A Non-Linear Lower Bound for Constant Depth Arithmetical Circuits via the Discrete Uncertainty Principle

Maurice J. Jansen Kenneth W. Regan*

February 27, 2008

Abstract

We prove a super-linear lower bound on the size of a bounded depth bilinear arithmetical circuit computing cyclic convolution. Our proof uses the strengthening of the Donoho-Stark uncertainty principle [DS89] given by Tao [Tao05], and a combinatorial lemma by Raz and Shpilka [RS03]. This combination and an observation on ranks of circulant matrices, which we use to give a much shorter proof of the Donoho-Stark principle, may have other applications.

Keywords. Computational complexity, arithmetical circuits, lower bounds, constant depth bilinear circuits.

1 Introduction

One of the central mysteries in arithmetic circuit complexity is the computational power conferred by the ability to perform arithmetic operations with arbitrary field elements at unit cost. Over the real numbers, for example, this assigns unit cost to manipulations with numbers of infinite precision and/or unbounded magnitude. Morgenstern [Mor73] argued that most algorithms used in practice use only constants of "reasonably" bounded magnitude. Possible exceptions are algorithms with constants obtained via de-randomization procedures or polynomial interpolation.

Restricting scalars in circuits to have bounded magnitude *does* make it easier to prove lower bounds. Examples are the volumetric lower bounds of [Mor73] for *bounded coefficient linear* circuits, and the $\Omega(N \log N)$ size lower bound of Raz [Raz03] in the *bounded coefficient bilinear* model for the mapping defined by multiplication of two $n \times n$ matrices, where $N = n^2$ [Raz03]. Bürgisser and Lotz [BL04], building on the work of Raz, proved a tight $\Omega(n \log n)$ size lower bound for the convolution of two n-vectors of variables.

For linear and bilinear circuits with unrestricted constants, however, no super-linear size lower bounds have been obtained despite four decades of attention. The question is whether this owes only to a current lack of lower bound techniques, or whether there is a real loss in computational power when restricting scalar magnitudes. The known results are mainly size-depth tradeoffs. For linear circuits of fixed depth d, Pudlák [Pud94] obtained size lower bounds of order $\Omega(n\lambda_d(n))$, where the functions $\lambda_d(n)$ for $d = 1, 2, \ldots$ are unbounded but extremely slow growing. These were partly based on lower bounds for depth-d superconcentrators. Shoup and Smolensky [SS91] gave lower bounds of order $\Omega(dn^{1+1/d})$ for the task of evaluating a univariate polynomial at some fixed set of complex numbers $p_1, p_2, \ldots, p_n$. This corresponds to computation of the linear map defined by the Vandermonde matrix with ijth entry p_i^j. Here, either $p_1, p_2, \ldots, p_n$ are required to be algebraically

*Part of this work by both authors was supported by NSF Grant CCR-9821040. Corresponding author: Dept. of CSE at UB, 201 Bell Hall, Buffalo, NY 14260-2000; (716) 645-3180 x114, fax 645-3464; regan@cse.buffalo.edu

independent over the field of rationals, or they have to grow very rapidly. This result can also be interpreted as giving a lower bound for a set of degree n polynomials, by considering $p_1, p_2, \ldots, p_n$ to be part of the input. Related to this, in a very recent paper, Raz has proved an $\Omega(n^{1+1/(2d)})$ lower bound on the size of depth-d circuits computing some explicitly defined polynomials of degree $5d+2$ [Raz08].

For bounded depth bilinear circuits, Raz and Shpilka proved that any depth d circuit for multiplying two $m \times m$ matrices is of size $\Omega(\frac{1}{d^2} m^2 \lambda_d(m^2))$ [RS03]. In this paper, building on the work of [RS03], we prove a size-depth tradeoff for the circular convolution mapping that was considered in [BL04]. We employ Tao's strengthening for prime n [Tao05] of the discrete form of the Heisenberg uncertainty principle obtained by Donoho and Stark [DS89]. The next section gives background and circuit definitions, a new and notably shorter proof of Donoho and Stark's result, a sketch of Tao's proof, and combinatorial information used by above-cited papers.

