On Multipartition Communication Complexity*

Pavol Ďuriš[1,2] Juraj Hromkovič[1] Stasys Jukna[3]
Martin Sauerhoff[4] Georg Schnitger[3]

[1] Lehrstuhl für Informatik I, RWTH Aachen,
Ahornstraße 55, 52074 Aachen, Germany.
jh@i1.informatik.rwth-aachen.de

[2] Department of Informatics, Comenius University,
Mlynska dolina, 84215 Bratislava, Slovakia.

[3] Fachbereich Informatik, Johann Wolfgang Goethe-Universität Frankfurt,
Robert-Mayer-Straße 11–15, 60054 Frankfurt am Main, Germany.
{georg;jukna}@thi.informatik.uni-frankfurt.de

[4] FB Informatik, LS 2, Universität Dortmund,
44221 Dortmund, Germany.
sauerhof@ls2.cs.uni-dortmund.de

Abstract. We study *k-partition communication protocols*, an extension of the standard two-party best-partition model to k input partitions. The main results are as follows.

1. A strong explicit *hierarchy* on the degree of non-obliviousness is established by proving that, using $k+1$ partitions instead of k may decrease the communication complexity from $\Theta(n)$ to $\Theta(\log k)$.
2. Certain linear codes are hard for k-partition protocols even when k may be exponentially large (in the input size). On the other hand, one can show that all characteristic functions of linear codes are *easy* for randomized OBDDs.
3. It is proven that there are subfunctions of the *triangle-freeness function* and the function $\oplus\,\textsc{Clique}_{n,3}$ which are hard for multipartition protocols. As an application, truly exponential lower bounds on the size of nondeterministic read-once branching programs for these functions are obtained, solving an open problem of Razborov [19].

Keywords: computational complexity, rectangular complexity, multipartition communication complexity, non-obliviousness, lower bounds, complexity hierarchy, read-once branching programs.

*The work of the first and second author has been supported by DFG grant Hr 14/3-2, and of the fourth author by DFG grant We 1066/9-1.

ISSN 1433-8092

1 Introduction

The communication complexity of two-party protocols was introduced by Yao [22]. The initial goal was to develop a method for proving lower bounds on the complexity of distributed and parallel computations. In the meantime, communication complexity has been successfully applied as a tool for proving lower bounds in various other models of computation (see, e. g., [9, 14] for a survey).

Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function defined on a set X of n Boolean variables, and let $\Pi = (X_1, X_2)$ be a balanced partition of X, i. e., a partition with $-1 \leqslant |X_1| - |X_2| \leqslant 1$. A *deterministic two-party communication protocol P for f according to* Π is an algorithm by which two players, called Alice and Bob, can evaluate f as follows. At the beginning of the computation, Alice obtains an input $x\colon X_1 \to \{0,1\}$ and Bob an input $y\colon X_2 \to \{0,1\}$. Then the players communicate according to P by exchanging messages. The players may use unbounded resources to compute their messages. At the end, one of them has to output $f(x, y)$. A *nondeterministic protocol* allows each player to access a (private) string of *nondeterministic bits* as an additional input. It is required that there is an assignment to the nondeterministic bits such that the protocol outputs 1 if and only if $f(x, y) = 1$.

The *complexity of a nondeterministic protocol P* is the maximum of the number of exchanged bits taken over all inputs, including the nondeterministic bits. The *nondeterministic communication complexity of f according to* Π, *ncc* (f, Π), is the minimum complexity of a nondeterministic protocol according to Π which computes f. Finally, the *(best-partition) nondeterministic communication complexity of f, ncc* (f), is defined as the minimum of *ncc* (f, Π) over all balanced partitions Π of the set of input variables of f.

A protocol is *oblivious* because it uses only one partition of the set of input variables for all inputs. Most applications of communication complexity are therefore restricted to oblivious models of computation. However, Borodin, Razborov, and Smolensky [5] succeeded in deriving exponential lower bounds for the non-oblivious model of computation of (syntactic) read-k-times branching programs. Their approach leads, from the perspective of communication protocols, to the following notion of *multipartition communication protocols* [10]:

Definition 1. Let f be a Boolean function defined on a set X of Boolean variables, and let k be a positive integer. A *k-partition protocol P for f* is a collection of k nondeterministic (sub-)protocols $P_1, \dots, P_k$, each P_i with its own balanced partition of X, such that $f = P_1 \vee P_2 \vee \cdots \vee P_k$, where we use P_i also to denote the function computed by protocol P_i. If m_i is the number of all-1 submatrices of P_i (i. e., m_i is the number of 1-leaves in the protocol tree of P_i), then the *complexity of P* is $\lceil \log\bigl(\sum_{i=1}^{k} m_i\bigr) \rceil$. The *$k$-partition communication complexity of f, k-pcc* (f), is the minimum complexity of a k-partition protocol computing f. The *multipartition communication complexity of f* is *mpcc* $(f) := \min\{k\text{-}pcc\,(f) \mid k \in \mathbb{N}\}$.

To better understand the model of multipartition communication, we compare *mpcc* (f) with the best-partition nondeterministic communication complexity *ncc* (f). Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function, $A \subseteq f^{-1}(1)$, and let Π be a partition of the variables of f.

Define the distribution μ_A on $\{0,1\}^n$ by $\mu_A(x) := |A|^{-1}$ if $x \in A$, and $\mu_A(x) := 0$ otherwise. Define $B^1_{A,\Pi}(f) := \log\bigl(1/\max_M \mu_A(M)\bigr)$, where the maximum extends over all all-1 submatrices M of the communication matrix of f according to Π.

We have $\mathit{ncc}\,(f,\Pi) = \max_{A\subseteq f^{-1}(1)} B^1_{A,\Pi}(f) + O(\log n)$ by the proof of Theorem 2.16 in [14], and consequently

$$\mathit{ncc}\,(f) = \min_{\Pi} \max_{A\subseteq f^{-1}(1)} B^1_{A,\Pi}(f) + O(\log n),$$

where the minimum extends over all balanced partitions Π of the variables of f. A similar argument yields:

Lemma 1. *For every Boolean function $f\colon \{0,1\}^n \to \{0,1\}$,*

$$\mathit{mpcc}\,(f) = \max_{A\subseteq f^{-1}(1)} \min_{\Pi} B^1_{A,\Pi}(f) + O(\log n).$$

When dealing with multipartition communication complexity, the notion of rectangles as introduced by Borodin, Razborov, and Smolensky [5] is useful. Let X be a set of n variables and let $\Pi = (X_1, X_2)$ be a balanced partition of X. A function $r\colon \{0,1\}^n \to \{0,1\}$ defined on X is called a *rectangle (with respect to* Π*)* if it can be written as $r = r^1 \wedge r^2$, where the functions r^i depend only on variables from X_i, $i = 1, 2$. Given a Boolean function f defined on X, its *rectangle complexity* $R(f)$ is the minimal number t for which there exist t rectangles $r_1, r_2, \ldots, r_t$ (each with its own partition of the variables in X) such that $f = r_1 \vee r_2 \vee \cdots \vee r_t$. The *k-partition rectangle complexity* $R_k(f)$ of f is the minimal number of rectangles needed to cover f under the restriction that these rectangles may use at most k different partitions. Note that

$$R_k(f) = \min_{f_1, f_2, \ldots, f_k} R_1(f_1) + R_1(f_2) + \cdots + R_1(f_k),$$

where the minimum is taken over all k-tuples of Boolean functions $f_1, f_2, \ldots, f_k$ with $f_1 \vee f_2 \vee \cdots \vee f_k = f$. Furthermore, $R(f) = \min_k R_k(f)$. We obtain:

Proposition 1. *For all Boolean functions f,*

$$\lceil \log R_k(f) \rceil = k\text{-}\mathit{pcc}\,(f), \quad \textit{and} \quad \lceil \log R(f) \rceil = \mathit{mpcc}\,(f).$$

The measure $R(f)$ can also be used to prove lower bounds on the size of nondeterministic read-once branching programs (1-n.b.p. for short): Borodin, Razborov, and Smolensky [5] have shown that every Boolean function f requires a 1-n.b.p. of size at least $R(f)^{1/4}$. In fact this lower bound is $R(f)/(2n)$ for n-input functions f due to an observation of Okolnishnikova [17].

The goal of this paper is to develop lower bounds for the fundamental measures $\mathit{mpcc}\,(f)$ and $R(f)$, resp., and apply these results to branching programs. In the following, we give an overview on the paper.

1. In [10], an exponential gap between $ncc\,(f) = 1\text{-}pcc\,(f)$ and $2\text{-}pcc\,(f)$ has been shown. In Section 2 (Theorem 1), we prove that for infinitely many n and for all $k = k(n)$, there is an explicitly defined function $f_{k,n} \colon \{0,1\}^n \to \{0,1\}$ such that,

   $$k\text{-}pcc\,(f_{k,n}) = \Omega(n), \quad \text{and} \quad (k+1)\text{-}pcc\,(f_{k,n}) = O(\log k).$$

   Thus, a small increase of the degree of non-obliviousness can result in an unbounded decrease of communication complexity.

2. In Section 3, we observe that an argument from [11, 17] yields a *linear* lower bound on the multipartition communication complexity of the characteristic function of a *random* linear code. Moreover, $mpcc\,(\mathrm{BCH}_n) \geqslant \log R(\mathrm{BCH}_n) = \Omega\big(n^{1/2}\big)$ for the characteristic function of a BCH-code of length n and designed distance $d = 2t + 1$ with $t \approx n^{1/2}$ (Theorem 2).

