Electronic Colloquium on Computational Complexity - Reports Series 1995 - available via:
ECCC
TR95-041
FTP: ftp.eccc.uni-trier.de:/pub/eccc/
WWW: http://www.eccc.uni-trier.de/eccc/
Email: ftpmail@ftp.eccc.uni-trier.de with subject 'help eccc'

Optimal Bounds for the Approximation of Boolean Functions and Some Applications

Alexander E. Andreev
Department of Mathematics
University of Moscow
Moscow

Andrea E. F. Clementi
Centre Universitaire d'Informatique
University of Geneva
Geneva

José D. P. Rolim
Centre Universitaire d'Informatique
University of Geneva
24, rue Général-Dufour - CH 1211 - Geneva
E-mail: rolim@cui.unige.ch

(Extended Abstract)

Abstract

We prove an optimal bound on the *Shannon* function $L(n, m, \epsilon)$ which describes the trade-off between the circuit-size complexity and the degree of approximation; that is

$$L(n, m, \epsilon) \;=\; \Theta\left(\frac{m\epsilon^2}{\log(2 + m\epsilon^2)}\right) + O(n).$$

Our bound applies to any partial boolean function and any approximation degree, and thus completes the study of boolean function approximation, introduced by Pippenger [11], concerning circuit-size complexity. As a consequence, we provide the approximation degree achieved by polynomial size circuits on a 'random' boolean function; that is

$$App_0(n, l(n) = n^k) = \Theta\left(\frac{\sqrt{(n^k - n)}}{(2^{\frac{n}{2}})}\right) \;,\; k > 1.$$

As an application, we obtain a non trivial upper bound on the *hardness* function $H(f)$ introduced by Nisan and Wigderson [10]; that is, for any boolean function $f : \{0,1\}^n \to \{0,1\}$:

$$H(f) \;\leq\; 2^{n/3} + n^2 + O(1).$$

The optimal bound for $L(n, m, \epsilon)$ gives also a general criterium for determining the quality of a given *learning* algorithm for partial boolean functions. The contribution in our proofs can be viewed as a new technique based on some particular algebraic properties of linear operator for approximating partial boolean function.

1 Introduction

We investigate the concept of approximation of boolean functions introduced by Pippenger [11]. The main result of this paper is an optimal bound on the trade-off between the circuit-size complexity and the degree of approximation. The obtained result holds for the general case, that is, for any partial boolean function and for any approximation degree function. As a consequence, we provide a rather precise answer to another central question: "Which is the degree of approximation achieved by polynomial-size circuits for any (and thus even random) boolean function?" or, equivalently, "How much information a polynomial-size circuit can give about a random boolean function?"

This question plays an important role in several topics such as pseudorandom generators [13, 9], randomized computations [4] and cryptography [5]. In particular, we show some consequences of our results for the relevant work, due to Nisan and Wigderson [10], concerning the notion of *hardness* of boolean functions and, more generally, the *hardness-randomness trade-offs.*

- Pippenger's concept of approximation, prior works and our results

Let $f : \mathcal{A} \to \{0,1\}$ ($\mathcal{A} \subseteq \{0,1\}^n$) be a partial boolean function and consider the uniform probability function defined on $\mathcal{A}$. The function $L(f,\epsilon)$ denotes the minimum positive integer l_{min} for which a boolean circuit[1] S of size l_{min} exists such that $\mathbf{Pr}\,(f = S) \geq \frac{1}{2} + \epsilon$ where ϵ (with $0 < \epsilon \leq 1/2$) denotes the *approximation degree*. The function $App(f,l)$ is the maximum value ϵ for which there exists a circuit S of size at most l such that the above inequality holds. These two functions describe the trade-off between circuit-size complexity and degree of approximation. We can then introduce the 'approximation' version of the Shannon functions, i.e., the function $L_0(n,\epsilon)$ defined as the maximum value of $L(f,\epsilon)$ achieved by any boolean function $f : \{0,1\}^n \to \{0,1\}$ and the function $L(n,m,\epsilon)$ as the maximum value of $L(f,\epsilon)$ achieved by any partial boolean function $f : \mathcal{A} \to \{0,1\}$ with $|\mathcal{A}| \leq m$. The functions $L(f)$, $L_0(n)$, $L(n,m)$ denote the corresponding Shannon functions for perfect constructions (i.e. for $\epsilon = 1/2$). Moreover, we can define the function $App_0(n,l)$ as the minimum value of $App(f,l)$ achieved by any function $f : \{0,1\}^n \to \{0,1\}$ and, equivalently, the function $App(n,m,l)$ for partial boolean functions.

Pippenger [11] introduced the function $L(n,m,\epsilon)$ and obtained the asymptotical behaviour for a restricted case, that is, when $m = \Omega(2^n)$ and ϵ is a positive constant independent of n:

$$L(n,m,\epsilon) \;\sim\; \left(1 + \left(\frac{1}{2}+\epsilon\right)\log\left(\frac{1}{2}+\epsilon\right) + \left(\frac{1}{2}-\epsilon\right)\log\left(\frac{1}{2}-\epsilon\right)\right)\frac{m}{n}.$$

In this case, we have $L(n,m,\epsilon) \;=\; \Theta\left(\frac{m}{n}\right)$. Notice that the function $\frac{m}{n}$ is also the asymptotical behaviour for $L(n,m)$ (see [12, 3]). Informally speaking, such results tell us that, when a constant degree of approximation (thus very high!) is required, it is then necessary to use boolean circuits having size equivalent to those required for perfect constructions. However no information can be derived from Pippenger's result about the degree of approximation achieved either by circuit of smaller size (in particular those having polynomial size) or when the domain size is not exponential in n. In some applications, such as the construction of pseudorandom generators and/or crypto-systems, such information is generally required. Our main contribution consists in determining the optimal bound for the functions $L_0(n,\epsilon)$ and $L(n,m,\epsilon)$ in the general case.

Theorem 1.1 *For any $n > 0$, $0 < m \leq 2^n$ and $\epsilon > 0$, we have:*

$$L(n,m,\epsilon) \;=\; \Theta\left(\frac{m\epsilon^2}{\log(2+m\epsilon^2)}\right) + O(n), \quad \textit{and} \quad L_0(n,\epsilon) \;= \Theta\left(\frac{2^n\epsilon^2}{\log(2+2^n\epsilon^2)}\right) + \Theta(n).$$

We can then apply this theorem for obtaining the behaviour of functions $App_0(n,l)$ and $App(n,m,l)$.

[1] we consider boolean circuits having any kind of gates of one or two inputs

Corollary 1.1 *A constant $c > 1$ exists such that if $l \geq cn$ then*

$$App_0(n,l) = \Theta\left(\frac{\sqrt{(l-n)}}{2^{\frac{n}{2}}}\right), \quad and \quad App_0(n, l(n) = n^k) = \Theta\left(\frac{\sqrt{(n^k - n)}}{(2^{\frac{n}{2}})}\right) \quad for\ k > 1. \qquad (1)$$

In [10], it is proved that if a function $F : \{0,1\}^* \rightarrow \{0,1\}$ exists such that *i)* $F \in EXP$ and *ii)* for any $n > 0$, the restriction of F to the finite domain $\{0,1\}^n$ is not approximable[2] by any polynomial-size circuit, then $BPP \subset \cap_{\alpha>0} DTIME(2^{\alpha n})$. Observe that Eq. 1 implies that for any n there is a boolean function f_n for which the approximation degree achieved by polynomial-size circuits is bounded by the inverse of an exponential function in n and thus it cannot be approximated. Although this fact does not imply that the hypothesis of Nisan and Wigderson's theorem is completely true (observe that condition (i) might not be verified), our lower bound provides further positive indications on the conjecture that the gap between deterministic and randomized computational power is not large.

On the other hand, the lower bound in Corollary 1.1 implies a non trivial upper bound on the *hardness* function $H(f) = \min\{l \ : \ App(f,l) \geq \frac{1}{l}\}$ [10].

Corollary 1.2 *For any $n > 0$ and for any boolean function $f : \{0,1\}^n \rightarrow \{0,1\}$, we have:*

$$H(f) \ \leq \ 2^{n/3} + n^2 + O(1).$$

The interest in the *hardness* function lies in the following theorem [10]. If a function $F : \{0,1\}^* \rightarrow \{0,1\}$ exists such that *i)* $F \in EXP$ and *ii)* for any $n > 0$, the restriction of F to the finite domain $\{0,1\}^n$ has *hardness* $2^{\alpha n}$ for some $\alpha > 0$, then $BPP = P$. In particular, Corollary 1.2 provides a new upper bound on the value α in condition (ii) (i.e. $\alpha \leq 1/3$). Since this bound holds also for a non recursive function F, our opinion here is that condition (ii) is very strong for a function belonging to EXP.

