Time-Space Tradeoff Lower Bounds for Integer Multiplication and Graphs of Arithmetic Functions

(Extended Abstract)

Martin Sauerhoff*
FB Informatik, LS 2, Univ. Dortmund, Germany
martin.sauerhoff@cs.uni-dortmund.de

Philipp Woelfel
FB Informatik, LS 2, Univ. Dortmund, Germany
philipp.woelfel@cs.uni-dortmund.de

ABSTRACT

We prove exponential size lower bounds for nondeterministic and randomized read-k BPs as well as a time-space tradeoff lower bound for unrestricted, deterministic multi-way BPs computing the middle bit of integer multiplication. The lower bound for randomized read-k BPs is superpolynomial as long as the error probability is superpolynomially small. For polynomially small error, we have a polynomial upper bound on the size of approximating read-once BPs for this function. The lower bounds follow from a more general result for the graphs of universal hash classes that is applicable to the graphs of arithmetic functions such as integer multiplication, convolution, and finite field multiplication.

Categories and Subject Descriptors

F.1.1 [**Models of Computation**]: Branching Programs, Random Access Machines; F.1.2 [**Modes of Computation**]: Nondeterminism, Probabilistic Computation; F.2.3 [**Analysis of Algorithms and Problem Complexity**]: Tradeoffs between Complexity Measures

General Terms

Theory

Keywords

Integer multiplication, branching program, random access machine, hash class, lower bound, time-space tradeoff.

1. INTRODUCTION

Branching programs (BPs) are the standard model for nonuniform, sequential computation (see [21] for a thorough introduction). We consider BPs for functions defined on variables taking values in the domain $D = \{0, \ldots, q-1\}$.

*Supported by DFG grant We 1066/9-2.

STOC'03, June 9–11, 2003, San Diego, California, USA.

DEFINITION 1. *A* (deterministic) q-way branching program *on the variable set* $X = \{x_1, \ldots, x_n\}$ *is a directed acyclic graph with one source and two sinks. The sinks are labeled by the constants* 0 *and* 1*, resp. Each interior node is labeled by a variable from* X*, has* q *outgoing edges, and each value from* D *is assigned to exactly one of these edges as a label. The BP computes a function* $f\colon D^n \to \{0,1\}$ *defined on* X *as follows. For an input* $a \in D^n$*,* $f(a)$ *is equal to the label of the sink reached by the* computation path for a*, which is the path from the source to a sink obtained by following the edge labeled by* a_i *for nodes labeled by* x_i*. The* size $|G|$ *of a BP* G *is the number of its nodes. The* space *is the logarithm of* $|G|$ *and the* length *(or* time*) is the maximum number of edges on a computation path.*

We call a graph *multi-way BP* if it is a q-way BP for some q. For $q = 2$, we obtain the usual model of *boolean BPs*. *Nondeterministic BPs* and *randomized BPs* are defined in the obvious way by introducing additional, unlabeled nodes at which nondeterministic or randomized decisions, resp., are taken. An *approximating BP for* f *with (two-sided) error* ε is a deterministic BP computing an *ε-approximation of* f, which is a function that differs from f on at most an ε-fraction of the inputs.

It is well-known [11] that space S and time T RAMs with an arbitrary instruction set can be simulated by size $2^{O(S)}$ and time T BPs. (Note that the space for a RAM whose registers are able to hold values from D includes a factor of $\log |D|$ for the width of the registers.) This explains the importance of proving superpolynomial lower bounds on the size of general BPs. So far, we still lack sufficiently powerful methods for obtaining such bounds for functions in P or NP. Nevertheless, considerable progress in the development of proof methods has been made by the investigation of less and less restricted BP models (see [21]).

In particular, there is an extensive amount of literature on several variants of oblivious BPs and read-k BPs. A BP is called *oblivious*, if the underlying graph of the BP is leveled and the nodes in each level are all labeled by the same variable or are all sinks. In a *(syntactic) read-k BP*, each variable may occur at most k times on each path in the graph. Note that in the latter model all paths have a length of at most k times the input length.

A major step towards time-space tradeoff lower bounds and towards the current proof methods has been the proof of exponential size lower bounds for nondeterministic read-k BPs due to Borodin, Razborov, and Smolensky [12]. In a se-

ries of recent breakthroughs, Beame, Jayram, and Saks [4], Ajtai [2, 3], and Beame, Saks, Sun, and Vee [5] have even managed to prove exponential size lower bounds for general BPs that are only restricted in the length of their computation paths. Some of the results in these papers are strong enough to give time-space tradeoff lower bounds for general BPs. The best tradeoffs are from [5] and are even valid for randomized BPs with small two-sided error. They are of the form $T = \Omega(n\sqrt{\log(n/S)/\log\log(n/S)})$ for input length n, space S, and time T in the case of boolean inputs and of the form $T = \Omega(n\log(n/S))$ for input variables taking values in a domain of linear size in n. Moreover, Beame and Vee [6] have obtained the even larger time lower bound $T = \Omega(n\log^2 n)$ for space $S \leq n^{1-\varepsilon}\log|D|$ and a function on a large domain D whose size is exponential in n.

So far, time-space tradeoff lower bounds for general BPs could only be achieved for a quite limited class of functions. For the boolean case, the only results are for quadratic forms based on Hankel matrices [3, 5], while the results for the large domain case are for the element distinctness problem, the related Hamming closeness problem, and again quadratic forms as well as tensor products [2, 4, 5, 6]. Apart from proving results for the most general BP model possible, it is therefore also important to apply the existing methods to a larger class of functions, preferably practically important ones. This is the line of research followed in the present paper.

It is natural that integer multiplication as one of the basic arithmetic functions has been in the focus of several complexity theoretical investigations. Let $\mathrm{MUL}_n\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^{2n}$ denote the (binary) integer multiplication function that for the inputs $x, y \in \{0,1\}^n$ computes the binary representation $z \in \{0,1\}^{2n}$ of the product of the integers represented by x and y. The boolean function $\mathrm{MUL}_{i,n}$ represents the output bit i of integer multiplication, i. e., $\mathrm{MUL}_{i,n}\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ is the function that maps the inputs x, y to z_i if $\mathrm{MUL}_n(x,y) = (z_{2n-1}, \ldots, z_0)$. Note that with respect to read-once projections, the function $\mathrm{MUL}_{n-1,n}$, called the *middle bit of integer multiplication*, is the hardest one for space-bounded models of computation [13, 21].

The branching program complexity of the middle bit of integer multiplication has been first investigated by Bryant [13] who has obtained an exponential lower bound for oblivious read-once BPs, better known as *OBDDs* (*ordered binary decision diagrams*). Gergov [17] has extended this to oblivious BPs of linear length. Ablayev and Karpinski [1] have applied Gergov's reduction to deduce that also randomized OBDDs require exponential size and they have shown that, in contrast to this, the graph of integer multiplication has randomized OBDDs of polynomial size. In 1995, Ponzio [20] has proved that even unrestricted read-once BPs for $\mathrm{MUL}_{n-1,n}$ require size $2^{\Omega(\sqrt{n})}$.

The fact that integer multiplication defines a universal hash class [15, 16, 23], called *multiplicative hash class*, has been used in [24] to derive an improved lower bound on the size of OBDDs for $\mathrm{MUL}_{n-1,n}$, and in [9], Ponzio's lower bound has been improved to the strongly exponential lower bound $2^{\Omega(n)}$ for the size of unrestricted read-once BPs. Since then, exponential lower bounds have been proved only for slightly more general read-once BP models that allow limited nondeterminism and for models where some but not all variables may be tested multiple times [7, 8, 10, 25].

However, even the nondeterministic read-once case remained open, and the lower bound methods did not seem to be strong enough for BPs that allow all variables to be read more than once. Moreover, while most of the more recent results are based on the observation that integer multiplication defines a universal hash class, nothing could be proved for single output bit functions belonging to other universal hash classes. This is not surprising, because for certain universal hash classes (using, e. g., convolution) it is possible to design even linear-size read-once BPs that compute an arbitrary output bit of the functions from this class. Here we apply universal hashing in a different way.

In Section 2.1 we define two properties of universal hash classes, called *linear c-universality* and *well-distributedness*. These are natural properties that important hash classes like those based on integer multiplication, convolution, or finite field multiplication enjoy. Then we prove that computing the graph of hash functions from linearly c-universal hash classes is hard for nondeterministic read-k BPs and for length-restricted q-way BPs, where q is large enough. Showing that the multiplicative hash class has these additional properties, we obtain that for these two BP models it is hard to verify whether the output bits $n-m, \ldots, n-1$ of integer multiplication have a fixed value, e. g. are all 1. This improves upon an earlier result in [18] saying that the verification of some carefully chosen, non-consecutive output bits requires exponential size for nondeterministic read-k BPs. In contrast to [18], our result allows us to construct a reduction to the middle bit of integer multiplication, which then yields exponential lower bounds on the size of nondeterministic read-k BPs and length-restricted q-way BPs for this function. Finally, using the result about the graph of universal hash classes, also hardness results for the graph of other arithmetic functions, i. e., convolution and finite field arithmetic, follow.

In the following, we state our main theorems about integer multiplication. The results about the graph of hash classes and its applications to convolution and finite field multiplication are stated in Section 2.2.

In addition to $\mathrm{MUL}_{n-1,n}$, we consider the extension $\mathrm{MUL}^q_{n-1,n}$ of this function to the q-ary case for q a power of 2. For $x, y \in \{0, \ldots, q-1\}^n$, let $\mathrm{MUL}^q_{n-1,n}(x,y) = 1$ if $z_{n-1} \geq q/2$ for the q-ary representation $z = (z_{2n-1}, \ldots, z_0) \in \{0, \ldots, q-1\}^{2n}$ of the product of the numbers represented by x and y, and $\mathrm{MUL}^q_{n-1,n}(x,y) = 0$ otherwise. Note that since q is a power of 2, $\mathrm{MUL}^q_{n-1,n}$ is equal to the middle bit of binary integer multiplication for $(n\log q)$-bit numbers.

Theorem 1.