2 Preliminaries

We define the discrete Fourier transform matrix DFT_n by $(DFT_n)_{st} = \omega^{st}$, for $s, t \in \{0, 1, \ldots, n-1\}$, and where $\omega = e^{2\pi i/n}$. Let $F_n = n^{-1/2} DFT_n$. The conjugate transpose of a matrix A will be denoted by A^*. The *cyclic convolution* $x \circ y$ of two n-vectors $x = (x_0, x_1, \ldots, x_{n-1})^T$ and $y = (y_0, y_1, \ldots, y_{n-1})^T$ is the n-vector $z = (z_0, \ldots, z_{n-1})^T$ with components

$$z_k = \sum_{i+j \equiv k \bmod n} x_i y_j,$$

for $0 \leq k < n$. In other words, thinking of x and y as representing univariate polynomials $f = \sum_{i=0}^{n-1} x_i t^i$ and $g = \sum_{i=0}^{n-1} y_i t^i$, $z = x \circ y$ represents the polynomial $f \cdot g$ computed modulo $t^n - 1$. For example with $n = 5$:

$$x \circ y = \begin{pmatrix} x_0y_0 + x_4y_1 + x_3y_2 + x_2y_3 + x_1y_4 \\ x_1y_0 + x_0y_1 + x_4y_2 + x_3y_3 + x_2y_4 \\ x_2y_0 + x_1y_1 + x_0y_2 + x_4y_3 + x_3y_4 \\ x_3y_0 + x_2y_1 + x_1y_2 + x_0y_3 + x_4y_4 \\ x_4y_0 + x_3y_1 + x_2y_2 + x_1y_3 + x_0y_4 \end{pmatrix}.$$

For vector $x = (x_0, \ldots, x_{n-1})^T$, the *circulant matrix* $Circ(x)$ is defined by

$$Circ(x) = \begin{pmatrix} x_0 & x_{n-1} & \cdots & x_2 & x_1 \\ x_1 & x_0 & \cdots & x_3 & x_2 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{n-2} & x_{n-3} & \cdots & x_0 & x_{n-1} \\ x_{n-1} & x_{n-2} & \cdots & x_1 & x_0 \end{pmatrix}.$$

We have that $x \circ y = Circ(x)y = Circ(y)x$. We write $diag(x)$ for the $n \times n$ matrix with x on the main diagonal and 0s elsewhere. Convolution can be computed using the Fourier transform, according to the following folklore result:

Theorem 2.1 (The Convolution Theorem) *For any n-vector $x = (x_0, x_1, \ldots, x_{n-1})^T$,*

$$Circ(x) = F_n^* diag(DFT_n x) F_n.$$

2.1 Discrete Uncertainty Principles

The following alternative proof exploits Theorem 2.1 and the relation it gives between rank and the *support* of an n-vector f, which is defined by $\mathrm{supp}(f) = \{i : f_i \neq 0\}$. The size of $\mathrm{supp}(f)$ is a crude measure of the amount of localization of the vector f. Analogous to the Heisenberg uncertainty principle, the following says that a nonzero vector f and its Fourier transform $\hat{f} =_{\mathrm{def}} F_n f$ cannot both be arbitrarily narrowly localized.

Theorem 2.2 ([DS89]) *For any n-vector $f \neq 0$, $|\mathrm{supp}(f)| \cdot |\mathrm{supp}(\hat{f})| \geq n$.*

Proof. Since by Theorem 2.1,

$$Circ(f) = \sqrt{n} F_n^* \mathrm{diag}(\hat{f}) F_n,$$

we have that $|\mathrm{supp}(\hat{f})| = rank(Circ(f))$. Now partition f into "blocks" consisting of a nonzero entry and the maximal string of zero entries following it, wrapping from the end of the vector to the beginning if needed. Take R to be the maximum length of a block. Then $R \geq n/|\mathrm{supp}(f)|$. Now consider the R rows of $Circ(f)$ corresponding to a size-R block—without loss of generality we may cycle these around to the first R positions. These contain an $R \times R$ upper-triangular matrix with nonzero main diagonal, and so are independent. Hence $rank(Circ(f)) \geq R \geq \frac{n}{|\mathrm{supp}(f)|}$. □

In case n is prime, Tao showed that Theorem (2.2) can be significantly improved [Tao05]. The point is that for prime p the matrix DFT_p is *totally regular*, i.e. every square submatrix is non-singular, a fact attributed to Chebotarëv in [SJ96]. Given this fact, for which [Tao05] gives an elementary proof, Tao's improvement follows readily:

Theorem 2.3 ([Tao05]) *For prime p, for any nonzero p-vector f and its Fourier transform $\hat{f} = F_p f$ we have that $|\mathrm{supp}(f)| + |\mathrm{supp}(\hat{f})| \geq p + 1$.*

