of \mathbf{c}_{μ_f} is zero. It means that the B-representation of $\{s_\tau\}$ does not have the term $\binom{n}{2^t-1}$. That is, as $c_{2^t-1} = 0$, then $\overline{x^{2^t-1}}$ is not part of the r-ANF of $\{s_\tau\}$. \square

As a consequence of the previous results, we can deduce the following corollary.

Corollary 1: *The linear complexity LC of a balanced sequence with period 2^t satisfies $LC < 2^t$.*

4.2 Relation between the representations of a sequence

In Theorem 2, we show the relation between between the r-ANF and the B-representation of a sequence. We can obtain directly one representation from the other one. Moreover, in Section 3, we analyse the relation between the ANF and the B-representation of a sequence, obtaining that the vector \mathbf{m}_{μ_f} can be considered as the sequence. This connection allows us to obtain the r-ANF of a sequence $\{s_\tau\}$ directly from the sequence.

From equation (12) and Theorem 2, we can obtain

$$\mathbf{c}_{\mu_f} = (s_0, s_1, \dots, s_{2^t-1}) \cdot H_t \text{ mod } 2.$$

We can rewrite this equation in terms of the ANF as follows

$$\mathbf{c}_{\mu_f} = \mathbf{m}_{\mu_f} \cdot H_t \text{ mod } 2; \tag{17}$$

that is, we can obtain the B-representation from the ANF, and vice versa. Moreover, we can also obtain the r-ANF from the ANF and vice versa. We check this relation from the following example.

Example 7: Consider again the Example 5. We have that

$$\mathbf{m}_{\mu_f} = (0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1).$$

Applying equation (17), we obtain that

$$\mathbf{c}_{\mu_f} = (0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0),$$

that is, $b(\{s_\tau\}) = \{3, 4, 6, 7, 8, 9, 10, 11, 12\}$. Therefore, the r-ANF of $\{s_\tau\}$ is

$$\begin{aligned} \mathcal{B}(\{s_\tau\}) &= \overline{x^3} + \overline{x^4} + \overline{x^6} + \overline{x^7} + \overline{x^8} + \overline{x^9} + \overline{x^{10}} + \overline{x^{11}} + \overline{x^{11}} \\ &= x_2x_1 + x_3 + x_3x_2 + x_3x_2x_1 + x_4 + x_4x_1 + x_4x_2 + x_4x_2x_1 + x_4x_3. \end{aligned}$$

Next result establishes that a balanced Boolean function cannot have maximum degree. Remember that a balanced function is determined by its truth table ξ_f , and, equivalently, through the vector \mathbf{c}_{μ_f} .

Theorem 6: The ANF of a balanced Boolean function of t variables cannot include the maximum monomial $x_1x_2 \cdots x_t$, that is, it does not have maximum degree.

PROOF: Consequence of Theorem 3. \square

Example 8: Consider a balanced Boolean function of 4 variables whose truth table is

$$\xi_f = (1, 0, 0, 1, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0).$$

Applying equation (8), we obtain that

$$\mathbf{m}_{\mu_f} = (1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0),$$

so, its ANF is

$$f(\mathbf{x}) = x^0 + x^1 + x^2 + x^5 + x^{10} + x^{14} = 1 + x_4 + x_3 + x_2x_4 + x_1x_3 + x_1x_2x_3,$$

which has not the maximum degree.

If we compute the r-ANF, as the vector ξ_f is equivalent to the vector \mathbf{c}_{μ_f} , we have that $b(\{s_\tau\}) = \{0, 3, 4, 5, 8, 10, 12, 13\}$. The element $2^4 - 1 = 15$ is not part of the B-representation, then from Theorem (3), the r-ANF

$$\begin{aligned} \mathcal{B}(\{s_\tau\}) &= \overline{x^0} + \overline{x^3} + \overline{x^4} + \overline{x^5} + \overline{x^8} + \overline{x^{10}} + \overline{x^{12}} + \overline{x^{13}} \\ &= 1 + x_1x_2 + x_3 + x_1x_3 + x_4 + x_2x_4 + x_3x_4 + x_1x_2x_4, \end{aligned}$$

has not maximum degree.

Corollary 2: Let f be the Boolean function associated with the sequence $\{s_\tau\}$ of period a power of two, that is, its ANF or r-ANF. If f is balanced, then f has not maximum degree.

Next, we present some important results which allows us to know more about the r-ANF through the B-representation. Both results follow immediately from the Definition 1.

Theorem 7: *The r-ANF of a sequence $\{s_\tau\}$ is $\mathcal{B}(\{s_\tau\}) = x_j$, if and only if $b(\{s_\tau\}) = \binom{n}{2^j-1}$. Particularly, $\mathcal{B}(\{s_\tau\}) = 1$, if and only if the B-representation is $b(\{s_\tau\}) = \binom{n}{0}$.*

Check, for example, the r-ANF of the sequences $\binom{n}{1}$, $\binom{n}{2}$, $\binom{n}{4}$ and $\binom{n}{8}$ in Table 3.

Theorem 8: *The r-ANF of a sequence $\{s_\tau\}$ is $\mathcal{B}(\{s_\tau\}) = x_j x_{j-1} \cdots x_2 x_1$ if and only if the B-representation is $b(\{s_\tau\}) = \binom{n}{2^j-1}$.*

Check, for example, the r-ANF of the sequences $\binom{n}{3}$, $\binom{n}{7}$, and $\binom{n}{15}$ in Table 3.

Next result allows us to obtain the r-ANF of a sequence $\{s_\tau\}$ from its B-representation.

Theorem 9: *Given a sequence $\{s_\tau\}$ with B-representation $b(\{s_\tau\}) = \sum_{i=0}^{2^t-1} c_i \binom{n}{i}$, we have that its r-ANF is given by*

$$\mathcal{B}(\{s_\tau\}) = \sum_{i=0}^{2^t-1} c_i \mathcal{B}\left(\binom{n}{i}\right)$$

PROOF: Note that the map considered in equation (14) is a linear operator, that is, for any two sequences $\{s_\tau\}$ and $\{u_\tau\}$, both with period power of two, we have that $\phi(\{s_\tau\} + \{u_\tau\}) = \phi(\{s_\tau\}) + \phi(\{u_\tau\})$. Therefore,

$$\mathcal{B}\left(\sum_{i=0}^{2^t-1} c_i \binom{n}{i}\right) = \phi\left(\sum_{i=0}^{2^t-1} c_i \binom{n}{i}\right) = \sum_{i=0}^{2^t-1} c_i \phi\left(\binom{n}{i}\right) = \sum_{i=0}^{2^t-1} c_i \mathcal{B}\left(\binom{n}{i}\right).$$