Theorem 1.1 has also an important consequence in learning partial boolean functions. In particular, the optimal bound for $L(n,m,\epsilon)$ gives a good criterium for determining the quality of a learning algorithm. A learning algorithm for a partial boolean function $f(x_1,\ldots,x_n)$ (with domain size m) can be reasonably seen as a boolean circuit $S^T(x_1,\ldots,x_n)$ which makes use of a set T of positive and negative examples (i.e. a table T of boolean vectors $\vec{x}$'s with the corresponding values $f(\vec{x})$'s) (see also [6]). Suppose now that S^T achieves an approximation degree equal to ϵ (i.e. $\mathbf{Pr}\left(f = S^T\right) \geq \frac{1}{2} + \epsilon$). If $|S^T| = \Omega(L(n,m,\epsilon))$, we can state that S^T is not so efficient since our upper bound for $L(n,m,\epsilon)$ implies that there is a circuit S of size $O(L(n,m,\epsilon))$ which gives the same approximation degree. In other words, the use of table T is not relevant. On the other hand, if $|S^T| = o(L(n,m,\epsilon))$, our lower bound for $L(n,m,\epsilon)$ implies that, in this case, the use of table T in deriving S^T is useful (although a further analysis is required to establish upper and lower bounds on the size of T).

- Techniques adopted: the probabilistic method and linear approximation

Although the lower bound in Theorem 1.1 is obtained by standard *counting arguments*, another contribution is the technique herein proposed to derive the upper bound. This technique consists in the use of the *probabilistic method* (see [1]) for deriving the existence of some particular *linear operators* which can be applied for approximating boolean functions. We first introduce a natural algebraic property on the set of boolean domains: a set $\mathcal{A} \subseteq \{0,1\}^n$ is *4-regular* if, for any choice of four pairwise different vectors in $\mathcal{A}$, their sum (i.e. the $\oplus$ operation performed component by component) yields a non *zero* vector. We then prove that if we choose randomly a linear operator $\vec{l}$ (i.e. a vectorial

[2] According to [10], a function f is not approximable by a circuit C if $\mathbf{Pr}\,(f \neq S) \geq \frac{1}{n^k}$, for some $k > 0$

linear function) defined on a 4-regular domain $\mathcal{A}$ then, with high probability, it is 'well-distributed' (i.e. there are not too many elements in $\mathcal{A}$ having the same image according to $\vec{l}$). From this algebraic result, we show how to use linear operators for approximating partial boolean functions defined on 4-regular domains. We thus obtain the upper bound of Theorem 1.1 in this restricted case. The next step consists in extending the upper bound for 4-regular domains to general domains. For achieving this aim, we prove the existence of an injective vectorial function $\vec{\mathbf{J}}_n$, having linear circuit complexity, which maps the set $\{0,1\}^n$ into the set $\{0,1\}^{cn}$ (for some constant $c > 1$), such that the resulting subset $\vec{\mathbf{J}}_n(\{0,1\}^n)$ is 4-regular. The technique for the case of general domains is still based on linear operators; indeed, the function $\vec{\mathbf{J}}_n$ is a convenient composition of linear operators which maps each nonzero element of $\{0,1\}^n$ to an element of $\{0,1\}^{cn}$ having a large (i.e. linear) number of 1's. We then show that this property yields a 4-regular domain. Finally, for any function f defined on $\mathcal{A}$, we use the injectivity of $\vec{\mathbf{J}}_n$ in order to construct a new function f^* defined on $\vec{\mathbf{J}}_n(\{0,1\}^n)$ such that $L(f,\epsilon) \leq L(f^*,\epsilon) + O(n)$. Since $\vec{\mathbf{J}}_n(\{0,1\}^n)$ is 4-regular, we obtain the upper bound for the general case.

We believe that our rather general approximation method, based on linear operators, will give potential tools also in providing non trivial upper bounds on the circuit-depth complexity for approximating boolean functions.

2 Upper Bounds

2.1 Linear approximation on 4-regular domains

In this section we show some interesting properties of linear boolean functions. In particular, we first introduce a particular class of boolean domains denoted in the sequel as *4-regular domains*. We then prove an upper bound for the function $L(n,m,\epsilon)$ restricted to partial boolean functions defined on 4-regular domains.

A boolean function $l(x_1,\ldots,x_n)$ is linear if it can be represented in the following way: $l(x_1,\ldots,x_n) = \alpha_1 x_1 \oplus \ldots \oplus \alpha_n x_n \oplus \beta$, where $\alpha_1,\ldots,\alpha_n,\beta$ are boolean constants. Moreover, the set of all linear functions with n variables is denoted as $\mathcal{L}_n$.

Definition 2.1 *A domain $\mathcal{A} \subseteq \{0,1\}^n$ is* 4-regular *if for any 4-tple $\langle \vec{a_1},\vec{a_2},\vec{a_3},\vec{a_4}\rangle$ of vectors in $\mathcal{A}$, such that $\vec{a_i} \neq \vec{a_j}$ $(i \neq j)$, the three vectors $\vec{a_1} \oplus \vec{a_2}$, $\vec{a_1} \oplus \vec{a_3}$ and $\vec{a_1} \oplus \vec{a_4}$ are linear independent.*

Observe that requiring the linear independence of the above three vectors is equivalent to the following statement: $\vec{a_1} \oplus \vec{a_2} \oplus \vec{a_3} \oplus \vec{a_4} \neq (0,\ldots,0)$. Our next goal is to extimate the degree of approximation achieved by a *random* linear function with respect to a boolean function $f : \mathcal{A} \to \{0,1\}$ $(\mathcal{A} \subseteq \{0,1\}^n)$. when its domain $\mathcal{A}$ is 4-regular. We thus consider a linear function as a random element selected from the space $\mathcal{L}_n$ with uniform probability. For any $\vec{a} \in \mathcal{A}$, we introduce the 'agreement' function defined on the space $\mathcal{L}_n$: $\xi_{\vec{a}}(l) = l(\vec{a}) \oplus f(\vec{a}) \oplus 1$ (notice that $\xi_{\vec{a}}(l) = 1$ iff $f(\vec{a}) = l(\vec{a})$). We then consider the following sum:

$$\Xi_{\mathcal{A}}(l) = \sum_{\vec{a}\in\mathcal{A}} \xi_{\vec{a}}(l) \ = |\{\vec{a} \in \mathcal{A} \ : \ f(\vec{a}) = l(\vec{a})\}|.$$

In the following lemma, we provide some properties of the expected value of $\Xi_{\mathcal{A}}$ (for the proof see Lemma A.1 in Appendix A).

Lemma 2.1 *Let $\mathbf{E}(\Xi_{\mathcal{A}})$ denote the expected value $\Xi_{\mathcal{A}}$ in $\mathcal{L}_n$. Then, for any set $\mathcal{A}$ of size m we have:*

$$\mathbf{E}(\Xi_{\mathcal{A}}) \;=\; \frac{m}{2} \quad \textit{and} \quad \mathbf{E}\left((\Xi_{\mathcal{A}} - \mathbf{E}(\Xi_{\mathcal{A}}))^2\right) \;=\; \frac{m}{4}. \tag{2}$$

Moreover, if $\mathcal{A}$ is 4-regular then:

$$\mathbf{E}\left((\Xi_{\mathcal{A}} - \mathbf{E}(\Xi_{\mathcal{A}}))^4\right) \;=\; \frac{3m^2}{16} - \frac{m}{8} \quad \text{and} \quad \mathbf{Pr}\left((\Xi_{\mathcal{A}} - \mathbf{E}(\Xi_{\mathcal{A}}))^2 < \frac{1}{2}\mathbf{E}\left((\Xi_{\mathcal{A}} - \mathbf{E}(\Xi_{\mathcal{A}}))^2\right)\right) \leq \frac{47}{48}. \tag{3}$$

The above results are now used in determining the portion of linear functions yielding an approximation degree not smaller than a positive value ϵ for a boolean function f defined on a 'small' 4-regular domain. Let us describe this fact in a formal way. For any boolean function f we denote as $\mathcal{L}_n(f,\epsilon)$ the set of all linear functions $l \in \mathcal{L}_n$ such that $Pr(f = l) \;\geq\; \frac{1}{2} + \epsilon$.

Lemma 2.2 *Let $\mathcal{A}$ be a 4-regular domain of size $m \leq \frac{1}{8}\epsilon^{-2}$, where $0 < \epsilon \leq 1/2$. Then, for any boolean function f defined on $\mathcal{A}$, we have*

$$|\mathcal{L}_n(f,\epsilon)| \;\geq\; \frac{1}{96}\;|\mathcal{L}_n| \;.$$

Sketch of the proof. From Lemma 2.1, we have

$$\mathbf{Pr}\left(\left|\Xi_{\mathcal{A}} - \frac{m}{2}\right| \;\geq\; \frac{\sqrt{m}}{2\sqrt{2}}\right) \geq \frac{1}{48}\;.$$

Since $\mathbf{Pr}\left(\Xi_{\mathcal{A}} - \frac{m}{2} \;=\; x\right) \;=\; \mathbf{Pr}\left(\Xi_{\mathcal{A}} - \frac{m}{2} \;=\; -x\right)$, we have

$$\mathbf{Pr}\left(\Xi_{\mathcal{A}} \;\geq\; m\left(\frac{1}{2} + \sqrt{\frac{1}{8m}}\right)\right) \;\geq\; \frac{1}{96}\;.$$