Proof. Let $k = p - |\mathrm{supp}(\hat{f})|$. There are k zeroes in $\hat{f}$. Let $I \subseteq \{0, 1, \ldots, p-1\}$ be the indices of these zeroes. Suppose $|\mathrm{supp}(f)| \leq k$. Let $J \subseteq \{0, 1, \ldots, p-1\}$ be a set of size k that contains all indices of non-zero entries of f. In the following $DFT^p_{I,J}$ denotes the minor of DFT_p with rows I and columns J. We have that $(DFT^p_{I,J}) f_J = (DFT_p f)_I = 0$, but $f_J \neq 0$ since $f \neq 0$. This is a contradiction since $DFT^p_{I,J}$ is non-singular. Hence $|\mathrm{supp}(f)| > k = p - |\mathrm{supp}(\hat{f})|$. □

2.2 Combinatorial Lemma

For a function $f : \mathbf{N} \to \mathbf{N}$, define $f^{(i)}$ to be the composition of f with itself i times—i.e., $f^{(0)}$ is the identity function, and for $i > 0$, $f^{(i)} = f \circ f^{(i-1)}$. Then provided $f(n) < n$ for all $n > 0$, define

$$f^*(n) = \min\{i : f^{(i)} \leq 1\}.$$

The labeling of the following set of extremely slow-growing functions $\lambda_d(n)$ follows [RS03]; each is a monotone non-decreasing function tending to infinity.

Definition 2.1 ([RS03]). Let

1. $\lambda_1(n) = \lfloor \sqrt{n} \rfloor$,
2. $\lambda_2(n) = \lceil \log n \rceil$,
3. $\lambda_d(n) = \lambda^*_{d-2}(n)$, for $d > 2$.

For a directed acyclic graph G, V_G denotes the set of all nodes, I_G those with in-degree 0, and O_G those with out-degree 0. The depth of G is the length in edges of the longest path from I_G to O_G. For subsets $A \subseteq I_G$, $B \subseteq O_G$ and $V \subset V_G$, let $\mathrm{P}[A, B, V]$ be the number of distinct paths from vertices in A to vertices in B that do not go over vertices in V.

Lemma 2.4 ([RS03]) *Let $0 < \beta < 1$, $0 < \epsilon < 1/400$ and $d \geq 2$. For any large enough n, if G is a leveled directed acyclic graph of depth d, with more than n vertices and less than $\epsilon n \lambda_d(n)$ edges, then there exists a set of vertices V and a set J of inputs and outputs such that:*

1. $\sqrt{n} \leq |V| \leq \beta n$,
2. $|J| \leq 5\epsilon dn$, *and*
3. $P_G[I_G \backslash J, O_G \backslash J, V] \leq \epsilon \frac{n^2}{|V|}$.

2.3 Bilinear Circuits

Let $\mathbf{C}$ denote the field of complex numbers. An arithmetical circuit over inputs $X = \{x_1, x_2, \ldots, x_n\}$ and $\mathbf{C}$ is given by a directed acyclic graph $G = (V, E)$. Vertices of in-degree zero are called *inputs*, and are labeled with variables from X or field constants from $\mathbf{C}$. Vertices with out-degree zero are called *outputs*. Any vertex of in-degree at least one is labeled with an element $\in \{+, \times\}$. These are called *gates*. Edges are labeled with field constants. A label $\alpha \in \mathbf{C}$ on an edge is intended to mean multiplication with α. Associated then, with each input or gate g is the *polynomial computed by* g, defined in the obvious way. *Linear* circuits are those without $\times$ gates.

Since we are working over a field of characteristic zero, for the computation of bilinear forms, we can assume our circuits to be *bilinear*, at the cost of a constant factor increase in size and depth (See Proposition 4.2 in [RS03]). A bilinear circuit over sets of variables $X = \{x_1, x_2, \ldots, x_n\}$ and $Y = \{y_1, y_2, \ldots, y_n\}$ has the following structure. First, there is a set S_1 of addition gates computing homogeneous linear forms in X. Second, there is a set S_2 disjoint from S_1 computing homogeneous linear forms in Y. Third, there is a set S_3 of multiplication gates of degree two, that take one input from S_1 and one from S_2. Finally, there is a set S_4 of addition gates that compute linear combinations of the bilinear forms computed by the multiplication gates in S_3. The outputs of the circuit form a subset of S_4. As in [RS03], we only count the number of edges present in the circuit above the multiplication gates.

Definition 2.2 ([RS03]). For a bounded depth bilinear circuit C, define its size $s(C)$ to be the number of edges in the circuit between the multiplication gates and the outputs, and define its depth $d(C)$ to be the length of a longest path in edges from a multiplication gate to an output.