Theorem 9 implies that the application ϕ given in (14) is an \mathbb{F}_2 -linear space isomorphism. Therefore, in order to obtain the r-ANF of a sequence, we can just sum over \mathbb{F}_2 the r-ANFs of the binomial sequences involved in the B-representation. As an example, observe that the r-ANF of the sequence $\binom{n}{7} + \binom{n}{6}$ is given by

$$\mathcal{B}\left(\binom{n}{7} + \binom{n}{6}\right) = \mathcal{B}\left(\binom{n}{7}\right) + \mathcal{B}\left(\binom{n}{6}\right) = x_3 x_2 x_1 + x_3 x_2.$$

4.3 reverse-ANF of shifted sequences

It is well known that, when a sequence is shifted a finite number of positions, the B-representation changes [26]. As a consequence, the same happens with the reverse-ANF. In this section, we study the r-ANF of shifted versions of a sequence.

Theorem 10: *Consider the binomial sequence $\binom{n}{2^j-1}$ represented by x_j , for $j \in \{2, 3, \dots\}$. If we shift cyclically such a sequence one bit to the left, then the r-ANF of the shifted version is $x_1 x_2 \cdots x_{j-1} + x_j$. When $j = 1$, we have that the r-ANF of its shifted version is $1 + x_1$.*

PROOF: Consequence of the isomorphism properties of (14), Theorem 8 and Lemma 9 in [26]. \square

Example 9: Consider the sequence

$$\binom{n}{8} = \{0000000011111111 \dots\}$$

whose r -ANF is x_4 . If we shift the sequence one bit to the left, we obtain the following shifted version of the same sequence (see [26, Lemma 9] for more details):

$$\binom{n}{8} + \binom{n}{7} = \{0000000111111110 \dots\},$$

whose r -ANF is $x_1x_2x_3 + x_4$.

We can generalize the result given in Theorem 10 for sequences whose r -ANFs are monomials.

Theorem 11: Consider the binary sequence represented by $x_{i_k}x_{i_{k-1}} \cdots x_{i_2}x_{i_1}$ with $i_k > i_{k-1} > \cdots > i_2 > i_1$. If we shift cyclically such a sequence one bit to the left, then the r -ANF of the resultant shifted version is $x_{i_k}x_{i_{k-1}} \cdots x_{i_3}x_{i_2} + x_{i_k}x_{i_{k-1}} \cdots x_{i_2}x_{i_1}$.

PROOF: Consequence of the isomorphism properties of (14) and Lemma 9 in [26]. \square

Example 10: Consider the sequence

$$\binom{n}{7} = \{0000000100000001 \dots\}$$

whose r -ANF is $x_3x_2x_1$. If we shift the sequence one bit to the left, we obtain the following shifted version of the same sequence (see [26, Lemma 9] for more details):

$$\binom{n}{7} + \binom{n}{6} = \{0000001000000010 \dots\},$$

whose r -ANF is $x_3x_2x_1 + x_3x_2$.

As a consequence of Theorems 10 and 11, we can introduce the following theorem.

Theorem 12: Consider the r -ANF of a sequence $\mathcal{B}(\{s_\tau\}) = \sum_{i=0}^{2^t-1} c_i \overline{x^i}$. If we shift the sequence one bit to the left, we obtain a shifted version of the same sequence with r -ANF given by

$$\mathcal{B}(\{s_\tau\}) = \sum_{i=1}^{2^t-1} c_i \left(\overline{x^i} + \Delta \left(\overline{x^i} \right) \right) + c_0.$$

where

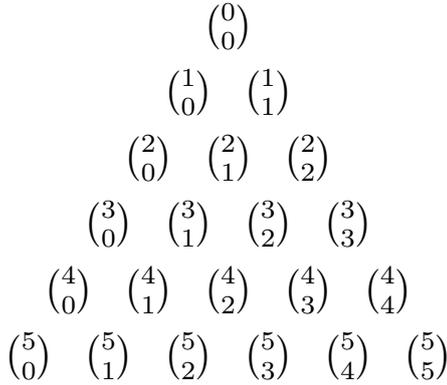
$$\Delta \left(\overline{x^i} \right) = \begin{cases} x_{i_j} \cdots x_{i_3}x_{i_2} & \text{if } \overline{x^i} = x_{i_j} \cdots x_{i_2}x_{i_1} \quad \text{for some } i_1 < i_2 < \cdots < i_j. \\ x_{j-1} \cdots x_2x_1 & \text{if } \overline{x^i} = x_j \quad \text{for some } j. \end{cases}$$

PROOF: Consequence of the isomorphism properties of (14) and Theorem 15 in [26]. \square

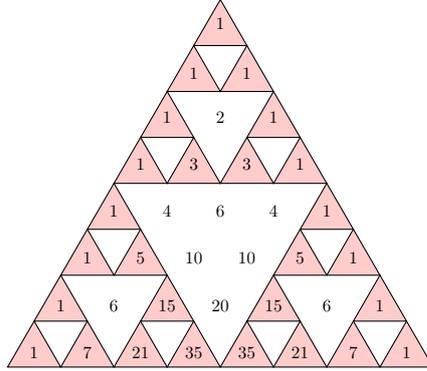
Example 11: Consider the sequence $\{s_\tau\} = \{01000100 \dots\}$ represented by $b(\{s_\tau\}) = \binom{n}{3} + \binom{n}{1}$. The r -ANF of this sequence is given by $\mathcal{B}(\{s_\tau\}) = x_2x_1 + x_1$. This sequence

Fig. 4: Binomial coefficients arranged as triangles

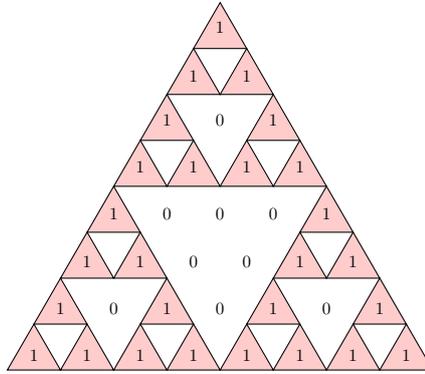
(a) Pascal's triangle



(b) Sierpinski's triangle



(c) Sierpinski's triangle mod 2



the Pascal's triangle modulo 2 (Figure 4c), we can observe that the diagonals are the binomial sequences starting in a different position (here they start with the first 1).

First, we need to introduce a minor result regarding the rows of the (binary) Sierpinski triangle.