1. *There is a constant $\gamma > 0$ such that for any $k \leq \gamma \log n$, each nondeterministic read-k q-way BP for $\mathrm{MUL}^q_{n-1,n}$ has size $2^{\Omega(n\log q \cdot k^{-2}3^{-4k})}$. In particular, for $q = 2$, this bound applies to boolean BPs and $\mathrm{MUL}_{n-1,n}$.*

2. *For any $q = n^{O(1)} \geq 2^{120}$ there is a constant $\gamma' > 0$ such that each q-way BP for $\mathrm{MUL}^q_{n-1,n}$ of length $kn \leq \gamma' n\log q$ has size $2^{\Omega(n\log q \cdot k^{-2}3^{-6k})}$.*

The second part implies the time-space tradeoff lower bound $T = \Omega(n\log((n\log q)/S))$ for space S and time T on RAMs with $(\log q)$-bit registers for any $q = n^{\Theta(1)}$. In particular, for space $S = (n\log q)^{1-\delta}$ for some constant $\delta > 0$, the time is $T = \Omega(n\log n)$.

For read-k BPs, we also get a lower bound in the randomized case. The bound is superpolynomially large as long as the error probability is superpolynomially small.

THEOREM 2. *Let $\varepsilon = \varepsilon(n)$ be a non-increasing function in n such that $\varepsilon \geq 2^{-n \log q \cdot 3^{-2k}}$. There is a constant $\gamma > 0$ such that for any $k \leq \gamma \log n$, each randomized read-k q-way BP for $\mathrm{MUL}^q_{n-1,n}$ with error ε requires size $2^{\Omega(\log(1/\varepsilon) \cdot k^{-2} 3^{-2k} - \log q)}$.*

Finally, complementing this result, we prove that if we are content with a deterministic BP that may err with polynomially small probability on inputs chosen uniformly at random, i. e., with an approximating BP, then $\mathrm{MUL}^q_{n-1,n}$ and its boolean variant $\mathrm{MUL}_{n-1,n}$ can be represented in small size even by read-once BPs.

THEOREM 3. *The function $\mathrm{MUL}^q_{n-1,n}$ can be approximated with error ε by read-k q-way BPs of size $2^{O((\log(1/\varepsilon) + \log n)/k + \log q)}$.*

We remark that although the case $q = 2^\ell$ is most natural, analogous results can be obtained for arbitrary prime powers q.

2. LOWER BOUNDS

2.1 Universal Hash Classes

Let U and R be two finite sets with $|U| \leq |R|$, called *universe* and *range*, resp. We call a family $\mathcal{H}$ of functions $U \to R$ *hash class*. The functions in $\mathcal{H}$ are called *hash functions* and the elements in the universe are called *keys*. We say that two keys $x \neq x'$ *collide* under a hash function $h \in \mathcal{H}$ if $h(x) = h(x')$. Carter and Wegman [14, 22] have introduced the following definitions. A hash class $\mathcal{H}$ is called *universal*, if any two distinct keys collide under a randomly chosen hash function from $\mathcal{H}$ with a probability of at most $1/|R|$. It is called *strongly universal*, if for any two keys $x \neq x'$ and a randomly chosen $h \in \mathcal{H}$ the hash values $h(x)$ and $h(x')$ are independently and uniformly distributed. Since then, many generalizations and modifications of the original definitions have been considered. An example are *c-universal* hash classes, where the collision probability of two keys is bounded by $c/|R|$ instead of $1/|R|$.

In conjunction with branching program complexity, strongly universal hash classes have been considered by Mansour, Nisan, and Tiwari [19]. Let $\mathcal{H}$ be a hash class with universe $U = \{0,1\}^n$ and range $R = \{0,1\}^m$. The function $f_\mathcal{H}$ takes a string describing a hash function $h \in \mathcal{H}$ and a key $x \in U$ as input and computes the m-bit string $h(x)$. An extended variant of the usual BP model, called *multi-output BP* here, allows to produce output bits at edges traversed during the computation (see, e. g., [11]). A multi-output BP can be used to compute the function $f_\mathcal{H}$. For strongly universal hash families $\mathcal{H}$, a time-space tradeoff lower bound of $TS = \Omega(mn)$ for multi-output BPs representing $f_\mathcal{H}$ has been shown in [19]. Note that the hash classes that are called *universal* in [19] are *strongly universal* according to our definition.

Clearly, it is not easier to compute multiple output bits of a function than to *verify* a given function value. In this paper, we are interested in the complexity of single output bit functions and thus consider the *graph* of hash classes.

DEFINITION 2. *Let $\mathcal{H}$ be a hash class with universe U and range R, where the elements from $\mathcal{H}$, U, and R can be encoded by vectors over $D = \{0, \ldots, q-1\}$. We define the* graph *of $\mathcal{H}$, denoted by $\mathrm{GRAPH}_\mathcal{H}$, for $x \in U$, $y \in R$, and $h \in \mathcal{H}$ by $\mathrm{GRAPH}_\mathcal{H}(x,y,h) = 1$ if $h(x) = y$ and $\mathrm{GRAPH}_\mathcal{H}(x,y,h) = 0$ otherwise. For fixed $y \in R$, we let $\mathrm{GRAPH}_{\mathcal{H},y}(x,h) = \mathrm{GRAPH}_\mathcal{H}(x,y,h)$.*

The inputs for $\mathrm{GRAPH}_\mathcal{H}$ are vectors over D encoding x, y, and h. We will make the encoding precise later on. We show hardness results for the graph of c-universal hash classes that fulfill two additional properties.

DEFINITION 3. *Let $(U,+)$ be an abelian group. A hash class $\mathcal{H}$ with universe U and range R is called* linearly c-universal*, if for any two distinct keys x, x' and a randomly chosen hash function $h \in \mathcal{H}$,*

$$\mathrm{Prob}\left(\exists z \in U :\ h(x+z) = h(x'+z)\right) \leq c/|R|.$$

A linearly 1-universal hash class is called linearly universal*. A hash class is called* well-distributed*, if for any fixed $h \in \mathcal{H}$ the hash value $h(x)$ of a randomly chosen $x \in U$ is uniformly distributed over the range.*

At the first glance, linear c-universality seems to be much harder to achieve than only c-universality. In the following we show, though, that several well-known c-universal hash classes are in fact linearly c-universal. The following observation is helpful.

REMARK 1. *A hash class with universe $(U,+)$ consisting entirely of homomorphisms is c-universal if and only if it is linearly c-universal.*

For a vector $v = (v_{N-1}, \ldots, v_0)$, let $[v]_i^j$ denote the vector $(v_j, \ldots, v_i)$. We consider the following hash classes.

The Finite Field Class. Let q be a prime power and let $\mathbb{F}_q = \{0, \ldots, q-1\}$ be the finite field of cardinality q and $U = \mathbb{F}_q^n$ and $R = \mathbb{F}_q^m$ be extension fields of $\mathbb{F}_q$. The *finite field class* is the family $\mathcal{F} = \mathcal{F}_{q,n,m,i}$, $0 \leq i \leq n-m$, which consists of the functions $f_a \colon U \to R$, $x \mapsto [a \cdot x]_i^{i+m-1}$, with $a \in \mathbb{F}_q^n - \{0\}$. It is well known that the hash class $\mathcal{F} \cup \{f_0\}$ is a universal hash class. Hence, $\mathcal{F}$ is also universal, since all pairs of keys collide under f_0. Since all its functions are homomorphisms, it is even linearly universal. (Note that the mapping $x \mapsto [x]_i^j$ is a homomorphism.) If $a \neq 0$, then $a \cdot x$ and $[a \cdot x]_i^{i+m-1}$ are uniformly distributed over U and R, resp. (for randomly chosen $x \in U$). Hence, $\mathcal{F}$ is well-distributed.

The Convolution Class. Let q be a prime power and let $U = \mathbb{F}_q^n$, $R = \mathbb{F}_q^m$ be the groups with componentwise addition in $\mathbb{F}_q$. Let $x = (x_{n-1}, \ldots, x_0) \in \mathbb{F}_q^n$ and $y = (y_{n'-1}, \ldots, y_0) \in \mathbb{F}_q^{n'}$. Assuming $x_j = 0$ or $y_{i-j} = 0$ for $j \geq n$ or $i - j \geq n'$, resp., the convolution of x and y, written $x \circ y$, is the string $z = (z_{n+n'-2}, \ldots, z_0) \in \mathbb{F}^{m+n'-1}$, where $z_i = \sum_{j=0}^{i} x_j y_{i-j}$. (The multiplications and additions are in $\mathbb{F}_q$.) It has been shown in [19] that for $q = 2$ the mappings $U \to R$, $x \mapsto [a \circ x]_{n-1}^{n+m-2} + b$ with $a \in \mathbb{F}_q^{n+m-1}$ and $b \in \mathbb{F}_q^m$ define a strongly universal hash class, and this is well known to be true also for all other prime powers q. By the proof in [19] it is easy to see that the hash class remains universal (but not strongly universal), if one omits the addition of the parameter b. While the resulting class is in fact linearly universal because all its functions are homomorphisms, it is not well-distributed, though.

Therefore, we use only a subset of the functions by shortening the parameter a in such a way that it has the same length as the keys and by fixing its least significant digit to 1. The *convolution class* $\mathcal{C}_{q,n,m}$ consists of all functions $g_a\colon U \to R$, $x \mapsto [a \circ x]_{n-m}^{n-1}$, where $a = (a_{n-1}, \ldots, a_1, 1) \in \mathbb{F}_q^n$. By similar arguments as in [19], it can be seen that $\mathcal{C}_{q,n,m}$ is universal. Since all its functions are homomorphisms, $\mathcal{C}_{q,n,m}$ is even linearly universal.

It is convenient to identify each $b = (b_{n-1}, \ldots, b_0) \in \mathbb{F}_q^n$ with the polynomial $\sum_{i=0}^{n-1} b_i X^i$ in $\mathbb{F}_q[X]$. Then the operation $\odot$ defined by $a \odot b = [a \circ b]_0^{n-1}$ corresponds to the polynomial multiplication modulo X^n and thus $\left(\mathbb{F}_q^n, \odot\right)$ is a monoid. It is easy to see that each $a = (a_{n-1}, \ldots, a_1, 1)$ is an invertible element of the monoid which implies directly that $x \odot a$ takes each value in $\mathbb{F}_q^n$ exactly once. Hence, $\mathcal{C}_{q,n,m}$ is well-distributed.

The Multiplicative Class. For a positive integer N, $\mathbb{Z}_N$ denotes the (additive) group $\{0, \ldots, N-1\}$ modulo N. Let q be a power of the prime p. The family $\mathcal{M}_{q,n,m}$ consists of all hash functions $f_a\colon \mathbb{Z}_{q^n} \to \mathbb{Z}_{q^m}$, $x \mapsto \left\lfloor ((a \cdot x) \bmod q^n)/q^{n-m} \right\rfloor$, where $a \in \mathbb{Z}'_{q^n} = \{ip + 1 \mid 0 \le i < q^n/p\}$. It has been proved in [16] that $\mathcal{M}_{2,n,m}$ is 2-universal. A generalization of the proof shows that $\mathcal{M}_{q,n,m}$ is also 2-universal if p is an arbitrary prime (see [23]). Following these proofs, it is easy to see that $\mathcal{M}_{q,n,m}$ is in fact linearly 2-universal and even well-distributed.