Moreover, the inequality $m \;\leq\; \frac{1}{8}\epsilon^{-2}$ implies that $\mathbf{Pr}\left(\Xi_{\mathcal{A}} \;\geq\; m\left(\frac{1}{2} + \epsilon\right)\right) \;\geq\; \frac{1}{96}$. Finally, from the definition of $\Xi_{\mathcal{A}}$ we obtain

$$\mathbf{Pr}\left(\Xi_{\mathcal{A}} \;\geq\; m\left(\frac{1}{2} + \epsilon\right)\right) \;=\; \frac{|\mathcal{L}_n(f,\epsilon)|}{|\mathcal{L}_n|} \geq \frac{1}{96}.$$

□

The above result can be interpreted as follows. A boolean function f having 4-regular domain can be approximated by a 'large' number of linear functions if the domain size is 'sufficiently small' with respect to the desired quality of approximation. Roughly speaking, our next step consists in defining a suitable partition for 4-regular domains. This partition will permit us to reduce the approximation problem for general 4-regular domains to the same problem restricted to 4-regular domains satisfying the 'size' condition required by Lemma 2.2. Our partition technique is based on the use of *linear operators*. Consider a linear operator $\vec{l} \;=\; (l_1, l_2, ..., l_s) \in (\mathcal{L}_n)^s$ and an element $\vec{b} \in \{0,1\}^s$. We then define the $\{0,1\}^n$-subset $\vec{l}^{-1}(\vec{b}) = \{\vec{a} \;:\; \vec{l}(\vec{a}) = \vec{b}\}$. Moreover, for any partial boolean function f with domain $\mathcal{A} \subseteq \{0,1\}^n$, we define the following partial boolean function: $f_{\vec{l},\vec{b}} = f(\vec{a})$ if $\vec{a} \in \mathcal{A} \cap \vec{l}^{-1}(\vec{b})$ and not defined otherwise. It is easy to verify that[3]: $f(\vec{a}) \;=\; \vee_{\vec{b}}(\wedge_{i=1}^{s}(l_i(\vec{a}) \oplus b_i \oplus 1)) \wedge f_{\vec{l},\vec{b}}(\vec{a})$.

[3] we assume that $0 \wedge * = 0$ and $1 \vee * = 1$, where $*$ is the undefined value

Lemma 2.3 *If the domain $\mathcal{A}$ of a boolean function f is 4-regular and a linear operator $\vec{l} \in (\mathcal{L}_n)^s$ exists such that, for any $\vec{b} \in \{0,1\}^s$ ($s \geq 1$), the domain size of function $f_{\vec{l},\vec{b}}$ is at most $\frac{1}{8}\epsilon^{-2}$ (i.e. $|\mathcal{A} \cap \vec{l}^{-1}(\vec{b})| \leq \frac{1}{8}\epsilon^{-2}$), then*

$$L(f,\epsilon) \ = O\left(\frac{2^s}{s}\right) + O(n) \ .$$

Sketch of the proof. Consider the set partition constructed as follows. Let $\mathcal{Q} \subseteq \{0,1\}^s$. Since for any $\vec{b} \in \mathcal{Q}$ the domain size $|\mathcal{A} \cap \vec{l}^{-1}(\vec{b})|$ of function $f_{\vec{l},\vec{b}}$ is at most $\frac{1}{8}\epsilon^{-2}$, we can apply Lemma 2.2 thus obtaining $|\mathcal{L}_n(f_{\vec{l},\vec{b}},\epsilon)|/|\mathcal{L}_n| \geq 1/96$. Consequently a (at least one) linear function $h_{\mathcal{Q}}$ exists such that:

$$\left|\left\{\vec{b} \ : \ \vec{b} \in \mathcal{Q} \ , \ \mathbf{Pr}\left(f_{\vec{l},\vec{b}} \ = \ h_{\mathcal{Q}}\right) \ \geq \ \frac{1}{2}+\epsilon\right\}\right| \ \geq \ \frac{1}{96}|\mathcal{Q}| \ . \tag{4}$$

Let us define $\mathcal{Q}_0^* = \{0,1\}^s$ and consider the corresponding linear function $h_{\mathcal{Q}_0^*}$ as defined above; by induction, we can then construct the following sequence of pairs $\langle \mathcal{Q}_i, h_i \rangle$ ($i > 0$) as follows:

$$\mathcal{Q}_i^* \ = \ \{0,1\}^s \setminus \left(\bigcup_{j=1}^{i} \mathcal{Q}_j\right) \ , \qquad h_{i+1} \ = \ h_{\mathcal{Q}_i^*} \ ,$$

and

$$\mathcal{Q}_{i+1} \ = \ \left\{\vec{b} \ : \ \vec{b} \in \mathcal{Q}_i^* \ , \ \mathbf{Pr}\left(f_{\vec{l},\vec{b}} \ = \ h_{i+1}\right) \ \geq \ \frac{1}{2}+\epsilon\right\} \ .$$

We have thus constructed a set partition $\vec{\mathcal{Q}} = (\mathcal{Q}_1, \mathcal{Q}_2, ..., \mathcal{Q}_k)$ having the following properties. For any $i \geq 0$ and for any $b \in \mathcal{Q}_i$, $\mathbf{Pr}\left(f_{\vec{l},\vec{b}} = h_i\right) \ \geq \ \frac{1}{2}+\epsilon$. Observe also that the process terminates on the first step $k > 0$ for which $\mathcal{Q}_i^* = \emptyset$ and, by Eq. 4, it is not hard to prove that $k = O(\log n)$. Furthermore, we have that $\mathcal{Q}_i \cap \mathcal{Q}_j = \emptyset$ ($i \neq j$) and observe also that Eq. 4 implies $|\mathcal{Q}_i| \geq \ \frac{1}{96}\sum_{j=i+1}^{k}|\mathcal{Q}_j|$.

Let us now consider the following 'selector' operator $\vec{F}_{\vec{\mathcal{Q}}}(x_1 \ldots, x_s) = (f_{\vec{\mathcal{Q}},1}, \ldots, f_{\vec{\mathcal{Q}},k})$, where $f_{\vec{\mathcal{Q}},i}(\vec{b})$ is equal to 1 if $\vec{b} \in \mathcal{Q}_i$ and 0 otherwise. The properties of the set partition $\vec{\mathcal{Q}}$ implies that the circuit complexity of $\vec{F}_{\vec{\mathcal{Q}}}$ is $O(\frac{2^s}{s})$. The proof of this fact is shown in Lemma A.2 in Appendix A. We can now define the function f_{apr} which approximates f. Indeed, for any $\vec{a} \in \mathcal{A}$, we define:

$$f_{apr}(\vec{a}) \ = \ \bigvee_{i=1}^{k}\left(f_{\mathcal{Q}_i}(\vec{l}(\vec{a})) \wedge h_i(\vec{l}(\vec{a}))\right) . \tag{5}$$

By construction of the components of f_{apr}, it follows that

$$\mathbf{Pr}(f = f_{apr}) \ \geq \ \frac{1}{2}+\epsilon.$$

From Definition (5), we have the following upper bound on the circuit complexity of f_{apr}:

$$L(f_{apr}) \ \leq \ L(\vec{F}_{\vec{\mathcal{Q}}}) + L(\vec{l}) + L((h_1, h_2, ..., h_k)) \ . \tag{6}$$

Moreover, the circuit complexity of linear operators satisfies the following upper bounds (see [8] for a proof):

$$L(\vec{l}) \ = \ O\left(\frac{ns}{\log n}\right) + O(n) \ , \qquad L((h_1, h_2, ..., h_k)) \ = \ O\left(\frac{nk}{\log n}\right) + O(n) \ . \tag{7}$$

Finally, Eq. (6), Eq. (7) and the bound $k = O(\log n)$ imply

$$L(f_{apr}) = O\left(\frac{2^s}{s}\right) + O\left(\frac{sn}{\log n}\right) + O\left(\frac{kn}{\log n}\right) = O\left(\frac{2^s}{s}\right) + O(n) .$$

□

For any $\vec{a} \in \{0,1\}^n$ and for any $\vec{b} \in \{0,1\}^s$, consider now the 'agreement' function $\xi_{\vec{a},\vec{b}} : (\mathcal{L}_n)^s \rightarrow \{0,1\}$ defined as $\xi_{\vec{a},\vec{b}}(\vec{l}) = \Pi_{i=1}^{s}(l_i(\vec{a}) \oplus b_i \oplus 1)$, where $\vec{l} = (l_1, l_2, ..., l_s)$. Consider also the sum of points in $\{0,1\}^n$ on which $\vec{l}$ is equal to a fixed value $\vec{b}$:

$$\Xi_{\mathcal{A},\vec{b}}(\vec{l}) = \sum_{\vec{a} \in \mathcal{A}} \xi_{\vec{a},\vec{b}}(\vec{l}) .$$

Notice that, according to the definitions adopted in Lemma 2.3, the value of $\Xi_{\mathcal{A},\vec{b}}(\vec{l})$ is the size of the domain of function $f_{\vec{l},\vec{b}}$. The linear operator required by Lemma 2.3 is thus given by the following result (which is proved in Lemmas A.3 and A.4 in Appendix A).

Lemma 2.4 *Let $\mathcal{A}$ be a 4-regular set of size m and let $s \geq 1$; then a linear operator $\vec{l} \in (\mathcal{L}_n)^s$ exists such that*

$$\sum_{\vec{d} \,:\, \Xi_{\mathcal{A},\vec{d}}(\vec{l}) \,>\, 2\ 2^{-s}m\ +1} \Xi_{\mathcal{A},\vec{d}}(\vec{l}) \leq 2^s .$$

In 1989, Andreev [3] proved the following result.