A circuit of depth d is *leveled*, if we can partition the vertices into sets $L_0, L_1, \ldots, L_d$, such that edge only go between consecutive levels L_i and L_{i+1}. A circuit of depth d can be leveled at the cost of increasing the size by factor of d.

Note that Cooley and Tukey [CT65] gave $O(n \log n)$ size, $O(\log n)$ depth linear circuits that compute DFT_n. So using Theorem 2.1, we obtain $O(n \log n)$ size bilinear circuits for computing circular convolution. These circuits have complex coefficients on the wires of norm 1. Bürgisser and Lotz proved that this is optimal for circuits that have their constants restricted to be of norm $O(1)$ [BL04].

3 Lower Bounds for Cyclic Convolution

For depth one we have the following result, which is tight due to Theorem 2.1.

Proposition 3.1 *Any leveled bilinear circuit C of depth 1 computing the circular convolution $x^T Circ(y)$ has size $s(C) \geq n^2$.*

Proof. A circuit of depth 1 has a very simple structure. There are some number r of multiplication gates M_r computing products $M_r = L_r(x)R_r(y)$, where $L_r(x)$ and $R_r(y)$ are homogeneous linear forms. Then there is one layer of output gates, each gate computing summation over some set of input multiplication gates.

We will argue that each output gate must be connected to at least n multiplication gates. For purpose of contradiction suppose that this is not the case. Say some output gate O_i takes input from $< n$ multiplication gates. Without loss of generality we may assume gate O_i computes $(Circ(y)x)_i$. Consider the subspace of dimension at least 1 defined by equations $L_j(x) = 0$, for each multiplication gate j attached to output O_i. We can select a non-zero vector a from this space such that for any assignment $y = b$,

$$(Circ(b)a)_i = 0.$$

This yields a contradiction, for example we can take b to be equal to a^* shifted by i, then $Circ(b)a)_i = ||a||_2^2$, which is non-zero, since a is a non-zero vector. □

Theorem 3.2 *There exists $\delta > 0$, such that for any d, for any large enough prime number p, any leveled bilinear circuit with inputs $x = (x_0, x_1, \ldots, x_{p-1})^T$ and $y = (y_0, y_1, \ldots, y_{p-1})^T$ of depth d computing cyclic convolution $Circ(y)x$ has size $s(C) \geq \delta \frac{1}{d} p \lambda_d(p)$.*

Proof. The result holds for $d = 1$ by Proposition 3.1. Assume $d \geq 2$. Write using Theorem 2.1,

$$Circ(y)x = F_p^* \mathrm{diag}(DFT_p(y)) F_p x.$$

We first apply substitutions $x := F_p^* x'$ and $y = \frac{1}{p} DFT_p^* y'$ at the inputs. This does not alter the circuit above the multiplication gates, but now we have a circuit computing

$$F_p^* \mathrm{diag}(y')x'.$$

For simplicity, let us rename x' by x and y' by y again. Let G be the leveled directed acyclic graph of depth d given by the part of circuit above the multiplication gates. The set I_G is the collection of multiplication gates $M_i = L_i(x)R_i(y)$, where $L_i(x)$ and $R_i(y)$ are homogeneous linear forms. Take $O_G = \{1, 2, \ldots, p\}$ to be the set of outputs of the circuit. Let $\delta > 0$ and $\beta > 0$ be small enough constants to be determined later. Let $\epsilon = \frac{\delta}{400d}$. Trivially G has at least p vertices. Suppose that G has strictly fewer than $\epsilon p \lambda_d(p)$ edges. Lemma 2.4 applies, and we obtain sets $I \subset I_G$, $O \subset O_G$ and $V \subset V_G$ such that

1. $|I|, |O| \leq 5\epsilon dp = \frac{5\delta}{400}p$,
2. $|V| = k$, with $\sqrt{p} \leq k \leq \beta p$, and
3. $P_G[I_G \backslash I, O_G \backslash O, V] \leq \epsilon \frac{p^2}{k}$.