We summarize the results about the hash classes in the following theorem

THEOREM 4. *The hash classes $\mathcal{F}_{q,n,m,i}$ and $\mathcal{C}_{q,n,m}$ are linearly universal, $\mathcal{M}_{q,n,m}$ is linearly 2-universal, and all these hash classes are well-distributed.*

There may be other known universal hash classes for which it can be shown that they are in fact linearly universal. Nevertheless the three classes suggested here are of prime interest for complexity theoretical investigations.

2.2 The Graphs of Arithmetic Functions

Let q be a prime power and $D = \{0, \ldots, q-1\}$. Let U be one of the groups $(\mathbb{F}_q^n, +)$ or $(\mathbb{Z}_{q^n}, +)$. Define the function $\mathrm{val}\colon D^n \to U$ for $x \in D^n$ by $\mathrm{val}(x) = x$ if $U = (\mathbb{F}_q^n, +)$ and $\mathrm{val}(x) = \sum_{i=0}^{n-1} x_i q^i$ if $U = (\mathbb{Z}_{q^n}, +)$.

We consider the function $\mathrm{GRAPH}_{\mathcal{H}}$ from the last section and make the encoding of the inputs over D more explicit. Let $\ell = \left\lceil \log_q |\mathcal{H}| \right\rceil$. For $x \in D^n$, $y \in D^m$, and $z \in D^\ell$, let $\mathrm{GRAPH}_{\mathcal{H}}(x, y, z) = 1$ if z is the code of a hash function $h_z \in \mathcal{H}$ and $h_z(\mathrm{val}(x)) = \mathrm{val}(y)$ and let $\mathrm{GRAPH}_{\mathcal{H}}(x, y, z) = 0$ otherwise.

The following two generic lower bounds for the graphs of hash classes allow us to obtain our main results for integer multiplication and the additional results for the graphs of convolution and finite field multiplication.

THEOREM 5. *Let $\mathcal{H}$ be a linearly c-universal hash class. Let k be a positive integer with $64k^2 3^{k+1} \le n$.*

1. *Let $644k^2 3^{k+1} \le m \le n/3^k$. Then each nondeterministic read-k q-way BP for $\mathrm{GRAPH}_{\mathcal{H}}$ has size $2^{\Omega((m \log q - 6 \log c) \cdot k^{-2} 3^{-k})}$.*
2. *Let $m = \lfloor n/3^k \rfloor \ge 60$. Then there is a constant $\lambda > 1/100$ such that for all $q \ge 2^{120}$, each q-way BP G of length $kn \le \lambda n \log q$ for $\mathrm{GRAPH}_{\mathcal{H}}$ has size $2^{\Omega((m \log q - 12 \log c) \cdot k^{-2} 3^{-k})}$.*

THEOREM 6. *Let $\mathcal{H}$ be a linearly c-universal, well-distributed hash class. Let k be a positive integer with $64k^2 3^{k+1} \le n$. Let $\varepsilon > 0$ be such that $\log_q(1/16\varepsilon) \ge m + 1$. Let $y \in D^m$ be arbitrarily chosen.*

1. *Let $644k^2 3^{k+1} \le m \le n/3^k$. Then each nondeterministic read-k q-way BP for $\mathrm{GRAPH}_{\mathcal{H},y}$ and each read-k BP approximating $\mathrm{GRAPH}_{\mathcal{H},y}$ with error ε has size $2^{\Omega((m \log q - 6 \log c) \cdot k^{-2} 3^{-k})}$.*
2. *Let $m = \lfloor n/3^k \rfloor \ge 60$. Then there is a constant $\lambda > 1/100$ such that for all $q \ge 2^{120}$, each q-way BP G of length $kn \le \lambda n \log q$ computing $\mathrm{GRAPH}_{\mathcal{H},y}$ has size $2^{\Omega((m \log q - 12 \log c) \cdot k^{-2} 3^{-k})}$.*

The proofs of these theorems are presented in Section 3. We now consider the families $\mathcal{C}_{q,n,m}$, $\mathcal{F}_{q,n,m,i}$, and $\mathcal{M}_{q,n,m}$ from Section 2.1. Since these families are all linearly c-universal for constant c and well-distributed, Theorems 5 and 6 yield hardness results for the verification of the underlying arithmetic functions.

Let $\mathrm{CONV}_n^{[i\ldots j]}$, $\mathrm{FMUL}_n^{[i\ldots j]}$, and $\mathrm{MUL}_n^{[i\ldots j]}$ be the functions mapping two n-digit strings to the output digits at the positions $i, \ldots, j$ of the convolution, finite field product, or integer product, resp. More precisely, for $x, y \in D^n$ and val the identity over $\mathbb{F}_q^n$, $\mathrm{CONV}_n^{[i\ldots j]} = [\mathrm{val}(x) \circ \mathrm{val}(y)]_i^j$. The functions $\mathrm{FMUL}_n^{[i\ldots j]}$ and $\mathrm{MUL}_n^{[i\ldots j]}$ are defined analogously. For any function $f\colon D^n \to D^m$, let $\mathrm{GRAPH}_{f,y}\colon D^n \to \{0,1\}$ map the input x to 1 if $f(x) = 1$ and to 0, otherwise.

COROLLARY 1. *Let k be a positive integer such that $64k^2 3^{k+1} \le n$. Let f be any of the functions $\mathrm{FMUL}_n^{[i\ldots i+m-1]}$, $0 \le i \le n-m$, $\mathrm{CONV}_n^{[n-m\ldots n-1]}$, or $\mathrm{MUL}_n^{[n-m\ldots n-1]}$ and let $y \in D^m$ be arbitrarily fixed.*

1. *Suppose that $644k^2 3^{k+1} \le m \le n/3^k$ and that $\varepsilon > 0$ is such that $\log_q(1/16\varepsilon) \ge m + 1$. Then each nondeterministic read-k q-way BP and each approximating read-k q-way BP with error ε for $\mathrm{GRAPH}_{f,y}$ has a size of $2^{\Omega(m \log q \cdot k^{-2} 3^{-k})}$.*
2. *Let $m = \lfloor n/3^k \rfloor \ge 60$ and $q \ge 2^{120}$. Then each general q-way BP of length $kn \le (n \log q)/100$ for $\mathrm{GRAPH}_{f,y}$ has a size of $2^{\Omega(m \log q \cdot k^{-2} 3^{-k})}$.*

PROOF. We prove the corollary for $f = \mathrm{MUL}_n^{n-m\ldots n-1}$. The results for the other two functions follow with analogous arguments. Assume the existence of a read-k q-way BP G of length ℓ and size s that computes $\mathrm{GRAPH}_{f,y}$ with error ε. The inputs for the BP are the strings $x = (x_{n-1}, \ldots, x_0)$ and $z = (z_{n-1}, \ldots, z_0)$ in D^n. We redirect all edges leaving a z_0-node and which are labeled with a value not equal to $jp+1$ for some $0 \le j < q/p$ in such a way that they point to a 0-sink. Hence, the resulting BP G' computes the function g with error ε, where $g(x, z) = \mathrm{GRAPH}_{f,y}(x, z)$ if $z_0 = jp + 1$ $(0 \le j < q/p)$, and $g(x, z) = 0$, otherwise. Note that $\mathrm{val}(z) \in \mathbb{Z}'_{q^n} = \{ip + 1 \mid 0 \le i < q^n/p\}$ if and only if $z_0 = jp + 1$ for some $j \in \{0, \ldots, q/p - 1\}$. Since the multiplicative class $\mathcal{M} = \mathcal{M}_{q,n,m}$ consists exactly of the hash functions $h_z\colon x \mapsto \mathrm{MUL}_n^{n-m\ldots n-1}(x, z)$ with $z \in \mathbb{Z}'_{q^n}$, we have $g = \mathrm{GRAPH}_{\mathcal{M},y}$. Clearly, G' is read-k, its length

is at most ℓ, and its size is bounded by s. Hence, applying Theorem 6 with the parameter $c = 2$ (because $\mathcal{M}$ is linearly 2-universal according to Theorem 4), we obtain the claimed bounds on s. □

Hence, it is hard to verify m consecutive output digits of these basic arithmetic functions for suitable $m = \Omega(n)$. We get no hardness result for $m = 1$, i. e., for computing only one output digit. For convolution, this is not surprising because it is easy to see that any single output digit of the convolution can be computed by read-once BPs of linear size. For finite field multiplication, we leave open whether a lower bound for single output digits can be proved by different means. In the remainder of the section, we deal with integer multiplication and show that in this case, single output digits are indeed hard to compute.

2.3 The Middle Bit of Integer Multiplication

In this section, we consider only the case where q is a power of 2. Nevertheless, all results can be generalized to powers of other primes. Let $D = \{0, \dots, q-1\}$. Let val now be the function $D^n \to \mathbb{Z}_{q^n}$ mapping a q-ary string to the integer it represents (as defined in the previous section). For the sake of readability, we write $|x|$ instead of $\mathrm{val}(x)$. Recall that the function $\mathrm{MUL}^q_{n-1,n}$ computes the middle bit of the product of two integers given as n-digit q-ary strings.

In order to apply the results about the graph of the multiplicative hash class from the previous section, we construct a reduction from $\mathrm{MUL}^q_{n-1,n}$ to this function.

LEMMA 1. *Suppose there is a sequence of (nondeterministic) read-k q-way BPs G_N of length $\ell(N)$ and size $s(N)$ that compute $\mathrm{MUL}^q_{N-1,N}$. Then for any n, m, there is a $y \in \mathbb{Z}_q^m$ and a (nondeterministic) read-$(2k)$ q-way BP of size $O(q \cdot s(2n))$ and length $\ell(n) + \ell(2n) + 1$ that computes $\mathrm{GRAPH}_{\mathcal{M},y}$, $\mathcal{M} = \mathcal{M}_{q,n,m}$. Analogously, if the G_N are randomized BPs with error $\varepsilon(N)$, then randomized BPs for $\mathrm{GRAPH}_{\mathcal{M},y}$ with the same restrictions as above and error $\varepsilon(n) + \varepsilon(2n)$ exist.*

PROOF. Suppose the claimed BPs G_N for $\mathrm{MUL}^q_{N-1,N}$ exist. Let n, m be fixed and let $y \in D^m$ be the unique q-ary m-digit string for which $|y| = q^m/2$. Consider the inputs $x, z \in D^n$ for the function $\mathrm{GRAPH}_{\mathcal{M},y}$, where z describes the hash function f_z defined by $f_z(x) = \left\lfloor \big((|x||z|) \bmod q^n\big)/q^{n-m} \right\rfloor$. Note that according to the definition of $\mathcal{M}$, f_z is in $\mathcal{M}$ if and only if $|z| \in \mathbb{Z}'_{q^n}$, where $\mathbb{Z}'_{q^n}$ is the set of odd integers in $\mathbb{Z}_{q^n}$ (due to the assumption that q is a power of 2).