   On the other hand, we prove that the characteristic function of the complement of a linear code can be computed by small *randomized OBDDs* with arbitrarily small one-sided error (Theorem 3). Thus we obtain the apparently best known tradeoff between randomized and nondeterministic branching program complexity.

3. In Section 4, we consider the problem of determining whether a given graph has no triangles. The corresponding *triangle-freeness function* Δ_n has $n = \binom{m}{2}$ Boolean variables (one for each potential edge) and accepts a given graph G on m vertices if and only if G has no triangles. We prove that there is a subfunction Δ'_n of Δ_n with $R(\Delta'_n) = 2^{\Omega(n)}$ (Theorem 4).

   Although this result does not imply a lower bound on the rectangle complexity (and thus the multipartition complexity) of the triangle-freeness function Δ_n itself, the result has an interesting consequence for nondeterministic read-once branching programs. Razborov ([19], Problem 11) asks whether a truly exponential lower bound holds for the function $\oplus\,\mathrm{CLIQUE}_{n,3}$ on $n = \binom{m}{2}$ variables which outputs the parity of the number of triangles in a graph on m vertices. In the case of *deterministic* read-once branching programs, such a lower bound for $\oplus\,\mathrm{CLIQUE}_{n,3}$ has been proven by Ajtai *et al.* in [2]. We solve this problem by proving that nondeterministic read-once branching programs for $\oplus\,\mathrm{CLIQUE}_{n,3}$ and for the triangle-freeness function Δ_n require size at least $2^{\Omega(n)}$. The only other truly exponential lower bounds for nondeterministic read-once programs have been proven for a class of functions based on quadratic forms in [3–5]. In the deterministic case, the recent celebrated result of Ajtai [1] gives a truly exponential lower bound for a function similar to $\oplus\,\mathrm{CLIQUE}_{n,3}$ even for linear time branching programs.

2 A Strong Hierarchy on the Degree of Non-Obliviousness

The goal of this section is to prove that allowing one more partition of the input variables can lead to an unbounded decrease of the communication complexity for explicitly defined functions.

Theorem 1. *For infinitely many n and all $k = k(n)$, there is an explicitly defined function $f_{k,n}\colon \{0,1\}^n \to \{0,1\}$ such that,*

$$k\text{-}pcc\,(f_{k,n}) = \Omega(n), \quad \text{and} \quad (k+1)\text{-}pcc\,(f_{k,n}) = O(\log k).$$

Furthermore, the upper bound can even be achieved by using $(k+1)$-partition protocols where each protocol is deterministic.

We describe how the functions used in the proof of Theorem 1 are constructed. The idea is to take some function h which is known to be "hard" even if *arbitrarily* many partitions are allowed. From h, a new function f_k is constructed which will be "easy" for $(k+1)$-partition protocols, but "hard" for k-partition protocols.

For $h\colon \{0,1\}^m \to \{0,1\}$, the respective function f_k is defined on vectors of variables $x = (x_1, \dots, x_{2m})$, $y = (y_0, \dots, y_{\ell-1})$, and $z = (z_0, \dots, z_{\ell-1})$, where $\ell := \lceil \log(k+1) \rceil$. We use a fixed set $\mathcal{P} = \{\Pi^*_1, \dots, \Pi^*_{k+1}\}$ of balanced partitions of the x-variables (described later on). For a given value i from $\{1, \dots, k+1\}$ represented by the y-variables, the vector x is divided into two halves $x^1(i)$, $x^2(i)$ of length m according to the partition Π^*_i. The function f_k is defined by $f_k(x, y, z) := h\bigl(x^1(i)\bigr)$. (Observe that the z-variables are only used for "padding" the input.)

It is obvious that f_k has $(k+1)$-partition protocols of small complexity:

Proof of Theorem 1 – Upper Bound. The protocol for f_k uses $k+1$ partitions which divide the x-vector according to the partitions in $\mathcal{P}$, and which give all y-variables to the first player and all z-variables to the second player. In the ith subprotocol, the first player outputs $h\bigl(x^1(i)\bigr)$ if i is the value represented by the y-variables, and 0 otherwise. The second player does nothing. The complexity of the whole protocol is obviously $\lceil \log(2(k+1)) \rceil = \lceil \log(k+1) \rceil + 1$. □

In the following, we can only give an outline of the proof of the lower bound. We first describe the main combinatorial idea. If we can ensure that all the sets occurring as halves of partitions in $\mathcal{P}$ (where $|\mathcal{P}| = k+1$) are "very different," then the partitions in $\mathcal{P}$ cannot be "approximated" by only k partitions, as the following lemma shows.

Lemma 2. *Define the (Hamming) distance between two sets $A, B \subseteq \{1, \dots, n\}$ by $d(A, B) := |A \cap \overline{B}| + |\overline{A} \cap B|$. Let $\mathcal{A}$ and $\mathcal{B}$ be families of subsets of $\{1, \dots, n\}$ with $|A| = n/2$ for all $A \in \mathcal{A}$, $D \leqslant d(A, A') \leqslant n - D$ for all different $A, A' \in \mathcal{A}$, and $\bigl||B| - n/2\bigr| \leqslant D/4$ for all $B \in \mathcal{B}$. If $|\mathcal{A}| \geqslant |\mathcal{B}| + 1$ then there exists an $A_0 \in \mathcal{A}$ such that $|A_0 \cap B| \geqslant D/8$ and $|A_0 \cap \overline{B}| \geqslant D/8$ for all $B \in \mathcal{B}$.*

Proof. We first show that there is an $A_0 \in \mathcal{A}$ such that $D/2 \leqslant d(A_0, B) \leqslant n - D/2$ for all $B \in \mathcal{B}$. Assume to the contrary that for each $A \in \mathcal{A}$ there is a $B \in \mathcal{B}$ such that $d(A, B) < D/2$ or $d(\overline{A}, B) = n - d(A, B) < D/2$. Since $|\mathcal{A}| \geqslant |\mathcal{B}| + 1$, the pigeonhole principle implies that there exists $B \in \mathcal{B}$ such that $d(S_1, B) < D/2$ *and* $d(S_2, B) < D/2$ for some $S_1 \in \{A_1, \overline{A_1}\}$, $S_2 \in \{A_2, \overline{A_2}\}$ and $A_1, A_2 \in \mathcal{A}$, $A_1 \neq A_2$. But then $d(S_1, S_2) \leqslant d(S_1, B) + d(B, S_2) < D$, a contradiction.

For any two sets A and B, we have $d(A, B) = |A| + |B| - 2|A \cap B|$. Thus, for the above A_0 and all $B \in \mathcal{B}$,

$$|A_0 \cap B| = \frac{1}{2}\left(|A_0| + |B| - d(A_0, B)\right) \geqslant \frac{1}{2}\left(\frac{n}{2} + \frac{n}{2} - \frac{D}{4} - \left(n - \frac{D}{2}\right)\right) = \frac{D}{8}.$$

Analogously, we get $|A_0 \cap \overline{B}| \geqslant D/8$ for all $B \in \mathcal{B}$. □

In order to meet the requirements of Lemma 2, we choose $\mathcal{P}$ such that the characteristic vectors of the Π_i^* form a code $C \subseteq \{0,1\}^{2m}$ with the following properties: (i) All $x \in C$ have exactly m ones and m zeros, i. e., C is a so-called *balanced code*. (ii) Any two different codewords have Hamming distance at least $D = 2\delta m$ and at most $2m - D = 2(1-\delta)m$, $\delta > 0$ a constant. To construct a code with these properties and exponentially many codewords, we start with a Justesen code (see, e. g., [15]), which is a linear code with appropriate lower *and* upper bounds on the weight of its codewords, and then “balance” the codewords by “padding.”

Let $\Pi_i^* = (\Pi_{i,1}^*, \Pi_{i,2}^*)$, for $i = 1, \ldots, k+1$. Let $\Pi_i = (\Pi_{i,1}, \Pi_{i,1})$, for $i = 1, \ldots, k$, be arbitrary balanced partitions. We apply Lemma 2 to $\mathcal{A} = \{\Pi_{i,1}^* \mid i = 1, \ldots, k+1\}$ and $\mathcal{B} = \{X \cap \Pi_{i,1} \mid i = 1, \ldots, k\}$, where $X = \{x_1, \ldots, x_{2m}\}$. This yields an index i_0 such that the first half of the partition $\Pi_{i_0}^*$ has at least $D/8$ variables on both sides of all partitions Π_i, $i = 1, \ldots, k$. It is now easy to prove the following.

Lemma 3. *Let $\beta := D/(8m) = \delta/4$. There are partitions $\Pi_1', \ldots, \Pi_k'$ of the variables of h which are β-balanced, i. e. $|\Pi_{i,1}'|, |\Pi_{i,2}'| \geqslant \lfloor \beta m \rfloor$ for $i = 1, \ldots, k$, and a k-partition protocol for h with these partitions which has complexity at most k-pcc(f_k).*

To obtain the desired lower bound for f_k, we require an explicitly defined function h which has large multipartition complexity even if the given partitions are only β-balanced for some small constant $\beta > 0$. A linear lower bound of this type is contained, e. g., in the results of Beame, Saks, and Thathachar ([4], Lemma 4) or in [13].