Lemma 1: Consider $d^n = (d_0^n, d_1^n, \dots, d_n^n)$, for $n = 0, 1, \dots$, the n -th row of the binary Sierpinski triangle (see Figure 4c). Then,

1. The first and the last elements of d^n are equal to 1, i.e., $d_0^n = d_n^n = 1$.
2. It is possible to generate the elements in the $(n+1)$ -th row, denoted by d^{n+1} , using the elements in d^n , in the following way: $d_{k+1}^{n+1} = d_k^n + d_{k+1}^n$, for $k = 1, \dots, n-1$.

PROOF: 1. The first and last elements are $\binom{n}{0} = \binom{n}{n} = 1$.

2. Consequence of the following combinatorial property:

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$$

Theorem 13: Consider the binomial sequence $\binom{n}{k}$ with period $T = 2^t$ and consider r a positive integer such that $r < T$. If we shift such sequence r positions to the left, then the r -ANF of the resultant shifted version is given by

$$\sum_{i=0}^{\min\{r,k\}} d_{r-i}^r \phi \left(\binom{n}{k-i} \right) \quad (18)$$

where $d^r = (d_0^r, d_1^r, \dots, d_r^r)$ is the r -th row of the binary Sierpinski triangle (see Figure 4c).

PROOF: Let us start with the binomial sequence $\binom{n}{k}$. If we shift this sequence one position to the left we obtain the sequence $\binom{n}{k-1} + \binom{n}{k}$ (see [26] for more details).

This means that the r -ANF of this shifted version is $d_0^1 \phi \left(\binom{n}{k-1} \right) + d_1^1 \phi \left(\binom{n}{k} \right)$, where $d^1 = (d_0^1, d_1^1) = (1, 1)$ is the 1st row of Sierpinski (Note that the 0-th row is $d^0 = (1)$).

Assume now that after r shifts (with $r < k$), the B-representation is

$$\sum_{i=0}^r d_{r-i}^r \binom{n}{k-i} = d_0^r \binom{n}{k-r} + d_1^r \binom{n}{k-r+1} + \dots + d_{r-1}^r \binom{n}{k-1} + d_r^r \binom{n}{k}$$

and the corresponding r -ANF is

$$\mathcal{B} \left(\sum_{i=0}^r d_{r-i}^r \binom{n}{k-i} \right) = \sum_{i=0}^r d_{r-i}^r \mathcal{B} \left(\binom{n}{k-i} \right).$$

Now, if we shift this r -shifted version one position to the left, we obtain that the B-representation is given by

$$\sum_{i=0}^r d_{r-i}^r \left(\binom{n}{k-i-1} + \binom{n}{k-i} \right) = d_0^r \binom{n}{k-r-1} + \sum_{i=0}^{r-1} (d_{r-i}^r + d_{r-i+1}^r) \binom{n}{k-i} + d_r^r \binom{n}{k} \quad (19)$$

(see [26] for more details). According to Lemma 1, we have that $d_0^r = d_0^{r+1} = 1$, $d_r^r = d_{r+1}^{r+1} = 1$ and $d_{r-i}^r + d_{r-i+1}^r = d_{r-i+1}^{r+1}$. Therefore, expression (19) is equal to

$$\sum_{i=0}^{r+1} d_{r+1-i}^{r+1} \binom{n}{k-i}$$

and the corresponding r -ANF for $r + 1$ shifts is given by

$$\mathcal{B} \left(\sum_{i=0}^{r+1} d_{r+1-i}^{r+1} \binom{n}{k-i} \right) = \sum_{i=0}^{r+1} d_{r+1-i}^{r+1} \mathcal{B} \left(\binom{n}{k-i} \right),$$

which proves that the expression (18) is correct for $r < k$.

When $r = k$, the r -ANF has the following form

$$\mathcal{B} \left(\sum_{i=0}^k d_{k-i}^k \binom{n}{k-i} \right) = \sum_{i=0}^k d_{k-i}^k \mathcal{B} \left(\binom{n}{k-i} \right).$$

Note that, in this case, the binomial sequence $\binom{n}{0}$ appears for the first time in the B-representation of the shifted version; this means that the term 1 appears for the first time in the r -ANF. It is worth mentioning that the binomial sequence $\binom{n}{0}$ is invariant for translations, i.e., it remains the same after shifts. Therefore, if we shift the binomial sequence $r = k + 1$ position to the left, the B-representation has the following form

$$\sum_{i=0}^{k-1} d_{k-i}^k \left(\binom{n}{k-i-1} + \binom{n}{k-i} \right) + \binom{n}{0} = \sum_{i=1}^k (d_{k-i}^k + d_{k-i+1}^k) \binom{n}{k-i} + d_k^k \binom{n}{k}. \quad (20)$$

Recalling that $d_{k+1}^{k+1} = d_k^k = 1$ and $d_{k-i}^k + d_{k-1-i}^k = d_{k+1-i}^{k+1}$, we have that the expression in (20) is equal to

$$\sum_{i=0}^k d_{k+1-i}^{k+1} \binom{n}{k-i}.$$

Therefore, the r -ANF in this case ($r = k + 1$) is

$$\mathcal{B} \left(\sum_{i=0}^k d_{k+1-i}^{k+1} \binom{n}{k-i} \right) = \sum_{i=0}^k d_{r-i}^r \mathcal{B} \left(\binom{n}{k-i} \right).$$

Using the same argument for any $k + 1 < r < T$, the theorem follows.

Example 12: Consider the binomial sequence

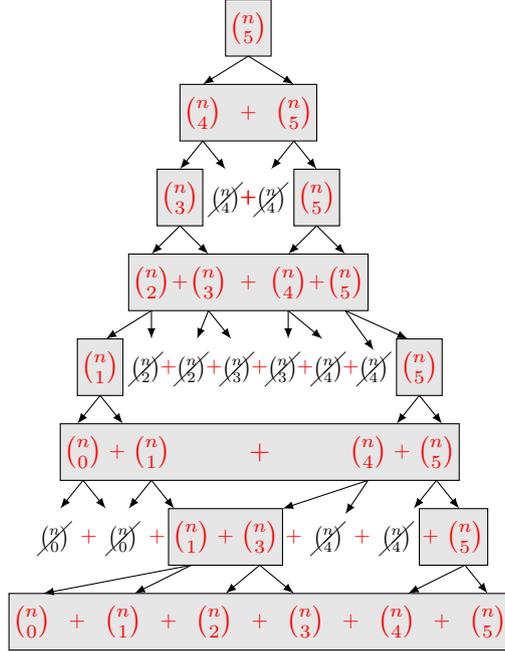
$$\binom{n}{k} = \binom{n}{5} = \{00000101\dots\}$$

whose r -ANF is x_3x_1 . Assume we shift this sequence $r = 4$ positions to the left, then the shifted version is given by $\{01010000\dots\}$. The 4th row of the Sierpinski triangle is $d^4 = (1, 0, 0, 0, 1)$. Therefore the r -ANF of the shifted version is given by