We may assume w. l. o. g. that the BP for $\mathrm{GRAPH}_{\mathcal{M},y}$ to be constructed in this proof first tests whether $|z|$ is odd by examining the least significant digit of z. If this is not the case, then $f_z \notin \mathcal{M}$, and the BP outputs 0 according to the definition of $\mathrm{GRAPH}_{\mathcal{M},y}$. It is easy to see that each read-k BP can be modified without destroying the read-k property in such a way that it performs this test. The length increases by at most 1, and the size only by a factor of q.

Assume now that $f_z \in \mathcal{M}$. Let $x', z' \in D^{2n}$ such that $|x'| = |x| \cdot q^n + 1$ and $|z'| = |z| + q^{2n}/2 - q^{2n-m}$. Note that $x' = (x'_{2n-1}, \dots, x'_0)$ where $x'_0 = 1$, $x'_{i+n} = x_i$ for $0 \le i \le n-1$, and all other digits are 0. Similarly, $z' = (z'_{2n-1}, \dots, z'_0)$ with $z'_i = z_i$ for $0 \le i \le n-1$, $z'_i = 0$ for $n \le i \le 2n-m-1$, $z'_i = q-1$ for $2n-m \le i \le 2n-2$, and $z_{2n-1} = q/2 - 1$.

CLAIM 1. $\mathrm{GRAPH}_{\mathcal{M},y}(x, z) = 1$ *if and only if* $\mathrm{MUL}^q_{n-1,n}(x, z) = MUL^q_{2n-1,2n}(x', z') = 1$.

From this claim, the statement of the lemma follows right away: One can easily construct a (nondeterministic) BP for $\mathrm{MUL}^q_{n-1,n}(x, z) \wedge \mathrm{MUL}^q_{2n-1,2n}(x', z')$ by replacing the 1-sink of the BP for $\mathrm{MUL}^q_{n-1,n}$ with the source of the BP for $\mathrm{MUL}^q_{2n-1,2n}$. This BP is read-$(2k)$ and has a size of at most $2s(2n)$ and length $\ell(n) + \ell(2n)$. According to the discussion above one can modify it in such a way that the resulting BP tests whether $f_z \in \mathcal{M}$, and the read-$(2k)$ restriction and the claimed size and length bounds are valid. Moreover, according to Claim 1 the resulting BP computes $\mathrm{GRAPH}_{\mathcal{M},y}$ if the input (x, z) and the transformed input (x', z') are plugged into it. If the BPs for the middle bit function are randomized with error $\varepsilon(n)$ and $\varepsilon(2n)$, resp., then the constructed BP for the graph errs with a probability of at most $\varepsilon(n) + \varepsilon(2n)$. Hence, it suffices to show Claim 1.

We have $\mathrm{GRAPH}_{\mathcal{M},y}(x, z) = 1$ if and only if

$$|y| \cdot q^{n-m} \;\le\; (|x||z|) \bmod q^n \;<\; (|y|+1)q^{n-m}.$$

Since $|y| = q^m/2$,

$$\begin{aligned} &\mathrm{GRAPH}_{\mathcal{M},y}(x, z) = 1 \\ &\Leftrightarrow\; q^n/2 \;\le\; (|x||z|) \bmod q^n \;<\; q^n/2 + q^{n-m} \\ &\Leftrightarrow\; q^{2n}/2 \;\le\; (q^n|x||z|) \bmod q^{2n} \;<\; q^{2n}/2 + q^{2n-m}. \end{aligned} \tag{1}$$

Similarly,

$$\mathrm{MUL}^q_{n-1,n}(x, z) = 1 \;\Leftrightarrow\; (q^n|x||z|) \bmod q^{2n} \;\ge\; q^{2n}/2. \tag{2}$$

Using

$$\begin{aligned} |x'||z'| \;&\equiv\; \big(q^n|x| + 1\big) \cdot \big(|z| + q^{2n}/2 - q^{2n-m}\big) \\ &\equiv\; q^n|x||z| + q^{2n}/2 - q^{2n-m} + |z| \pmod{q^{2n}}, \end{aligned}$$

we obtain

$$\begin{aligned} &\mathrm{MUL}^q_{2n-1,2n}(x', z') \;=\; 1 \\ &\quad\Leftrightarrow \big(q^n|x||z| + q^{2n}/2 - q^{2n-m} + |z|\big) \bmod q^{2n} \ge q^{2n}/2 \\ &\quad\Leftrightarrow \big(q^n|x||z| - q^{2n-m} + |z|\big) \bmod q^{2n} < q^{2n}/2. \end{aligned}$$

Together with (2) this means that $\mathrm{MUL}^q_{n-1,n}(x, z) = \mathrm{MUL}^q_{2n-1,2n}(x', z') = 1$ if and only if

$$\begin{aligned} &\big(q^n|x||z| - q^{2n-m} + |z|\big) \bmod q^{2n} \\ &\qquad\qquad < \; q^{2n}/2 \;\le\; \big(q^n|x||z|\big) \bmod q^{2n}. \end{aligned} \tag{3}$$

It can be easily seen that for any $a, b \in \mathbb{Z}$ and $0 \le b < N/2$,

$$\begin{aligned} &(a-b) \bmod N \;<\; N/2 \;\le\; a \bmod N \\ &\qquad\Leftrightarrow\quad N/2 \;\le\; a \bmod N \;<\; (N/2+b) \bmod N. \end{aligned}$$

Hence, (3) is equivalent to

$$q^{2n}/2 \;\le\; (q^n|x||z|) \bmod q^{2n} \;<\; q^{2n}/2 + q^{2n-m} - |z|.$$

This is equivalent to (1), because $q^n|x||z|$ is a multiple of q^n and $|z| < q^n$. □

We now combine the fact that the multiplicative hash class is linearly 2-universal and well-distributed, Theorem 6, and the above reduction to prove our main results about integer multiplication.

PROOF OF THEOREM 1. We first consider part 2 and deal with part 1 afterwards.

Part 2: Let a sequence of BPs G_N of length kN and size $s(N)$ for $\mathrm{MUL}^q_{N-1,N}$ be given. We use these BPs to construct a BP for the graph of the multiplicative hash class and apply Theorem 6(2) to the latter function.

For any sufficiently large n, let $K = K(n) = 3k+1$, $m = m(n) = \lfloor n/3^K \rfloor$, and $\mathcal{H} = \mathcal{M}_{q,n,m}$. By the hypothesis of part 2 of Theorem 1, $\log q \le \alpha \log n$ for sufficiently large n and a constant $\alpha > 0$. We claim that the constant parameter $\gamma' > 0$ in the theorem can be chosen such that for $k \le \gamma' \log q$,

(i) $K \le \lambda \log q$, where $\lambda > 0$ is the constant from Theorem 6(2);

(ii) $64K^2 3^{K+1} \le 2048K^3 3^{2K} \le n$; and

(iii) $m = \lfloor n/3^K \rfloor \ge 60$.

We prove this first. For (i), observe that $K = 3k+1 \le 4k$ for each positive integer k, and thus $k \le (\lambda/4)\log q$ implies the desired bound on K. Now consider (ii). The first inequality is obviously true for each integer $K \ge 1$ and thus in particular for each integer $k \ge 1$. Furthermore, for each integer $k \ge 1$,

$$2048K^3 3^{2K} \;\le\; 2^{11}(4k)^3 3^{6k+2} \;=\; 2^{17}\cdot 3^2 \cdot k^3 \cdot 3^{6k} \;\le\; 3^{19k}.$$

Hence, for $k \le (\log q)/(19 \log 3 \cdot \alpha)$, the second inequality of (ii) is satisfied. Finally, for sufficiently large n, the latter bound on k also implies (iii). Altogether, (i)–(iii) are satisfied for $1 \le k \le \gamma' \log q$ where $\gamma' = \min\{\lambda/4, 1/(19\log 3 \cdot \alpha)\}$.

Due to Lemma 1, we get a $y \in D^m$ and a BP of length $3kn + 1 \le Kn$ and size $s'(n) = O(q \cdot s(2n))$ for $\mathrm{GRAPH}_{\mathcal{H},y}$. We want to apply Theorem 6(2) to $\mathrm{GRAPH}_{\mathcal{H},y}$ and first check that the assumptions in the hypothesis are satisfied. We have $m = \lfloor n/3^K \rfloor \ge 60$ by the definition of m and (iii), $64K^2 3^{K+1} \le n$ by (ii), and $Kn \le \lambda n \log q$ by (i). We have $q \ge 2^{120}$ by the hypothesis of Theorem 1. Furthermore, $\mathcal{H} = \mathcal{M}_{q,n,m}$ is linearly 2-universal and well-distributed. Hence, the theorem is applicable and yields

$$s'(n) \;\ge\; 2^{\,c' m \log q \cdot K^{-2} 3^{-K}}$$

for a constant $c' > 0$ and n sufficiently large. Hence,

$$s'(n)/q \;\ge\; 2^{\,c'(m - (K^2 3^K)/c') \log q \cdot K^{-2} 3^{-K}}.$$

Now $m = \lfloor n/3^K \rfloor$ and $n \ge 2^{11} K^3 3^{2K}$ by (ii). Thus, assuming $k \ge (1/(3\cdot 2^{11})) \cdot (2/c' + 1)$ and using $K \ge 3k$,

$$\begin{aligned} m \;&\ge\; n/3^K - 1 \;\ge\; 2^{11} K^3 3^K - 1 \\ &\ge\; K^2 3^K \cdot (2^{11} K - 1) \;\ge\; 2 \cdot (K^2 3^K)/c'. \end{aligned}$$

It follows that

$$\begin{aligned} s'(n)/q \;&\ge\; 2^{\,(1/2) c' m \log q \cdot K^{-2} 3^{-K}} \\ &\ge\; 2^{\,(1/2) c' m \log q \cdot k^{-2} 3^{-6k}} \end{aligned}$$

for $k \ge \max\{1, (1/(3\cdot 2^{11})) \cdot (2/c' + 1)\}$. Since $s(2n) = \Omega(s'(n)/q)$, this gives the claimed result.