Lemma 2.5 *The circuit complexity of any partial boolean function with n variables and domain size equal to m satisfies the following upper bound*

$$L(n,m) = O(n) + (1 + o(1))\frac{m}{\log m}.$$

The above lemmas permit us to achieve the main result of this section.

Theorem 2.1 *If a partial boolean function f of n variables is defined on a 4-regular domain $\mathcal{A}$ of size m then, for any positive ϵ, we have:*

$$L(f,\epsilon) = O\left(\frac{m\epsilon^2}{\log(2 + m\epsilon^2)}\right) + O(n) .$$

Sketch of the proof. If ϵ is greater than some constant independent of n, then the thesis is an immediate consequence of Lemma 2.5. Consequently, we can assume that $\epsilon \leq \frac{1}{4}$. Let us consider an integer s such that $32m\epsilon^2 \leq 2^s \leq 64m\epsilon^2$. Since $(1/16)\ \epsilon^{-2} \geq 1$, the definition of s implies that $2\ 2^{-s}m + 1 \leq (1/8)\ \epsilon^{-2}$. We can apply Lemma 2.4 thus obtaining a linear operator $\vec{l} \in (\mathcal{L}_n)^s$ such that:

$$\sum_{\vec{d} \,:\, \Xi_{\mathcal{A},\vec{d}}(\vec{l}) \,>\, \frac{1}{8}\epsilon^{-2}} \Xi_{\mathcal{A},\vec{d}}(\vec{l}) \leq 2^s . \tag{8}$$

We can then construct the functions f_1 and f_2 as follows:

- For any $a \in \mathcal{A}$: $f_1(\vec{a}) = f(\vec{a})$ if $\Xi_{\mathcal{A},\vec{l}(\vec{a})} \leq \frac{1}{8}\epsilon^{-2}$ and undefined otherwise.

- For any $a \in \mathcal{A}$: $f_2(\vec{a}) = f(\vec{a})$ if $\Xi_{\mathcal{A},\vec{l}(\vec{a})} > \frac{1}{8}\epsilon^{-2}$ and undefined otherwise.

We now need to introduce another 'selector', that is the function $g : \{0,1\}^s \to \{0,1\}$ defined as $g(\vec{b}) = 1$ if $\Xi_{\mathcal{A},\vec{b}} \leq (1/8)\ \epsilon^{-2}$ and $g(\vec{b}) = 0$ otherwise. It is then easy to prove that the function f can be defined as $f(\vec{a}) = [f_1(\vec{a}) \wedge g(\vec{l}(\vec{a}))] \vee [f_2(\vec{a}) \wedge (\neg g(\vec{l}(\vec{a})))]$. Consequently, the following inequality holds:

$$L(f,\epsilon) \ \leq \ L(f_1,\epsilon) + L(f_2) + L(g) + L(\vec{l}) + O(1)\ . \tag{9}$$

Let $\mathcal{A}_1$ be the domain of function f_1. By definition, we have that $\Xi_{\mathcal{A}_1,\vec{b}} \ \leq \ \frac{1}{8}\epsilon^{-2}$, for any $\vec{b} \in \{0,1\}^s$. Consequently, by Lemma 2.3 we have $L(f_1,\epsilon) = O(\frac{2^s}{s}) + O(n)$. Moreover, Eq. (8) implies that the size of the domain of function f_2 is at most 2^s. Consequently, by Lemma 2.5 we have $L(f_2) \ = O(\frac{2^s}{s}) + O(n)$. From Lemma 2.5 it follows that $L(g) = O(\frac{2^s}{s}) + O(n)$. Concerning the circuit complexity of linear operators, we have (see [8] for a proof) $L(\vec{l}) = O(\frac{ns}{\log n}) = O(\frac{2^s}{s}) + O(n)$. Finally, by introducing the above bounds in Eq. (9), we obtain the thesis

$$L(f,\epsilon) \ = \ O\left(\frac{2^s}{s}\right) + O(n).$$

□

2.2 Upper bound for general domains

The aim of this section is to extend the upper bound for 4-regular domains stated in Theorem 2.1 to general domains. The main technical result consists in proving the existence of an injective function $\vec{\mathbf{J}}_n : \{0,1\}^n \to \{0,1\}^{c_1 n}$ (for some constant $c > 1$) having linear circuit complexity and such that the resulting subset $\vec{\mathbf{J}}_n(\{0,1\}^n)$ is 4-regular. The technique adopted here is still based on linear operators.

Lemma 2.6 *For any $n \geq 1$, a function $\vec{\mathbf{J}}_n : \{0,1\}^n \to \{0,1\}^{c_1 n}$ exists such that $L(\vec{\mathbf{J}}_n) \ \leq \ c_2\, n$ and the set $\vec{J}_n(\{0,1\}^n)$ is 4-regular, where c_1 and c_2 are some positive constants.*

Sketch of the proof. We first observe that if $\vec{\mathbf{J}}_n$ is injective then it is not hard to see that the set $\vec{\mathbf{J}}_n(\{0,1\}^n)$ is 4-regular if and only if for any different choice of $\langle \vec{a_1}, \vec{a_2}, \vec{a_3}, \vec{a_4} \rangle$ $(\vec{a_i} \neq \vec{a_j},\ i \neq j)$[4] we have: $\vec{\mathbf{J}}_n(\vec{a_1}) \oplus \vec{\mathbf{J}}_n(\vec{a_2}) \oplus \vec{\mathbf{J}}_n(\vec{a_3}) \oplus \vec{\mathbf{J}}_n(\vec{a_4}) \neq \vec{0}$. Furthermore, it is possible to prove that, for any $n \geq 1$, a linear operator $\vec{\mathbf{Q}}_n = (Q_1, \ldots, Q_{6n}) \in (\mathcal{L}_n)^{6n}$ exists such that: *i)* $\vec{\mathbf{Q}}_n(\vec{0}) = \vec{0}$; *ii)* for any non zero vector $\vec{a} \in \{0,1\}^n$, $\left|\vec{\mathbf{Q}}_n(\vec{a})\right| \ \geq \ \delta n$, for some $\delta > 0$; *iii)* the circuit complexity of $\vec{\mathbf{Q}}_n$ is linear in n (more precisely $L(\vec{\mathbf{Q}}_n) \leq 10n$). The existence of $\vec{\mathbf{Q}}_n$ is shown in Appendix B (see Lemmas B.1 and B.2). Let $r > 0$ and consider a random index sequence $(i(1), i(2), ..., i(r))$ such that $1 \leq i(1) < i(2) < ... < i(r) \leq 6n$. If $\vec{a_1}, \vec{a_2} \in \{0,1\}^n$ $(\vec{a_1} \neq \vec{a_2})$, then it is easy to prove that

$$\mathbf{Pr}\left(Q_{i(j)}(\vec{a_1}) = Q_{i(j)}(\vec{a_2})\ ,\ j = 1,2,...,r\right) \ \leq \ \left(\begin{array}{c} 6n - \delta n \\ r \end{array}\right)\left(\begin{array}{c} 6n \\ r \end{array}\right)^{-1} \ \leq \ e^{-\frac{1}{6}\delta r}. \tag{10}$$

If g is a random boolean function of r variables then, for any $\langle \vec{a_1}, \vec{a_2}, \vec{a_3}, \vec{a_4} \rangle$, we have that

$$\mathbf{Pr}\left(\bigoplus_{k=1}^{4} g\left(Q_{i(1)}(\vec{a_k}), Q_{i(2)}(\vec{a_k}), ..., Q_{i(r)}(\vec{a_k})\right) = \ 0\right) \ \leq \ \left(\begin{array}{c} 4 \\ 2 \end{array}\right) e^{-\frac{1}{6}\delta r} + \frac{1}{2}\ .$$

[4] in the sequel, we will always consider 4-tple satisfying this condition

We can thus choose a positive constant r such that the above probability is bounded by $\frac{3}{4}$. We now apply the same reasoning on a 'vectorial' system having R components. Let $i(j_1, j_2)$ ($j_1 = 1, \ldots, R$ and $j_2 = 1, \ldots, r$) be a set of random sequences such that $1 \leq i(j,1) < \ldots < i(j,r) \leq 6n$ ($j = 1, \ldots, R$) and consider a set of R random boolean functions $g_j(x_1, \ldots, x_r)$ ($j = 1, \ldots, R$); notice that the random choices are made independently and uniformly. Hence, we can define the random boolean operator $\vec{G} = (G_1, \ldots, G_R)$ where, for any $\vec{a} \in \{0,1\}^n$, we have $G_j(\vec{a}) = g_j(Q_{i(j,1)}(\vec{a}), \ldots, Q_{i(j,r)}(\vec{a}))$. In this case, for any $\langle \vec{a_1}, \vec{a_2}, \vec{a_3}, \vec{a_4} \rangle$ we have that $\mathbf{Pr}\left(\bigoplus_{k=1}^{4} \vec{G}(\vec{a}_k) = \vec{0}\right) \leq (\frac{3}{4})^R$. Furthermore, if we choose $R = 16n$ we obtain the following inequalities

$$\sum_{\langle \vec{a_1}, \vec{a_2}, \vec{a_3}, \vec{a_4} \rangle} \mathbf{Pr}\left(\bigoplus_{k=1}^{4} \vec{G}(\vec{a}_k) \;=\; \vec{0}\right) \leq \left(\frac{3}{4}\right)^R (2^n)^4 \;\leq\; \left(\frac{3}{4}\right)^n \;\leq\; \frac{3}{4}\,.$$