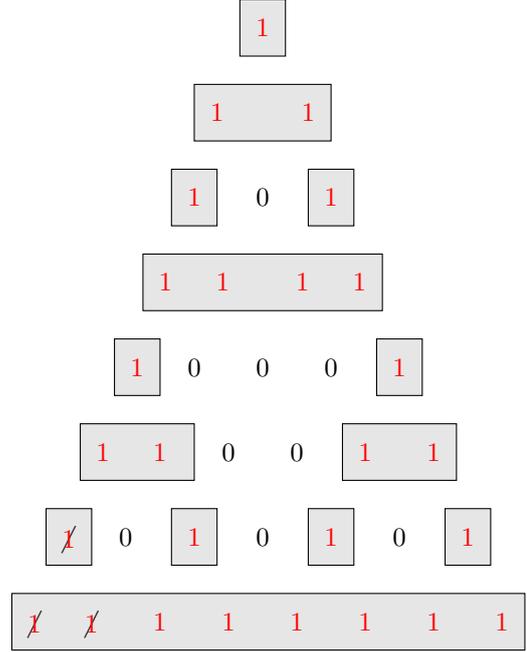
$$\sum_{i=0}^4 d_{4-i}^4 \mathcal{B} \left(\binom{n}{5-i} \right) = \mathcal{B} \left(\binom{n}{5} \right) + \mathcal{B} \left(\binom{n}{1} \right) = x_3x_1 + x_1.$$

Fig. 5: Binomial representations of $\binom{n}{5}$ and the Sierpinski's triangle

(a) Binomial representations of $\binom{n}{5}$



(b) First 8 rows of the Sierpinski triangle



Now, assume we shift the sequence $r = 6$ positions to the left, then the shifted version is given by $\{0100001 \dots\}$. The 6th row of the triangle is $d^6 = (1, 0, 1, 0, 1, 0, 1)$. Therefore the r -ANF of the shifted version is given by

$$\sum_{i=0}^5 d_{6-i}^6 \mathcal{B} \left(\binom{n}{5-i} \right) = \mathcal{B} \left(\binom{n}{5} \right) + \mathcal{B} \left(\binom{n}{3} \right) + \mathcal{B} \left(\binom{n}{1} \right) = x_3 x_1 + x_2 x_1 + x_1.$$

In this case, since $r > k$, we only used the last $k + 1 = 6$ elements of row $d^6 = (1, 0, 1, 0, 1, 0, 1)$. Check Figure 5 to understand more deeply the relation between the shifted versions and the Sierpinski triangle.

Corollary 3: Assume $\sum_{i=0}^{2^t-1} c_i \binom{n}{i}$ is the B -representation of $\{s_\tau\}$, then the r -ANF of the sequence cyclically shifted r bits to the left is

$$\sum_{i=1}^{2^t-1} c_i \sum_{j=0}^{\min\{r,i\}} d_{r-j}^r \mathcal{B} \left(\binom{n}{k-j} \right),$$

where d^r is the r -th row of the binary Sierpinski triangle (see Figure 4c).

5 The reverse sequence

Now, we are ready to introduce the definition of reverse sequence and its relation with the shifted versions of a sequence.

Definition 2: Given the sequence $\{s_\tau\}$ with period T , the reverse sequence $\{s_\tau\}^*$ is a sequence of period T as well, where its τ -th term satisfies the equality $s_\tau^* = s_{T-1-\tau}$ for $\tau = 0, 1, \dots, T-1$.

For instance, the reverse sequence of the binomial sequence $\binom{n}{5} = \{0\ 0\ 0\ 0\ 0\ 1\ 0\ 1\}$ is the sequence given by $\binom{n}{5}^* = \{1\ 0\ 1\ 0\ 0\ 0\ 0\}$, which is also a shifted version of $\binom{n}{5}$.

Theorem 14: The reverse sequence $\binom{n}{k}^*$ of the binomial sequence $\binom{n}{k}$ is a shifted version $\binom{n}{k}$ with shift k .

PROOF: This is consequence of the following facts:

1. The first k elements of the sequence $\binom{n}{k}$ are zeros and the $(k+1)$ -th element is 1.
2. If the period of the sequence $\binom{n}{k}$ is T , then the subsequence $\binom{n}{k}_{k+1 \leq n \leq T-1}$ is symmetric, in the sense that such a subsequence is equal to its own reverse sequence (for more detail on the properties of the binomial sequences please check Figure 8 in [26]). \square

As a consequence of the previous theorem and Theorem 13, we have the following result.

Corollary 4: The B -representation of $\binom{n}{k}^*$ can be obtained by using the k -th row of the Sierpinski triangle, denoted by d^k , as follows:

$$b\left(\binom{n}{k}^*\right) = \sum_{i=0}^k d_i^k \binom{n}{i}.$$

As a consequence of the previous corollary we have that the r-ANF of the sequence $\binom{n}{k}^*$ is given by

$$\mathcal{B}\left(\binom{n}{k}^*\right) = \sum_{i=0}^k d_i^k \mathcal{B}\left(\binom{n}{i}\right).$$

Example 13: Given the sequence $\binom{n}{5}$, if we want to compute both representations of the reverse sequence, then we use the 5th row of the Sierpinski's triangle, that is, $d^5 = (1, 1, 0, 0, 1, 1)$. In fact, we have that

$$b\left(\binom{n}{5}^*\right) = \sum_{i=0}^5 d_i^5 \binom{n}{i} = \binom{n}{0} + \binom{n}{1} + \binom{n}{4} + \binom{n}{5}.$$

$$\mathcal{B}\left(\binom{n}{5}^*\right) = \mathcal{B}\left(\binom{n}{0} + \binom{n}{1} + \binom{n}{4} + \binom{n}{5}\right) = 1 + x_1 + x_3 + x_3x_1$$

It is easy to check that the reverse sequence is a shifted version of the own sequence:

$$\binom{n}{5} : 0\ 0\ 0\ 0\ 0\ 1\ 0\ 1\ \dots$$

$$\binom{n}{5}^* = \binom{n}{0} + \binom{n}{1} + \binom{n}{4} + \binom{n}{5} : 1 0 1 0 0 0 0 \dots$$

