New Results on the Density of Irreducible NFSRs

Daoyuan Zhang and Dongdai Lin

Abstract

Nonlinear feedback shift registers (NFSRs) are fundamental building blocks for modern stream-cipher constructions. An n-stage NFSR f is classified as irreducible when the output sequence set of f does not contain the output family of any NFSR of order less than n. Existing research has established upper and lower bounds for the density of irreducible NFSRs, confining this value within the range of 0.4461 to 0.4834. This study tightens these bounding intervals with high accuracy, reducing the original 0.04 gap down to only 8×10^{-6}.

Index Terms

Nonlinear feedback shift register (NFSR), irreducible NFSR, density estimation, Boolean Nullstellensatz, Gröbner basis, stream ciphers.

I. Introduction

FEEDBACK shift registers serve as fundamental components in the design of stream cipher systems. For a long time, linear feedback shift registers (LFSRs) have been the most popular building block because they are well studied, offer efficient implementations, and their output sequences can be guaranteed to have excellent statistical properties. However, LFSR-based stream ciphers are susceptible to cryptanalytic attacks, such as algebraic attacks and correlation attacks. Consequently, many recently proposed stream ciphers adopt nonlinear feedback shift registers (NFSRs) as an important type of nonlinear sequence generator. For example, Grain and Trivium, two eSTREAM hardware-oriented finalists, both use NFSRs as their main building blocks [1], [2].

After almost five decades of intensive study, the basic algebraic framework and theoretical foundations of Nonlinear Feedback Shift Registers (NFSRs) are still largely obscure. A central difficulty is ascertaining the cycle structure of the sequences they generate. In the linear case, an LFSR is fully described by a single characteristic polynomial $f \in \mathbb{F}_2[x]$, and the cycle structure of its output is immediately obtained from the factorization of f [3], [4]. By contrast, for generic NFSRs the cycle structure appears computationally intractable. Only restricted classes like symmetric NFSRs [5]–[8] have been adequately analyzed.

A particular aspect of cycle structure concerns sub-families. Given two NFSRs, say F_1 and F_2, if every output sequence producible by F_2 is also among the outputs of F_1, then F_1 is termed reducible and F_2 is referred to as a sub-family of F_1. Conversely, if no such smaller NFSR exists whose entire output set is contained within that of F_1, then F_1 is declared irreducible. Given F_1, F_2, it is known that the subfamily relation could be tested by a generalized division algorithm proposed in [9].

For a given NFSR f, finding NFSR sub-families of f is a challenging research problem. Although the affine case has been solved [10], no feasible algorithm is known for nonlinear instances, so the irreducibility of an arbitrary large NFSR remains infeasible to decide. Consequently, researchers quantify the likelihood that a random n-stage NFSR is irreducible by studying the density of irreducible NFSRs among all n-stage registers, denoted $\mu(n)$. In 2013, Tian and Qi [11] proved that the density of irreducible NFSRs is larger than 0.39. Later, Jiang and Lin [12] improved the lower bound and provided an upper bound, proving that $0.4461 < \mu(n) < 0.4834$.

In this paper, we determine the density of irreducible NFSRs by exploiting a geometric description of NFSR sub-families. We design the first algorithm for this density-estimation task and refine the bounds to

$$0.4701455 < \mu(n) < 0.4701535, \quad \text{for } n \geq 12.$$

with additional explicit bounds for $6 \leq n \leq 11$, see Table III. We note that for $n \leq 5$, n is so small that the density of irreducible NFSRs can be explicitly computed.

The rest of the paper is structured as follows. Section II provides a concise review of Boolean functions and nonlinear feedback shift registers. Section III presents the basic estimation framework. Section IV details the theory developed for density analysis. Section V details the algorithm used to derive tight lower and upper bounds on the density of irreducible NFSRs. The Conclusion is presented in Section VI.

Daoyuan Zhang and Dongdai Lin are with the Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China. E-mail: zhangdaoyuan@iie.ac.cn; ddlin@iie.ac.cn.

II. Preliminaries

In this section, we briefly review basic facts about Boolean functions and nonlinear feedback shift registers. We use the operations "$+$" and "$-$" to denote ordinary integer addition and subtraction, and "$\oplus$" to denote addition modulo 2. $\vee$ and $\wedge$ are the logical disjunction and conjunction, respectively. For a finite set X, the cardinality of X is denoted by $|X|$. The finite field with two elements is denoted by $\mathbb{F}_2$, the n-dimensional vector space over $\mathbb{F}_2$ is denoted by $\mathbb{F}_2^n$, and the set of positive integers is denoted by $\mathbb{N}^*$.

A. Boolean Functions

Let $n \in \mathbb{N}^*$. An n-variable Boolean function $f(x_0, x_1, \ldots, x_{n-1})$ is a mapping from $\mathbb{F}_2^n$ to $\mathbb{F}_2$. The set of all n-variable Boolean functions is denoted by $\mathcal{B}(x_0, \ldots, x_{n-1})$, or simply $\mathcal{B}_n$. An n-variable Boolean function can be uniquely expressed as a multivariate polynomial of the form:

$$f(x_0, \ldots, x_{n-1}) = \bigoplus_{a=(a_0,\ldots,a_{n-1})\in\{0,1\}^n} c_a \prod_{j=0}^{n-1} x_j^{a_j},$$

where $c_a \in \mathbb{F}_2$. This unique representation is called the algebraic normal form (ANF) of f.

The algebraic degree of f, denoted by $\deg(f)$, is the global degree of the ANF of f. If $\deg(f) \leq 1$, then f is called affine. If the constant term of an affine Boolean function is 0, then f is called linear. The *order* of a nonconstant Boolean function f, denoted by $\mathrm{ord}(f)$, is the highest subscript i for which x_i occurs in the ANF of f. Since the lowest subscript is always 0, a Boolean function with order n has $n+1$ variables. A product of the form $x_0^{a_0} x_1^{a_1} \cdots x_{n-1}^{a_{n-1}}$ is called a term.

We use the inverse lexicographic (invlex) order on terms, denoted by $\prec$. For $\alpha, \beta \in \mathbb{Z}_{\geq 0}^n$, define $\alpha \prec \beta$ if the rightmost nonzero entry of $\alpha - \beta$ is negative. This induces an order on monomials $x^\alpha = \prod_{i=0}^{n-1} x_i^{\alpha_i}$ by declaring $x^\alpha \prec x^\beta$ if and only if $\alpha \prec \beta$. For a nonconstant Boolean polynomial $f \in \mathfrak{B}(x_0, \ldots, x_{n-1})$, let $\mathrm{HT}(f)$ and $\mathrm{LT}(f)$ be the largest and smallest terms of f with respect to $\prec$.

B. Nonlinear Feedback Shift Registers

Fix a feedback function $f_1(x_0, x_1, \ldots, x_{n-1}) \in \mathcal{B}_n$. The corresponding n-stage NFSR is described by the characteristic function

$$f(x_0, x_1, \ldots, x_n) = f_1(x_0, x_1, \ldots, x_{n-1}) \oplus x_n \in \mathcal{B}_{n+1}.$$

A binary sequence $\underline{s} = (s_t)_{t\geq 0}$ is produced by this register if and only if it satisfies

$$s_{i+n} = f_1(s_i, s_{i+1}, \ldots, s_{i+n-1}), \quad \text{for all } i \geq 0.$$

Given an initial state $(s_0, s_1, \ldots, s_{n-1})$, the output sequence is uniquely determined by this recursion. Let $S(f)$ denote the set of all 2^n sequences obtained by running the NFSR over all initial states.