3 The Multipartition Communication Complexity of Linear Codes

A (binary) code of length n and distance d is a subset of vectors $C \subseteq \{0,1\}^n$ for which the Hamming distance between any two vectors in C is at least d. The following lemma is implicit in [11, 17], where a stronger version has been used to show that linear codes are hard for read-k-times branching programs:

Lemma 4 ([11, 17]). *Let $C \subseteq \{0,1\}^n$ be a code of distance $2t+1$. Let P be a multipartition protocol computing the characteristic function of C. Then P uses at least $\log\left(|C| \cdot \binom{\lfloor n/2 \rfloor}{t}^2 \cdot 2^{-n}\right)$ bits of communication.*

The number of codewords and the distance of *random* linear codes are known to meet the Gilbert-Varshamov bound [15]. As a consequence, the above lemma gives *linear* lower bounds for the characteristic functions of such codes. To give a constructive example, we consider binary BCH-codes with length $n = 2^m - 1$ and designed distance $d = 2t + 1$; such a code has at least $2^n/(n+1)^t$ vectors and distance at least d. Let BCH_n be the characteristic function of such a BCH code with $t \approx n^{1/2}$. Using Lemma 4, we obtain:

Theorem 2. *Each multipartition protocol for* BCH_n *has complexity at least* $\Omega\bigl(n^{1/2}\bigr)$.

On the other hand, all linear codes have small randomized communication complexity even in the fixed-partition model (we omit the easy proof):

Proposition 2. *Let f_C be a characteristic function of a linear binary code of length n. Then the two-party fixed-partition one-round bounded error communication complexity of f_C is $O(1)$ with public coins and $O(\log n)$ with private coins.*

The characteristic functions f_C of linear codes are known to be hard for different models of branching programs, including k-n.b.p.'s – nondeterministic read-k-times branching programs where along any path no variable appears more than k times [11], and $(1, +k)$-b.p.'s – deterministic branching programs where along each *consistent* path at most k variables are allowed to be tested more than once [12]. On the other hand, the negation $\neg f_C$ is just an OR of at most n scalar products of an input vector with the rows of the corresponding parity-check matrix. Hence, for every linear code, the characteristic function $\neg f_C$ of its complement has a small *nondeterministic OBDD* (an OBDD is a read-once branching program where the variables along every path appear according to a fixed order). We can strengthen this observation even to *randomized OBDDs with one-sided error*.

Theorem 3. *Let $C \subseteq \{0,1\}^n$ be a linear code and let f_C be its characteristic function. Then, for every integer $r \geqslant 2$, $\neg f_C$ can be computed by a randomized OBDD of size $O\bigl(n^{4r}\bigr)$ with one-sided error at most 2^{-r}.*

Sketch of Proof. Let H be the $m \times n$ parity-check matrix of C. Let $\mathbf{w}$ be chosen uniformly at random from $\{0,1\}^n$. The essence of the construction is the simple fact that $\mathbf{w}^\top Hx \equiv 0 \bmod 2$ for $x \in C$, whereas $\mathrm{Prob}\bigl[\mathbf{w}^\top Hx \not\equiv 0 \bmod 2\bigr] = 1/2$ for $x \notin C$. We cannot use this representation of f_C directly to construct a randomized OBDD, since this OBDD would require exponentially many probabilistic nodes to randomly choose the vector $\mathbf{w}$.

To reduce the number of random bits, we apply an idea which has appeared in different disguises in several papers (see, e. g., Newman [16]): By a probabilistic argument it follows that, for all δ with $0 < \delta < 1/2$, there is a set $W \subseteq \{0,1\}^n$ with $|W| = O\bigl(n/\delta^2\bigr)$ such that for $\mathbf{w}$ chosen uniformly at random from W and all $x \notin C$, $\mathrm{Prob}\bigl[\mathbf{w}^\top Hx \not\equiv 0 \bmod 2\bigr] \geqslant 1/2 - \delta$. Choose $\delta = 1/5$ and let W be the obtained set of vectors.

Let G be the randomized OBDD which starts with a tree on $\lceil \log |W| \rceil$ probabilistic variables at the top by which an element $w \in W$ is chosen uniformly at random. At the leaf of the tree belonging to the vector w, append a deterministic sub-OBDD which checks whether $w^\top Hx \equiv 0 \bmod 2$. By the above facts, this randomized OBDD computes $\neg f_C$ with one-sided error at most $7/10$. The size of G is bounded by $O\bigl(n^2\bigr)$.

To decrease the error probability, we regard G as a deterministic OBDD on all variables (deterministic and probabilistic ones). Applying the known OBDD-algorithms, we obtain an OBDD G' for the OR of $2r$ copies of G with different sets of probabilistic variables. This OBDD G' has one-sided error at most $(7/10)^{2r} < 2^{-r}$ and size $O\big(n^{4r}\big)$. □

Apparently, this result gives the strongest known tradeoff between nondeterministic and randomized branching program complexity.

4 A Lower Bound for Triangle-Freeness

The *triangle-freeness function* Δ_n is a function on $n = \binom{m}{2}$ Boolean variables (encoding the edges on an m-vertex graph) which, given a graph G on m vertices, accepts it if and only if G has no triangles. The function $\oplus\,\textsc{Clique}_{n,3}$ has the same set of variables and outputs the parity of the number of triangles in G.

Theorem 4. *There is a subfunction Δ'_n of Δ_n such that $R(\Delta'_n) = 2^{\Omega(n)}$. The same holds also for $\oplus\,\textsc{Clique}_{n,3}$.*

This result is sufficient to prove that each nondeterministic read-once branching program detecting the triangle-freeness of a graph requires truly exponential size. Since by assigning constants to some variables, we can only decrease the branching program size, the desired lower bound on the size of any 1-n.b.p. computing Δ_n follows directly from Theorem 4 and the fact that each Boolean function f on n variables requires a 1-n.b.p. of size at least $R(f)/(2n)$ (as mentioned in the introduction). We obtain the following main result which also answers Problem 11 of Razborov from [19].

Theorem 5. *Nondeterministic read-once branching programs for the triangle-freeness function Δ_n as well as for $\oplus\,\textsc{Clique}_{n,3}$ require size $2^{\Omega(n)}$.*

Remark. Using a similar probabilistic argument, the following has recently been proven in [13]: (i) $R(\Delta_n) = 2^{\Omega\left(n^{3/4}\right)}$; (ii) $R_k(\Delta_n) = 2^{\Omega(n)}$ provided $k \leqslant 2^{c\sqrt{n}}$ for a sufficiently small constant $c > 0$; and (iii) there is a constant $C > 0$ such that syntactic nondeterministic read-k-times branching programs, detecting the absence of 4-cliques in a graph on m vertices, require size at least $2^{\Omega\left(m^2/C^k\right)}$. Moreover, it is shown that Theorem 4 remains true also for β-balanced partitions, for all constants β with $0 < \beta \leqslant 1/2$.

4.1 Outline of the Proof of Theorem 4

We give the details only for Δ_n and discuss the changes required for $\oplus\,\textsc{Clique}_{n,3}$ at the end of this section. To define the desired subfunction of Δ_n, we consider graphs on m vertices partitioned into sets $U = \{1, \dots, m/2\}$ and $V = \{m/2 + 1, \dots, m\}$. The subfunction Δ'_n will depend only on variables corresponding to the edges in the bipartite graph $U \times V$; the variables corresponding to the edges within the parts U and V will be fixed. Hence, Δ'_n will still have $m^2/4$ variables.

The proof consists essentially of two parts: First, we probabilistically construct an assignment which fixes the subgraphs G_U and G_V on the vertex sets U and V. After fixing these graphs, we obtain a subfunction Δ'_n of Δ_n which depends only on variables belonging to edges in the bipartite graph $G_B = U \times V$. We then consider only those partitions Π which are balanced with respect to the bipartite (non-fixed) part. Our goal is to choose the graphs G_U and G_V such that none of them contains a triangle and the resulting graph $G = G_U \cup G_V \cup G_B$ contains many triangles whose bipartite edges belong to different halves of a partition.

A pair of edges in $U \times V$ is called a *test*, if they form a triangle together with an edge from G_U or G_V. Two tests are said to *collide*, if a triangle can be formed by picking one edge from the first test, one edge from the second test and an edge from $G_U \cup G_V$. In particular, tests collide if they share an edge.

Given a balanced partition $\Pi = (E_1, E_2)$ of the edges in $U \times V$, say that a test is *hard for* Π, if each part E_i of the partition contains one edge of the test. The following lemma about graph partitions is the core of our argument.

Lemma 5. *There exist triangle-free graphs G_U and G_V such that for all balanced partitions $\Pi_1, \dots, \Pi_k$ of $U \times V$, where $k \leqslant 2^{\alpha m^2}$ and $\alpha > 0$ is a sufficiently small constant, the graph $G = G_U \cup G_V \cup G_B$ has a set T of tests such that T does not contain any colliding pairs, and T contains a subset T_i of $\Omega\big(m^2\big)$ hard tests for each Π_i, $i = 1, \dots, k$.*

Let us first show how this lemma implies the theorem; we will then sketch the proof of the lemma itself.

Choose G_U and G_V according to the lemma and let Δ'_n be the resulting subfunction on $U \times V$. Let functions $f_1, \dots, f_k$ be given with $\Delta'_n = f_1 \vee \cdots \vee f_k$, $k \leqslant 2^{\alpha m^2}$, and $\sum_{i=1}^{k} R_1(f_i) = R_k(\Delta'_n)$, and let $\Pi_1, \dots, \Pi_k$ be the partitions corresponding to optimal covers of $f_1, \dots, f_k$ by rectangles.