Notice that Eq. 10 implies also the injectivity of $\vec{G}$. Consequently, an injective operator $\vec{G} : \{0,1\}^n \to \{0,1\}^{16n}$ exists such that, for any $\langle \vec{a_1}, \vec{a_2}, \vec{a_3}, \vec{a_4} \rangle$, we have $\bigoplus_{k=1}^{4} \vec{G}(\vec{a}_k) \neq \vec{0}$. This implies that the subset $\vec{G}(\{0,1\}^n)$ is 4-regular. Finally, we observe that the operator $\vec{G}$ has linear complexity since R is linear and r is a constant. The thesis is then proved by defining $\vec{\mathbf{J}}_n = \vec{G}$. □

Theorem 2.2 *For any $n > 0$, $0 < m \leq 2^n$ and $\epsilon > 0$, we have:*

$$L(n, m, \epsilon) \;=\; O\left(\frac{m\epsilon^2}{\log(2 + m\epsilon^2)}\right) + O(n), \quad \textit{thus } L_0(n, \epsilon) \;= O\left(\frac{2^n\epsilon^2}{\log(2 + 2^n\epsilon^2)}\right) + O(n).$$

Sketch of the proof. Let $f : \mathcal{A} \to \{0,1\}$ be a boolean function with $|\mathcal{A}| = m$. Consider the operator $\vec{\mathbf{J}}_n : \{0,1\}^n \to \{0,1\}^R$ defined in Lemma 2.6 where R is linear in n. Then, we define the partial boolean function $f^* : \{0,1\}^R \to \{0,1\}$ as follows: for any $\vec{b} \in \{0,1\}^R$, $f^*(\vec{b}) = f(\vec{a})$ if $\mathbf{J}_n(\vec{a}) = \vec{b}$ for some $\vec{a} \in \mathcal{A}$ and undefined otherwise. From Lemma 2.6, f^* has a 4-regular domain and, thus, by Theorem 2.1 we have (notice that the size of the domain of f^* is also equal to m):

$$L(f^*, \epsilon) \;= O\left(\frac{m\epsilon^2}{\log(2 + m\epsilon^2)}\right) + O(R) \;=\; O\left(\frac{m\epsilon^2}{\log(2 + m\epsilon^2)}\right) + O(n)\,. \tag{11}$$

Since $\vec{\mathbf{J}}_n$ is bijective from the f-domain to the f^*-domain, we have that $L(f, \epsilon) \;\leq\; L(f^*, \epsilon) + L(\vec{\mathbf{J}}_n)$. Finally, by (11) and Lemma 2.6 we obtain

$$L(f, \epsilon) \;= O\left(\frac{m\epsilon^2}{\log(2 + m\epsilon^2)}\right) + O(n).$$

□

2.3 Upper bound for the hardness function

The upper bound shown in Theorem 2.2 gives a non trivial upper bound for the *hardness function* introduced in [10]: $H(f) = \min\{l \;:\; App(f, l) \geq \frac{1}{l}\}$. We first observe that Theorem 2.2 implies that, for any $l \geq n^2$,

$$App_o(n, l) \;\geq\; \sqrt{\frac{l - O(n)}{2^n}}. \tag{12}$$

Then, for the 'general' hardness function $H(n) \;=\; \max_{f(x_1, x_2, \ldots, x_n)} H(f)$ we have the following result.

Corollary 2.1 $H(n) \leq 2^{n/3} + n^2 + O(1)$.

Sketch of the proof. From Eq. 12, for any boolean function $f(x_1, \ldots, x_n)$ we have that $App(f,l) \geq App_0(n,l) \geq \sqrt{(l - O(n))2^{-n}}$. Thus, for $l = 2^{n/3} + n^2$ and for almost every n, we obtain

$$App(f,l) \geq 2^{-n/3} \geq \frac{1}{l}.$$

Consequently, $H(f) \leq l$.

□

3 Lower bounds for $L_0(n,\epsilon)$ and $L(n,m,\epsilon)$

Theorem 3.1 *For any $n > 0$, $0 < m \leq 2^n$ and $0 < \epsilon \leq 1/2$, we have:*

1)

$$L(n,m,\epsilon) = \Omega\left(\frac{m\epsilon^2}{\log(2 + m\epsilon^2)}\right) + O(n) ;$$

2)

$$L_0(n,\epsilon) = \Omega\left(\frac{2^n\epsilon^2}{\log(2^n\epsilon^2)}\right) + \Omega(n) .$$

Sketch of the proof.
1) Observe first that if $m\epsilon^2 \leq n \log n$ then $(m\epsilon^2)(\log(2 + m\epsilon^2))^{-1} = O(n)$ and, thus, we obtain the thesis. We now assume that $m\epsilon^2 \geq n \log n$. Let $\mathcal{A} \subseteq \{0,1\}^n$ with $|\mathcal{A}| = m$, then from Stirling's formula, the number of partial boolean function defined on $\mathcal{A}$ which can be approximated with degree ϵ by a fixed circuit satisfies the following bound.

$$\sum_{k \leq \left(\frac{1}{2} - \epsilon\right)m} \binom{m}{k} \leq 2^m 2^{-c\epsilon^2 m} . \tag{13}$$

where c is some positive constant. Consequently, by Eq. (13), at least $(2^m)(2^m 2^{-c\epsilon^2 m})^{-1} = 2^{c\epsilon^2 m}$ circuits are required for approximating all boolean functions defined on $\mathcal{A}$. Furthermore, the number of circuits with size not bigger than l is at most $(c_1 n l)^{l+n}$, for some positive constant c_1 (see [7]). Consequently, the required number of circuits, with size not bigger than l, must satisfy the following inequality: $(c_1 n l)^{l+n} \geq 2^{c\epsilon^2 m}$ and we finally obtain

$$l = \Omega\left(\frac{m\epsilon^2}{\log(2 + m\epsilon^2)}\right) .$$

2) We first observe that $L_0(n,\epsilon) \geq n - 1$. Indeed, consider the boolean function $f = x_1 \oplus x_2 \oplus \ldots \oplus x_n$; if a boolean function g (and thus a circuit) has at most $n-1$ 'significative' variables, then $\mathbf{Pr}(f = g) = \frac{1}{2}$. The bound

$$L_0(n,\epsilon) = \Omega\left(\frac{2^n\epsilon^2}{\log(2^n\epsilon^2)}\right) + \Omega(n) .$$

is thus a consequence of the above fact and of the lower bound for $L(n,m,\epsilon)$ proved in part (1) of this proof.

□

References

[1] Alon N. and Spencer J.H. (1992), *The Probabilistic Method*, Wiley-Interscience Publication.

[2] Andreev A.E. (1985), "The universal principle of self-correction", Mat. Sbornik 127(169), N 6, 147-172 (in Russian). English transl. in Math. USSR Sbornik, 55 (1), 145-169 (1986).

[3] Andreev A.E. (1989), "On the complexity of the realization of partial Boolean functions by circuits of functional elements", Diskret. mat. 1, pp.36-45 (in Russian). English translation in: J. of Discrete Mathematics and Applications 1, 251-262 (1989).

[4] Boppana R. and Hirshfield (1989), "Pseudorandom generators and complexity classes", in *Randomness and Computation* (S. Micali Ed.), 5, 1-26, Adv. in Comput. Res., JAI Press.

[5] Blum M. and Micali S. (1984), "How to generate cryptographically strong sequences of pseudorandom bits", J. SIAM, 13(4), 850-864.

[6] Kearns M. and Vazirani U. (1994), *Topics in Computational Learning Theory*, MIT Press.

[7] Lupanov, O.B. (1965), "About a method circuits design – local coding principle", Problemy Kibernet. 10, 31-110 (in Russian). English Translation in Systems Theory Res., 10, 1963.

[8] Nechiporuk E.I. (1965), "About the complexity of gating circuits for the partial boolean matrix", Dokl. Akad. Nauk SSSR, 163, 40-42 (in Russian). English translation in Soviet Math. Docl.

[9] Nisan N. (1992), *Using Hard Problems to Create Pseudorandom Generators*, ACM Distinguished Dissertation, MIT Press.

[10] Nisan N. and Wigderson A. (1994), "Hardness vs Randomness", J. Comput. System Sci. 49, 149-167 (presented also at the 29th IEEE FOCS, 1988).

[11] Pippenger N. (1977), "Information theory and the complexity of Boolean functions", Math. Systems Theory 10, 129-167.

[12] Shannon, C.E. (1949), "The synthesis of two-terminal switching circuits", Bell. Syst. Tech. J. 28, 59-98.

[13] Yao A.C. (1982), "Theory and applications of trapdoor functions", Proc. of the 23th IEEE FOCS, 80-91.