Since the internal state space has size 2^n, the state evolution starting from any initial state must eventually repeat, and therefore every output sequence is ultimately periodic. Moreover, all sequences in $S(f)$ are purely periodic if and only if the characteristic function $f(x_0, x_1, \ldots, x_n)$ is *nonsingular* [13], i.e.,

$$f(x_0, x_1, \ldots, x_n) = x_0 \oplus f_2(x_1, \ldots, x_{n-1}) \oplus x_n.$$

Throughout the paper, we restrict to nonsingular characteristic functions. For $n \geq 1$, let $\Omega(n)$ denote the set of all nonsingular characteristic functions of n-stage NFSRs, then $|\Omega(n)| = 2^{2^{n-1}}$.

If f is a linear characteristic function, i.e.,

$$f(x_0, \ldots, x_n) = x_0 \oplus c_1 x_1 \oplus \cdots \oplus c_{n-1} x_{n-1} \oplus x_n,$$

then the corresponding register is an LFSR. For two LFSRs with characteristic functions f and g, the inclusion $S(f) \subseteq S(g)$ holds if and only if the associated polynomials satisfy $\phi(f) \mid \phi(g)$ in $\mathbb{F}_2[x]$.

For nonlinear characteristic functions, Mykkeltveit et al. [9] extended this divisibility viewpoint via a generalized division algorithm on Boolean functions, which can be stated as follows.

Lemma 1 (Generalized division)**.** *Let $m, n \in \mathbb{N}^*$ with $m \leq n$. Let $f(x_0, \ldots, x_m)$ and $g(x_0, \ldots, x_n)$ be Boolean functions, and assume*

$$f(x_0, \ldots, x_m) = f_1(x_0, \ldots, x_{m-1}) \oplus x_m.$$

Then there exist unique Boolean functions $h_i(x_0, \ldots, x_{m+i-1})$ for $0 \leq i \leq n-m$ and a unique Boolean function $r(x_0, \ldots, x_{m-1})$ such that

$$g(x_0, \ldots, x_n) = \bigoplus_{i=0}^{n-m} h_i(x_0, \ldots, x_{m+i-1})\, f(x_i, \ldots, x_{m+i}) \oplus r(x_0, \ldots, x_{m-1}).$$

If the remainder in Lemma 1 is zero, we write $f \parallel g$ and say that g is divisible by f under the generalized division algorithm. Mykkeltveit et al. [9] proved that $f \parallel g$ holds if and only if $S(f) \subseteq S(g)$. We will prove this result in a more natural way later in Corollary 1.

Lemma 1 provides an effective way to count n-stage NFSRs that are reducible by a fixed m-stage NFSR with $m \leq n$. However, the analysis becomes substantially more intricate when reducibility with respect to multiple smaller-stage NFSRs must be taken into account.

III. Estimation Framework

In this section, we generalize the estimation framework proposed in [12]. The key idea is to use information from a limited collection of small irreducible NFSRs to bound the density of irreducible NFSRs in larger stages. We begin by introducing the basic notation needed for this framework.

Definition 1 (Irreducible NFSRs)**.** *An NFSR with characteristic function $f \in \Omega(n)$ is said to be irreducible if there does not exist an NFSR $g \in \Omega(m)$, with $m < n$, such that $S(g) \subsetneq S(f)$. If such a g exists, then the NFSR is called reducible.*

Let $\mathcal{N}_0(n)$ be the set of irreducible elements of $\Omega(n)$, and write $N_0(n) = |\mathcal{N}_0(n)|$. Similarly, let $\mathcal{N}_1(n)$ be the set of reducible elements of $\Omega(n)$, and write $N_1(n) = |\mathcal{N}_1(n)|$. Clearly,

$$N_0(n) + N_1(n) = |\Omega(n)| = 2^{2^{n-1}}.$$

Proposition 1. *An NFSR $f \in \Omega(n)$ is either irreducible or can be reduced by an irreducible NFSR.*

Proof. We prove this by induction on the order of f. For $\mathrm{ord}(f) = 1$, f can only be irreducible by definition. Suppose the statement is true for all NFSRs of order less than n. For an NFSR f of order n, if f is reducible, then there is an NFSR $g \in \Omega(m)$ with $m < n$ that $S(g) \subsetneq S(f)$. If g is irreducible, the statement is true. Otherwise by induction, there is an irreducible NFSR $g_1 \in \Omega(m_1)$ with $m_1 < m < n$ that $S(g_1) \subsetneq S(g) \subsetneq S(f)$. □

Definition 2 (Density of Irreducible NFSRs)**.** *The density of irreducible n-stage NFSRs is defined by*

$$\mu(n) = \frac{N_0(n)}{|\Omega(n)|} = 1 - \frac{N_1(n)}{2^{2^{n-1}}}. \tag{1}$$

Now we introduce the estimation framework. For NFSRs $f_1, f_2, \ldots, f_t$, define

$$\mathcal{P}_n(f_1, \ldots, f_t) = \{g \in \Omega(n) \mid S(f_i) \subseteq S(g) \text{ for all } i\},$$

and denote the cardinality of this set by $P_n(f_1, \ldots, f_t)$. By definition, an n-stage NFSR is reducible precisely when it lies in $\mathcal{P}_n(f)$ for some f of smaller order, hence

$$\mathcal{N}_1(n) = \bigcup_{m=1}^{n-1} \bigcup_{f \in \Omega(m)} \mathcal{P}_n(f).$$

Moreover, every reducible NFSR is divisible by an irreducible one, so we may restrict the union to irreducible f and obtain

$$\mathcal{N}_1(n) = \bigcup_{m=1}^{n-1} \bigcup_{f \in \mathcal{N}_0(m)} \mathcal{P}_n(f). \tag{2}$$

A. Upper Bound Estimation

To get an upper bound for $\mu(n) = 1 - N_1(n)/2^{2^{n-1}}$, we need to estimate a lower bound for $N_1(n)$. By Equation (2), for some $k \leq n-1$,

$$\mathcal{N}_1(n) \supseteq \bigcup_{m=1}^{k} \bigcup_{f \in \mathcal{N}_0(m)} \mathcal{P}_n(f). \tag{3}$$

Assume there are n_k irreducible NFSRs of order at most k, denoted $f_1, \ldots, f_{n_k}$. Then $\mathcal{N}_1(n) \supseteq \bigcup_{i=1}^{n_k} \mathcal{P}_n(f_i)$.