Part 1: For each N, let a nondeterministic read-k BP G_N of size $s(N)$ for $\mathrm{MUL}^q_{N-1,N}$ be given. Let n be any sufficiently large integer. Let $K = 2k$ and let $\mathcal{H} = \mathcal{M}_{q,n,m}$ with $m = \lfloor n/3^K \rfloor$. Lemma 1 yields a $y \in D^m$ and a read-K BP of size $s'(n)$ with $s'(n) = O(q \cdot s(2n))$ for $\mathrm{GRAPH}_{\mathcal{H},y}$.

We choose $\gamma = 1/(19 \log 3)$ for the constant in Theorem 1(1). Then, by the hypothesis of this theorem, $k \le \gamma \log n$. We verify that the assumptions in the hypothesis of Theorem 6(1) are satisfied.

By Fact (ii) of the first part, $64K^2 3^{K+1} \le n$. Also by Fact (ii),

$$n \ge 2048K^3 3^{2K} \ge 2047K^3 3^{2K} + 3^K \ge 644K^2 3^{2K+1} + 3^K$$

and hence

$$m = \lfloor n/3^K \rfloor \;\ge\; n/3^K - 1 \;\ge\; 644K^2 3^{K+1}.$$

By Theorem 6(1), $s'(n) \ge 2^{c' m \log q \cdot K^{-2} 3^{-K}}$ for a constant $c' > 0$ and n sufficiently large. The rest of the proof is done in the same way as for the first part. □

PROOF OF THEOREM 2. We use the lower bound from Theorem 6(1) for approximating read-k BPs. By Yao's principle, each lower bound for approximating BPs yields a lower bound of the same size and for the same error probability for the randomized variant of the considered BP model.

For each N, let a randomized read-k BP G_N of size $s(N)$ and error $\varepsilon(N)$ for $\mathrm{MUL}^q_{N-1,N}$ be given. Let n be any sufficiently large integer, $K = 2k$, $\varepsilon'(n) = 2\varepsilon(n) \ge \varepsilon(n) + \varepsilon(2n)$, $m = \lfloor \log_q(1/(16\varepsilon'(n))) \rfloor - 1$, and $\mathcal{H} = \mathcal{M}_{q,n,m}$. Lemma 1 yields a $y \in D^m$ and a randomized read-K BP for $\mathrm{GRAPH}_{\mathcal{H},y}$ with size $s'(n) = O(q \cdot s(2n))$ and error $\varepsilon'(n)$.

We now make sure that the assumptions in the hypothesis of Theorem 6(1) are satisfied. Since $\varepsilon(n) \ge 2^{-n \log q \cdot 3^{-2k}}$ by the hypothesis of Theorem 2,

$$m \;\le\; \big(\log(1/\varepsilon(n)) - 5\big)/(\log q) - 1 \;\le\; n/3^{2k} \;\le\; n/3^K,$$

as required for Theorem 6(1). Setting $\gamma = 1/(19\log 3)$ for the constant in Theorem 2, $1 \le k \le \gamma \log n$ implies $64K^2 3^{K+1} \le n$ and $644K^2 3^{K+1} \le m$ analogously to the proof of Theorem 1. By Theorem 6(1), for a constant $c' > 0$ and n sufficiently large,

$$s'(n) \;\ge\; 2^{\,c' m \log q \cdot K^{-2} 3^{-K}} \;=\; 2^{\Omega(\log(1/\varepsilon(2n)) \cdot k^{-2} 3^{-2k})}.$$

Hence, due to the definitions of ε, K, and m and the fact that $\varepsilon(n)$ is non-increasing, $s(n) = 2^{\Omega(\log(1/\varepsilon)\cdot k^{-2} 3^{-2k} - \log q)}$ as claimed. □

Finally, we sketch the proof of the upper bound on the size of approximating BPs for $\mathrm{MUL}^q_{n-1,n}$ in Theorem 3. We need the following lemma from [16]. By $\gcd(x,y)$ we denote the greatest common divisor of two positive integers x and y.

LEMMA 2 ([16]). *Let N be a positive integer, $a \in \mathbb{Z}_N - \{0\}$, and $\gamma = \gcd(a, N)$. If x is chosen randomly from $\mathbb{Z}_N$, then $(ax) \bmod N$ is uniformly distributed over $\{i\gamma \mid 0 \le i < N/\gamma\}$.*

PROOF OF THEOREM 3 (SKETCH). We only consider the case $k = 1$ and construct a read-once BP with error ε. Let $m = \lceil \log_q(2n/\varepsilon) \rceil + 1$, which is only logarithmic in n for polynomially small error. We compute the output digits $(z_{n-1}, \ldots, z_{n-m})$ of integer multiplication correctly for the situation where the carry from the first $n-m$ digits is zero. Then we show that the output of $\mathrm{MUL}^q_{n-1,n}$ is not influenced by this carry for too many inputs.

More formally, for $x, y \in D^n$ let

$$s \;=\; s(x,y) \;=\; \sum_{i=0}^{n-1} y_i \cdot |(x_{n-1-i}, \ldots, x_{n-m-i})|,$$

where $x_i = 0$ for $i < 0$. Let $\mathrm{MUL}^*(x,y) = 1$ if $s \bmod q^m \geq q^m/2$ and 0 otherwise. Observe that the x-vector in the ith term of s is obtained from that in the $(i-1)$-th term by removing x_{n-i} in the front and appending x_{n-m-i} to the end. It is easy to see how an oblivious read-once BP can compute $s \bmod q^m$ and thus also MUL^* by adding the terms of s for $i = 0, \ldots, n-1$, storing only the subtotal, the current digit y_i, and m x-digits. The size is bounded by $nq^{2m+O(1)} = 2^{O(\log(1/\varepsilon)+\log n+\log q)}$ as claimed.

It is more difficult to show that MUL^* approximates $\mathrm{MUL}^q_{n-1,n}$ with error ε. Let $c(x,y)$ be the carry from the computation of the output digits $0, \ldots, n-m-1$ of the product of x and y. More precisely,

$$c \;=\; c(x,y) \;=\; \left\lfloor \frac{\sum_{i=0}^{n-m-1} q^i \cdot y_i \cdot |(x_{n-m-i-1}, \ldots, x_0)|}{q^{n-m}} \right\rfloor .$$

Further, let $c' = c'(x,y) = q^{n-m} \cdot c$, $p = (|x||y|) \bmod q^n = (q^{n-m}s + c') \bmod q^n$, and $p^* = (q^{n-m} \cdot s) \bmod q^n$. Then $\mathrm{MUL}^*(x,y) \neq \mathrm{MUL}(x,y)$ if and only if $p \geq q^n/2$ and $p^* < q^n/2$ or vice versa. Since $(p - p^*) \bmod q^n = c'$, $\mathrm{MUL}(x,y) = 1$ and $\mathrm{MUL}^*(x,y) = 0$ implies $q^n/2 \leq p < q^n/2 + c'$. Analogously, $\mathrm{MUL}(x,y) = 0$ and $\mathrm{MUL}^*(x,y) = 1$ implies $0 \leq p < c'$. Hence, in both cases p is in the set $I := \{0, \ldots, c'-1\} \cup \{q^n/2, \ldots, q^n/2 + c' - 1\}$. Therefore, it suffices to show that for randomly chosen $x, y \in D^n$ the probability of $p \in I$ is bounded by ε.

We show that even if $x \in D^n$ is fixed arbitrarily and y is chosen randomly from D^n the probability of $p \in I$ is at most ε. Let $x \in D^n$ and $\gamma = \gcd(|x|, q^n)$. If $\gamma \geq q^{n-m}$, then $|x|$ is a multiple of q^{n-m}, and thus $(x_{n-m-1}, \ldots, x_0) = (0, \ldots, 0)$. It is easy to see that in this case the carry c equals 0 and thus also $c' = 0$. Hence, $I = \emptyset$ and the probability of $p \in I$ equals 0. Now assume $\gamma < q^{n-m}$ and note that in this case γ divides q^{n-m} because q is a prime power. Since c' is a multiple of q^{n-m} we have $\lceil c'/\gamma \rceil = c'/\gamma$. By Lemma 2 the random value $p = (|x||y|) \bmod q^n$ is uniformly distributed over $\{i\gamma \mid 0 \leq i < q^n/\gamma\}$. Hence, the probability that $p \in I$ is exactly

$$\frac{|I \cap \{i\gamma \mid 0 \leq i < q^n/\gamma\}|}{q^n/\gamma} \;=\; \frac{2\lceil c'/\gamma \rceil}{q^n/\gamma} \;=\; \frac{2c'}{q^n} \;=\; \frac{2c}{q^m}.$$

It is easy to see that c is bounded by qn, and therefore the probability that $p \in I$ is bounded by $2nq^{1-m} \leq \varepsilon$. □

3. PROOFS OF THE LOWER BOUNDS

Here we prove Theorems 5 and 6 from Section 2.2 on top of which our remaining lower bounds are built.

3.1 Proof Method

In this section, we describe the method used for proving lower bounds on the size of q-way BPs of bounded length for large q. This is a variant of a method due to Beame, Saks, and Thathachar [4].

First, we introduce some definitions required in the following. We consider functions defined on variables from the set V with values in $D = \{0, \ldots, q-1\}$. For any $S \subseteq V$, D^S denotes the set of mappings from S to D, which are called *assignments to S* and are usually identified with vectors from $D^{|S|}$. For $a \in D^V$, let $a_{|S}$ be the assignment obtained by projecting a to S. For $A \subseteq D^V$, let $A_{|S} = \{a_{|S} \mid a \in A\}$. For assignments a_1 and a_2 to disjoint sets of variables $S_1, S_2 \subseteq V$, let $a_1 \circ a_2 = a_1 a_2$ denote the assignment to $S_1 \cup S_2$ that agrees with a_1 on S_1 and with a_2 on S_2. Extend this to sets $A_1, \ldots, A_k$ of assignments to disjoint sets of variables $S_1, \ldots, S_k \subseteq V$ by setting $A_1 \circ \cdots \circ A_k = \{a \mid \exists\, a_1 \in A_1, \ldots, a_k \in A_k \colon a = a_1 \cdots a_k\}$ (the order of the factors does not matter). For $S \subseteq V$, an assignment $b \in D^S$, and $A \subseteq D^V$, let $A_{|b}$ be the set of assignments in D^{V-S} that are completed to assignments in A by b, i. e., $A_{|b} = \{x \in D^{V-S} \mid xb \in A\}$. For a function $f \colon D^V \to \{0,1\}$ let $f_{|b} \colon D^{V-S} \to \{0,1\}$ denote the *subfunction with respect to b* defined by $f_{|b}(x) = f(xb)$ for all $x \in D^{V-S}$.