Appendix

A Proof of lemmas in Section 2

- Linear functions

Lemma A.1 *For any set $\mathcal{A}$ of size m we have:*

$$\mathbf{E}(\Xi_{\mathcal{A}}) = \frac{m}{2} \text{ and } \mathbf{E}\left((\Xi_{\mathcal{A}} - \mathbf{E}(\Xi_{\mathcal{A}}))^2\right) = \frac{m}{4}. \tag{14}$$

Moreover, if $\mathcal{A}$ is 4-regular, we then have

$$\mathbf{E}\left((\Xi_{\mathcal{A}} - \mathbf{E}(\Xi_{\mathcal{A}}))^4\right) = \frac{3m^2}{16} - \frac{m}{8}, \tag{15}$$

$$\mathbf{Pr}\left((\Xi_{\mathcal{A}} - \mathbf{E}(\Xi_{\mathcal{A}}))^2 < \frac{1}{2}\mathbf{E}\left((\Xi_{\mathcal{A}} - \mathbf{E}(\Xi_{\mathcal{A}}))^2\right)\right) \leq \frac{47}{48}. \tag{16}$$

Proof.
- For any $\vec{a} \in \mathcal{A}$ we have $\mathbf{Pr}(\xi_{\vec{a}} = 0) = \mathbf{Pr}(\xi_{\vec{a}} = 1) = \frac{1}{2}$ and also

$$\mathbf{E}(\xi_{\vec{a}}) = \frac{1}{2}, \qquad \mathbf{E}\left((\xi_{\vec{a}} - \mathbf{E}(\xi_{\vec{a}}))^2\right) = \frac{1}{4}. \tag{17}$$

This implies $\mathbf{E}(\Xi_{\mathcal{A}}) = \frac{m}{2}$. Observe that if $\vec{a}, \vec{b} \in \mathcal{A}$ and $\vec{a} \neq \vec{b}$ then the random functions $\xi_{\vec{a}}$ and $\xi_{\vec{b}}$ are independent. Then, Eq. (17) implies that

$$\mathbf{E}\left((\Xi_{\mathcal{A}} - \mathbf{E}(\Xi_{\mathcal{A}}))^2\right) = \sum_{\vec{a} \in \mathcal{A}} \mathbf{E}\left((\xi_{\vec{a}} - \mathbf{E}(\xi_{\vec{a}}))^2\right) = \sum_{\vec{a} \in \mathcal{A}} \frac{1}{4} = \frac{m}{4}.$$

- Observe that if any four pairwise distinct random functions in the sum $\Xi_{\mathcal{A}}$ are independent then Eq. (15) holds (see for example [1])). Consequently, our next step consists in proving that if a set $\mathcal{A}$ is 4-regular then, for any choice of four pairwise distinct elements $\vec{a}_1, \ldots, \vec{a}_4 \in \mathcal{A}$, the corresponding random functions $\xi_{\vec{a}_1} \ldots \xi_{\vec{a}_4}$ are independent. Since these functions are binary, the probability of the event $\xi_{\vec{a}_1} = u_1 \wedge \ldots \wedge \xi_{\vec{a}_4} = u_4$ is equal to the portion of those linear functions $\alpha_1 x_1 \oplus \ldots \oplus \alpha_n x_n \oplus \beta$ such that the following equation is true

$$\begin{pmatrix} \vec{a}_1 & 1 \\ \vec{a}_2 & 1 \\ \vec{a}_3 & 1 \\ \vec{a}_4 & 1 \end{pmatrix} \times \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_n \\ \beta \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} \tag{18}$$

Since $\mathcal{A}$ is 4-regular it follows that the four vectors $\langle \vec{a}_1, 1 \rangle, \ldots, \langle \vec{a}_4, 1 \rangle$ are linear independent and, consequently, the portion of linear functions satisfying Eq. (18) is equal to $\frac{1}{16}$, i.e.

$$\mathbf{Pr}(\xi_{\vec{a}_1} = u_1 \wedge \ldots \wedge \xi_{\vec{a}_4} = u_4) = \frac{1}{16} = \mathbf{Pr}(\xi_{\vec{a}_1} = u_1) \times \mathbf{Pr}(\xi_{\vec{a}_2} = u_2) \times \mathbf{Pr}(\xi_{\vec{a}_3} = u_3) \times \mathbf{Pr}(\xi_{\vec{a}_4} = u_4).$$

This proves Eq. (15).

- Consider the values $\Theta = (\Xi_{\mathcal{A}} - \mathbf{E}(\Xi_{\mathcal{A}}))^2$ and $\alpha = 1 - \mathbf{Pr}\left(\Theta < \frac{1}{2}\mathbf{E}(\Theta)\right)$. If $\alpha > \frac{1}{4}$ then Eq. 16 is true. In the sequel we then assume that $\alpha \leq \frac{1}{4}$. Then Eq. 14 implies that $\Theta = (\Xi_{\mathcal{A}} - \frac{m}{2})^2$. It follows that the values of Θ can be written either as i^2 in the case of m even or as $i^2/4$ in the case of m odd, for some integer i. Then, for any positive integer i, we define $V(i) = i^2$ in the 'even' case and $V(i) = \frac{i^2}{4}$ otherwise. Moreover, consider the probability $\mathbf{p}(i) = \mathbf{Pr}(\Theta = V(i))$. We obtain $\mathbf{E}(\Theta) = \sum_{i=1}^{\infty} \mathbf{p}(i)V(i)$. Consider the following 'partition' of $\mathbf{E}(\Theta)$:

$$\Sigma_1 = \sum_{i\,:\,1 \leq V(i) < S} \mathbf{p}(i)V(i)\ ,\ \Sigma_2 = \sum_{i\,:\,S \leq V(i) < N} \mathbf{p}(i)V(i) \text{ and } \Sigma_3 = \sum_{i\,:\,V(i) \geq N} \mathbf{p}(i)V(i)$$

where $S = \frac{1}{2}\mathbf{E}(\Theta)$ and $N = \frac{1}{4\alpha}\mathbf{E}(\Theta)$. By the definition of α, we obtain that $\mathbf{Pr}(\Theta < S) = 1 - \alpha$ and thus $\Sigma_1 \leq (1-\alpha)S = \left(\frac{1}{2} - \frac{\alpha}{2}\right)\mathbf{E}(\Theta)$ and also $\mathbf{Pr}(S \leq \Theta < N) \leq \mathbf{Pr}(\Theta \geq S) = 1 - \mathbf{Pr}(\Theta < S) = \alpha$. Consequently, we have

$$\Sigma_2 \leq \alpha N = \alpha \frac{1}{4\alpha}\mathbf{E}(\Theta) = \frac{1}{4}\mathbf{E}(\Theta)\ .$$

Since $\mathbf{E}(\Theta) = \Sigma_1 + \Sigma_2 + \Sigma_3$ we have that $\Sigma_3 \geq \frac{1}{4}\mathbf{E}(\Theta)$ and thus

$$\mathbf{E}(\Theta^2) \geq \sum_{i\,:\,V(i) \geq N} p(i)V(i)^2 \geq N \sum_{i\,:\,V(i) \geq N} p(i)V(i) = N\Sigma_3 \geq N\frac{1}{4}\mathbf{E}(\Theta) = \frac{1}{16\alpha}(\mathbf{E}(\Theta))^2\ .$$

We have just proved that $\mathbf{E}(\Theta) = \frac{m}{4}$ and $\mathbf{E}\left((\Theta)^2\right) = \frac{3m^2}{16} - \frac{m}{8} \leq \frac{3m^2}{16}$. Finally, since $\frac{3m^2}{16} \geq \frac{1}{16\alpha}(\frac{m}{4})^2$ we obtain $\alpha \geq \frac{1}{48}$ which proves Eq. (16). □

- Linear operators

In the proof of Lemma 2.3, we use the fact that the 'selector' operator $\vec{F}_{\vec{\mathcal{Q}}}$ has circuit complexity $O(\frac{2^s}{s})$. The following lemma provide a formal proof of this fact. Let $\vec{\mathcal{Q}} = (\mathcal{Q}_1, \ldots, \mathcal{Q}_k)$ be a system of pairwise disjoint subsets of $\{0,1\}^s$ (i.e. a partition of some subset of $\{0,1\}^s$). We introduce the boolean operator $\vec{F}_{\vec{\mathcal{Q}}}(x_1 \ldots, x_s) = (f_{\vec{\mathcal{Q}},1}, \ldots, f_{\vec{\mathcal{Q}},k})$, where $f_{\vec{\mathcal{Q}},i}(\vec{b})$ is equal to 1 if $\vec{b} \in \mathcal{Q}_i$ and 0 otherwise.

Lemma A.2 *If $\vec{\mathcal{Q}}$ is a partition of some subset of $\{0,1\}^s$ and a constant C $(0 < C < 1)$ exists such that*

$$|\mathcal{Q}_i| \geq C \sum_{j=i+1}^{k} |\mathcal{Q}_j|\ , \qquad i = 1, 2, ..., k-1\ , \tag{19}$$

then $L(\vec{F}_{\vec{\mathcal{Q}}}) \leq O\left(\frac{2^s}{s}\right)$.