Theorem 15: Consider a sequence with period 2^t and B-representation $\sum_{k=0}^{2^t-1} c_k \binom{n}{k}$, then the B-representation of its reverse sequence is given by

$$b\left(\left\{\sum_{k=0}^{2^t-1} c_k \binom{n}{k}\right\}^*\right) = \sum_{k=0}^{2^t-1} c_k \left(\sum_{i=0}^k d_i^k \binom{n}{i}\right).$$

and the r-ANF is

$$\mathcal{B}\left(\left\{\sum_{k=0}^{2^t-1} c_k \binom{n}{k}\right\}^*\right) = \sum_{k=0}^{2^t-1} c_k \left(\sum_{i=0}^k d_i^k \mathcal{B}\left(\binom{n}{i}\right)\right).$$

PROOF: This is a consequence of the fact that the reverse version of a given sequence is the linear combination of the reverse sequences of the corresponding binomial sequences in the B-representation of such a sequence, that is:

$$\left\{\sum_{k=0}^{2^t-1} c_k \binom{n}{k}\right\}^* = \sum_{k=0}^{2^t-1} c_k \binom{n}{k}^*.$$

□

Example 14: Consider the sequence $\{s_\tau\}$ with B-representation $b(\{s_\tau\}) = \binom{n}{2} + \binom{n}{5} + \binom{n}{7}$. We can compute the B-representation of its reverse sequence $\{s_\tau\}^*$ summing up the reverse sequences of the corresponding binomial sequences involved in the B-representation:

$$\begin{array}{r} \left\{\binom{n}{2}\right\}^* : \left\{\binom{n}{0} + \binom{n}{2}\right\} \\ \left\{\binom{n}{5}\right\}^* : \left\{\binom{n}{0} + \binom{n}{1} + \binom{n}{4} + \binom{n}{5}\right\} \\ \left\{\binom{n}{7}\right\}^* : \left\{\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3} + \binom{n}{4} + \binom{n}{5} + \binom{n}{6} + \binom{n}{7}\right\} \\ \hline \left\{\binom{n}{2} + \binom{n}{5} + \binom{n}{7}\right\}^* : \left\{\binom{n}{0} + \binom{n}{3} + \binom{n}{6} + \binom{n}{7}\right\} \end{array}$$

Therefore, the B-representation and the r-ANF of the reverse sequence will be:

$$\begin{aligned} b(\{s_\tau\}^*) &= \binom{n}{0} + \binom{n}{3} + \binom{n}{6} + \binom{n}{7} \\ \mathcal{B}(\{s_\tau\}^*) &= x_3 x_2 x_1 + x_3 x_2 + x_2 x_1 + 1. \end{aligned}$$

It is easy to check that both sequences are reverses:

$$\begin{aligned} \left\{\binom{n}{2} + \binom{n}{5} + \binom{n}{7}\right\} &: 0 0 1 1 0 1 1 1 \\ \left\{\binom{n}{0} + \binom{n}{3} + \binom{n}{6} + \binom{n}{7}\right\} &: 1 1 1 0 1 1 0 0 \end{aligned}$$

6 Reverse-ANF of generalized sequences

In [27], the authors introduced different representations of generalized self-shrunk sequences and analysed some cryptographic properties of them through the representations given. In this section, we study this new Boolean representation for sequences, the reverse-ANF, in the particular case of the generalized sequences with the aim to complete the work done in [27]. We present some interesting results which could help us in the research of some open problems related with these sequences.

Next, we introduce the definition of the generalized sequences and some necessary results required for the understanding of this section.

Consider a PN-sequence $\{u_\tau\}_{\tau \geq 0}$ obtained from a maximal-length LFSR with L stages, an L -dimensional vector $\mathcal{G} = [g_0, g_1, g_2, \dots, g_{L-1}] \in \mathbb{F}_2^L$ and let $\{v_\tau\}_{\tau \geq 0}$ be the sequence defined as:

$$v_\tau = g_0 u_\tau + g_1 u_{\tau-1} + g_2 u_{\tau-2} + \dots + g_{L-1} u_{\tau-L+1} \quad \text{for } \tau \geq 0.$$

Now, the decimation rule to generate new sequences $\{s_j\}_{j \geq 0}$ is given by:

$$\begin{cases} \text{If } u_\tau = 1, & \text{then } s_j = v_\tau, \\ \text{If } u_\tau = 0, & \text{then } v_\tau \text{ is discarded.} \end{cases}$$

The sequence $\{s_j\}_{j \geq 0}$, denoted by $S(\mathcal{G})$, is called the *generalized self-shrunk sequence*, GSS-sequence or simply **generalized sequence** associated with \mathcal{G} (see [30]); and the sequence generator is called the **generalized self-shrinking generator** (GSSG).

Note that when \mathcal{G} runs over $\mathbb{F}_2^L \setminus \{\mathbf{0}\}$ we obtain all the shifted versions of $\{u_\tau\}_{\tau \geq 0}$. The set of sequences $\mathcal{F} = \{S(\mathcal{G}) \mid \mathcal{G} \in \mathbb{F}_2^L \setminus \{\mathbf{0}\}\}$ is called the *family of generalized sequences* based on the PN-sequence $\{u_\tau\}_{\tau \geq 0}$. It is easy to check that the family $\mathcal{F} \cup \{S(\mathbf{0})\}$, where $S(\mathbf{0}) = \{0\ 0\ 0\ 0\ \dots\}$, is an additive group with the operation addition modulo 2 [30]. In particular, the opposite of any sequence $S(\mathcal{G})$ is the sequence itself. Moreover, the period of every generalized sequence is a divisor of 2^{L-1} (the number of ones in the PN-sequence) and every sequence of this family is balanced except for the identically 1 sequence and the null sequence [30, Theorem 1]. The generalized sequences of period 1 and 2, that is, $\{111111\dots\}$, $\{010101\dots\}$ and $\{101010\dots\}$, are referred as the trivial sequences of the family of generalized sequences. The null sequence is not considered as a generalized sequence.