We denote $\eta_{k,n}$ as the cardinality of the union $\bigcup_{i=1}^{n_k} \mathcal{P}_n(f_i)$. By the inclusion-exclusion principle,

$$\begin{aligned} \eta_{k,n} &= \Big|\bigcup_{i=1}^{n_k} \mathcal{P}_n(f_i)\Big| \\ &= \sum_{j=1}^{n_k} (-1)^{j+1} \sum_{1 \leq i_1 < \cdots < i_j \leq n_k} P_n(f_{i_1}, \ldots, f_{i_j}). \end{aligned} \tag{4}$$

Then,

$$\mu(n) < 1 - \frac{\eta_{k,n}}{2^{2^{n-1}}}. \tag{5}$$

B. Lower Bound Estimation

To get an lower bound for $\mu(n)$, we need to estimate a upper bound for $N_1(n)$. By Equation (2), we know

$$N_1(n) \leq \eta_{k,n} + \sum_{m=k+1}^{n-1} \sum_{f \in \mathcal{N}_0(m)} P_n(f). \tag{6}$$

Now we handle the later part of summation. For $f \in \mathcal{N}_0(m)$, we have $P_n(f) = |\Omega(n)|/2^{2^m}$, as stated formally in Lemma 2.

Lemma 2. *[lemma 4 in [11]] Let $m, n \in \mathbb{N}^*$ with $m < n$ and let $f \in \Omega(m)$. Then*

$$P_n(f) = |\mathcal{P}_n(f)| = \frac{|\Omega(n)|}{2^{2^m}} = 2^{2^{n-1}-2^m}. \tag{7}$$

As a result, we can write

$$N_1(n) < \eta_{k,n} + |\Omega(n)| \sum_{m=k+1}^{n-1} \frac{N_0(m)}{2^{2^m}}. \tag{8}$$

Therefore,

$$\mu(n) > 1 - \frac{\eta_{k,n}}{2^{2^{n-1}}} - \sum_{m=k+1}^{n-1} \frac{N_0(m)}{2^{2^m}}. \tag{9}$$

Thus, we can give an lower bound of $\mu(n)$ by providing upper bounds of each $N_0(m)$ for $m = k+1, \ldots, n-1$. For $m \leq 5$, the values of $N_0(m)$ can be computed exactly via exhaustive search; the corresponding densities $\mu(m)$ are summarized in Table I.

TABLE I
VALUES OF $N_0(m)$ AND DENSITY $\mu(m)$ OF IRREDUCIBLE NFSRS FOR SMALL STAGES

m	1	2	3	4	5
$N_0(m)$	2	3	9	127	31233
$\mu(m)$	1	$\frac{3}{4}$	$\frac{9}{16} \approx 0.56$	$\frac{127}{256} \approx 0.50$	$\frac{31233}{65536} \approx 0.48$

For $m > 5$, we obtain upper bounds on $N_0(m)$ from the lower bounds on $N_1(m)$ derived in the previous subsection, using the identity $N_0(m) = |\Omega(m)| - N_1(m)$.

The central challenge of the upper and lower bound estimation is to evaluate $\eta_{k,n}$, which ultimately requires a systematic method for computing $P_n(f_{i_1}, \ldots, f_{i_j})$.

Jiang *et al.* [12] took $k = 2$, so that $n_k = 5$ and the inclusion–exclusion expansion involves $2^5 = 32$ term. These terms were computed by hand using the division algorithm. This approach does not scale to larger k. In the next section, we introduce an automated method to compute $P_n(f_{i_1}, \ldots, f_{i_j})$ on a computer, which allows us to increase the parameter to $k = 4$ and work with $n_k = 141$ irreducible NFSRs.

IV. THEORY

A. Boolean Nullstellensatz

Theorem 1. *There exists a one-to-one correspondence among subsets of $\mathbb{F}_{2^n}$, Boolean functions in n variables, and ideals of the Boolean polynomial ring $\mathfrak{B}(x_0, \ldots, x_{n-1})$.*

Proof. Define a map from Boolean functions in n variables to subsets of $\mathbb{F}_{2^n}$ by

$$f \longmapsto V_n(f) = \{x \in \mathbb{F}_{2^n} \mid f(x) = 0\}.$$

This map is injective: if $V_n(f) = V_n(g)$, then $f(x) = 0$ if and only if $g(x) = 0$ for every x, and since f, g take values in $\{0, 1\}$ we obtain $f = g$. It is also surjective: given any subset $S \subseteq \mathbb{F}_{2^n}$, the Boolean function f_S defined by $f_S(x) = 0$ for $x \in S$ and $f_S(x) = 1$ for $x \notin S$ satisfies $V_n(f_S) = S$.

Next, consider the map from Boolean functions to ideals of $\mathfrak{B}(x_0, \ldots, x_{n-1})$ given by $f \mapsto I(f)$. It is injective: if $I(f_1) = I(f_2)$, then $f_1 = f_2 h$ and $f_2 = f_1 g$ for some Boolean functions g, h, so $f_2 f_1 = f_2^2 h = f_2 h = f_1$ and similarly $f_1 f_2 = f_2$, which implies $f_1 = f_2$. It is surjective because every ideal in a Boolean ring is principal. Indeed, if $J = I(f_1, \ldots, f_k)$ and

$$u = \bigvee_{i=1}^{k} f_i = 1 + \prod_{i=1}^{k} (1 + f_i),$$

then expanding the product shows that u is a Boolean sum of terms, each divisible by some f_i, hence $u \in J$. Moreover, for each i we have $f_i u = f_i$, so $f_i \in I(u)$ and therefore $J \subseteq I(u)$. Since also $u \in J$ implies $I(u) \subseteq J$, we conclude that $J = I(u)$. □

To further exploit the relation between subsets of $\mathbb{F}_{2^n}$ and Boolean functions in n variables, we have the following proposition:

Proposition 2. *For subsets $S_1, S_2 \subseteq \mathbb{F}_{2^n}$ with corresponding Boolean functions f_{S_1} and f_{S_2}:*

$$S_1 \subseteq S_2 \iff f_{S_1} \mid f_{S_2} \iff I(f_{S_1}) \supseteq I(f_{S_2}).$$

Proof. For the second equivalence, note that $f_{S_1} \mid f_{S_2}$ holds if and only if $f_{S_2} = f_{S_1} h$ for some Boolean function h, which is equivalent to $f_{S_2} \in I(f_{S_1})$, and hence to $I(f_{S_2}) \subseteq I(f_{S_1})$.

For the first equivalence, recall that for any subset $S \subseteq \mathbb{F}_{2^n}$, the associated Boolean function satisfies $f_S(x) = 0$ if and only if $x \in S$. Moreover, let any element of S be denoted by $a = (a_0, \ldots, a_{n-1})$. Then f_S admits the conjunctive normal form

$$f_S(x) = \bigwedge_{a \in S} \bigvee_{i=0}^{n-1} (x_i \oplus a_i).$$

If $S_1 \subseteq S_2$, then the above representation implies

$$f_{S_2}(x) = \Big(\bigwedge_{a \in S_1} \bigvee_{i=0}^{n-1} (x_i \oplus a_i) \Big) \wedge \Big(\bigwedge_{a \in S_2 \setminus S_1} \bigvee_{i=0}^{n-1} (x_i \oplus a_i) \Big),$$

so $f_{S_2} = f_{S_1} h$ for some Boolean function h, i.e., $f_{S_1} \mid f_{S_2}$. Conversely, if $f_{S_1} \mid f_{S_2}$, then for every $x \in S_1$ we have $f_{S_1}(x) = 0$, hence $f_{S_2}(x) = 0$, which implies $x \in S_2$. Therefore $S_1 \subseteq S_2$.