We construct a set A of hard 1-inputs for Δ'_n which will already require many rectangles to be covered according to the partitions $\Pi_1, \dots, \Pi_k$. Let T be the set of tests obtained by Lemma 5. Edge variables outside of T are fixed to 0 for all inputs in A. For each test in T, we then choose exactly one edge and set the respective variable to 1, the second one is set to 0. Thus, the graph corresponding to an input in A has precisely one of the two edges of each test in T, and two graphs differ only on edges in T. Since no two tests in T collide, the graphs are triangle-free and we obtain a total of $2^{|T|}$ graphs. Hence, $|A| = 2^{|T|}$.

Now observe that there is at least one function f_i with $|f_i^{-1}(1) \cap A| \geqslant |A|/k = 2^{|T|}/k$. By Lemma 5, there is a set $T_i \subseteq T$ of $h = \Omega\big(m^2\big)$ tests which are hard for the partition Π_i. Let $B \subseteq f_i^{-1}(1) \cap A$ be a set of maximum size such that two different inputs from B differ in at least one bit corresponding to a test in T_i. Then $|B| \geqslant |f_i^{-1}(1) \cap A|/2^{|T|-h} \geqslant 2^h/k$.

Since all the inputs from B are accepted by f_i, it remains to show that no rectangle $r \leqslant f_i$ with the underlying partition Π_i can accept more than one input from B. Assume that (a, b) and (a', b') are two different inputs in B accepted by r. By the choice of B, they differ in a test $t = \{e_1, e_2\}$ which is hard for Π_i, i. e., whose edges belong to different halves of the partition Π_i. By the definition of A, exactly one of the two edges e_1 and e_2 is present in each of the graphs belonging to (a, b) and (a', b'), resp., and these edges are different.

Now, if $r(a, b) = 1$, then $r(a, b') = 0$ or $r(a', b) = 0$ because either the graph corresponding to (a, b') or to (a', b) will contain *both* edges e_1, e_2, which, together with the corresponding edge of G_U or G_V, forms a triangle. This is a contradiction to the fact that r is a rectangle. Altogether, we have completed the proof of the lower bound for Δ'_n.

Changes for $\oplus\,\text{CLIQUE}_{n,3}$. We consider the subfunction $\oplus\,\text{CLIQUE}'_{n,3}$ which is obtained from $\oplus\,\text{CLIQUE}_{n,3}$ in same way as Δ'_n from Δ_n. Let $t := |T|$. For $x, y \in \{0,1\}^t$, define $\text{IP}_t(x, y) := \sum_{i=1}^{t} x_i y_i \bmod 2$. Define the set A of hard inputs for $\oplus\,\text{CLIQUE}'_{n,3}$ as follows: For all $(x, y) \in \text{IP}_t^{-1}(1)$, include the input obtained by setting variables outside of T to 0 and setting the two edge variables of the ith test in T to x_i and y_i, resp. Then $|A| = |\text{IP}_t^{-1}(1)| \geqslant 2^{2t-1}$ and $A \subseteq \oplus\,\text{CLIQUE}_{n,3}^{-1}(1)$.

Following the proof for Δ'_n, we obtain a set B of at least $2^{t+h-1}/k$ inputs from A which are hard for one of the partitions Π_i in a cover of $\oplus\,\text{CLIQUE}'_{n,3}$. Using the well-known fact that $|r^{-1}(1)| \leqslant 2^t$ for each rectangle $r \leqslant \text{IP}_t$ or $r \leqslant \neg\,\text{IP}_t$, one easily proves that no rectangle $r' \leqslant \oplus\,\text{CLIQUE}'_{n,3}$ can contain more than 2^t inputs from B. Thus, at least $2^{h-1}/k$ rectangles are needed to cover B. □

4.2 Sketch of Proof for Lemma 5

Recall that a test is a pair of edges in $U \times V$ which form a triangle together with an edge in G_U or G_V, and that a test is hard with respect to a partition Π if its two edges lie in different halves of Π.

Lemma 6. *There exist graphs G_U and G_V such that:*

(i) each of the graphs G_U and G_V has $\Theta(m)$ edges, at most $O(1)$ triangles, and at most $O(m)$ paths of length 2 or 3; and

(ii) for every balanced partition Π of $U \times V$, there are $h = \Omega\big(m^2\big)$ tests which are hard for Π.

Sketch of Proof. We prove the existence of the desired graphs by a probabilistic argument. In what follows, let $\mathbf{G_U}$ ($\mathbf{G_V}$) stand for the random graph on U (resp., on V) obtained by inserting the edges independently at random with probability $p = \Theta(1/m)$ each[1]. Using Markov's inequality, it is easy to show that the graphs $\mathbf{G_U}$ and $\mathbf{G_V}$ have the properties described in Part (i) of the lemma with probability at least $1/2$. It remains to prove that, with probability larger than $1/2$, for every balanced partition of $U \times V$, there are at least $\Omega\big(m^2\big)$ hard tests.

Let Π be such a balanced partition. The partition Π distributes the edges in $U \times V$ to two sets of size $m^2/8$ each which are given to the players Alice and Bob. Call a vertex *mixed* if each of the two players has at least $\frac{1}{8} \cdot \frac{m}{2}$ bipartite edges incident to it.

Claim 1. There are $\Omega(m)$ mixed vertices in each of the sets U and V.

[1] For the sake of simplicity, we omit the exact constant in the definition of p here.

Proof of the Claim. We use essentially the same argument as Papadimitriou and Sipser in [18]. W. l. o. g., assume that we have at most εm mixed vertices in V, where $\varepsilon > 0$ is a sufficiently small constant ($\varepsilon < 1/112$ works fine). Call a vertex v an *A-vertex* (resp. *B-vertex*) if Alice (resp. Bob) has at least $\frac{7}{8} \cdot \frac{m}{2}$ edges incident to v. Thus, vertices which are neither A- nor B-vertices are mixed. Observe first that the number of A-vertices as well as the number of B-vertices in each of the sets U and V is at most $b_{max} := \frac{4}{7} \cdot \frac{m}{2}$, since otherwise Alice or Bob would have more than $m^2/8$ edges. On the other hand, the number of A-vertices as well as the number of B-vertices in U (in V) is bounded *from below* by $b_{min} := \frac{3}{7} \cdot \frac{m}{2} - \varepsilon m$, since otherwise there would be more than εm mixed vertices in U (in V), contrary to the assumption.

Now *more* than half of the edges from A-vertices in U to B-vertices in V belong to Alice, because otherwise there will be an A-vertex $u \in U$ such that Alice has at most half of the edges from u to B-vertices in V, and thus altogether at most $\frac{1}{2} \cdot b_{max} + |V| - b_{min} = \frac{1}{2} \cdot \frac{4}{7} \cdot \frac{m}{2} + \frac{m}{2} - \left(\frac{3}{7} \cdot \frac{m}{2} - \varepsilon m\right) \leqslant \frac{6}{7} \cdot \frac{m}{2} + \varepsilon m < \frac{7}{8} \cdot \frac{m}{2}$ edges incident to u. With the same reasoning, however, *more* than half of all edges from A-vertices in U to B-vertices in V belong to Bob. Contradiction. □

For each mixed vertex $u \in U$, let $V_A(u)$ ($V_B(u)$) be the set of vertices $v \in V$ for which Alice (resp. Bob) has the edge $\{u, v\}$. Since u is mixed, $|V_A(u)|, |V_B(u)| \geqslant \frac{1}{8} \cdot \frac{m}{2}$. Observe that each edge between $V_A(u)$ and $V_B(u)$ leads to a hard test with respect to the given partition Π.

Claim 2. The following event has probability larger than $1/2$ *with respect to the random choices of* $\mathbf{G_V}$*: For all pairs of disjoint sets* $S_1, S_2 \subseteq V$ *of size at least* $m/16$ *each, the number of edges in* $\mathbf{G_V}$ *between* S_1 *and* S_2 *is at least* $p|S_1||S_2|/2$.

Proof of the Claim. The expected number of edges between fixed sets of vertices S_1 and S_2 is $p|S_1||S_2|$. By Chernoff bounds, the true number of edges is at least $p|S_1||S_2|/2$ with probability at least $1 - e^{-cm}$, where the constant $c > 0$ can be adjusted by the choice of the constant in the definition of p. Since there are at most $\left(2^{m/2}\right)^2 = 2^m$ choices for the sets $S_1, S_2 \subseteq V$, the probability of the described event is at least $1 - 2^m \cdot e^{-cm}$, which is larger than $1/2$ for appropriate c. □

We apply the claim to the sets $V_A(u)$ and $V_B(u)$, where u is a mixed vertex. Due to the claim, the event that, for all partitions Π and all $\Omega(m)$ mixed vertices u with respect to Π, the respective sets $V_A(u)$ and $V_B(u)$ are connected by at least $p|V_A(u)||V_B(u)|/2 = \Omega(m)$ edges, has probability larger than $1/2$. Thus, with probability larger than $1/2$, for each partition Π there are $\Omega\left(m^2\right)$ hard tests. This completes the proof of the lemma. (Observe that it does not matter whether we carry out the above argument for mixed vertices in U or in V.) □

We apply Lemma 6 and fix graphs G_U and G_V with the described properties. Since there are only $O(1)$ triangles, we can remove these triangles without destroying the other properties. Especially, we still have linearly many edges. By Property (ii), this pair of graphs produces a set of $h = \Omega\left(m^2\right)$ hard tests T_i for each of the partitions Π_i ($i = 1, \ldots, k$) from a given multipartition protocol for Δ_n.