Theorem 16: *The r-ANF of the trivial sequences in the family of generalized sequences are:*

$$\mathcal{B}(\{111111\dots\}) = 1, \quad \mathcal{B}(\{010101\dots\}) = x_1, \quad \text{and} \quad \mathcal{B}(\{101010\dots\}) = 1 + x_1.$$

Example 15: *Consider the primitive polynomial $p(x) = 1 + x + x^4$. In Table 5 we can find the generalized sequences obtained with $p(x)$ and the corresponding B-representation and r-ANF for each sequence. Note that all sequences (except for the trivial ones) contain the monomials x_3 or x_3x_1 , which indicates the linear complexity*

Table 5: B-representation and r-ANF of generalized sequences

\mathcal{G}	Generalized sequence	B-representation	r-ANF
0001	0 1 0 1 0 1 0 1	(1)	x_1
0010	1 0 1 1 0 0 0 1	(0,1,3,4,5)	$1 + x_1 + x_2x_1 + x_3 + x_3x_1$
0011	1 1 1 0 0 1 0 0	(0,3,4,5)	$1 + x_2x_1 + x_3 + x_3x_1$
0100	0 1 1 1 0 0 1 0	(1,2,3,5)	$x_1 + x_2 + x_2x_1 + x_3x_1$
0101	0 0 1 0 0 1 1 1	(2,3,5)	$x_2 + x_2x_1 + x_3x_1$
0110	1 1 0 0 0 0 1 1	(0,2,4)	$1 + x_2 + x_3$
0111	1 0 0 1 0 1 1 0	(0,1,2,4)	$1 + x_1 + x_2 + x_3$
1000	1 1 1 1 1 1 1 1	(0)	1
1001	1 0 1 0 1 0 1 0	(0,1)	$1 + x_1$
1010	0 1 0 0 1 1 1 0	(1,3,4,5)	$x_1 + x_2x_1 + x_3 + x_3x_1$
1011	0 0 0 1 1 0 1 1	(3,4,5)	$x_2x_1 + x_3 + x_3x_1$
1100	1 0 0 0 1 1 0 1	(0,1,2,3,5)	$1 + x_1 + x_2 + x_2x_1 + x_3x_1$
1101	1 1 0 1 1 0 0 0	(0,2,3,5)	$1 + x_2 + x_2x_1 + x_3x_1$
1110	0 0 1 1 1 1 0 0	(2,4)	$x_2 + x_3$
1111	0 1 1 0 1 0 0 1	(1,2,4)	$x_1 + x_2 + x_3$

(6 and 5 in this case). Notice also that the r-ANF of the first half of the sequences is the same as the inferior half except for the term 1.

Next theorem provides some important properties of the r-ANF of the generalized sequences.

Theorem 17: *The r-ANF $\mathcal{B}(\{s_\tau\})$ of any generalized sequence $\{s_\tau\}$ obtained from a primitive polynomial of degree L satisfies the following properties:*

- $\mathcal{B}(\{s_\tau\})$ is composed of $|b(\{s_\tau\})|$ monomials, where $|b(\{s_\tau\})|$ denotes the support of the B-representation of $\{s_\tau\}$.
- The maximum sub-index j in the variables x_j of $\mathcal{B}(\{s_\tau\})$ is $L - 1$.
- $\mathcal{B}(\{s_\tau\})$ does not include the maximum term $x_{L-1} \cdots x_2x_1$.
- There exists another generalized sequence in the same family with r-ANF $\mathcal{B}(\{s_\tau\}) + 1$.

PROOF: a) This item follows immediately from the definition of r-ANF.

- It is well-known that the period of the generalized sequences is a divisor of 2^{L-1} [30]. The best case scenario is when the period is exactly 2^{L-1} . In this case the last term of the B-representation of the sequence is $\binom{n}{k}$ with $2^{L-2} \leq k \leq 2^{L-1} - 1$. The binary representation of k has length $L - 1$, therefore, the r-ANF have $L - 1$ variables: x_1, x_2, \dots, x_{L-1} .
- Except for the trivial sequence $\{11111111 \dots\}$, the generalized sequences are always balanced [30]. Therefore, according to Theorem 6, the generalized sequences cannot have maximum degree, i.e., they cannot include the maximum term $x_{L-1} \cdots x_2x_1$.
- According to [26, 27], for any generalized sequence $\{s_\tau\}$ with B-representation $b(\{s_\tau\})$, there exists another generalized sequence in the same family with B-representation $b(\{s_\tau\}) + \binom{n}{0}$. The results follows from the definition of r-ANF. \square

As a consequence of item b) of Theorem 17, we can deduce that the B-representation of a generalized sequence cannot contain the term $\binom{n}{2^{L-1}-1}$, and then the linear complexity satisfies $LC < 2^{L-1}$.

Table 6: r-ANF and B-representation of the shifted versions of a generalized sequence of period 8.

Shifted versions	B-representation	r-ANF
1 0 1 1 0 0 0 1	(0,1,3,4, 5)	$1 + x_1 + x_2x_1 + x_3 + x_3x_1$
0 1 1 0 0 0 1 1	(1,2,5)	$x_1 + x_2 + x_3x_1$
1 1 0 0 0 1 1 0	(0,2,4,5)	$1 + x_2 + x_3 + x_3x_1$
1 0 0 0 1 1 0 1	(0,1,2,3,5)	$1 + x_1 + x_2 + x_2x_1 + x_3x_1$
0 0 0 1 1 0 1 1	(3,4,5)	$x_2x_1 + x_3 + x_3x_1$
0 0 1 1 0 1 1 0	(2,5)	$x_2 + x_3x_1$
0 1 1 0 1 1 0 0	(1,2,4,5)	$x_1 + x_2 + x_3 + x_3x_1$
1 1 0 1 1 0 0 0	(0,2,3,5)	$1 + x_2 + x_2x_1 + x_3x_1$

Example 16: Consider, for example, the generalized sequence $\{s_\tau\}$ in Example 15 with r-ANF $\mathcal{B}(\{s_\tau\}) = 1 + x_1 + x_2x_1 + x_3 + x_3x_1$ and B-representation $b(\{s_\tau\}) = (0, 1, 3, 4, 5)$. Now, Note that the maximum sub-index in the r-ANF is $L - 1 = 3$ and it does not include the maximum term $x_3x_2x_1$. Also, the r-ANF is composed of 5 monomials, which coincides with the support of the B-representation. It is worth noticing that there exists another generalized sequence in the same family, with r-ANF $\mathcal{B}(\{s_\tau\}) + 1 = x_1 + x_2x_1 + x_3 + x_3x_1$ and B-representation $(1, 3, 4, 5)$.

From Theorem 17, we know that the number of monomials of the r-ANF of a binary sequence is the support of its B-representation. Next result gives us an upper bound of this value for generalized sequences, which is related with the linear complexity of them. In [38], authors proved that, for these sequences, $LC \leq 2^{L-1} - (L - 2)$. Note that a similar bound was found in [39, 40] for the self-shrinking generator.

