Submitted to: *Inf. Processing Letters* (November 11, 1998)

Linear codes are hard for oblivious read-once parity branching programs

S. JUKNA †

Department of Computer Science, University of Trier, D-54286 Trier, Germany
Institute of Mathematics, 2600 Vilnius, Lithuania

We show that the characteristic functions of linear codes are exponentially hard for the model of oblivious read-once branching programs with parity accepting mode, known also as Parity OBDDs. The proof is extremely simple, and employs a particular combinatorial property of linear codes – their universality.

Keywords: Computational complexity, parity branching programs, lower bounds, linear codes, computer aided design

Introduction Interesting aspect of linear codes is that their characteristic functions appear to be hard for all known "reading-restricted" models of branching programs: syntactic read-k times branching programs, where along every path (be it consistent or not) every variable appears at most k times, and $(1,+s)$-branching programs, where along every consistent path at most s variables are tested more than once. Namely, it is known that for some explicit linear codes $C \subseteq \{0,1\}^n$, their characteristic functions f_C require (deterministic [6] and non-deterministic [2]) syntactic read-k times branching programs and deterministic $(1,+s)$-b.p. [4] of super-polynomial size, as long as $k = o(\log n)$ or $s = o(n/\log n)$.

On the other hand, looking at parity-check matrix of C we see that each such function f_C is just an And of $m \leqslant n$ (negations of) parity functions $\bigoplus_{j\in S_i} x_j$, for particular subsets $S_1,\ldots,S_m \subseteq \{1,\ldots,n\}$. Because of their intimate relation to Parity, it is natural to ask if linear codes can be computed more efficiently using the parity accepting mode? So far, this is open even for read-once parity branching programs (1-p.b.p.).

In this note we answer this question negatively under additional restriction that branching programs are *oblivious*. The main contribution, however, is the extreme simplicity of the proof.

An *oblivious read-once parity branching program* † (or an *oblivious 1-p.b.p.* for short) is a rooted ordered graph whose nodes are partitioned into at most n levels. Edges must go only from one level to the next, but neither in-degree nor out-degree of nodes is restricted.

† Supported by DFG grant Me 1077/10–1.

† This model is known in CAD community as a "Parity-OBDD". Practical interest in this model is stipulated by the fact that, as a data structure, Parity-OBDDs have similar properties as well-known model of ordered binary decision diagrams (OBDD), and hence, can be also used for a practical verification of chips (see. e.g. [1, 7]).

All the nodes of one level (except for the last level, whose nodes are *sinks*) are labelled by one and the same variable, and different levels have different variables. If an edge leaves a node labelled by a variable x_i then the edge itself is labelled either by x_i or by its negation $\bar{x}_i$. Such a graph G computes a Boolean function in the following sense: given an input $a \in \{0,1\}^n$, $G(a) = 1$ iff the number of paths from the root to a sink, which are consistent with a, is odd.

General lower bound We will employ one specific property of linear codes which was already used in [4] to show that linear codes are hard for $(1, +s)$-b.p.'s.

The *minimal distance* of a code C is a minimal Hamming distance between any pair of distinct vectors in C. It is well known (and easy to show) that minimal distance of C coincides with the minimum weight of (i.e. the number of 1's in) a non-zero vector form C. The *dual* of C is the set

$$C^{\perp} = \{x : \langle x, y\rangle = 0 \text{ for all } y \in C\}$$

of all those vectors $x \in \{0,1\}^n$, which are orthogonal to all the vectors from C; here $\langle x, y\rangle = x_1 y_1 \oplus \ldots \oplus x_n y_n$ is the standard scalar product over GF(2). A set of vectors $C \subseteq \{0,1\}^n$ is *k-universal* if for any subset of k coordinates $I \subseteq \{1, \ldots, n\}$ the projection of vectors from C onto this set I gives the whole cube $\{0,1\}^k$. A nice property of linear codes is that their duals are universal.

$(*)$*If C is a linear code of minimal distance $k+1$ then its dual $C^{\perp}$ is k-universal.*

Indeed, take a set $I \subseteq \{1, \ldots, n\}$ with $|I| \leqslant k$. The set of all projections of strings in C onto I is a linear subspace in $\{0,1\}^I$, and this subspace is proper if and only if all strings $a \in C$ satisfy a non-trivial linear relation $\sum_i \xi_i a_i = 0 \bmod 2$ whose support $\{i : \xi_i = 1\}$ is contained in I. But, by definition, $C^{\perp}$ consists exactly of all relations ξ satisfied by C, and its minimal distance is exactly the minimal possible cardinality of a set I for which the projection of C onto $\{0,1\}^I$ is proper.