Let T_0 be the set of all tests induced by G_U and G_V, and let $t = |T_0|$ be its size. Since both graphs G_U and G_V have $\Theta(m)$ edges, $t = \Omega\big(m^2\big)$. Using the properties of these graphs stated in Lemma 6 (i), it is easy to show (by case analysis) that at most $O(t)$ of all $\binom{t}{2}$ pairs of tests in T_0 will collide:

Lemma 7. *There are at most $O(t)$ pairs of colliding tests in T_0.*

To finish the proof of Lemma 5, it remains to find a subset $T \subseteq T_0$ such that: (i) there is no pair of tests from T which collide; and (ii) $|T \cap T_i| = \Omega\big(m^2\big)$ for all $i = 1, \ldots, k$. We again use a probabilistic construction. Let $\mathbf{T}$ be a set of s tests picked uniformly at random from the set T_0, where $s = \gamma t$ and γ is a constant with $0 < \gamma < 1$ chosen later on.

Lemma 8.

(i) With probability at least $1/2$, the set $\mathbf{T}$ contains at most $O\big(s^2/t\big)$ pairs of colliding tests (where $t = |T_0|$ is the total number of tests).

(ii) With probability larger than $1/2$, $|\mathbf{T} \cap T_i| \geqslant \frac{s\cdot h}{2t}$ for all $i = 1, \ldots, k$.

Proof. Part (i): We define the *collision graph* to have tests as vertices and edges for each collision. Let c be the number of edges in the collision graph. By Lemma 7, we know that $c = O(t)$.

Let $\mathbf{c_T}$ be the number of edges in the subgraph of the collision graph induced by the randomly chosen set $\mathbf{T}$. Since we pick tests uniformly at random, the expected number of edges is $\mathrm{E}\left[\mathbf{c_T}\right] = \frac{s(s-1)}{t(t-1)} \cdot c$. By Markov's inequality, it follows that the actual number of edges is at most $2 \cdot \mathrm{E}\left[\mathbf{c_T}\right]$ with probability at least $1/2$. Hence, the number of pairs of colliding tests in $\mathbf{T}$ is at most $2 \cdot \mathrm{E}\left[\mathbf{c_T}\right] = O\big((s/t)^2 \cdot c\big) = O\big(s^2/t\big)$ with probability at least $1/2$.

Part (ii): Consider a fixed partition Π_i. The probability to choose a hard test from T_i is h/t, $t = \Omega\big(m^2\big)$ the total number of tests. Thus the expected number of elements in $\mathbf{T} \cap T_i$ for a randomly chosen set $\mathbf{T}$ of s tests is $s \cdot h/t$. Let $\lambda := h/(2t)$. By Chernoff bounds, it follows that $\mathrm{Prob}\left[|\mathbf{T} \cap T_i| < \lambda \cdot s\right] \leqslant 2e^{-\lambda^2 s} = e^{-\Omega(s)}$. Hence, the probability that $\mathbf{T}$ contains at least $\lambda \cdot s = sh/(2t)$ hard tests for each of the partitions at least $1 - k \cdot 2^{-\Omega(s)}$. Since $s = \gamma t = \Theta\big(m^2\big)$, this probability is larger than $1/2$ for $k \leqslant 2^{\alpha m^2}$ with $\alpha > 0$ sufficiently small. □

Lemma 8 yields the existence of a set $T \subseteq T_0$ with the following properties: (i) $|T| = s = \gamma t$; (ii) there are at most $\delta s^2/t$ pairs of tests in T which collide, $\delta > 0$ some constant; and (iii) for all $i = 1, \ldots, k$, $|T \cap T_i| \geqslant sh/(2t)$.

By deleting at most $\delta s^2/t$ tests from T, we remove all collisions, obtaining a smaller set T'. The number of hard tests for each Π_i in T' is still $sh/(2t) - \delta s^2/t = (s/t) \cdot (h/2 - \delta s) = \gamma \cdot (h/2 - \delta \gamma t)$. Since this number is of the order $\Omega\big(m^2\big)$ for $\gamma = h/(4\delta t) = O(1)$, we have completed the proof of Lemma 5. □

Acknowledgment

Thanks to Ingo Wegener for critical comments on an earlier version of this paper.

Appendix

A1: The Hierarchy for k-Partition Protocols

Here we prove the lower bound from Theorem 1. We first repeat the definition of the type of functions used in the hierarchy result. Some details of the definition will be filled in later on.

Definition A.1: Let k and m be positive integers, and let $h\colon \{0,1\}^m \to \{0,1\}$ be an arbitrary function. Let $X = \{x_1, \dots, x_{2m}\}$ and let $\mathcal{P} = \{\Pi_1^*, \dots, \Pi_{k+1}^*\}$, where Π_i^* is a balanced partition of X. For $i = 1, \dots, k+1$, define $x^1(i)$ and $x^2(i)$ as the vectors of m variables corresponding to the halves of Π_i^*. Let $\ell = \lceil \log(k+1) \rceil$, $Y = \{y_0, \dots, y_{\ell-1}\}$, and $Z = \{z_0, \dots, z_{\ell-1}\}$. Define $f_{h,\mathcal{P}}\colon \{0,1\}^{2(m+\ell)} \to \{0,1\}$ by

$$f_{h,\mathcal{P}}(x,y,z) := \bigvee_{1 \leqslant i \leqslant k+1} [|y|_2 = i] \wedge h\bigl(x^1(i)\bigr),$$

where $|y|_2$ denotes the value of y as a binary number.

As already shown in the main text, $(k+1)$-*pcc* $(f_{h,\mathcal{P}}) \leqslant \lceil \log(k+1) \rceil + 1$. It remains to prove a linear lower bound on the k-partition communication complexity for suitably chosen h and $\mathcal{P}$. We will first prove a technical lemma which replaces Lemma 3 from the main text. Then we apply the lemma to an explicit example.

Definition A.2: For $x, y \in \{0,1\}^n$, let $d(x,y)$ denote the *(Hamming) distance* between x and y. By the *weight* of x, denoted by $w(x)$, we mean the number of 1-entries in x. A set $C \subseteq \{0,1\}^n$ is called a *balanced code* if the $w(x) = \lfloor n/2 \rfloor$ for all $x \in C$ [1].

Lemma A.3: *Let $C_n \subseteq \{0,1\}^n$ be a family of balanced codes which is defined for infinitely many n such that for all different $x, y \in C_n$, $D \leqslant d(x,y) \leqslant n - D$.*

Let m be a positive integer, and let $n = 2m$ be such that C_n is defined. Suppose that $D \geqslant 8\lfloor \alpha m \rfloor$. Let $N = |C_n|$, and suppose that $\ell := \lceil \log N \rceil \leqslant D/8$. Let $C_n = \{c_1, \dots, c_N\}$.

Let $X = \{x_1, \dots, x_{2m}\}$. For $i = 1, \dots, N$, let $\Pi_i^* = (\Pi_{i,1}^*, \Pi_{i,2}^*)$, with $\Pi_{i,1}^* := \{x_j \mid c_{i,j} = 1\}$ and $\Pi_{i,2}^* := X - \Pi_{i,1}^*$. Let $\mathcal{P} = \{\Pi_1^*, \dots, \Pi_N^*\}$. Let $f_{h,\mathcal{P}}$ be the function obtained according to Definition A.1 using an arbitrary function $h\colon \{0,1\}^m \to \{0,1\}$ and the set of partitions $\mathcal{P}$ defined here.

Then there are α-balanced partitions $\Pi_1', \dots, \Pi_{N-1}'$ of the input variables of h, i.e., $\Pi_i' = (\Pi_{i,1}', \Pi_{i,2}')$ with $|\Pi_{i,1}'|, |\Pi_{i,2}'| \geqslant \lfloor \alpha m \rfloor$, and an $(N-1)$-partition protocol P' for h according to $\Pi_1', \dots, \Pi_{N-1}'$ such that the complexity of P' is bounded from above by $(N-1)$-pcc $(f_{h,\mathcal{P}})$.

[1] This is the definition used in coding theory. In fact, we only need $w(x) \in \{\lfloor n/2 \rfloor, \lceil n/2 \rceil\}$ here.

Proof. Let P be an optimal $(N-1)$-partition protocol for $f_{h,\mathcal{P}}$ according to the balanced partitions $\Pi_1, \dots, \Pi_{N-1}$ of $X \cup Y \cup Z$, where $\Pi_i = (\Pi_{i,1}, \Pi_{i,2})$.

For $i \in \{1, \dots, N\}$, define $S_i := \{x^1(i)\}$ and $\overline{S_i} := \{x^2(i)\} = X - S_i$. For $i \in \{1, \dots, N-1\}$, define $T_i := X \cap \Pi_{i,1}$ and $\overline{T_i} = X \cap \Pi_{i,2} := X - T_i$. Since there are $2\lceil \log N \rceil$ y- and z-variables, the number of x-variables in each half of Π_i is at least $n/2 - 2\lceil \log N \rceil \geqslant n/2 - D/4$. Hence, $|T_i|, |\overline{T_i}| \geqslant n/2 - D/4$.

We apply Lemma 2 from the main text (page 5) to $\mathcal{A} := \{S_i \mid i = 1, \dots, N\}$ and $\mathcal{B} := \{T_i \mid i = 1, \dots, N-1\}$. This yields an index $i_0 \in \{1, \dots, N\}$ with $|S_{i_0} \cap T_j| \geqslant D/8$ and $|S_{i_0} \cap \overline{T_j}| \geqslant D/8$ for all $j \in \{1, \dots, N-1\}$. Since $D \geqslant 8\lfloor \alpha m \rfloor$ by assumption, we have $|S_{i_0} \cap T_j| \geqslant \lfloor \alpha m \rfloor$ and $|S_{i_0} \cap \overline{T_j}| \geqslant \lfloor \alpha m \rfloor$ for all $j \in \{1, \dots, N-1\}$.