Theorem 18: Let $\{s_\tau\}$ be a generalized sequence. The number of monomials \mathcal{N} of its r-ANF $\mathcal{B}(\{s_\tau\})$ satisfies

$$\mathcal{N} = |b(\{s_\tau\})| \leq 2^{L-1} - (L - 2).$$

PROOF: It is an immediate consequence from the definition B-representation and the bound of the LC for generalized sequences. The maximum value in the B-representation of a generalized sequence is the term $\binom{n}{i_k}$, as we saw in Section 2.4. Moreover, $LC = i_k + 1 \leq 2^{L-1} - (L - 2)$, that is, $i_k \leq 2^{L-1} - (L - 1)$. But, in the worst case, we could have in the B-representation of the sequence all the elements $\binom{n}{i}$, for $i = 0, \dots, 2^{L-1} - (L - 1)$. Therefore, $|b(\{s_\tau\})| \leq 2^{L-1} - (L - 2)$. \square

Theorem 19: Let $\{s_\tau\}$ be a generalized sequence of period 2^{L-1} generated by a primitive polynomial of degree L . Then, 2^{L-2} shifted versions of $\{s_\tau\}$ have the term $\binom{n}{0}$ in their B-representation, or, in other words, they have the monomial 1 in their r-ANF.

PROOF: If $\{s_\tau\}$ is balanced, then it will be composed of 2^{L-2} zeros and 2^{L-2} ones. This means that half of the shifted versions (2^{L-2} versions) start with 1. As a consequence of Theorem 7, the binomial representation of those shifted versions includes the term $\binom{n}{0}$, that is, the r-ANF includes the monomial 1. \square

Example 17: In Table 6 we can find the B-representation and the r-ANF of all shifted versions of $\{s_\tau\} = \{10110001\}$, one of the generalized sequences in Example 15. Note that the r-ANF of all the shifted versions have the same maximum monomial x_3x_1 , which corresponds to the term $\binom{n}{5}$ in the B-representation. It is worth noticing that

some shifted versions correspond to other generalized sequences in the same family (check Table 5). This usually happens when the degree of the polynomial is small. Note that half of the shifted versions have the term $\binom{n}{0}$ in the B-representation, as proved in Theorem 19.

7 Conclusions

In this paper, we define a bijection between the set of Boolean functions and the set of binary sequences of period a power of two. This connection allows us to analyse properties of the binary sequences through Boolean functions. Moreover, we define the reverse-ANF of a sequence, that is, a new representation of binary sequences based on Boolean functions. We introduce this new Boolean representation, instead of using the ANF, in order to our representation of a sequence, through Boolean functions, is unique. We show the relation between the different representations presented in terms of binary sequences and Boolean functions. We study the reverse-ANF in the family of generalized sequences, its relation to the Sierpinski triangle, and the representation of the reverse sequences. As future work, we would like to go in depth in the analysis of more important cryptographic properties of Boolean functions, as strict avalanche criterion, correlation immunity, non-linearity, among others, in terms of this r-ANF and how it can be interpreted in function of the binary sequences and other representations. Our aim is to use this connection to try to solve difficult problems in the field of binary sequences using Boolean functions, and vice versa. For instance, an open problem for the generalized sequences, obtained from a primitive polynomial of degree L , is to prove that their period is always 2^{L-1} , except for the trivial sequences with period 1 and 2; and that the linear complexity is lower bounded by 2^{L-2} .

Acknowledgements

This work is part of the R+D+i grant P2QProMeTe (PID2020-112586RB-I00) funded by MICIU/AEI/10.13039/501100011033 (Spain). The work of the first author was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil, project 405842/2023-6 and by Fapesp with process 2024/05051-7. The work of the third author was supported by the I+D+i project VIGROB-287 of the Universitat d'Alacant.

References

- [1] Crama, Y., Hammer, P.L.: Boolean functions: Theory, algorithms, and applications. In: Encyclopedia of Mathematics and Its Applications. Cambridge University Press, Cambridge (2011)

- [2] Krause, M., Wegener, I.: Circuit complexity. In: Crama, Y., Hammer, P.L. (eds.) *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, pp. 506–530. Cambridge University Press, Cambridge (2010)
- [3] Papadimitriou, C.H.: *Computational Complexity*. Addison Wesley Publishing Co., MA (1994)
- [4] MacWilliams, F.J., Sloane, N.J.A.: *The Theory of Error-Correcting Codes*, 6th edn. North-Holland, Amsterdam (1988)
- [5] Carlet, C.: *Boolean Functions for Cryptography and Coding Theory*. Cambridge University Press, Cambridge (2021)
- [6] Ragin, C.C.: *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. University of California Press, California (1987). <http://www.jstor.org/stable/10.1525/j.ctt1pnx57>
- [7] Feldman, J.: Minimization of boolean complexity in human concept learning. *Nature* **407**, 630–633 (2000) <https://doi.org/10.1038/35036586>
- [8] Bonates, T., Hammer, P.L.: Logical analysis of data: From combinatorial optimization to medical applications. *Annals of Operations Research* **148**, 203–225 (2006)
- [9] Shmulevich, I., Zhang, W.: Binary analysis and optimization-based normalization of gene expression data. *Bioinformatics* **18**(4), 555–565 (2002) <https://doi.org/10.1093/bioinformatics/18.4.555> <https://academic.oup.com/bioinformatics/article-pdf/18/4/555/669285/180555.pdf>
- [10] Braeken, A., Nikov, V., Nikova, S., Preneel, B.: On Boolean functions with generalized cryptographic properties. In: Canteaut, A., Viswanathan, K. (eds.) *Progress in Cryptology – INDOCRYPT 2004*. Lecture Notes in Computer Science, vol. 3348, pp. 120–135. Springer, Berlin (2004)
- [11] Meier, W., Staffelbach, O.: Nonlinearity criteria for cryptographic functions. In: Quisquater, J.J., Vandewalle, J. (eds.) *Advances in Cryptology – EUROCRYPT’89*. Lecture Notes in Computer Science, vol. 434, pp. 549–562. Springer, Berlin (1990)
- [12] Adams, C.M.: Constructing symmetric ciphers using the CAST design procedure. *Designs, Codes and Cryptography* **12**, 283–316 (1997) <https://doi.org/10.1023/A:1008229029587>
- [13] Gupta, K.C., Sarkar, P.: Improved construction of nonlinear resilient S-boxes. *IEEE Transactions on Information Theory* **51**(1), 339–348 (2005) <https://doi.org/10.1109/TIT.2004.839524>