□

B. NFSR Subfamily Characterization

In this subsection, we use the Boolean Nullstellensatz to characterize subfamily inclusion for NFSRs. For a set X of binary sequences, define

$$\mathcal{V}_N(X) = \{(a_0, \ldots, a_{N-1}) \mid a = (a_0, a_1, \ldots) \in X\} \subseteq \mathbb{F}_2^N.$$

For an n-variable boolean function, define $V_n(f)$ as the zero set of f. It is easy to see that for any characteristic function f of order n, $V_{n+1}(f) = \mathcal{V}_{n+1}(S(f))$.

Theorem 2. *Let $g = x_0 + g_1(x_1, \ldots, x_{n-1}) + x_n$ and $f = x_0 + f_1(x_1, \ldots, x_{m-1}) + x_m$ be characteristic functions of two NFSRs. Then $S(f) \subseteq S(g)$ if and only if $\mathcal{V}_{n+1}(S(f)) \subseteq \mathcal{V}_{n+1}(S(g))$.*

Proof. The forward implication is immediate from the definition of $\mathcal{V}_{n+1}(\cdot)$. For the reverse implication, assume $\mathcal{V}_{n+1}(S(f)) \subseteq \mathcal{V}_{n+1}(S(g))$ and let $a = (a_0, a_1, \ldots) \in S(f)$. Since $S(f)$ is shift-invariant, for each $k \geq 0$ the block $(a_k, \ldots, a_{k+n})$ belongs to $\mathcal{V}_{n+1}(S(f))$, hence also to $\mathcal{V}_{n+1}(S(g))$. Therefore there exists $b \in S(g)$ whose first $n+1$ bits equal $(a_k, \ldots, a_{k+n})$, and in particular $g(a_k, \ldots, a_{k+n}) = g(b_0, \ldots, b_n) = 0$. As this holds for all $k \geq 0$, we have $a \in S(g)$, and thus $S(f) \subseteq S(g)$. □

Remark 1. *The equivalence in Theorem 2 extends directly to arbitrary shift-invariant sets: for any shift-invariant set X of binary sequences, one has $X \subseteq S(g)$ if and only if $\mathcal{V}_{n+1}(X) \subseteq \mathcal{V}_{n+1}(S(g))$.*

For $i \geq 0$, define the shift operator δ^i on Boolean functions by

$$\delta^i(f) = f(x_i, x_{i+1}, \ldots, x_{m+i}).$$

Then $\mathcal{V}_{n+1}(S(f))$ is exactly the common zero set of $f, \delta(f), \ldots, \delta^{n-m}(f)$ in $\mathbb{F}_2^{n+1}$, while $\mathcal{V}_{n+1}(S(g))$ is the zero set of g. By the Boolean Nullstellensatz, the inclusion $\mathcal{V}_{n+1}(S(f)) \subseteq \mathcal{V}_{n+1}(S(g))$ is equivalent to

$$I(g) \subseteq I(f, \delta(f), \ldots, \delta^{n-m}(f)),$$

or equivalently, to $g \in I(f, \delta(f), \ldots, \delta^{n-m}(f))$.

Corollary 1. *Let $g = x_0 + g_1(x_1, \ldots, x_{n-1}) + x_n$ and $f = x_0 + f_1(x_1, \ldots, x_{m-1}) + x_m$ be characteristic functions. Then $(f, \delta(f), \ldots, \delta^{n-m}(f))$ is a Gröbner basis under inverse lexicographic order, and $S(f) \subseteq S(g)$ if and only if g reduces to 0 with respect to this Gröbner basis.*

Proof. Under inverse lexicographic order, the leading term of $\delta^i(f)$ is x_{m+i}, so these leading terms are pairwise coprime. Hence $(f, \delta(f), \ldots, \delta^{n-m}(f))$ is a Gröbner basis by Buchberger's gcd criterion [14]. By the preceding discussion, $S(f) \subseteq S(g)$ holds if and only if $g \in I(f, \delta(f), \ldots, \delta^{n-m}(f))$. Therefore this holds if and only if the remainder of g upon division by this Gröbner basis is zero, i.e., g reduces to 0. □

The process of performing this reduction is the generalized division presented in Lemma 1.

C. Regular Sets and Capacity

The Boolean Nullstellensatz identifies n-variable Boolean functions with subsets of $\mathbb{F}_2^n$ via their zero sets. Thus, any restricted family of Boolean functions corresponds to a restricted family of subsets.

Here we restrict to $\Omega(n)$, the set of nonsingular characteristic functions of stage-n NFSRs, i.e., Boolean functions of the form

$$f = x_0 + f_1(x_1, \ldots, x_{n-1}) + x_n.$$

Since f is determined by its inner function f_1, its zero set in $\mathbb{F}_2^{n+1}$ is determined by its "inner set", as stated formally in Proposition 3.

Proposition 3. *Let $f = x_0 + f_1(x_1, \ldots, x_{n-1}) + x_n$. Then $V_{n+1}(f)$ is constructed from $\mathbb{F}_{2^{n-1}}$ by appending (x_0, x_n) to each element in $\mathbb{F}_{2^{n-1}}$ under the restriction that $x_0 + x_n = f_1(x_1, \ldots, x_{n-1})$. In particular, for each $(x_1, \ldots, x_{n-1}) \in \mathbb{F}_2^{n-1}$ there are exactly two points of $V_{n+1}(f)$ with these middle coordinates.*

We now consider the following counting problem: How many nonsingular characteristic functions of order n satisfy that its zero set contain a prescribed subset $S \subseteq \mathbb{F}_2^{n+1}$? As a warm-up, we first treat the unrestricted case of an arbitrary Boolean function.

Lemma 3. *Let $\mathcal{B}_n$ denote the set of all n-variable Boolean functions. For any $S \subseteq \mathbb{F}_2^n$, the number of $f \in \mathcal{B}_n$ such that $S \subseteq V_n(f)$ is $2^{2^n - |S|}$.*

Proof. By the Boolean Nullstellensatz, each $f \in \mathcal{B}_n$ corresponds bijectively to its zero set $T = V_n(f) \subseteq \mathbb{F}_2^n$. The condition $S \subseteq V_n(f)$ is equivalent to $S \subseteq T$. The number of subsets T containing S is $2^{2^n - |S|}$, so the number of such f is equal to $2^{2^n - |S|}$. □

To proceed to the restricted case of nonsingular characteristic functions, we need the following definition:

Definition 3. *A subset S of $\mathbb{F}_2^{n+1}$ is called* regular *if there do not exist two elements $x, y \in S$ such that $x_0 + x_n \neq y_0 + y_n$, but $x_i = y_i$ for all $0 < i < n$.*

Definition 4. *Suppose S is an regular subset of $\mathbb{F}_2^{n+1}$. The* capacity *of S, denoted by $\operatorname{cap}(S)$, is the number of distinct middle $(n-1)$-bit segments in S.*

Theorem 3. *Let $\Omega(n)$ be the set of nonsingular characteristic functions of stage n. For any $S \subseteq \mathbb{F}_2^{n+1}$, if S is regular, then the number of $f \in \Omega(n)$ such that $S \subseteq V_{n+1}(f)$ is $|\Omega(n)| \cdot 2^{-\operatorname{cap}(S)}$. If S is not regular, then the count is 0.*

Proof. Write $x = (x_0, \ldots, x_n) \in \mathbb{F}_2^{n+1}$ and denote its middle block by $\bar{x} = (x_1, \ldots, x_{n-1}) \in \mathbb{F}_2^{n-1}$. Let

$$M(S) = \{\bar{x} \mid x \in S\},$$

so $|M(S)| = \operatorname{cap}(S)$. For $\bar{u} \in M(S)$, regularity of S is equivalent to the requirement that $x_0 + x_n$ takes a unique value among all $x \in S$ with $\bar{x} = \bar{u}$; denote this value by $\varepsilon(\bar{u}) \in \mathbb{F}_2$. If S is not regular then some $\bar{u}$ yields two contradictory values of $x_0 + x_n$, so no f of the form $x_0 + f_1 + x_n$ can vanish on all of S, and the count is 0.