As a preparation of the following, we give an outline of the proof method due to Beame, Jayram, and Saks [4]. Call a set of input assignments R an *(embedded) rectangle* if it can be written in the form $R = A \circ B \circ \{c\}$ for two sets A, B of assignments to disjoint sets of variables $X_1, X_2 \subseteq V$ and an assignment c to the variables in $V - X_1 - X_2$. Call c the *fixed part* of R. Given a short, small q-way BP G, the method of Beame, Jayram, and Saks guarantees the existence of a large rectangle R that only contains inputs accepted by G. In the case of approximations, one obtains a disjoint cover of a large fraction of the inputs by rectangles that contain only a small fraction of non-accepted inputs. It then remains to achieve a good upper bound on the size of the respective type of rectangles using the properties of the function under consideration.

Next, we describe our version of this method. For the whole proof method, fix a set $X \subseteq V$ of *important variables* with $|X| = n$, the set $W = V - X$, and an integer $d \geq 2$. Different from [4], we consider rectangles whose unfixed parts only consist of assignments to important variables. Furthermore, we work with d-dimensional rectangles instead of 2-dimensional ones. Later on, we will set $d = 3$ for concrete applications.

DEFINITION 4. *Let $X_1, \ldots, X_d \subseteq X$ be disjoint sets of important variables, let $X_0 = X - (X_1 \cup \cdots \cup X_d)$, and let $B \subseteq D^{X \cup W}$. Call $\{X_1, \ldots, X_d\}$ a* (d-dimensional) variable partition. *Call $R = (B, X_1, \ldots, X_d)$ a* (d-dimensional) rectangle (in $D^{X \cup W}$) *if there are sets $B_i \subseteq D^{X_i}$, $i = 1, \ldots, d$, and a $\rho \in D^{X_0 \cup W}$ such that $B = B_1 \circ \cdots \circ B_d \circ \{\rho\}$. Call $\{X_1, \ldots, X_d\}$ the* variable partition of R *and call R an s-*rectangle *if $|X_1| = \cdots = |X_d| = s$. Let $\alpha(R) = |B|/|D|^{|X_1|+\cdots+|X_d|}$ be the* density of R.

For simplicity, we identify rectangles with their associated sets of inputs B if the variable partition is clear or does not matter. Exploiting the ideas from [4], one can prove the following lemma.

LEMMA 3. *Let V be a finite set of variables, let $X \subseteq V$ be a set of important variables, and let $W = V - X$. Let d, k, r be integers such that $d \geq 2$, $2 \leq k \leq n = |X|$, and $r = 64k^2d^{k+1} \leq n$. Let $\beta = (3/4)d^{-k}$ and let $s \leq \beta n$ be a positive integer. Let $p_{\max}$ be an arbitrarily chosen positive integer. For each family $\mathcal{P}$ of at most $p_{\max}$ variable partitions $\{X_1, \ldots, X_d\}$ with $X_1, \ldots, X_d \subseteq X$ and $|X_1| = \cdots = |X_d| = s$, let a set $A(\mathcal{P}) \subseteq D^W$ be given. Let A be the union of all $A(\mathcal{P})$. Let f be a 0-1-valued function defined on $D^{X \cup W}$. Let $\eta = \min_{w \in A} |(f_{|w})^{-1}(1)|/|D|^n$.*

1. *Let G be a deterministic $|D|$-way BP for f of length $(k-1)n$. Suppose that $t = |G|^r \cdot 2^{d(k \log d + 2)\beta n + r \log d} \leq p_{\max}$. Then there is a $w \in A$ and an s-rectangle $R \subseteq (f_{|w})^{-1}(1)$ with $\alpha(R) \geq (1/t) \cdot |(f_{|w})^{-1}(1)|/|D|^n$.*

2. *Let G be a nondeterministic read-k $|D|$-way BP for f with $t = (|D||G|)^r \leq p_{\max}$. Then there is a $w \in A$ and an s-rectangle $R \subseteq (f_{|w})^{-1}(1)$ with $\alpha(R) \geq (1/t) \cdot |(f_{|w})^{-1}(1)|/|D|^n$.*

3. *Let G be a read-k $|D|$-way BP that approximates f with error ε, $0 \leq \varepsilon \leq \eta/(16|D|)$, and satisfies $t = (|D||G|)^r \leq p_{\max}$. Suppose that for all $A(\mathcal{P})$, $|A(\mathcal{P})| \geq |D|^{|W|}/(2|D|)$. Then there is a $w \in A$ and an s-rectangle R with $|R \cap (f_{|w})^{-1}(0)|/|R| \leq 8|D|\varepsilon/\eta$ and $\alpha(R) \geq (1/t) \cdot (\eta/4)$.*

Due to the space restrictions, we can only give a rough outline of the technically involved proof of this lemma for the case of length-restricted BPs (part 1). The details will be given in the forthcoming full version of the paper.

A key notion required for the proof are so called pseudo-rectangles. Let $X_0, X_1, \ldots, X_d \subseteq X$ be disjoint sets of important variables as in Definition 4 above and let $B \subseteq D^{X \cup W}$. Call $Q = (B, X_1, \ldots, X_d)$ a *d-dimensional pseudo-rectangle* if for each assignment $\rho \in D^{X_0 \cup W}$, $R = (Q_{|\rho} \circ \{\rho\}, X_1, \ldots, X_d)$ is a rectangle. It is easy to see that, equivalently, one can require that the characteristic function χ_B of the set B can be written as $\chi_B = \chi_{B,1} \wedge \cdots \wedge \chi_{B,d}$ where $\chi_{B,i}$ only depends on variables in $X_i \cup X_0 \cup W$. Call Q an *s-pseudo-rectangle* if $|X_1| = \cdots = |X_d| = s$.

For technical reasons, we consider BPs that have length $(k-1)n$ with $k \geq 2$ instead of length kn. This does not matter for the case of superlinear time bounds that we are mainly interested in. In the first stage of the proof of Lemma 3, it is shown that given a BP G of length at most $(k-1)n$ and a parameter r as in the lemma, G can be decomposed into at most $|G|^r$ well-structured sub-BPs whose sets of accepted inputs form a partition of $f^{-1}(1)$. The set of inputs accepted by each of the sub-BPs is then again partitioned into at most $2^{d(k \log d + 2)\beta n + r \log d}$ s-pseudo-rectangles, where the parameter s is chosen such that $s \leq \beta = (3/4)d^{-k}$. Altogether, we obtain a family $\mathcal{Q}$ of s-pseudo-rectangles with $|\mathcal{Q}| \leq t$ that disjointly cover $f^{-1}(1)$.

More precisely, given a BP with at most kn accesses to the important variables and a parameter $r \in \{1, \ldots, kn\}$, each of the sub-BPs can be described as a forest of decision trees that each have at most $\lceil kn/r \rceil$ important variables on its paths. The function computed by the forest is the conjunction of the functions of the respective decision trees. It can be shown that for each forest F and each input a, the set of trees in F can be partitioned into subsets $F_1, \ldots, F_d$ such that the set of important variables $X_i(a) \subseteq X$ read exclusively in trees in F_i during the computation for a is at least $\lfloor (3/4)d^{-k}n \rfloor \geq s$. Furthermore, it is easy to verify that by grouping together the inputs a with the same sets $X_1(a), \ldots, X_d(a)$, one obtains a pseudo-rectangle.

In the second stage, a particular, good pseudo-rectangle is picked from the family $\mathcal{Q}$. By averaging, it follows that for any $w \in D^W$ there is a pseudo-rectangle $Q \in \mathcal{Q}_{|w} = \{Q'_{|w} \mid Q' \in \mathcal{Q}, Q'_{|w} \neq \emptyset\}$ with respect to a variable partition $\{X_1, \ldots, X_d\}$ where $|X_1| = \cdots = |X_d| = s$ such that $|Q \cap (f_{|w})^{-1}(1)| \geq |(f_{|w})^{-1}(1)|/t$. Again by averaging, we get a $\rho \in D^{(X - (X_1 \cup \cdots \cup X_d)) \cup W}$ such that $R = Q_{|\rho} \circ \{\rho\}$ is an s-rectangle in D^X (not a pseudo-rectangle) and satisfies $\alpha(R) \geq (1/t) \cdot |(f_{|w})^{-1}(1)|/|D|^n$.

3.2 Application of the Proof Method

For the whole section, let q be a prime power and $D = \{0, \ldots, q-1\}$. Furthermore, let $\mathcal{H}$ be a linearly c-universal hash class of functions $U \to D^m$, where $U = (\mathbb{F}_q^n, +)$ or $U = (\mathbb{Z}_{q^n}, +)$. Let $\mathrm{GRAPH}_{\mathcal{H}}$ be defined on the sets of variables X, Y, and Z encoding the universe, the range, and the hash class $\mathcal{H}$, resp., where $|X| = n$, $|Y| = m$, and $|Z| = \lceil \log_q |H| \rceil$. Let $V = X \cup Y \cup Z$.

We extend the function val from Section 2.2 to assignments $a \in D^{X'}$ with $X' \subseteq X$ by setting $\mathrm{val}(a) = \mathrm{val}(a \circ z)$, where z is the assignment in $D^{X - X'}$ that sets all variables in $X - X'$ to 0.

In the following, we consider subfunctions $\mathrm{GRAPH}_{h,y} = (\mathrm{GRAPH}_{\mathcal{H}})_{|h,y}$ of $\mathrm{GRAPH}_{\mathcal{H}}$ for carefully chosen $h \in \mathcal{H}$ and arbitrary $y \in D^Y$ describing a value in the range. Our aim is to derive a good upper bound on the density of rectangles in D^X that mainly consist of inputs accepted by such a subfunction.

Call $h \in \mathcal{H}$ *good for* $X' \subseteq X$ if for all distinct $a, b \in D^{X'}$ and for all $z \in U$, $h(\mathrm{val}(a) + z) \neq h(\mathrm{val}(b) + z)$. Let $\mathcal{P}$ be a family of variable partitions $\{X_1, \ldots, X_d\}$ with $X_1, \ldots, X_d \subseteq X$. Call h *good for* $\mathcal{P}$ if h is good for all sets X_i, $i = 1, \ldots, d$, for each $\{X_1, \ldots, X_d\} \in \mathcal{P}$.