We construct the desired $(N-1)$-partition protocol P' for h by setting variables to constants in the given protocol P for $f_{h,\mathcal{P}}$. Let $f_{h,\mathcal{P}} = P_1 \vee \dots \vee P_{N-1}$, where P_i is the function computed by the ith subprotocol P_i of P. We fix the y-variables such that y represents the value i_0. Furthermore, we fix the variables in $\overline{S_{i_0}}$ and the z-variables in an arbitrary way.

Let P' and $P'_1, \dots, P'_{N-1}$ be the protocols obtained from P and $P_1, \dots, P_{N-1}$, resp., by the above variable assignments. The new protocols only work on variables from S_{i_0}, and we have

$$P'_1 \vee \dots \vee P'_{N-1} = h\big(x^1(1)\big).$$

By restricting the partitions $\Pi_1, \dots, \Pi_{N-1}$ to the remaining variables in S_{i_0}, we obtain new partitions $\Pi'_1, \dots, \Pi'_{N-1}$, where $\Pi'_i = (\Pi'_{i,1}, \Pi'_{i,2})$, such that $|\Pi'_{i,1}|, |\Pi'_{i,2}| \geqslant \lfloor \alpha m \rfloor$ for all $i = 1, \dots, N-1$. Each protocol P'_i is a nondeterministic two-party protocol according to Π'_i.

Altogether, P' is a protocol of the desired type for h (defined on S_{i_0}), and the complexity of P' is bounded from above by the complexity of P. □

Application to an Explicit Example

Definition A.4: Let $C \subseteq \{0,1\}^n$. The *rate* of C is defined as $(\log |C|)/n$.

Proposition A.5: *Let $C \subseteq \{0,1\}^n$ be a code with rate α and $\delta n \leqslant d(x,y) \leqslant (1-\delta)n$ for all different $x, y \in C$, where $\alpha, \delta > 0$. Let $N := 2n$ and define*

$$C^{\mathrm{b}} := \{(x,y) \mid x \in C, y \in \{0,1\}^n \text{ with } w(y) = n - w(x)\} \subseteq \{0,1\}^N.$$

Then C^{b} is a balanced code with rate at least $\alpha/2$ and $(\delta/2)N \leqslant d(x,y) \leqslant (1-\delta/2)N$ for all different $x, y \in C^{\mathrm{b}}$.

Definition A.6: Let m be a positive integer, $N = 2^m - 1$, and let α be a primitive element of $\mathbb{F}_{2^m}$. Let K be an integer with $1 \leqslant K \leqslant N-1$, and define $D := N - K + 1$. Let $\mathcal{R}_{N,K}$ be the $[N,K]$-*Reed-Solomon code* i. e., the linear code of length N over $\mathbb{F}_{2^m}$ with parity-check matrix

$$H_{N,K} = \begin{pmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{N-1} \\ 1 & \alpha^2 & \alpha^4 & \cdots & \alpha^{2(N-1)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & \alpha^{D-1} & \alpha^{(D-1)\cdot 2} & \cdots & \alpha^{(D-1)(N-1)} \end{pmatrix}$$

This code has dimension K and distance D (see, e. g., [15]).

For $x \in \mathbb{F}_{2^m}$ and $1 \leqslant i \leqslant N$, define $c_i(x) := (x, \alpha^i \cdot x)$. For $x = (x_1, \dots, x_N) \in \mathbb{F}_{2^m}^N$, define $c(x) := (c_1(x_1), \dots, c_N(x_N))$ and regard this as a vector from $(\mathbb{F}_2)^{2mN}$. The binary code $\mathcal{J}_{N,K}$ of length $2mN$ defined by

$$\mathcal{J}_{N,K} := \{c(x) \mid x \in \mathcal{R}_{N,K}\}$$

is called ($[N, K]$-)*Justesen code*. This code is linear and has dimension mK (both facts follow directly from the definition).

Theorem A.7 (Justesen): *Let m be a positive integer and $0 < R < 1/2$. Let $N = 2^m - 1$ and $n = 2mN$. Define $K := \lceil R \cdot 2N \rceil$. Then the Justesen code $\mathcal{J}_m^R := \mathcal{J}_{N,K}$ has rate at least $K/(2N) \geqslant R$, and for each constant $\varepsilon > 0$ and m sufficiently large,*

$$w(x) \ \geqslant \ (1 - \varepsilon)(1 - 2R)cn, \quad \textit{for all } x \in \mathcal{J}_m^R,$$

where $c := H^{-1}(1/2) > 0.11002$ (H is the binary entropy function).

Observation A.8: *There is a deterministic polynomial time algorithm which, given $x \in \{0, 1\}^n$, checks whether $x \in J_m^R$.*

Proof. This follows from the facts that (i) an irreducible polynomial $p(x) \in \mathbb{F}_2[x]$ of degree m can be found by a deterministic algorithm with polynomial time in m (see, e. g., [21]); and (ii) addition and multiplication in $\mathbb{F}_{2^m}$ can be carried out efficiently by algorithms on polynomials in $\mathbb{F}_2[x]/(p(x))$, given the irreducible polynomial $p(x)$. □

Observation A.9: *For each constant $\varepsilon > 0$ and m sufficiently large,*

$$w(x) \ \leqslant \ n - (1 - \varepsilon)(1 - 2R)cn, \quad \textit{for all } x \in \mathcal{J}_m^R.$$

Proof. This follows in the same way as the lower bound on the weight in the standard proof of Theorem A.7 (see, e. g., [15]).

The standard proof exploits the fact that there are only few Boolean vectors of fixed length with small weight. The same holds for vectors of large weight, as stated below.

Claim. Let $0 < \gamma,\ \delta < 1$. Suppose that $M = M(L) = \big(2^{\delta L} - 1\big)(\gamma + o(1))$ (for $L \to \infty$), and let $W = W(L)$ be the sum of all weights of M different vectors from $\{0, 1\}^L$. Then $W \leqslant LM - \big(H^{-1}(\delta) - o(1)\big)LM$.

Proof of the Claim. The number of vectors from $\{0, 1\}^L$ with weight at least $(1 - \lambda)L$ is

$$N_{\text{large}} \ = \sum_{(1-\lambda)L \leqslant i \leqslant L} \binom{L}{i} \ = \sum_{0 \leqslant i \leqslant \lambda L} \binom{L}{i} \ \leqslant \ 2^{H(\lambda)L},$$

where the inequality at the end follows by well-known estimates of the binomial coefficients. This yields

$$\begin{aligned} W &\leqslant LN_{\text{large}} + (1-\lambda)L(M - N_{\text{large}}) \\ &= LN_{\text{large}} + LM - \lambda LM - LN_{\text{large}} + \lambda LN_{\text{large}} \\ &= LM - \lambda LM(1 - N_{\text{large}}/M) \\ &\leqslant LM - \lambda LM\big(1 - 2^{H(\lambda)L}/M\big). \end{aligned}$$

We set $\lambda := H^{-1}\big(\delta - 1/\log L\big) = H^{-1}(\delta) \pm o(1)$. Then

$$\begin{aligned} W &\leqslant LM - \lambda LM\big(1 - 2^{H(\lambda)L}/M\big) \\ &= LM - \lambda LM\left(1 - \frac{2^{\delta L - L/\log L}}{\big(2^{\delta L}-1\big)(\gamma + o(1))}\right) \\ &= LM - \lambda LM(1 - o(1)) \\ &= LM - \big(H^{-1}(\delta) - o(1)\big)LM \end{aligned}$$

□

By the following claim, each codeword $x = (x_1, \ldots, x_N)$ of the Reed-Solomon code with a fixed number of non-zero entries (i. e., fixed weight) leads to a codeword $c(x)$ in the Justesen code with the same number of different components $c_i(x_i)$ from $\{0,1\}^{2m}$.

Claim. For $1 \leqslant i < j \leqslant N = 2^m - 1$ *and* $u, v \in \mathbb{F}_{2^m} - \{0\}$, $(u, \alpha^i u) \neq (v, \alpha^j v)$.

Proof of the Claim. If $(u, \alpha^i u) = (v, \alpha^j v)$, then $\alpha^i u = \alpha^j u$, and thus, by division in $\mathbb{F}_{2^m}$, $\alpha^i = \alpha^j$. It follows that $i \equiv j \bmod \big(2^m - 1\big)$. □

Now we prove the claimed upper bound on the weight of codewords in $\mathcal{J}_m^R$. The distance of the Reed-Solomon code $\mathcal{R}_{N,K}$ is $D = N - K + 1$. Thus, each code word $x = (x_1, \ldots, x_N) \in \mathbb{F}_{2^m}^N$ has at least D non-zero entries. By the second claim, the corresponding codeword $c(x) = (c_1(x_1), \ldots, c_N(x_N))$ in the Justesen code contains at least D different vectors from $\{0,1\}^{2m}$. We have

$$D = N - K + 1 \geqslant N - K = N - \lceil R \cdot 2N \rceil = \big(2^m - 1\big)(1 - 2R + o(1)).$$

We apply the first claim to bound the weight of $c(x)$. We set $L := 2m$, $\delta := 1/2$, $\gamma := 1 - 2R$, and $M := \big(2^m - 1\big)(1 - 2R + o(1))$. By the claim, the weight of $c(x)$ is bounded by

$$\begin{aligned} W &\leqslant LM - \big(H^{-1}(\delta) - o(1)\big)LM \\ &\leqslant 2m\big(2^m - 1\big) - \big(H^{-1}(\delta) - o(1)\big)2m\big(2^m - 1\big)(1 - 2R - o(1)) \\ &= n - (1 - 2R)H^{-1}(\delta)(1 - o(1))n, \end{aligned}$$

where we have used the bound $M \leqslant 2^m - 1$ for the second line. □

Lemma A.10 (Beame, Saks, Thathachar [4]): *Let $n = 2^d$, and let S_n be the $n \times n$ Sylvester matrix defined by $S_n(x, y) := (-1)^{x^\top y \bmod 2}$, $x, y \in \{0, 1\}^n$. Define $\mathrm{BQF}_n \colon \{0, 1\}^n \to \{0, 1\}$ by $\mathrm{BQF}_n(x) := \left[x^\top S_n x \equiv 0 \bmod 3\right]$.*

(i) $\left|\mathrm{BQF}_n^{-1}(1)\right| \geqslant 2^{n-24\log n/\sqrt{n}}$.