Proof. Let g_i denote the restriction of function $f_{\vec{\mathcal{Q}},i}$ on the set $\cup_{j=i}^{k}\mathcal{Q}_j$ and define also the functions $h_i = \vee_{j=1}^{i} f_{\vec{\mathcal{Q}},j}$. We have that $f_{\vec{\mathcal{Q}},i} = (\neg h_{i-1}) \wedge g_i$ and, consequently, we can compute the operator $\vec{F}_{\vec{\mathcal{Q}}}$ using the following recursion. Let $h_1 = f_{\vec{\mathcal{Q}},1}$ then

$$\begin{aligned}
f_{\vec{\mathcal{Q}},2} &= (\neg h_1) \wedge g_2 \qquad & h_2 &= h_1 \vee f_{\vec{\mathcal{Q}},2}\ ;\\
f_{\vec{\mathcal{Q}},3} &= (\neg h_2) \wedge g_3\ , \qquad & h_3 &= h_2 \vee f_{\vec{\mathcal{Q}},3}\ ;\\
&\ldots & &\\
f_{\vec{\mathcal{Q}},k} &= (\neg h_{k-1}) \wedge g_k\ , \qquad & h_k &= h_{k-1} \vee f_{\vec{\mathcal{Q}},k}\ .
\end{aligned}$$

Observe that this construction immediately implies that $L(\vec{F}_{\vec{\mathcal{Q}}}) \leq O(k) + \sum_{i=1}^{k} L(g_k)$. If m_i denotes the size of domain of g_i, we have that $m_i = \sum_{j=i}^{k} |\mathcal{Q}_j|$ and condition (19) implies that $|\mathcal{Q}_1| \geq C^{i-1} \sum_{j=i}^{k} |\mathcal{Q}_j|$ or, equivalently, $m_i \leq 2^s C^{1-i}$. It follows that $k = O(s)$ since $1 \leq m_k \leq 2^s C^{1-k}$.

By Lemma 2.5 we finally have

$$L(\vec{F}_{\vec{\mathcal{Q}}}) \leq \sum_{i=1}^{k} \left(\frac{2^s C^{1-i}}{\log(2^s C^{1-i})} + O(s) \right) = O\left(\frac{2^s}{s}\right).$$

□

For any $\vec{a} \in \{0,1\}^n$ and for any $\vec{d} \in \{0,1\}^s$, consider now the 'agreement' function $\xi_{\vec{a},\vec{d}} : (\mathcal{L}_n)^s \to \{0,1\}$ defined as $\xi_{\vec{a},\vec{d}}(\vec{l}) = \Pi_{i=1}^{s}(l_i(\vec{a}) \oplus d_i \oplus 1)$, where $\vec{l} = (l_1, l_2, ..., l_s)$. Consider also the sum of points in $\{0,1\}^n$ on which $\vec{l}$ is equal to a fixed value $\vec{d}$:

$$\Xi_{\mathcal{A},\vec{d}}(\vec{l}) = \sum_{\vec{a} \in \mathcal{A}} \xi_{\vec{a},\vec{d}}(\vec{l}) .$$

In order to prove Lemma 2.4, we will use the following preliminary result which can be proved in a way similar to that for Eq. 2 of Lemma A.1.

Lemma A.3 *Let $\mathcal{A} \in \{0,1\}^n$ of size m; then we have*

$$\mathbf{E}\left(\Xi_{\mathcal{A},\vec{d}}\right) = 2^{-s}m\,, \quad \mathbf{E}\left(\left(\Xi_{\mathcal{A},\vec{d}} - \mathbf{E}\left(\Xi_{\mathcal{A},\vec{d}}\right)\right)^2\right) = (1-2^{-s})2^{-s}m\,.$$

Lemma A.4 *If the set $\mathcal{A}$ of size m is 4-regular and $s \geq 1$ then a linear operator $\vec{l} \in (\mathcal{L}_n)^s$ exists such that*

$$\sum_{\vec{d}\,:\,\Xi_{\mathcal{A},\vec{d}}(\vec{l}) \,>\, 2\ 2^{-s}m+1} \Xi_{\mathcal{A},\vec{d}}(\vec{l}) \leq 2^s .$$

Proof. Consider the set $\mathcal{A}(\vec{a}) = \mathcal{A} - \{a\}$ and define $(\mathcal{L}_n)^s(\vec{a},\vec{d}) = \{\vec{l} \in (\mathcal{L}_n)^s : \xi_{\vec{a},\vec{d}}(\vec{l}) = 1\}$. For any $\vec{b} \in \mathcal{A}(\vec{a})$ we denote as $\xi^{\vec{a}}_{\vec{b},\vec{d}}$ the random function $\xi_{\vec{b},\vec{d}}$ restricted to the subspace $(\mathcal{L}_n)^s(\vec{a},\vec{d})$. We also define the sum $\Xi^{\vec{a}}_{\mathcal{A},\vec{d}} = \sum_{\vec{b} \in \mathcal{A}(\vec{a})} \xi^{\vec{a}}_{\vec{b},\vec{d}}$. Since for any $\vec{b} \in \mathcal{A}(\vec{a})$ the random functions $\xi_{\vec{a},\vec{d}}$, $\xi_{\vec{b},\vec{d}}$ are independent, we have that $\mathbf{E}\left(\xi^{\vec{a}}_{\vec{b},\vec{d}}\right) = \mathbf{E}\left(\xi_{\vec{b},\vec{d}}\right) = 2^{-s}$ and $\mathbf{E}\left(\left(\xi^{\vec{a}}_{\vec{b},\vec{d}} - \mathbf{E}\left(\xi^{\vec{a}}_{\vec{b},\vec{d}}\right)\right)^2\right) = \mathbf{E}\left(\left(\xi_{\vec{b},\vec{d}} - \mathbf{E}\left(\xi_{\vec{b},\vec{d}}\right)\right)^2\right) = (1-2^{-s})2^{-s}$. Observe also that $\mathbf{E}\left(\Xi^{\vec{a}}_{\mathcal{A},\vec{d}}\right) = 2^{-s}(m-1)$.

Since $\mathcal{A}$ is 4-regular, for any different $\vec{b_1}, \vec{b_2} \in \mathcal{A}(\vec{a})$, the random functions $\xi_{\vec{a},\vec{d}}$, $\xi_{\vec{b_1},\vec{d}}$ and $\xi_{\vec{b_2},\vec{d}}$ are independent and, thus, also the random functions $\xi^{\vec{a}}_{\vec{b_1},\vec{d}}$ and $\xi^{\vec{a}}_{\vec{b_2},\vec{d}}$ are independent. Consequently, the function $\Xi^{\vec{a}}_{\mathcal{A},\vec{d}}$ satisfies the following equation

$$\mathbf{E}\left(\left(\Xi^{\vec{a}}_{\mathcal{A},\vec{d}} - \mathbf{E}\left(\Xi^{\vec{a}}_{\mathcal{A},\vec{d}}\right)\right)^2\right) = \sum_{\vec{b} \in \mathcal{A}(\vec{a})} \mathbf{E}\left(\left(\xi^{\vec{a}}_{\vec{b},\vec{d}} - \mathbf{E}\left(\xi^{\vec{a}}_{\vec{b},\vec{d}}\right)\right)^2\right) = (1-2^{-s})2^{-s}(m-1)$$

which implies that

$$\mathbf{Pr}\left(\left|\Xi^{\vec{a}}_{\mathcal{A},\vec{d}} - 2^{-s}(m-1)\right| \geq 2^{-s}m\right) = \mathbf{Pr}\left(\left|\Xi^{\vec{a}}_{\mathcal{A},\vec{d}} - \mathbf{E}\left(\Xi^{\vec{a}}_{\mathcal{A},\vec{d}}\right)\right| \geq 2^{-s}m\right) \leq \frac{\mathbf{E}\left(\left(\Xi^{\vec{a}}_{\mathcal{A},\vec{d}} - \mathbf{E}\left(\Xi^{\vec{a}}_{\mathcal{A},\vec{d}}\right)\right)^2\right)}{\left(2^{-s}m\right)^2} \leq \frac{1}{2^{-s}m}.$$

Thus, $\mathbf{Pr}\left(|\Xi_{\mathcal{A},\vec{l}(\vec{a})} - 2^{-s}m| \geq 2^{-s}m + 1\right) \leq \frac{1}{2^{-s}m}$. We now define the function $\theta_{\vec{a}}(\vec{l}) = 1$ if $\Xi_{\mathcal{A},\vec{l}(\vec{a})} \leq 2\ 2^{-s}m + 1$ and $\theta_{\vec{a}}(\vec{l}) = 0$ otherwise. We then have that $\mathbf{E}(\theta_{\vec{a}}) \leq \frac{1}{2^{-s}m}$ and thus $\mathbf{E}\left(\sum_{a\in\mathcal{A}} \theta_{\vec{a}}\right) \leq 2^s$. It follows that, for any linear operator $\vec{l}$, we have

$$\sum_{\vec{a}\in\mathcal{A}} \theta_{\vec{a}}(\vec{l}) \ = \sum_{\vec{d}\ :\ \Xi_{\mathcal{A},\vec{d}}\ >\ 2\ 2^{-s}m+1} \Xi_{\mathcal{A},\vec{d}}.$$

This immediately implies that there exists at least one linear operator $\vec{l}$ for which

$$\sum_{\vec{d}\ :\ \Xi_{\mathcal{A},\vec{d}}\ >\ 2\ 2^{-s}m+1} \Xi_{\mathcal{A},\vec{d}}(\vec{l}) \ \leq \ 2^s \ .$$

□

B Proof of lemmas in Section 2.2

In the proof of Lemma 2.6, we have assumed the existence of the linear operator $\vec{\mathbf{Q}}_n$. In this section, we provide a formal proof of this fact. Let us introduce the uniform probability function on the set Γ_n of all permutations of n elements. If $\gamma \in \Gamma_n$ and $\vec{a} = \langle a_1, \ldots, a_n\rangle \in \{0,1\}^n$, we can define the new vector: $\gamma(\vec{a}) = (a_{\gamma(1)}, a_{\gamma(2)}, ..., a_{\gamma(n)})$. We introduce the following linear operator $\vec{H}^n(x_1, \ldots, x_n) = (H_1, \ldots, H_n)$ where $H_i(x_1, x_2, ..., x_n) = \oplus_{j=1}^{i} x_j$. In the sequel, the terms $|\vec{a}|$ will denote the number of 1's in $\vec{a}$.