Lemma 4. *Let $\mathcal{P}$ be a family of variable partitions $\{X_1, \ldots, X_d\}$ with $X_1, \ldots, X_d \subseteq X$ and $|X_1| = \cdots = |X_d| = s$. Then $h \in \mathcal{H}$ chosen uniformly at random is good for $\mathcal{P}$ with probability at least $1 - d|\mathcal{P}| \cdot cq^{2s-m}$.*

Proof. Let $X' \subseteq X$, $|X'| = s$, and $M = \big\{(\mathrm{val}(a), \mathrm{val}(b)) \;\big|\; a, b \in D^{X'}, a \neq b\big\}$. For fixed $(x, x') \in M$, the probability that there is a $z \in U$ such that $h(x + z) = h(x' + z)$ is bounded above by cq^{-m} since $\mathcal{H}$ is linearly c-universal. Since $|M| \leq q^{2s}$, $\mathrm{Prob}_{h \in \mathcal{H}}(h \text{ is not good for } X') \leq cq^{2s-m}$. Hence, the probability that a random h is not good for any of the at most $d|\mathcal{P}|$ sets of variables occurring as parts of variable partitions in $\mathcal{P}$ is bounded above by $d|\mathcal{P}| \cdot cq^{2s-m}$. □

The following lemma yields the desired upper bound on the rectangle density.

Lemma 5. *Let $R = (B, X_1, \ldots, X_d)$ be a rectangle in D^X. Then for each $h \in \mathcal{H}$ that is good for $X_1, \ldots, X_d$ and any $y \in D^Y$, $|B \cap \mathrm{GRAPH}_{h,y}^{-1}(0)|/|B| \geq 1 - |B|^{-1/d}$. In particular, $B \subseteq \mathrm{GRAPH}_{h,y}^{-1}(1)$ implies $|B| \leq 1$.*

Proof. We claim that for all $u, v \in \mathrm{GRAPH}_{h,y}^{-1}(1)$ with $u \neq v$, there are different i, j such that $u_{|X_i} \neq v_{|X_i}$ and $u_{|X_j} \neq v_{|X_j}$. We prove this first. Suppose that, w. l. o. g., $u_{|X_1} \neq v_{|X_1}$ and $u_{|X_i} = v_{|X_i}$ for $i = 2, \ldots, d$. Let $x = \mathrm{val}(u_{|X_2 \cup \cdots \cup X_d})$, $x_u = \mathrm{val}(u_{|X_1})$, and $x_v = \mathrm{val}(v_{|X_1})$. Due to the definition of val for the relevant universes U, it follows that $x + x_u = \mathrm{val}(u)$ and $x + x_v = \mathrm{val}(v)$. Since $x_u \neq x_v$ and h is good for X_1, $h(x + x_u) \neq h(x + x_v)$, and thus $\mathrm{GRAPH}_{h,y}(u) \neq 1$ or $\mathrm{GRAPH}_{h,y}(v) \neq 1$. Hence, the claim is true.

Now let $B = B_1 \circ \cdots \circ B_d \circ \{\rho\}$, where $B_i \subseteq D^{X_i}$ for $i = 1, \ldots, d$ and $\rho \in D^{X - (X_1 \cup \cdots \cup X_d)}$. Consider the $q^{|X_1|} \times \cdots \times q^{|X_d|}$ matrix $M = (m(a_1, \ldots, a_d))_{a_1, \ldots, a_d}$ with $m(a_1, \ldots, a_d) = \mathrm{GRAPH}_{h,y}(a_1 \cdots a_d \rho)$ for $a_i \in D^{X_i}$ and $i = 1, \ldots, d$. Note that $|B \cap \mathrm{GRAPH}_{h,y}^{-1}(1)|$ is equal to the number of 1-entries in the submatrix $M(B_1, \ldots, B_d)$ of M consisting of the entries with indices in $B_1 \times \cdots \times B_d$. For $a \in D^{X_i}$, define the matrix $M_{i,a} = (m_{i,a}(b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_d))$ with

$b_j \in D^{X_j}$, $j \neq i$, by $m_{i,a}(b_1,\ldots,b_{i-1},b_{i+1},\ldots,b_d) = m(b_1,\ldots,b_{i-1},a,b_{i+1},\ldots,b_d)$. The claim from the beginning of the proof implies that for different a, a', the matrices $M_{i,a}$ and $M_{i,a'}$ do not have a 1-entry at the same position. Thus, $M_i = \sum_{a\in B_i} M_{i,a}$ is a boolean matrix and the number of 1-entries in $M(B_1,\ldots,B_d)$ is equal to the number of 1-entries in the submatrix of M_i with index set $B_1 \times \cdots \times B_{i-1} \times B_{i+1} \times \cdots \times B_d$, which is trivially upper bounded by $\prod_{j\neq i}|B_j|$. It follows that $|B \cap \mathrm{GRAPH}_{h,y}^{-1}(1)| \leq \min_{1\leq i\leq d}\prod_{j\neq i}|B_j| \leq |B|^{(d-1)/d}$. This yields the bound in the lemma. □

We are now ready to prove the main theorems from Section 2.2.

Proof of Theorem 5. We deal with the second part for general BPs first. Read-k BPs are handled similarly afterwards.

Part 2: Let X, Y, and Z be the sets from the definition of $\mathrm{GRAPH}_{\mathcal{H}}$, choose X as the set of important variables, and let $W = Y \cup Z$. Let $d = 3$. For the application of Lemma 3, choose $r = 64k^2 3^{k+1} \leq n$, $s = \lfloor (1/5)(2m - \log_q(6c)\rfloor$, and $p_{\max} = \lfloor 1/(6c)\cdot q^{-2s+m}\rfloor$. Since $m \leq n/3^k$, $s \leq (3/4)3^{-k}n$ as required for Lemma 3. Let $\mathcal{P}$ be a family of variable partitions $\{X_1, X_2, X_3\}$ with $|X_1| = |X_2| = |X_3| = s$ such that $|\mathcal{P}| \leq p_{\max}$. By Lemma 4, the probability that a randomly chosen $h \in \mathcal{H}$ is good for $\mathcal{P}$ is at least $1 - 3p_{\max}cq^{2s-m} \geq 1/2$. Hence, we can fix some $h \in \mathcal{H}$ which is good for $\mathcal{P}$. Furthermore, by the pigeonhole principle we find a $y \in D^Y$ such that $|h^{-1}(y)| \geq |U|/|D^Y| = q^{n-m}$. Let $A(\mathcal{P}) = \{yz\}$, where $z \in D^Z$ is the code of h, and let A be the union of all $A(\mathcal{P})$. Then for each $w \in A$, $|((\mathrm{GRAPH}_{\mathcal{H}})_{|w})^{-1}(1)|/q^n \geq q^{-m}$.

Our aim is to apply Lemma 3 to BPs of length $(k-1)n$ for $\mathrm{GRAPH}_{\mathcal{H}}$. The size lower bound that we obtain is still good enough to imply the claimed result for length kn. Thus, let a BP G of length $(k-1)n$ for $\mathrm{GRAPH}_{\mathcal{H}}$ be given and set $t = |G|^r \cdot 2^{3(k\log 3+2)\beta n + r\log 3}$ as in Lemma 3(1). We derive a lower bound on t as follows. We distinguish two cases. In the first case,

$$t \;\geq\; p_{\max} + 1 \;\geq\; 1/(6c)\cdot q^{-2s+m}. \tag{1}$$

In the second case, $t \leq p_{\max}$ and Lemma 3 yields a $w \in A$ and an s-rectangle $R = (B, X_1, X_2, X_3)$ such that $B \subseteq ((\mathrm{GRAPH}_{\mathcal{H}})_{|w})^{-1}(1)$ and $\alpha(R) \geq (1/t)\cdot q^{-m}$. Due to the definitions, the hash function encoded in w is good for X_1, X_2, and X_3. By Lemma 5, $|B| \leq 1$ and thus $\alpha(R) = |B|/|D|^{3s} \leq q^{-3s}$. This implies the second lower bound on t,

$$t \;\geq\; q^{3s-m}. \tag{2}$$

Since $s \leq (1/5)(2m - \log_q(6c))$, it follows that $3s - m \leq -2s + m - \log_q(6c)$. Hence, the lower bound for the first case is always at least large as that for the second case, and we have $t \geq q^{3s-m}$ in any case. Since $\beta n = (3/4)3^{-k}n \leq 3^{-k}n - 1 \leq m = \lfloor n/3^k\rfloor$ (taking into account that $64k^2 3^{k+1} \leq n$), we have

$$t \;=\; |G|^r \cdot 2^{3(k\log 3+2)\beta n + r\log 3} \;\leq\; |G|^r \cdot 2^{3(k\log 3+2)m + r\log 3}.$$

Substituting this into (2), taking logarithms, and rearranging yields

$$r\log|G| \;\geq\; -3(k\log 3+2)m - r\log 3 + 3s\log q - m\log q$$

By the definition of s, $s \geq (1/5)(2m - \log_q(6c)) - 1$. Hence,

$$\begin{aligned} r\log|G| \;&\geq\; -3(k\log 3+2)m - r\log 3 + \frac{6}{5}m\log q \\ &\quad - \frac{3}{5}\log(6c) - 3\log q - m\log q \\ &=\; m\left(\frac{1}{5}\log q - 3k\log 3 - 6\right) \\ &\quad - r\log 3 - \frac{3}{5}\log(6c) - 3\log q. \end{aligned}$$

Choose $\lambda = 1/(60\log 3)$ as the constant parameter in the theorem. We have $k \leq \lambda \log q$ and $\log q \geq 120$ due to the hypothesis. For such k and q,

$$3k\log 3 + 6 \;\leq\; \frac{1}{20}\log q + 6 \;\leq\; \frac{1}{10}\log q.$$

It follows that

$$\begin{aligned} r\log|G| \;&\geq\; \frac{1}{10}m\log q - r\log 3 - \frac{3}{5}\log(6c) - 3\log q \\ &\geq\; \frac{1}{20}m\log q - \frac{3}{5}\log c - O(r) \end{aligned}$$

for $m \geq 60$. Hence, since $r = 64k^2 3^{k+1}$,

$$\log|G| = \Omega\big(m\log q - 12\log c)/(k^2 3^k)\big)$$

as required.