(ii) Let $r \colon \{0, 1\}^n \to \{0, 1\}$ be a rectangle with respect to a δ-balanced partition of the variables of BQF_n with $r \leqslant \mathrm{BQF}_n$. Then $\left|r^{-1}(1)\right| \leqslant 2^{(1-\delta^2)n}$.

Especially, this lemma implies that BQF_n has multipartition communication complexity $\Omega(n)$ even with respect to partitions which are only δ-balanced for some constant $\delta > 0$.

Proof of Theorem 1 – Lower Bound. We prove that for all functions k with $k(n) \leqslant 2^{\alpha n-1} - 1$, $\alpha := (1/32)H^{-1}(1/2)$, there is a function $f_{h,\mathcal{P}}$ constructed according to Definition A.1 with $|\mathcal{P}| = k + 1$, input size $\Theta(n)$, and $k\text{-}pcc\,(f_{h,\mathcal{P}}) = \Omega(n)$. We do not make any attempt to optimize the constants here.

Let d be a positive integer and $R = 1/4$. Plug $\mathcal{J}_d^R$ into Proposition A.5. Let $m = 2d\big(2^d - 1\big)$, and let $C_{2m}^{\mathrm{b}} \subseteq \{0, 1\}^{2m}$ be the balanced code obtained according to the proposition, which has rate at least $R/2 = 1/8$ and $D \leqslant d(x, y) \leqslant 2m - D$ for all different $x, y \in C_{2m}^{\mathrm{b}}$, where $D := 8\delta(2m)$, $\delta := (1/64)H^{-1}(1/2) \in (0.01, 0.02)$.

Define $n := 2^{d+\lceil\log d\rceil-1}$. Then $n \leqslant m = 2d\big(2^d - 1\big)$ for sufficiently large d. Define $h \colon \{0, 1\}^m \to \{0, 1\}$ by $h(x_1, \dots, x_m) := \mathrm{BQF}_n(x_1, \dots, x_n)$. By Lemma A.10 and Proposition 1, the multipartition complexity of h with respect to δ-balanced partitions of the input variables is at least $\delta^2 n - o(1)$. Since $n = (m/2)(1 + o(1))$, this is of order $\Omega(m)$.

We have $|C_{2m}^{\mathrm{b}}| \geqslant 2^{m/4}$. We use h and a subset of C_{2m}^{b} of size $k + 1 \leqslant 2^{2\delta m-1}$ in Lemma A.3. Observe that the assumption $\lceil\log(k + 1)\rceil \leqslant D/8 = 2\delta m$ is fulfilled for the chosen k. For the function $f_{h,\mathcal{P}} \colon \{0, 1\}^{2m+2\lceil\log(k+1)\rceil} \to \{0, 1\}$ constructed from h and the set of partitions $\mathcal{P}$ corresponding to the chosen subset of C_{2m}^{b}, we obtain $k\text{-}pcc\,(f_{h,\mathcal{P}}) = \Omega(m)$. □

A2: Linear Codes

Proof of Theorem 2

Let $t := \lceil n^{1/2} \rceil$. Using Stirling's formula, one can easily prove the following estimate for the binomial coefficients occurring in Lemma 4:

$$\binom{\lfloor n/2 \rfloor}{t} = e^{-1}(2\pi)^{-1/2} \cdot n^{-1/4} \cdot \Big((e/2)n^{1/2}\Big)^{n^{1/2}} \cdot (1 + o(1)).$$

Thus, $\binom{\lfloor n/2 \rfloor}{t} \geqslant 2^{\alpha n^{1/2}} \cdot n^{(1/2)n^{1/2}}$, for some positive constant $\alpha < \log(e/2)$ $(\log(e/2) > 0.442)$.

By Lemma 4, we obtain the following lower bound on the multipartition communication complexity of the characteristic function of the considered BCH-code:

$$\log\left(|C|\cdot\binom{\lfloor n/2\rfloor}{t}^2\cdot 2^{-n}\right)\geqslant\log\left(\frac{2^{2\alpha n^{1/2}}\cdot n^{n^{1/2}}}{(n+1)^{\lceil n^{1/2}\rceil}}\right)=\Omega\big(n^{1/2}\big).$$

□

Proof of Proposition 2

Checking whether a given input is accepted reduces to checking whether the two strings, obtained by Alice and Bob by multiplying the parts of the input they see with the corresponding parts of the parity-check matrix, are equal. Hence, if H_1 and H_2 are the parts of the parity-check matrix corresponding to the parts of the inputs string (x, y) given to Alice and Bob, then testing whether $f_C(x, y) = 1$ is the same as testing the equality $H_1 \cdot x = H_2 \cdot y$ of two strings. □

Proof of Theorem 3

It remains to prove that the number of random bits can be reduced as described in the sketch of proof. A similar argument has been used by several authors, e. g., in [6, 7, 16, 20].

Although the main trick is quite simple, it is usually hidden behind the technical details of a particular model of computation. Since the argument may be of independent interest, it makes sense to formulate it as a separate combinatorial lemma about the average density of Boolean matrices.

Lemma A.11: *Let M, N be positive integers with $M \geqslant 5$ and $M = 2^{o(\sqrt{N})}$. Let A be a Boolean $M \times N$ matrix with the property that the* average density, *i. e. the average number of 1's, in each row does not exceed p, $0 \leqslant p < 1$. Then, for every constant $\delta > 0$, there is a set $I \subseteq \{1, \dots, N\}$ with $|I| = \lceil \log M/\delta^2 \rceil$ such that in the submatrix of A consisting of the columns with index in I, each row has average density at most $p + \delta$.*

Proof. Let $\boldsymbol{\xi_1}, \dots, \boldsymbol{\xi_t}$ be independent random variables which are uniformly distributed over $\{1, \dots, N\}$, where $t := \lceil \log M/\delta^2 \rceil$. First, observe that with probability $1 - \binom{t}{2}/N = 1 - o(1)$, all $\boldsymbol{\xi_1}, \dots, \boldsymbol{\xi_t}$ are distinct. Next, fix a row $x = (x_1, \dots, x_N)$ of A and consider the 0-1 random variables $\mathbf{X}_i = x_{\boldsymbol{\xi_i}}$, for $i = 1, \dots, t$. We have $\text{Prob}\,[\mathbf{X}_i = 1] \leqslant p$ for all i. By Chernoff bounds, the average density $\big(\sum_{i=1}^t \mathbf{X}_i\big)/t$ of 1's in x exceeds $p + \delta$ with probability at most $2e^{-\delta^2 t} \leqslant 2M^{-\log e}$. Thus, with probability at least $1 - 2M^{1-\log e}$, the restriction of *each* row of A to the columns with indices $\boldsymbol{\xi_1}, \dots, \boldsymbol{\xi_t}$ has density at most $p + \delta$. This probability is larger than 0 for $M \geqslant 5$. Altogether, the probability that the submatrix consisting of the columns with indices $\boldsymbol{\xi_1}, \dots, \boldsymbol{\xi_t}$ has the claimed properties is larger than 0. □

We apply this lemma as follows. Choose the set of all $x \in \{0,1\}^n$ with $\neg f_C(x) = 1$, i. e. $x \notin C$, as the row indices, and all vectors $w \in \{0,1\}^n$ as the column indices. Define the $2^n \times 2^n$ matrix $A = (a_{x,w})$ by setting $a_{x,w} := \left[w^\top Hx \not\equiv 0 \bmod 2\right]$. Then each row of A has density $1/2$. The lemma gives us a set $W \subseteq \{0,1\}^n$ with $|W| = \lceil \log M/\delta^2 \rceil = O\big(n/\delta^2\big)$ such that, for all x with $\neg f_C(x) = 1$ and $\mathbf{w}$ chosen uniformly at random from W, we have $\operatorname{Prob}\left[\mathbf{w}^\top Hx \not\equiv 0 \bmod 2\right] \geqslant 1/2 - \delta$.

This completes the proof of Theorem 3. □

A3: Triangle-Freeness

Proof of Lemma 6

It only remains to supply the details of the proof of Part (i) of the lemma.

Let **G** be a random graph on $m/2$ vertices where the edges are inserted independently at random with probability $p = \Theta(1/m)$. We claim that, with probability at least $3/4$, **G** has $\Theta(m)$ edges, $O(1)$ triangles, and $O(1)$ paths of length 2 and 3.

(a) The expected number of edges in **G** is $E = p \cdot \binom{m/2}{2} = \Theta(m)$. Using Chernoff bounds, we get that the actual number of edges is smaller than $E/2$ or larger than $(3/2)E$ only with exponential small probability.