For each output node $i \in O_G \backslash O$, define $P(i)$ to be the number of multiplication gates in $I_G \backslash I$ for which there exists a directed path that bypasses V and reaches node i. Let R be a set of $r = 10k$

output gates with lowest $P(i)$ values. This restricts $10\beta \le 1 - \frac{5\delta}{400}$. By averaging we get that

$$\sum_{i\in R} P(i) \le \frac{r}{|O_G \backslash O|} \sum_{i \in O_G \backslash O} P(i) \le \frac{r}{p-5\epsilon dp} \cdot \frac{\epsilon p^2}{k} = \frac{10\epsilon p}{1-5\epsilon d}.$$

Let I' be the set of all multiplication gates in $I_G \backslash I$ for which there exist directed paths to nodes in R that bypass V. We can conclude that

$$|I'| \le \frac{10\epsilon p}{1-5\epsilon d} = p\frac{10\delta}{400d-5\delta}.$$

Define a linear subspace W by the set of equations

$$R_i(y) = 0 \text{ for all } i \in I \cup I'.$$

For any fixed substitution for $y \in W$, the resulting circuit has all of the gates computing linear function in the x variables. Relative to a fixed choice for y, define linear subspace W_y by equations $g_v(x) = 0$ for all $v \in V$, where $g_v(x)$ denotes the linear form computed at gate v. Note that

$$\dim(W) \ge p\left(1 - \frac{5\delta}{400} - \frac{10\delta}{400d - 5\delta}\right), \tag{1}$$

and, for each y,

$$\dim(W_y) \ge p - k \ge p(1-\beta).$$

For small enough δ and β, both $\dim(W) > 0$ and $\dim(W_y) > 0$. Now we have arranged that for each $y \in W$, and each $x \in W_y$,

$$(F_p^* diag(y) x)_i = 0, \tag{2}$$

for each $i \in R$. In order to reach a contradiction, we will now argue that it is possible to select $y \in W$ and $x \in W_y$ such that some output in R is non-zero.

First of all, fix a vector $y \in W$ that has at most $p(\frac{5\delta}{400} + \frac{10\delta}{400d-5\delta})$ zeroes. This can be done because of Equation (1). Let A be the set of indices i for which $y_i = 0$. Let $m = |A|$. Let W'_y be a subspace of W_y of dimension 1 obtained by adding equations to a defining set S of equations of W_y in two steps as follows:

1. Add $x_i = 0$ to S, for each $i \in A$.
2. One-by-one, for each $i \notin A$, add the equation $x_i = 0$ to S, as long as the dimension of the solution space of (S) is bigger than one.

Observe that, since the starting space W_y has dimension at least $p - k \ge p(1-\beta)$, at the end of the first stage, the dimension will be cut down to at most $p - k - m$, provided $m \le p(1-\beta)$. The latter holds provided $1 - \beta \ge \frac{5\delta}{400} + \frac{10\delta}{400d-5\delta}$. This can easily be arranged for absolute constants δ and β close enough to zero. Hence we will be able to add the equation $x_i = 0$ in the second stage for at least $p - k - m - 1$ many i with $i \notin A$, and still have the final solution space W'_y to be of dimension at least one.

Select an arbitrary non-zero vector x from W'_y. Observe that of the $p - m$ indices i not in A, x_i is non-zero for at most $k+1$ entries, and that x_i is zero for all $i \in A$. So x_i is zero for each i for which $y_i = 0$. Since x itself is a nonzero vector there must be some place i where x_i and y_i are both nonzero. Let $f = diag(y)x$ and $\hat{f} = F_P^* f$. We thus conclude that f is a non-zero vector, but that

$|\text{supp}(f)| \leq k+1$. By the discrete uncertainty principle for cyclic groups of prime order [Tao05], stated in Theorem 2.3, we have that

$$\text{supp}(f) + \text{supp}(\hat{f}) \geq p + 1.$$

Hence the output vector of the circuit $\hat{f}$ is non-zero in at least $p+1-(k+1) = p-k$ places. Since R is of size $10k$, by the pigeonhole principle, there must be some output in R that is non-zero. This is in contradiction with Equation (2). $\square$

Theorem 3.2 extends to non-prime lengths, as pointed out by an anonymous referee of our original draft.

Corollary 3.3 *There exists $\delta > 0$, so that for any d, for any large enough n, any leveled bilinear arithmetical circuit over variables $\{x_0, x_1, \ldots, x_{n-1}\}$ and $\{y_0, y_1, \ldots, y_{n-1}\}$ of depth d computing $Circ(y)x$ requires size at least $\delta \frac{1}{d} n \lambda_d(n)$.*

Proof. By Chebyshev's proof of Bertrand's Postulate, for all $n \geq 6$ there exists a prime p with $\lfloor n/4 \rfloor < p < \lfloor n/2 \rfloor$. Given p-vectors x and y, extend them to n-vectors x' and y' by setting $x'_i = 0$ and $y'_i = y_{i \bmod p}$ for $p \leq i < n$. Then $x \circ y$ is given by the first p places of $x' \circ y'$, and since this reduction does not change the depth of the underlying circuits, the statement follows from Theorem 3.2. $\square$