Assume S is regular. Then $S \subseteq V_{n+1}(f)$ is equivalent to the constraints

$$f_1(\bar{u}) = \varepsilon(\bar{u}) \quad \text{for all } \bar{u} \in M(S).$$

Let $A = \{\bar{u} \in M(S) \mid \varepsilon(\bar{u}) = 0\}$ and $B = \{\bar{u} \in M(S) \mid \varepsilon(\bar{u}) = 1\}$. By the Boolean Nullstellensatz, choosing f_1 is equivalent to choosing its zero set $T = V_{n-1}(f_1) \subseteq \mathbb{F}_2^{n-1}$, and the above constraints become $A \subseteq T$ and $T \cap B = \varnothing$. Thus T can be any subset of $\mathbb{F}_2^{n-1} \setminus B$ containing A, so there are $2^{2^{n-1} - |A| - |B|} = 2^{2^{n-1} - \operatorname{cap}(S)}$ choices of T, hence of f_1. Since $|\Omega(n)| = 2^{2^{n-1}}$, this count is equal to $|\Omega(n)| \cdot 2^{-\operatorname{cap}(S)}$. □

D. Counting methods

Section III introduced our density-estimation framework. The computational core is the exact evaluation of $\eta_{k,n}$. Equation (4) shows that this hinges on computing $P_n(f_{i_1}, \ldots, f_{i_j})$ for characteristic functions $f_{i_1}, \ldots, f_{i_j}$ whose stages are strictly smaller than n.

Subsection IV.B presented two complementary formulations of this problem: Theorem 2 (the geometric approach) and Corollary 1 (the algebraic approach). The work [12] relies on the algebraic approach, i,e, the generalized division algorithm. Here we instead develop the geometric viewpoint. In particular, Theorem 2 extends directly to multiple functions:

Proposition 4. *Let $f_{i_1}, \ldots, f_{i_j}$ be characteristic functions. Then $g \in \mathcal{P}_n(f_{i_1}, \ldots, f_{i_j})$ if and only if*

$$\mathcal{V}_{n+1}\Big(\bigcup_{t=1}^{j} S(f_{i_t})\Big) \subseteq \mathcal{V}_{n+1}(S(g)).$$

By definition we have $\mathcal{V}_{n+1}(S(g)) = V_{n+1}(g)$. Combining Proposition 4 with Theorem 3 yields the following counting result.

Theorem 4. *Let* $U = \mathcal{V}_{n+1}\big(\bigcup_{t=1}^{j} S(f_{i_t})\big)$. *Then*

$$P_n(f_{i_1}, \ldots, f_{i_j}) = \begin{cases} \dfrac{|\Omega(n)|}{2^{\mathrm{cap}(U)}}, & \text{if } U \text{ is regular,} \\ 0, & \text{otherwise.} \end{cases}$$

Using Theorem 4 in the inclusion–exclusion expansion in Equation (4), we can compute $\eta_{k,n}$. To simplify this computation for large n, we study the asymptotic behavior of

$$\mathrm{cap}\Big(\mathcal{V}_{n+1}\Big(\bigcup_{t=1}^{j} S(f_{i_t})\Big)\Big) \quad \text{as } n \to \infty.$$

We begin with simple facts about $\mathcal{V}_n$.

Lemma 4. *For any set of binary sequences* X, *the sequence* $|\mathcal{V}_n(X)|$ *is nondecreasing in* n, *and* $|\mathcal{V}_n(X)| \le |X|$ *for all* $n \ge 1$.

Proof. Let $\pi : \mathbb{F}_2^{n+1} \to \mathbb{F}_2^n$ be the truncation map $\pi(a_0, \ldots, a_n) = (a_0, \ldots, a_{n-1})$. By definition of $\mathcal{V}_N$, we have $\pi(\mathcal{V}_{n+1}(X)) = \mathcal{V}_n(X)$, so $|\mathcal{V}_n(X)| \le |\mathcal{V}_{n+1}(X)|$.

Also, the map $X \to \mathcal{V}_n(X)$ given by $a \mapsto (a_0, \ldots, a_{n-1})$ is surjective, hence $|\mathcal{V}_n(X)| \le |X|$. □

Lemma 5. *Let* X *be a shift-invariant set of periodic sequences. There exists an integer* N_1 *such that for all* $n \ge N_1$, $|\mathcal{V}_n(X)| = |X|$.

Proof. Let L be the least common multiple of the periods of all sequences in X, and take $N_1 = L$. If $x, y \in X$ satisfy $\mathcal{V}_n(x) = \mathcal{V}_n(y)$ for some $n \ge N_1$, then x and y have the same first L bits. Since L is a multiple of the period of each sequence, these L bits determine the entire sequence, so $x = y$. Thus $\mathcal{V}_n$ is injective on X, which implies $|\mathcal{V}_n(X)| = |X|$. □

For a set of periodic sequences X, we define the separation threshold $\mathrm{sep}(X)$ as the smallest integer N_1 in Lemma 5, we can see from proof of Lemma 5 that $\mathrm{sep}(X) \le L$ for L the least common multiple of the periods of the sequences in X.

Theorem 5. *Let* X *be a shift-invariant set of periodic sequences. There exists an integer* N_2 *such that for all* $n \ge N_2$, *the set* $\mathcal{V}_n(X)$ *is regular and*

$$\mathrm{cap}\big(\mathcal{V}_n(X)\big) = |X|.$$

Proof. Set $N_2 = \mathrm{sep}(X) + 2$. Fix $n \ge N_2$ and suppose $u, v \in \mathcal{V}_n(X)$ have the same middle block. Choose $x, y \in X$ with $\mathcal{V}_n(x) = u$ and $\mathcal{V}_n(y) = v$. Then the shifted sequences $\sigma(x)$ and $\sigma(y)$ agree on their first $n-2$ bits. In particular, they agree on the first $\mathrm{sep}(X)$ bits. By Lemma 5, this forces $\sigma(x) = \sigma(y)$, so $x = y$.