(b) The expected number of triangles in **G** is $E = \binom{m/2}{3} \cdot p^3$. Hence, **G** has more than $16 \cdot E$ triangles with probability less than $1/16$ by Markov's inequality.

(c) The expected number of paths of length k in **G** is $E = \binom{m/2}{k+1} \cdot p^k$, and **G** has more than $32 \cdot E$ paths of length k with probability less than $1/32$. Thus the bound on the number of paths of length two and three is exceeded with probability at most $1/16$.

Altogether, the conjunction of (a), (b) and (c) holds with probability at least $1 - 3/16 > 3/4$. It follows that, with probability larger than $1/2$, *both* of the random graphs $\mathbf{G_U}$ and $\mathbf{G_V}$ considered in the main text have $\Theta(m)$ edges, $O(1)$ triangles, and $O(1)$ paths of length 2 and 3.

Proof of Lemma 7

Recall that our goal is to prove that there are at most $O(t)$ pairs of colliding pairs in the set T_0 of tests induced by the graphs G_U and G_V. We prove the claim by case inspection of all possible situations in which tests may collide.

A test is a pair of edges of the bipartite graph $G_B = U \times V$ which together with an edge from G_U or G_V form a triangle. Thus, a test is defined by a pair (e, v), where e is an edge in G_U (G_V) and a vertex $v \in V$ ($v \in W$, resp.).

Claim 1. Let (e_1, w_1) and (e_2, w_2) describe two colliding tests. Assume that e_1 and e_2 both belong to G_U (resp. that they both belong to G_V). Then at least one of the following conditions applies.

(a) $\{w_1, w_2\}$ *is an edge of* G_V *(resp. of* G_U*) and* e_1 *and* e_2 *belong to a* G_U*-path (resp. to a* G_V*-path) of length two;*

(b) $w_1 = w_2$ *and* e_1 *and* e_2 *belong to a* G_U*-path (resp. to a* G_V*-path) of length two or three.*

Proof of Claim 1. Assume first that a triangle is formed by picking a G_V-edge (resp. a G_U-edge) as the third edge. In this case the two bipartite edges originate from the same vertex in U (resp. V) which has to be a common endpoint of e_1 and e_2. Thus e_1 and e_2 belong to a G_U-path (resp. G_V-path) of length two and $\{w_1, w_2\}$ is the G_V-edge (resp. the G_U-edge) in question. (See Figure 1 a.)

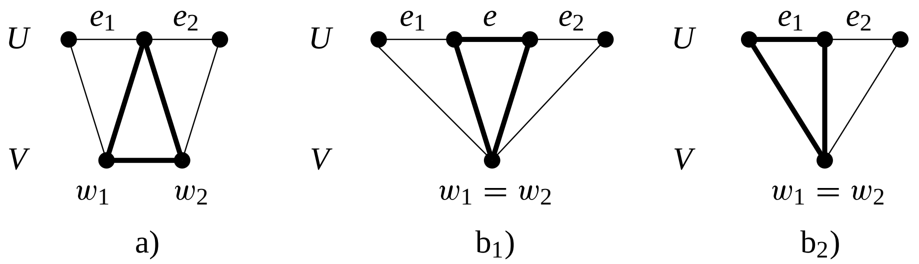

Figure 1.

Now assume that the triangle is formed by picking a G_U-edge (resp. a G_V-edge) e. Thus the triangle consists of e and the two bipartite edges: $w_1 = w_2$ follows. If e_1 and e_2 do not share an endpoint, then (e_1, e, e_2) is a G_U-path (resp. G_V-path) of length three (Figure 1 b_1). Finally, if e_1 and e_2 share an endpoint, then (e_1, e_2) is a G_U-path (resp. G_V-path) of length two (Figure 1 b_2). □

Claim 2. Assume that e_1 *belongs to* G_U *and that* e_2 *belongs to* G_V *(the situation where* e_1 *belongs to* G_V *and* e_2 *to* G_V *is completely symmetric). Then at least one of the following conditions applies.*

(c) w_1 *is an endpoint of* e_2 *and* w_2 *is an endpoint of* e_1*;*

(d) w_1 *is an endpoint of* e_2 *and* e_1 *belongs to a* G_U*-path of length two that begins in* w_2*;*

(e) w_2 *is an endpoint of* e_1 *and* e_2 *belongs to a* G_V*-path of length two that begins in* w_1*;*

Proof of Claim 2. There are essentially three different possible situations which are shown in Figure 2. Obviously, this is exactly what is described in conditions (c)–(e). Condition (e) is symmetric to (d).

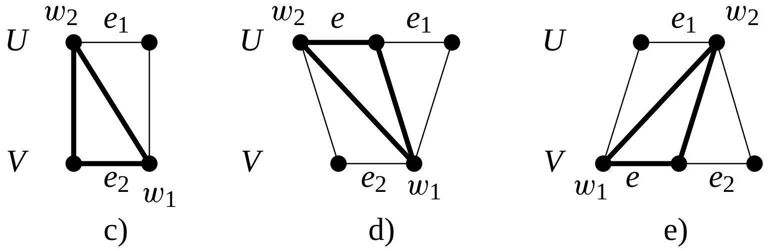

Figure 2.

We now estimate the number of colliding pairs of tests by using the above results and Lemma 6, Part (i). We show that there are only $O\big(m^2\big)$ pairs of tests for which one of the conditions (a)–(e) applies. Since $t = \Theta\big(m^2\big)$, this also proves that the number of colliding pairs is of order $O(t)$.

(a) There are only $O(m)$ edges $\{w_1, w_2\}$ in G_U (resp. G_V) and $O(m)$ G_V-paths (G_U-paths) of length two.

(b) There are only $m/2$ vertices w_1 and $O(m)$ G_U-paths (G_V-paths) of length three.

(c) The number of collisions of this type is $2|G_U||G_V| = O\big(m^2\big)$, since there are $|G_U||G_V|$ choices for e_1 and e_2 and two ways to place the endpoints w_1 and w_2 for each of these choices.

(d) There are $O(m)$ G_U-paths of length two and $2|G_V|$ choices for the pair (e_2, w_1).

(e) This is symmetric to (d).

□

References

[1] M. Ajtai, A non-linear time lower bound for Boolean branching programs, *Proc. of 40th FOCS*, 1999, pp. 60–70.

[2] M. Ajtai, L. Babai, P. Hajnal, J. Komlos, P. Pudlák, V. Rödl, E. Szemeredi, and Gy. Turán, Two lower bounds for branching programs, in: *Proc. 18th ACM STOC*, 1986, pp. 30–38.

[3] P. Beame, M. Saks, X. Sun, and E. Vee, Super-linear time-space tradeoff lower bounds for randomized computation, Technical Report **25**, *Electr. Coll. on Comp. Compl.*, 2000.

[4] P. Beame, M. Saks, and J. S. Thathachar, Time-space tradeoffs for branching programs, in: *Proc. of 39th FOCS*, 1998, pp. 254–263.

[5] A. Borodin, A. Razborov, and R. Smolensky, On lower bounds for read-k-times branching programs, *Computational Complexity* **3** (1993), pp. 1–18.

[6] R. Canetti and O. Goldreich, Bounds on tradeoffs between randomness and communication complexity, *Computational Complexity* **3** (1993), pp. 141–167.

[7] R. Fleischer, H. Jung, and K. Mehlhorn, A communication-randomness tradeoff for two-processor systems, *Information and Computation* **116** (1995), pp. 155–161.

[8] A. Hajnal, W. Maass, and G. Turán, On the communication complexity of graph properties, in: *Proc. of 20th ACM STOC*, 1988, pp. 186–191.

[9] J. Hromkovič, *Communication Complexity and Parallel Computing*, EATCS Texts in Theoretical Computer Science, Springer-Verlag, 1997.

[10] J. Hromkovič and M. Sauerhoff, Tradeoffs between nondeterminism and complexity for communication protocols and branching programs, in: *Proc. of STACS 2000*, LNCS 1770, pp. 145–156.

[11] S. Jukna, A note on read-k-times branching programs, *RAIRO Theor. Inf. and Applications* **29**:1 (1995), pp. 75–83.

[12] S. Jukna and A. Razborov, Neither reading few bits twice nor reading illegally helps much, *Discrete Appl. Math.* **85**:3 (1998), pp. 223–238.

[13] S. Jukna and G. Schnitger, On the complexity of graphs which lack small cliques, manuscript.

[14] E. Kushilevitz and N. Nisan, *Communication Complexity*, Cambridge University Press, 1997.

[15] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland, 1998.

[16] I. Newman, Private vs. common random bits in communication complexity, *Information Processing Letters* **39** (1991), pp. 67–71.

[17] E. A. Okol'nishnikova, On Lower Bounds for Branching Programs, *Siberian Advances in Mathematics* **3**:1 (1998), pp. 152–166.

[18] Ch. H. Papadimitriou and M. Sipser, Communication complexity, *J. Comput. Syst. Sci.* **28** (1984), pp. 260–269.

[19] A. Razborov, Lower bounds for deterministic and nondeterministic branching programs, in: *Proc. of FCT '91*, Lecture Notes in Computer Science **529**, Springer-Verlag 1991, pp. 47–60.

[20] M. Sauerhoff, Complexity theoretical results for randomized branching programs, PhD thesis, Univ. of Dortmund, Shaker 1999.

[21] I. E. Shparlinski, *Computational and Algorithmic Problems in Finite Fields*, Kluwer, 1992.

[22] A. Yao, The entropic limitations of VLSI computations, in: *Proc. 13th ACM STOC* (1981), pp. 308–311.

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'