Theorem: *Let $C \subseteq \{0,1\}^n$ be a linear code of distance d, and let its dual $C^{\perp}$ has distance $k+1$. If $d \geqslant k+1$ then any oblivious 1-p.b.p. computing the characteristic function f_C of C has size at least 2^k.*

Proof. Let P be an oblivious 1-p.b.p. computing f, and let $I \subseteq \{1, \ldots, n\}$ be the set of bits tested on the first $k = |I|$ levels of P. Every assignment $a : I \to \{0,1\}$ (treated for this purpose as a restriction) defines the subfunction f_a of f in $n - |I|$ variables which is obtained from f by setting x_i to $a(i)$ for all $i \in I$. Let $\mathcal{F}$ be a subspace of the 2^{n-k}-dimensional space of all Boolean functions on $n-k$ variables, generated by the subfunctions f_a of f with $a : I \to \{0,1\}$. It is easy to see that $size(P) \geqslant \dim(\mathcal{F})$. Indeed, if $v_1, \ldots, v_r$ are the nodes at the k-th level of P, then for every assignment $a : I \to \{0,1\}$, the subfunction f_a is a linear combination of the functions computed by an oblivious 1-p.b.p. 's with source-nodes $v_1, \ldots, v_r$. Hence, we need at least $r \geqslant \dim(\mathcal{F})$ such functions to get all the subfunctions in $\mathcal{F}$.

Now we can finish the proof as follows. Since the dual of C has distance $k+1$, we have by $(*)$, that the code C itself is k-universal. This, in particular, means that for every assignment $a : I \to \{0,1\}$ there is an assignment $x_a : \bar{I} \to \{0,1\}$ such that $(a, x_a) \in C$. Moreover, since C has distance $d > k = |I|$, we have that $(b, x_a) \notin C$ for every other assignment

$b : I \to \{0,1\}$, $b \neq a$. Thus, if we look the subfunctions f_a, $a : I \to \{0,1\}$, as rows of a $2^k \times 2^{n-k}$ matrix, then this matrix contains a diagonal $2^k \times 2^k$ submatrix with entries $f(a,x)$ such that $f(a,x) = 1$ iff $x = x_a$. So, the matrix has full row-rank equal 2^k, which means that the subfunctions in $\mathcal{F}$ are linearly independent (over any field, including GF(2)). Thus, $size(P) \geqslant \dim(\mathcal{F}) = |\mathcal{F}| \geqslant 2^k$, as desired. □

Explicit lower bound Recall that the r-th order binary Reed-Muller code $R(r,\ell)$ of length $n = 2^\ell$ is the set of graphs of all polynomials in ℓ variables over GF(2) of degree at most r. This code is linear and has minimal distance $2^{\ell-r}$.

Corollary: *Let $n = 2^\ell$ and $r = \lfloor(\ell-3)/2\rfloor$. Then every oblivious 1-p.b.p. computing the characteristic function of the Reed-Muller code $R(r,\ell)$ has size at least $2^{\Omega(\sqrt{n})}$.*

Proof. It is known (see, e.g. [5, p. 375]) that the dual of $R(r,\ell)$ is $R(\ell-r-1,\ell)$. Hence in the notation of Theorem, we have that $2^{\ell-r} = d \geqslant k+1 = 2^{r+1}+1 = \Omega(\sqrt{n})$. The desired bound follows. □

Acknowledgement: I am thankful to Stephan Waack for asking me the question answered in this note.

References

[1] J. Gergov and Ch. Meinel, Mod–2–OBDDs: a data structure that generalizes EXOR–sum–of–products and ordered binary decision diagrams, *Formal Methods in System Design*, **8** (1996) 273–282.

[2] S. Jukna, A note on read-k-times branching programs, *RAIRO Theoretical Informatics and Applications*, vol. 29, Nr. 1 (1995) 75-83.

[3] S. Jukna, *The graph of integer multiplication is hard for read-k times networks*, Forschungsbericht Nr. 10, Dept. of Computer Sci., University of Trier, 1995.

[4] S. Jukna and A. Razborov, Neither reading few bits twice nor reading illegally helps much, *Discrete Applied Mathematics* **85** (1998) 223–238.

[5] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes*. Elsevier, North-Holl., 1977.

[6] E. A. Okolnishnikova, Lower bounds for branching programs computing characteristic functions of binary codes. In: *Metody Discretnogo Analiza*, **51** (1991), 61–83 (in Russian).

[7] S. Waack, On the descriptive and algorithmic power of parity binary decision diagrams. In: *Proc. of 14th Annual Symp. on Theoret. Aspects of Computer Science* (STACS'97), Lect. Notes in Comput. Sci. **1200** (1997) 201–212.