This shows that the map $\mathcal{V}_n(X) \to \mathbb{F}_2^{n-2}$ sending $(a_0, \ldots, a_{n-1})$ to its middle block $(a_1, \ldots, a_{n-2})$ is injective, so $\mathcal{V}_n(X)$ is regular and $\mathrm{cap}(\mathcal{V}_n(X)) = |\mathcal{V}_n(X)|$. By Lemma 5, $|\mathcal{V}_n(X)| = |X|$ for all $n \ge N_2$, completing the proof.

□

For a set of periodic sequences X, we define the regular threshold $\mathrm{reg}(X)$ as the smallest integer N_2 in Theorem 5.

Corollary 2. *Let* X *be a shift-invariant set of periodic sequences. Then* $\mathrm{reg}(X) = \mathrm{sep}(X) + 2$.

Proof. By the proof of Theorem 5, we can see that $\mathrm{reg}(X) \le \mathrm{sep}(X)+2$. To show the reverse, suppose $\mathrm{reg}(X) \le \mathrm{sep}(X)+1$. That means for $n = \mathrm{sep}(X)+1$, map $\psi : \mathcal{V}_n(X) \to \mathbb{F}_2^{n-2}$ sending $(a_0, \ldots, a_{n-1})$ to its middle block $(a_1, \ldots, a_{n-2})$ is injective. Consider the map $\mathcal{V}_{n-2}$. By definition of the separation threshold, $\mathcal{V}_{n-2}$ is *not* injective. However, $\mathcal{V}_{n-2}$ factors as $\psi \circ \mathcal{V}_n \circ \sigma$, where σ is the left shift. Since both $\mathcal{V}_n$ and σ are bijections on X, the non-injectivity of $\mathcal{V}_{n-2}$ contradicts the assumed injectivity of ψ. □

Combine Theorem 5 with Theorem 4, we get the following result.

Theorem 6. *For* $n \ge \mathrm{reg}(\bigcup_{t=1}^{j} S(f_{i_t})) - 1$,

$$P_n(f_{i_1}, \ldots, f_{i_j}) = \frac{|\Omega(n)|}{2^{|(\bigcup_{t=1}^{j} S(f_{i_t}))|}}$$

Remark 2. *Theorems 4 and 6 generalize to arbitrary shift-invariant sets of periodic sequences. For such a set* X, *define* $P_n(X)$ *as the number of functions* $f \in \Omega(n)$ *satisfying* $X \subseteq S(f)$. *Letting* $U = \mathcal{V}_{n+1}(X)$, *we have*

$$P_n(X) = \begin{cases} \dfrac{|\Omega(n)|}{2^{\mathrm{cap}(U)}}, & \text{if } U \text{ is regular,} \\ 0, & \text{otherwise,} \end{cases}$$

and for $n \ge \mathrm{reg}(X) - 1$, *this simplifies to* $P_n(X) = |\Omega(n)|/2^{|X|}$.

V. Computation and Bounds

In this section we apply the theory represented in Section IV to estimate the density $\mu(n)$. By Equation 5 and Equation 9, we see that the central task is to calculate $\eta_{k,n}$. We first set up notation and the truncated inclusion–exclusion computation for a general parameter k, and then specialize to the case $k = 4$.

A. General computation of $\eta_{k,n}$

In this subsection, we introduce notations and the general idea on how to use the theory we developed in Section IV to compute $\eta_{k,n}$.

Before computation of $\eta_{k,n}$, we must obtain the set of irreducible NFSRs of order up to k. This is achieved by a brute-force approach. For each stage $m \le k$, we iterate over all possible NFSRs of order m. By Corollary 1 and Proposition 1, to test if an NFSR g of order m is irreducible, we use the generalized division algorithm from Lemma 1 to check whether g is divisible by any known irreducible NFSR f of order less than m.

If g is not divisible by any irreducible NFSR of lower order, we conclude that g is irreducible. Using this procedure, we compute the exact set of irreducible NFSRs of order $\le k$; denote its size by n_k and enumerate it as $f_1, \ldots, f_{n_k}$.

For each subset $I \subseteq \{1, \ldots, n_k\}$ with $I = \{i_1, \ldots, i_j\}$ we write

$$S_I = \bigcup_{t=1}^{j} S(f_{i_t}).$$

Define $a_I^{(k,n)}$ to be $2^{-\operatorname{cap}(\mathcal{V}_{n+1}(S_I))}$ when $\mathcal{V}_{n+1}(S_I)$ is regular, and 0 otherwise. For $j \ge 1$, define

$$A_j^{(k,n)} = \sum_{I \subset \{1,\ldots,n_k\}, |I|=j} a_I^{(k,n)}.$$

By Theorem 4, we have $P_n(f_{i_1}, \ldots, f_{i_j}) = |\Omega(n)| a_I^{(k,n)}$, and hence the j-th term in the alternating sum in Equation (4) is $|\Omega(n)| A_j^{(k,n)}$. Therefore

$$\eta_{k,n} = |\Omega(n)| \sum_{j=1}^{n_k} (-1)^{j-1} A_j^{(k,n)}.$$

We use Theorem 6 to accelerate computations for large n. For each k, we first calculate the number $R_k = \operatorname{reg}\left(\bigcup_{i=1}^{n_k} S(f_i)\right)$ by a brute-force search. The values of R_k for $k \le 4$ is shown in Table II.

TABLE II
Values of R_k for $k \le 4$

k	1	2	3	4
R_k	4	7	12	21

By monotonicity of the regular threshold, for every $S_I \subseteq \bigcup_{i=1}^{n_k} S(f_i)$, we have $\operatorname{reg}(S_I) \le \operatorname{reg}(\bigcup_{i=1}^{n_k} S(f_i))$. As a result, when $n \ge R_k - 1$, Theorem 6 applies to every S_I with $I \subseteq \{1, \ldots, n_k\}$, and the computation of $a_I^{(k,n)}$ simplifies to $a_I^{(k,n)} = 2^{-|S_I|}$.

B. Estimation of $\eta_{4,n}$

For our primary computation we take $k = 4$, where $n_k = 141$. The 141 irreducible NFSRs of order ≤ 4 form our dataset. Since enumerating all j-subsets for $j = 1, \ldots, 141$ is computationally infeasible, we are unable to calculate the precise number of $\eta_{4,n}$. However, we can truncate the inclusion–exclusion computation at $j \le 8$ to obtain tight lower and upper bounds for $\eta_{4,n}$.

Lemma 6 (Bonferroni inequality at $j \le 8$)**.** *For every positive integer n we have*

$$|\Omega(n)| \sum_{j=1}^{8} (-1)^{j-1} A_j^{(4,n)} \le \eta_{4,n} \le |\Omega(n)| \sum_{j=1}^{7} (-1)^{j-1} A_j^{(4,n)}.$$

Table II shows that for the 141 irreducible NFSRs of order ≤ 4 we have $\operatorname{reg}(\bigcup_{t=1}^{141} S(f_t)) = 21$. Thus in our dataset, the regular threshold is 21 , so we split the estimation of $\eta_{4,n}$ into two cases: $n \ge 20$ and $6 \le n \le 19$.

1) Case $n \geq 20$: For $n \geq 20$, the $(n+1)$-window projection $\mathcal{V}_{n+1}(S_I)$ is always regular, and the computation simplifies to $a_I^{(4,n)} = 2^{-|S_I|}$ by Theorem 6.