Part 1: We use the same parameters for Lemma 3 as in the first part. Since $m \leq n/3^k$, again $s \leq \beta n$ with $\beta = (3/4)3^{-k}$, as required for Lemma 3. As in part 2 we obtain the lower bound $t \geq q^{3s-m}$. Then we can substitute $t = (q|G|)^r$ according to Lemma 3(2) for the read-k case. This yields

$$r\log|G| \;\geq\; (3s-m)\log q - r\log q.$$

Substituting again $s \geq (1/5)\cdot(2m - \log_q(6c)) - 1$ as above, we get

$$\begin{aligned} r\log|G| \;&\geq\; \left(\frac{1}{5}m - \frac{3}{5}\log_q(6c) - 3\right)\log q - r\log q \\ &\geq\; \frac{1}{5}m\log q - \frac{3}{5}\log(6c) - (r+3)\log q. \end{aligned}$$

Since $m \geq 644k^2 3^{k+1} \geq 640k^2 3^{k+1} + 30$ by hypothesis and $r = 64k^2 3^{k+1}$, it follows that $(1/10)m\log q \geq (r+3)\log q$. Hence,

$$\begin{aligned} \log|G| \;&\geq\; \frac{1}{10}(m\log q - 6\log(6c))/r \\ &=\; \Omega((m\log q - 6\log c)/(k^2 3^k)). \end{aligned}$$

□

Proof of Theorem 6. The proof follows the same pattern as that for Theorem 5. We now consider the function $\mathrm{GRAPH}_{\mathcal{H},y}$ where y is an arbitrarily fixed assignment to the variables in Y. We apply the proof method from Section 3 with X as the set of important variables and $W = Z$. Let $d = 3$ and $r = 64k^2 3^{k+1} \leq n$ as in the proof of Theorem 5. Let $s = \lfloor (1/5)\cdot(2m - \log_q(6c))\rfloor$ for the case of nondeterministic read-k BPs and deterministic general BPs, and let $s = \lfloor (1/5)\cdot(2m - \log_q(6c) - \log_q c')\rfloor$ with $c' = (1/4)\cdot(1 - 8\varepsilon q^{m+1})^3$ for approximating read-k BPs with error ε. Let $p_{\max} = \lfloor 1/(6c)\cdot q^{-2s+m}\rfloor$ and for a family $\mathcal{P}$ of variable partitions $\{X_1, X_2, X_3\}$ with $|X_1| = |X_2| = |X_3| = s$ such that $|\mathcal{P}| \leq p_{\max}$, let

$A(\mathcal{P}) = \{z_h \mid h \text{ is good for } \mathcal{P}\}$, where $z_h \in D^Z$ denotes the code for h. Since a random $h \in \mathcal{H}$ is good for $\mathcal{P}$ with probability at least $1/2$ by Lemma 4, $|A(\mathcal{P})| \geq |\mathcal{H}|/2$. Furthermore, at least a $1/q$-fraction of all inputs in D^Z encode functions in $\mathcal{H}$. Thus, $|A(\mathcal{P})| \geq |D|^{|Z|}/(2q)$. Since $\mathcal{H}$ is well-distributed, $|h^{-1}(y)| \geq q^{n-m}$ for all $h \in \mathcal{H}$ and the y from the hypothesis. Let A be the union of all $A(\mathcal{P})$. Then $\eta = \min_{z \in A} |((\text{GRAPH}_{\mathcal{H}})_{|y,z})^{-1}(1)|/q^n \geq q^{-m}$. We now distinguish the case of deterministic and nondeterministic BPs from the case of approximating BPs.

Deterministic BPs of length kn and nondeterministic read-k BPs: Analogously to the proof of part 1 and part 2 of Theorem 5, we get the lower bound $t \geq q^{3s-m}$, where t is defined according to Lemma 3 depending on the considered type of BPs. The lower bounds on the size of length-restricted and read-k BPs follow in the same way as above.

Approximating read-k BPs: Let G be the given read-k BP and let $t = (q|G|)^r$. Since $\log_q(1/16\varepsilon) \geq m+1$ by the hypothesis of Theorem 6, we have $\varepsilon \leq \eta/(16q) = (1/16)q^{-(m+1)}$ as required for Lemma 3(3). Analogously to the proof of Theorem 5, either $t \geq p_{\max}+1 \geq 1/(6c) \cdot q^{-2s+m}$ or $t \leq p_{\max}$ and Lemma 3(3) is applicable. The lemma yields a $z \in A \subseteq D^Z$ and an s-rectangle $R \subseteq D^X$ with $|R \cap ((\text{GRAPH}_{\mathcal{H}})_{|y,z})^{-1}(0)|/|R| \leq \varepsilon' = 8q\varepsilon/\eta \leq 8\varepsilon q^{m+1}$ and $\alpha(R) \geq (1/t) \cdot (\eta/4) \geq (1/t) \cdot (1/4) \cdot q^{-m}$. By Lemma 5, $|R \cap ((\text{GRAPH}_{\mathcal{H}})_{|y,z})^{-1}(0)|/|R| \geq 1 - |R|^{-1/3}$ which implies $|R| \leq (1-\varepsilon')^{-3}$. Using the resulting lower bound on the density $\alpha(R) = |R|/q^{3s}$ of R, we get

$$\begin{aligned} t &\geq (1/4) \cdot (1-\varepsilon')^3 \cdot q^{3s-m} \\ &\geq (1/4) \cdot \left(1 - 8\varepsilon q^{m+1}\right)^3 \cdot q^{3s-m} = c' q^{3s-m}, \end{aligned}$$

where $c' = (1/4) \cdot (1 - 8\varepsilon q^{m+1})^3$ as defined above. Since $s \leq (1/5) \cdot (2m - \log_q(6c) - \log_q(c'))$ for this part, the bound $t \geq p_{\max}+1$ is at least as large as the above bound and it suffices to consider the latter. By the hypothesis, $\log_q(1/16\varepsilon) \geq m+1$, which implies $1 - 8\varepsilon q^{m+1} \geq 1/2$ and thus $c' \geq 1/32$. The lower bound on the size of G now follows analogously to the proof of part 1 of Theorem 5 using that $c' = \Omega(1)$. □

Acknowledgment

Thanks to Ingo Wegener for proofreading of an earlier version and for helpful discussions.

4. REFERENCES

[1] F. Ablayev and M. Karpinski. A lower bound for integer multiplication on randomized ordered read-once branching programs. In *Proc. of 1st CSIT*, *Electronic Edition*, 1999.

[2] M. Ajtai. Determinism versus non-determinism for linear time RAMs with memory restrictions. In *Proc. of 31st STOC*, pages 632–641, 1999.

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

[4] P. Beame, T. S. Jayram, and M. Saks. Time-space tradeoffs for branching programs. *Journal of Computer and System Sciences*, 63(4):542–572, 2001.

[5] P. Beame, M. Saks, X. Sun, and E. Vee. Super-linear time-space tradeoff lower bounds for randomized computation. In *Proc. of 41st FOCS*, pages 169–179, 2000. To appear in *Journal of the ACM*. See also www.cs.washington.edu/homes/beame/publications.html.

[6] P. Beame and E. Vee. Time-space tradeoffs, multiparty communication complexity, and nearest neighbor problems. In *Proc. of 34th STOC*, pages 688–697, 2002.

[7] B. Bollig. Restricted nondeterministic read-once branching programs and an exponential lower bound for integer multiplication. In *Proc. of 25th MFCS*, volume 1893 of *Lecture Notes in Computer Science*, pages 222–231. Springer, 2000.

[8] B. Bollig, S. Waack, and P. Woelfel. Parity graph-driven read-once branching programs and an exponential lower bound for integer multiplication. In *Proc. of 2nd TCS*, pages 83–94, 2002.

[9] B. Bollig and P. Woelfel. A read-once branching program lower bound of $\Omega(2^{n/4})$ for integer multiplication using universal hashing. In *Proc. of 33rd STOC*, pages 419–424, 2001.

[10] B. Bollig and P. Woelfel. A lower bound technique for nondeterministic graph-driven read-once-branching programs and its applications. In *Proc. of 27th MFCS*, pages 131–142, 2002.

[11] A. Borodin and S. Cook. A time-space tradeoff for sorting on a general sequential model of computation. *SIAM J. Comp.*, 11(2):287–297, 1982.

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

[13] R. E. Bryant. On the complexity of VLSI implementations and graph representations of boolean functions with applications to integer multiplication. *IEEE Transactions on Computers*, 40(2):205–213, 1991.

[14] J. L. Carter and M. N. Wegman. Universal classes of hash functions. *Journal of Computer and System Sciences*, 18(2):143–154, 1979.

[15] M. Dietzfelbinger. Universal hashing and k-wise independent random variables via integer arithmetic without primes. In *Proc. of 13th STACS*, pages 569–580, 1996.

[16] M. Dietzfelbinger, T. Hagerup, J. Katajainen, and M. Penttonen. A reliable randomized algorithm for the closest-pair problem. *Journal of Algorithms*, 25:19–51, 1997.

[17] J. Gergov. Time-space tradeoffs for integer multiplication on various types of input oblivious sequential machines. *Information Processing Letters*, 51:265–269, 1994.

[18] S. Jukna. The graph of integer multiplication is hard for read-k-times networks. Technical Report 95-10, Universität Trier, 1995. Available under ftp://ftp.informatik.uni-trier.de/pub/Users-Root/reports/95-10.ps.

[19] Y. Mansour, N. Nisan, and P. Tiwari. The computational complexity of universal hashing. *Theoretical Computer Science*, 107:121–133, 1993.

[20] S. Ponzio. A lower bound for integer multiplication with read-once branching programs. *SIAM Journal on Computing*, 28:798–815, 1998.

[21] I. Wegener. *Branching Programs and Binary Decision Diagrams—Theory and Applications*. Monographs on Discrete and Applied Mathematics. SIAM, Philadelphia, PA, 2000.

[22] M. N. Wegman and J. L. Carter. New classes and applications of hash functions. In *Proc. of 20th FOCS*, pages 175–182, 1979.

[23] P. Woelfel. Efficient strongly universal and optimally universal hashing. In *Proc. of 24th MFCS*, pages 262–272, 1999.

[24] P. Woelfel. New bounds on the OBDD-size of integer multiplication via universal hashing. In *Proc. of 18th STACS*, pages 563–574, 2001.

[25] P. Woelfel. On the complexity of integer multiplication in branching programs with multiple tests and in read-once branching programs with limited nondeterminism. In *Proc. of 17th Comp. Compl.*, pages 80–89, 2002.