Applying the observation ascribed to Pitassi and Wigderson in [RS03], also noted to us by the referees, these tradeoffs extend to families of polynomials that compute a single scalar output, over fields of characteristic zero. This follows because the construction in the Baur-Strassen Derivative Lemma [BS82] can be performed while maintaining constant bounded depth. For example, it can be concluded the polynomial $f = z^T Circ(y)x$ does not have linear size bounded depth circuits over the complex numbers. It is also worth remarking that a similar combination of Theorem 2.3 and Lemma 2.4 yields lower bounds for linear circuits, in the case of DFT:

Theorem 3.4 (case of [Pud94]) *There exists $\delta > 0$, such that for any $d \geq 1$, for any large enough prime number p, any leveled linear circuit of depth d with inputs $x = (x_0, x_1, \ldots, x_{p-1})^T$ computing the linear transformation $\lambda x.DFT_p x$ has size $s(C) \geq \delta \frac{1}{d} p \lambda_d(p)$.*

Theorem 3.4 likewise extends to arbitrary n, with the same application of Bertand's Postulate, albeit weakening the constants involved. This follows via Rader's FFT algorithm [Rad68] and some padding, reducing DFT_p to two applications of DFT_n at the cost of doubling the depth. Of course this result is already known via the lower bounds for superconcentrators given in [Pud94] (and also [DDPW83] for even d), and the well-known correspondence between superconcentrators and linear circuits computing the map of a totally regular matrix.

4 Conclusion

We have demonstrated that the discrete uncertainty principle, in its strongest form at least, can be used as a convenient tool to prove circuit lower bounds for bounded depth linear and bilinear arithmetical circuits. In this area the central open problem still is to obtain any kind of non-linear lower bound for unrestricted linear circuits. This problem has remained elusive for over 35 years.

Acknowledgments We thank the anonymous reviewers for comments on an earlier draft of this paper, including improvements referenced at the end of Section 3.

References

[BL04] P. Bürgisser and M. Lotz. Lower bounds on the bounded coefficient complexity of bilinear maps. *J. Assn. Comp. Mach.*, 51:464–482, 2004.

[BS82] W. Baur and V. Strassen. The complexity of partial derivatives. *Theor. Comp. Sci.*, 22:317–330, 1982.

[CT65] J.W. Cooley and J.W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Math. Comp.*, 19:297–301, 1965.

[DDPW83] D. Dolev, C. Dwork, N. Pippenger, and A. Wigderson. Superconcentrators, generalizers, and generalized connectors (preliminary version). In *Proc. 15th Annual ACM Symposium on the Theory of Computing*, pages 42–51, 1983.

[DS89] D.L. Donoho and P.B. Stark. Uncertainty principles and signal recovery. *SIAM J. Appl. Math.*, 49:906–931, 1989.

[Mor73] J. Morgenstern. Note on a lower bound of the linear complexity of the fast Fourier transform. *J. Assn. Comp. Mach.*, 20:305–306, 1973.

[Pud94] P. Pudlák. Communication in bounded-depth circuits. *Combinatorica*, 14:203–216, 1994.

[Rad68] C.M. Rader. Discrete Fourier transforms when the number of data samples is prime. *Proc. IEEE*, 56(6):1107–1108, 1968.

[Raz03] R. Raz. On the complexity of matrix product. *SIAM Journal of Computing*, 32(5):1356–1369, 2003.

[Raz08] R. Raz. Elusive functions and lower bounds for arithmetic circuits. Technical Report TR08-001, Electronic Colloq. on Computational Complexity, 2008.

[RS03] R. Raz and A. Shpilka. Lower bounds for matrix product, in arbitrary circuits with bounded gates. *SIAM J. Comput.*, 32:488–513, 2003.

[SJ96] P. Stevenhagen and H.W. Lenstra Jr. Chebotarëv and his density theorem. *Mathematical Intelligencer*, 18:26–37, 1996.

[SS91] V. Shoup and R. Smolensky. Lower bounds for polynomial evaluation and interpolation. In *Proc. 32nd Annual IEEE Symposium on Foundations of Computer Science*, pages 378–383, 1991.

[Tao05] T. Tao. An uncertainty principle for cyclic groups of prime order. *Mathematical Research Letters*, 12:121–127, 2005.