Algorithm 1 Estimation of $\eta_{4,n}$ for $n \geq 20$

Require: $k = 4$, hence $n_k = 141$; truncation level $J = 8$; stage $n \geq 20$
Collect all irreducible NFSRs $f_1, \ldots, f_{n_k}$ of order $\leq k$
Compute the cycle structure of each f_i and encode $S(f_i)$ as a bitset over the cycle universe
for $j = 1$ **to** J **do**
  Initialize $A_j^{(4,n)} \leftarrow 0$
  for each j-subset $I = \{i_1 < \cdots < i_j\} \subseteq \{1, \ldots, n_k\}$ **do**
    Compute $a_I^{(4,n)} = 2^{-|\bigcup_{t=1}^{j} S(f_{i_t})|}$
    Accumulate $A_j^{(4,n)} \leftarrow A_j^{(4,n)} + a_I^{(4,n)}$
  end for
end for
Compute upper bound $U = |\Omega(n)| \sum_{j=1}^{7} (-1)^{j-1} A_j^{(4,n)}$
Compute lower bound $L = |\Omega(n)| \sum_{j=1}^{8} (-1)^{j-1} A_j^{(4,n)}$
return Lower bound L and upper bound U for $\eta_{4,n}$

2) Case $6 \leq n \leq 19$: For $6 \leq n \leq 19$, we cannot simply use $|S_I|$. Instead we work with $(n+1)$-bit windows: each window $(a_0, \ldots, a_n)$ contributes its middle block $b = (a_1, \ldots, a_{n-1}) \in \mathbb{F}_2^{n-1}$ together with the parity label $p = a_0 + a_n \in \mathbb{F}_2$. A union S_I is regular at stage $n+1$ precisely when no middle block b appears with two different parity labels; in that case $\mathrm{cap}(\mathcal{V}_{n+1}(S_I))$ equals the number of distinct middle blocks occurring in the union.

Algorithm 2 Estimation of $\eta_{4,n}$ for $6 \leq n \leq 19$

Require: $k = 4$, hence $n_k = 141$; truncation level $J = 8$; stage $6 \leq n \leq 19$
Collect all irreducible NFSRs $f_1, \ldots, f_{n_k}$ of order $\leq k$
Precompute the $(n+1)$-window data of each $S(f_i)$ as middle blocks in $\mathbb{F}_2^{n-1}$ with parity labels
for $j = 1$ **to** J **do**
  Initialize $A_j^{(4,n)} \leftarrow 0$
  for each j-subset $I = \{i_1 < \cdots < i_j\} \subseteq \{1, \ldots, n_k\}$ **do**
    Check if $\mathcal{V}_{n+1}(\bigcup_{t=1}^{j} S(f_{i_t}))$ is regular using window constraints
    if regular **then**
      Compute $a_I^{(4,n)} = 2^{-\mathrm{cap}(\mathcal{V}_{n+1}(\bigcup_{t=1}^{j} S(f_{i_t})))}$
      Accumulate $A_j^{(4,n)} \leftarrow A_j^{(4,n)} + a_I^{(4,n)}$
    end if
  end for
end for
Compute upper bound $U = |\Omega(n)| \sum_{j=1}^{7} (-1)^{j-1} A_j^{(4,n)}$
Compute lower bound $L = |\Omega(n)| \sum_{j=1}^{8} (-1)^{j-1} A_j^{(4,n)}$
return Lower bound L and upper bound U for $\eta_{4,n}$

Once we have obtained lower and upper bounds for $\eta_{4,n}$, denoted as L and U respectively (so that $L \leq \eta_{4,n} \leq U$), we can substitute these into the density formulas to bound $\mu(n)$.

For the upper bound of $\mu(n)$, we recall from Equation (5) that $\mu(n) < 1 - \eta_{k,n}/2^{2^{n-1}}$. By applying our lower bound L for $\eta_{4,n}$, we get:

$$\mu(n) < 1 - \frac{L}{2^{2^{n-1}}}. \tag{10}$$

For the lower bound of $\mu(n)$, we recall from Equation (8) that $N_1(n) < \eta_{4,n} + |\Omega(n)| \sum_{m=5}^{n-1} \frac{N_0(m)}{2^{2^m}}$. Since $\mu(n) = 1 - N_1(n)/|\Omega(n)|$, we apply the upper bound U for $\eta_{4,n}$ to deduce:

$$\mu(n) > 1 - \frac{U}{2^{2^{n-1}}} - \sum_{m=5}^{n-1} \frac{N_0(m)}{2^{2^m}}. \tag{11}$$

The value $N_0(5)$ is known exactly as 31233. For $m > 5$, we simply use the upper bound for $\mu(m)$ to get upper bounds for $N_0(m)$. From Table III, we see that 0.472 is a safe universal upper bound of $\mu(m)$ for all $m > 5$. So $N_0(m) < 0.472 \cdot 2^{2^{m-1}}$ for all $m > 5$.

TABLE III
BOUNDS ON $\mu(n)$ FOR $n > 5$ WITH $k = 4$

n	Lower bound	Upper bound
6	0.4711846	0.4711942
7	0.4702690	0.4702776
8	0.4701412	0.4701492
9	0.4701279	0.4701358
10	0.4701639	0.4701719
11	0.4701458	0.4701538
≥ 12	0.4701455	0.4701535

Error Analysis: Now we investigate the error term in our estimation of $\mu(n)$. The difference between the upper bound and the lower bound is given by:

$$A_8^{(4,n)} + \sum_{m=5}^{n-1} \frac{N_0(m)}{2^{2^m}}.$$

the tail sum term $\sum_{m=5}^{n-1} \frac{N_0(m)}{2^{2^m}}$ is calculated to be approximately 7×10^{-6}. For the inclusion–exclusion truncation, the Bonferroni upper and lower bounds for $\eta_{4,n}$ at $J = 8$ differ by exactly $|\Omega(n)|A_8^{(4,n)}$, hence the corresponding contribution to the density gap is $A_8^{(4,n)}$. By calculation, $A_8^{(4,n)} \approx 6.20 \times 10^{-7}$, which is an order of magnitude smaller than the tail sum. Therefore, at truncation level $J = 8$ the dominant source of uncertainty in our bounds for $\mu(n)$ comes from the tail-sum term (i.e., contributions of irreducible NFSRs of order > 4), rather than from truncating the inclusion–exclusion expansion.

Let $T_{4,n}$ be the tail-sum term $T_{4,n} = \sum_{m=5}^{n-1} \frac{N_0(m)}{2^{2^m}}$. For sufficiently large n, using $N_0(5) = 31233$ and the uniform bound $N_0(m) < 0.472 \cdot 2^{2^{m-1}}$ for $m \geq 6$, we obtain

$$T_{4,n} = \frac{31233}{2^{32}} + 0.472 \sum_{m=6}^{n-1} \frac{1}{2^{2^{m-1}}}.$$

The latter sum is approximately $0.472/2^{32} \approx 10^{-10}$, so $T_{4,n}$ is essentially dominated by the leading term $\frac{31233}{2^{32}}$.