Lemma B.1 *Let $\vec{a} \in \{0,1\}^n$ such that $|\vec{a}| = k \geq 1$ and let $t \leq \frac{n}{16}$; then*

$$\mathbf{Pr}\left(\left|\vec{H}^n(\gamma(\vec{a}))\right| \ \leq \ t\right) \ \leq \ \left(\frac{8t}{n}\right)^{k/2}.$$

Proof. Observe first that

$$\mathbf{Pr}\left(\left|\vec{H}^n(\gamma(\vec{a}))\right| \ = \ r\right) \ = \ \frac{\left|\left\{\vec{b} \in \{0,1\}^n \ : \ \left|\vec{b}\right| = k \text{ and } \left|\vec{H}^n(\vec{b})\right| = r\right\}\right|}{\binom{n}{k}}. \tag{20}$$

Let $\vec{b} \in \{0,1\}^n$ such that $|\vec{b}| = k \geq 1$ and assume that $b_{i(1)} = \ldots = b_{i(k)} = 1$, where $i(1) < \ldots < i(k)$. It is not hard to verify that, for k even, we have $|\vec{H}^n(\vec{b})| = \sum_{j=1}^{k/2}(i(2j) - i(2j-1))$ and, for k odd, $|\vec{H}^n(\vec{b})| = n + 1 - i(k) + \sum_{j=1}^{(k-1)/2}(i(2j) - i(2j-1))$.

If k is even we can then choose the integer numbers $1 \leq i(1) < i(3) < \ldots < i(2k-1) \leq n$ in at most $\binom{n}{k/2}$ distinct ways. If $\left|\vec{H}^n(\vec{b})\right| \leq \ t$ we then have that, for any fixed sequence $1 \leq i(1) < i(3) < ..., < i(2k-1) \leq n$, the number of choices for the numbers $i(2), i(4), \ldots, i(2k)$ is bounded by the number of positive integer sequences $n_1, n_2, ..., n_{k/2}$ satisfying the condition $\sum_{j=1}^{k/2} n_j \ \leq \ t$. The number of such sequences is at most $\binom{t}{k/2}$. From the above facts we obtain (here we omit some computations):

$$\left|\left\{\vec{b}\in\{0,1\}^n \ : \ \left|\vec{b}\right|=k \ , \ \left|\vec{H}^n(\vec{b})\right|\le t\right\}\right| \le \left(\begin{array}{c} n \\ k/2 \end{array}\right)\left(\begin{array}{c} t \\ k/2 \end{array}\right) \le \left(\begin{array}{c} n \\ k \end{array}\right)\left(\frac{8t}{n}\right)^{k/2}.$$

By applying the same counting arguments, we have the following upper bound for the case of k odd:

$$\left(\begin{array}{c} n \\ k \end{array}\right)\left(\frac{8t}{n}\right)^{(k+1)/2} \le \left(\begin{array}{c} n \\ k \end{array}\right)\left(\frac{8t}{n}\right)^{k/2}.$$

From Eq. (20), the lemma is proved. □

Lemma B.2 *Let $\mathcal{L}_n^0$ denote the set of all linear function $l \in \mathcal{L}_n$ such that $l(\vec{0}) = 0$. Then, for any $n \ge 1$, there exists a positive constant δ such that a linear operator $\vec{\mathbf{Q}}_n \in (\mathcal{L}_n^0)^{6n}$ exists such that $|\vec{\mathbf{Q}}_n(\vec{a})| \ge \delta n$ for any vector $\vec{a} \ne \vec{0}$. Furthermore, $L(\vec{\mathbf{Q}}_n) \le 10n$.*

Proof. Suppose, that $\gamma_1, \gamma_2, \ldots, \gamma_5$ are permutations selected at random and independently from Γ_n and γ is, instead, a permutation selected at random from Γ_{5n}. For any positive constant $\delta \le \frac{1}{256}$, define $k_0 = \lfloor \delta n \rfloor$ and, for any $\vec{a} \in \{0,1\}^n$, consider

$$p(\vec{a},k) \ = \ \mathbf{Pr}\left(\left|\vec{H}^n(\gamma_i(\vec{a}))\right| \ < \ \frac{1}{16}k\sqrt{\frac{n}{k}} \text{ ,for any } i=1,2,3,4,5\right) \ , \ k=1,\ldots,k_0 \ .$$

We use the notation $p(k)$ for any $\vec{a}$ with $|\vec{a}| = k$ since in this case the above probability depends only on k. By Lemma B.1, we have (here we omit some computation)

$$p(k) \le \left(\left(\frac{8\frac{1}{16}k\sqrt{\frac{n}{k}}}{n}\right)^{k/2}\right)^5 \le \left(\begin{array}{c} n \\ k \end{array}\right)^{-1}\left(\frac{k}{4n}\right)^{k/4}.$$

Consequently,

$$\sum_{k=1}^{k_0}\left(\begin{array}{c} n \\ k \end{array}\right)p(k) \le \sum_{k=1}^{k_0}\left(\frac{k_0}{4n}\right)^{k/4} \le \sum_{k=1}^{k_0}\left(\frac{1}{256}\right)^{k/4} \le \frac{1}{3}.$$

It follows that there exists at least one sequence of five permutations $\gamma_1, \gamma_2, ..., \gamma_5$ such that, for any $\vec{a} \in \{0,1\}^n$ with $|\vec{a}| = k$ $(1 \le k \le k_0)$, there is at least one index $i \in \{1,\ldots,5\}$ for which $|\vec{H}_n(\gamma_i(\vec{a}))| \ge \frac{1}{16}k\sqrt{\frac{n}{k}}$. It follows that $|\vec{Z}(\vec{a})| \ge \frac{1}{16}k\sqrt{\frac{n}{k}}$ where $\vec{Z}(\vec{a}) = (\vec{H}^n(\gamma_1(\vec{a})), \vec{H}^n(\gamma_2(\vec{a})), \ldots, \vec{H}^n(\gamma_5(\vec{a})))$. Since $\frac{1}{16}k\sqrt{\frac{n}{k}} \le \frac{1}{16}k_0\sqrt{\frac{n}{k_0}} \le k_0$, Lemma B.1 implies that

$$\mathbf{Pr}\left(\left|\vec{H}^{5n}(\gamma(\vec{Z}(\vec{a})))\right| \ \le \ k_0\right) \ \le \ \left(\frac{8k_0}{5n}\right)^{\frac{1}{32}k\sqrt{\frac{n}{k}}} \ \le \ \left(\frac{8\delta}{5}\right)^{\frac{1}{32}k\sqrt{\frac{n}{k}}}.$$

We can then choose a constant δ such that $\left(\frac{8\delta}{5}\right)^{\frac{1}{32}k\sqrt{\frac{n}{k}}} \le \left(\begin{array}{c} n \\ k \end{array}\right)^{-1} 4^{-k}$. It follows that

$$\sum_{k=1}^{k_0}\mathbf{Pr}\left(\exists\vec{a}\in\{0,1\}^n \ : \ |\vec{a}|=k \ , \ \left|\vec{H}^{5n}(\gamma(\vec{Z}(\vec{a})))\right| \ \le \ k_0\right) \ \le \ \sum_{k=1}^{k_0}4^{-k} \ \le \ \frac{1}{3} \ .$$

Consequently, there exists at least one permutation $\gamma \in \Gamma_{5n}$ such that, for any $\vec{a} \in \{0,1\}^n$ with $|\vec{a}| = k$ $(1 \le k \le k_0)$, we have $|\vec{H}^{5n}(\gamma(\vec{Z}(\vec{a})))| \ge k_0$. Finally, we define

$$\vec{\mathbf{Q}}_n(\vec{a}) \;=\; \left(\vec{a}, \vec{H}^{5n}\left(\gamma\left(\vec{H}^n(\gamma_1(\vec{a})), \vec{H}^n(\gamma_2(\vec{a})), ..., \vec{H}^n(\gamma_5(\vec{a}))\right)\right)\right),$$

thus obtaining $\left|\vec{\mathbf{Q}}_n(\vec{a})\right| \geq \delta n$ for any $\vec{a} \in \{0,1\}^n$. Observe that the upper bound $10n$ for $L(\vec{\mathbf{Q}}_n)$ is an immediate consequence of the construction of $\vec{\mathbf{Q}}_n$. □