- [14] Nyberg, K.: Perfect nonlinear S-boxes. In: Davies, D.W. (ed.) *Advances in Cryptology – EUROCRYPT’91*. Lecture Notes in Computer Science, vol. 547, pp. 378–386. Springer, Berlin (1991)
- [15] Al Shehhi, M.A., Baek, J., Yeun, C.Y.: The use of boolean functions in stream ciphers. In: *2011 International Conference for Internet Technology and Secured Transactions*, pp. 29–33 (2011)
- [16] Wei, Y., Hu, Y.: Maximum autocorrelation analysis of nonlinear combining functions in stream ciphers. In: *2007 IEEE International Symposium on Information Theory*, pp. 176–180 (2007). <https://doi.org/10.1109/ISIT.2007.4557083>
- [17] Zhang, M., Chan, A.: Maximum correlation analysis of nonlinear s-boxes in stream ciphers. In: Bellare, M. (ed.) *Advances in Cryptology — CRYPTO 2000*, pp. 501–514. Springer, Berlin, Heidelberg (2000)
- [18] Rueppel, R.A.: *Linear Complexity and Random Sequences*. Lecture Notes in Computer Science, vol. 219, pp. 167–188. Springer, - (1986)
- [19] Rueppel, R.A., Staffelbach, O.J.: Products of linear recurring sequences with maximum complexity. *IEEE Transactions on Information Theory* **33**(1), 124–131 (1987) <https://doi.org/10.1109/TIT.1987.1057268>
- [20] Fúster-Sabater, A.: Computation of filtering functions for cryptographic applications. *Procedia Computer Science* **29**, 2013–2023 (2014) <https://doi.org/10.1016/j.procs.2014.05.185>
- [21] Golomb, S.W.: *Shift Register-Sequences*. Aegean Park Press, Laguna Hill, California (1982)
- [22] Chowdhury, S., Maitra, S.: Efficient software implementation of lfsr and boolean function and its application in nonlinear combiner model. In: Zhou, J., Yung, M., Han, Y. (eds.) *Applied Cryptography and Network Security*, pp. 387–402. Springer, Berlin, Heidelberg (2003)
- [23] Sadkhan, S.B., Reza, D.M.: Investigation of the best structure for the nonlinear combining function. In: *2017 Annual Conference on New Trends in Information & Communications Technology Applications (NTICT)*, pp. 180–185 (2017). <https://doi.org/10.1109/NTICT.2017.7976126>
- [24] Siegenthaler, T.: Correlation-immunity of nonlinear combining functions for cryptographic applications. *IEEE Transactions on Information Theory* **30**(5), 776–780 (1984) <https://doi.org/10.1109/TIT.1984.1056949>
- [25] Wagner, U.: Detection and exploitation of small correlations in stream ciphers. Diploma thesis, Institut für Mathematics, Universität Zürich, Zürich, Switzerland (2008)

- [26] Cardell, S.D., Fúster-Sabater, A.: Binomial representation of cryptographic binary sequences and its relation to cellular automata. *Complexity* **2019**, 1–13 (2019) <https://doi.org/10.1155/2019/2108014>
- [27] Cardell, S.D., Climent, J.-J., Fúster-Sabater, A., Requena, V.: Representations of generalized self-shrunk sequences. *Mathematics* (8), 1–26 (2020) <https://doi.org/10.3390/math8061006>
- [28] Pirsic, G., Winterhof, A.: Boolean functions derived from pseudorandom binary sequences. In: Helleseth, T., Jedwab, J. (eds.) *Sequences and Their Applications – SETA 2012*, pp. 101–109. Springer, Berlin, Heidelberg (2012)
- [29] Cardell, S.D., Fúster-Sabater, A., Orúe, A.B., Requena, V.: Randomness study of the concatenation of generalized sequences. *Logic Journal of the IGPL* **30**(6), 993–1004 (2022) <https://doi.org/10.1093/jigpal/jzac008>
- [30] Hu, Y., Xiao, G.: Generalized self-shrinking generator. *IEEE Transactions on Information Theory* **50**(4), 714–719 (2004) <https://doi.org/10.1109/TIT.2004.825256>
- [31] Olejár, D., Stanek, M.: On cryptographic properties of random boolean functions. *Journal of Universal Computer Science* **4**(8), 705–717 (1998)
- [32] Pasalic, E., Johansson, T.: Further results on the relation between nonlinearity and resiliency for boolean functions. *Lecture Notes in Computer Science* **1746**, 35–44 (1999)
- [33] Rota, G.C.: On the foundations of combinatorial theory i. theory of möbius functions. *Probability Theory and Related Fields* **2**(4), 340–368 (1964)
- [34] Pieprzyk, J., Wang, H., Zhang, X.M.: Möbius transforms, coincident boolean functions and non-coincidence property of boolean functions. *International Journal of Computer Mathematics* **88**(7), 1398–1416 (2011)
- [35] Fúster-Sabater, A., Requena, V., Cardell, S.D.: An efficient algorithm to compute the linear complexity of binary sequences. *Mathematics* **10**(5), 794–816 (2022) <https://doi.org/10.3390/math10050794>
- [36] Fúster-Sabater, A., Requena, V., Cardell, S.D.: A Hadamard Matrix-Based Algorithm to Evaluate the Strength of Binary Sequences. In: Groen, D., Mulatier, C., Paszynski, M., Krzhizhanovskaya, V.V., Dongarra, J.J., Sloot, P.M.A. (eds.) *Computational Science – ICCS 2022*, pp. 139–145. Springer, - (2022). https://doi.org/10.1007/978-3-031-08754-7_19
- [37] Climent, J.-J., García, F.J., Requena, V.: Computing the degree of a boolean function from its support. In: *ISITA/ISSSTA 2010 - 2010 International Symposium on Information Theory and Its Applications*, pp. 123–128 (2010). https://doi.org/10.1007/978-3-031-08754-7_19

[//doi.org/10.1109/ISITA.2010.5649426](https://doi.org/10.1109/ISITA.2010.5649426)

- [38] Fúster-Sabater, A., Cardell, S.D.: Linear complexity of generalized sequences by comparison of pn-sequences. *RACSAM* **114**(79) (2020) <https://doi.org/10.1007/s13398-020-00807-5>
- [39] Blackburn, S.R.: The linear complexity of the self-shrinking generator. *IEEE Transactions on Information Theory* **45**(6), 2073–2077 (1999) <https://doi.org/10.1109/18.782139>
- [40] Blackburn, S.R., Etzion, T., Paterson, K.G.: Permutation polynomials, de bruijn sequences and linear complexity. *Journal of Combinatorial Theory, Series A* **76**, 55–82 (1996)