The error analysis shows that the error comes mainly from the tail-sum term, in particular, the term $\frac{31233}{2^{32}}$. Let $R_i^{(k,n)} = A_i^{(k,n)} - A_i^{(k-1,n)}$, One could attempt to replace the $\frac{31233}{2^{32}}$ tail-sum term with $R_1^{(5,n)} - R_2^{(5,n)} + R_3^{(5,n)}$, but this yields very little overall improvement to the bounds due to the small magnitude of $R_2^{(5,n)}$ and $R_3^{(5,n)}$.

C. Choice of the Parameter k

In this subsection, we discuss choices of different values of the parameter k and explain why we chose the value $k = 4$. The choice of k reflects a trade-off in the estimation framework between the tail sum error (from uncomputed higher-order irreducible NFSRs) and the truncation error (from stopping the inclusion–exclusion principle early).

For $k = 1$, we only consider 2 irreducible NFSRs, $x_0 + x_1$ and $x_0 + x_1 + 1$, but we can still get a bound. According to our theory, we easily get $0.337 < \mu(n) \leq 0.5625, n \geq 3$. Tian's original result [11] made use of these two nfsrs together with other arguments to get a better lower bound 0.39.

For $k = 2$, there are only 5 irreducible NFSRs, making the full inclusion–exclusion sum trivial to compute without truncation. Jiang and Lin [12] previously computed this case to establish the bounds $0.4461 < \mu(n) < 0.4834, n \geq 5$, yielding a density gap of roughly 0.037.

For $k = 3$, there are 14 irreducible NFSRs. Again, the full inclusion–exclusion sum can be evaluated exactly without truncation. As shown in Table IV, the bounds for $n \geq 11$ narrow to $0.4688 < \mu(n) < 0.4708$, reducing the gap to approximately 0.002. However, the tail sum error is still prominent.

TABLE IV
BOUNDS ON $\mu(n)$ FOR $n > 5$ WITH $k = 3$

n	Lower bound	Upper bound
6	0.4697	0.4718
7	0.4689	0.4710
8–10	0.4688	0.4709
≥ 11	0.4688	0.4708

For $k = 4$, we work with 141 irreducible NFSRs. As demonstrated in the error analysis above, the tail sum error is heavily suppressed (to roughly 7×10^{-6}), and by computing the inclusion–exclusion sum up to $J = 8$, the truncation error is brought down to approximately 6.20×10^{-7}. This proves to be the optimal parameter configuration.

One might consider advancing to $k = 5$ to further reduce the tail sum error, which drops to roughly 10^{-10}. However, for $k = 5$, there are 31, 247 irreducible NFSRs. The sheer combinatorial explosion means it is only computationally feasible to evaluate the sum up to $J = 3$ (i.e., computing $A_1^{(5,n)}$, $A_2^{(5,n)}$, and $A_3^{(5,n)}$). Because $A_j^{(k,n)}$ monotonically increases with k, we know that $A_3^{(5,n)} > A_3^{(4,n)} \approx 4.88 \times 10^{-2}$. Truncating the Bonferroni sum at $J = 3$ for $k = 5$ would therefore leave a massive truncation error on the order of 10^{-2}, which completely overshadows the tail sum improvement from 7×10^{-6} to 10^{-10}. One could compromise to consider using only a subset of the order-5 irreducible NFSRs rather than all of them, but this does not lead to a significant improvement. For $k \geq 6$, the set of irreducible NFSRs of order k is too large to be computed.

Therefore, $k = 4$ remains the optimal choice for providing the tightest possible bounds.

VI. Conclusion

In this paper, we proved that the density of irreducible NFSRs of a given stage satisfies highly tight bounds $0.4701455 < \mu(n) < 0.4701535$ for $n \geq 12$, and additional explicit bounds for $6 \leq n \leq 11$, see Table III. Our theory supports automated large-scale inclusion-exclusion estimation beyond what is feasible with manual division-algorithm arguments. Testing irreducibility of arbitrary NFSRs remains an open problem for future research.

Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable comments. This work was partially supported by the National Natural Science Foundation of China under Grant 62372449.

References

[1] M. Hell, T. Johansson, A. Maximov, and W. Meier, "The grain family of stream ciphers," in *New Stream Cipher Designs*. Berlin, Germany: Springer, 2008, pp. 179–190.
[2] C. De Cannière and B. Preneel, "Trivium," in *New Stream Cipher Designs*. Berlin, Germany: Springer, 2008, pp. 244–266.
[3] N. Zierler, "Linear recurring sequences," *Journal of the Society for Industrial and Applied Mathematics*, vol. 7, no. 1, pp. 31–48, 1959. [Online]. Available: https://doi.org/10.1137/0107003
[4] R. Lidl and H. Niederreiter, *Finite Fields*, 2nd ed., ser. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1996.
[5] K. Kjeldsen, "On the cycle structure of a set of nonlinear shift registers with symmetric feedback functions," *Journal of Combinatorial Theory, Series A*, vol. 20, no. 2, pp. 154–169, 1976. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0097316576900133
[6] J. Søreng, "The periods of the sequences generated by some symmetric shift registers," *Journal of Combinatorial Theory, Series A*, vol. 21, no. 2, pp. 164–187, 1976. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0097316576900613
[7] J. Sø reng, "Symmetric shift registers." *Pacific J. Math.*, vol. 85, no. 2, pp. 201–229, 1979. [Online]. Available: http://dml.mathdoc.fr/item/1102784091
[8] U. Cheng, "On the cycle structure of certain classes of nonlinear shift registers," *Journal of Combinatorial Theory, Series A*, vol. 37, no. 1, pp. 61–68, 1984. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0097316584900190
[9] J. Mykkeltveit, M.-K. Siu, and P. Tong, "On the cycle structure of some nonlinear shift register sequences," *Information and Control*, vol. 43, no. 2, pp. 202–215, 1979. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0019995879907083
[10] J. Zhang, T. Tian, W. Qi, and Q. Zheng, "A new method for finding affine sub-families of NFSR sequences," *IEEE Trans. Inf. Theory*, vol. 65, no. 2, pp. 1249–1257, 2019. [Online]. Available: https://doi.org/10.1109/TIT.2018.2858769
[11] T. Tian and W. Qi, "On the density of irreducible nfsrs," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 4006–4012, 2013. [Online]. Available: https://doi.org/10.1109/TIT.2013.2247093
[12] Y. Jiang and D. Lin, "Lower and upper bounds on the density of irreducible nfsrs," *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3944–3952, 2018. [Online]. Available: https://doi.org/10.1109/TIT.2018.2796127
[13] S. W. Golomb, *Shift Register Sequences*. San Francisco, CA, USA: Holden-Day, 1967.
[14] D. A. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms*, 5th ed., ser. Undergraduate Texts in Mathematics. Springer Cham, 2025.

Daoyuan Zhang Daoyuan Zhang is currently pursing the Ph.D. degree at the Institute of Information Engineering of Chinese Academy of Sciences. His main research interest is in stream cipher.

Dongdai Lin Dongdai Lin received the M.S. degree and the Ph.D. degree in fundamental mathematics from the Institute of Systems Science, Chinese Academy of Sciences in 1987 and 1990 respectively. He is currently working on post quantum cryptography, sequences and stream cipher, Boolean functions.