Journal of Computer and System Sciences **63**, 449–473 (2001)
doi:10.1006/jcss.2001.1786, available online at http://www.idealibrary.com on IDEAL®

Spectral Methods for Matrix Rigidity with Applications to Size–Depth Trade-offs and Communication Complexity[1]

Satyanarayana V. Lokam[2]

School of Mathematics, Institute for Advanced Study, Princeton, New Jersey 08540
E-mail: satya@math.ias.edu

Received March 15, 2000; revised March 1, 2001

The rigidity of a matrix measures the number of entries that must be changed in order to reduce its rank below a certain value. The known lower bounds on the rigidity of explicit matrices are very weak. It is known that stronger lower bounds would have important consequences in complexity theory. We consider some restricted variants of the rigidity problem over the complex numbers. Using spectral methods, we derive lower bounds on these variants. Two applications of such restricted variants are given. First, we show that our lower bound on a variant of rigidity implies lower bounds on size–depth trade-offs for arithmetic circuits with bounded coefficients computing linear transformations. These bounds generalize a result of Nisan and Wigderson. The second application is conditional; we show that it would suffice to prove lower bounds on certain restricted forms of rigidity to conclude several separation results such as separating the analogs of PH and PSPACE in the model of two-party communication complexity. Our results complement and strengthen a result of Razborov.

We introduce a combinatorial complexity measure, called AC^0-*dimension*, of sets of Boolean functions. While high rigidity implies large AC^0-dimension, large AC^0-dimension for explicit sets would already give explicit languages outside the analog of PH in two-party communication complexity. Moreover, the concept of AC^0-dimension allows us to formulate interesting combinatorial problems which may be easier than rigidity and which would still have consequences to separation questions in communication complexity.

[1] A preliminary version of this paper appeared in Proceedings of the 36th IEEE Symposium on Foundations of Computer Science (FOCS) 1995.

[2] Part of the work done while the author was a student at the University of Chicago. On leave from Department of Mathematical and Computer Sciences, Loyola University Chicago. Partially supported by NSF Grants CCR-9988359 and DMS-9729992.

0022-0000/01 $35.00

1. INTRODUCTION

The rigidity of a matrix A over a field $\mathbb{F}$, denoted by $\mathcal{R}_A^{\mathbb{F}}(r)$, is the number of entries of A that must be changed to reduce its rank to at most r.

Proving lower bounds on the rigidity, and related functions, of *explicit* matrices is a fundamental question with applications in algebraic complexity [Va77, Pu94, PRS97, SS91], communication networks [Pu94], branching programs [BRS93], threshold circuits [KW91], and communication complexity [Ra89].

Valliant [Va77] introduced the concept of rigidity and showed that almost all $n \times n$ matrices have a rigidity of $(n-r)^2$ over an infinite field and $\Omega((n-r)^2/\log n)$ over a finite field[3]. Pudlák and Rödl [PR94] showed a similar result for (0, 1)-matrices over $\mathbb{R}$. Valiant proposed the problem of finding *explicit* matrices with high rigidity in view of its application to algebraic complexity: a lower bound of $\mathcal{R}_A^{\mathbb{F}}(\varepsilon n) \geqslant n^{1+\delta}$, for some constants $\varepsilon, \delta > 0$, would imply that the linear trans formation defined by A cannot be computed by linear size, log-depth arithmetic circuits consisting of gates that compute linear functions over $\mathbb{F}$. We note that proving superlinear lower bounds on the arithmetic circuit size of explicit linear transformations is a major challenge in algebraic complexity theory [BoMu75, BCS97].

Proving superlinear lower bounds on the rigidity of explicit matrices (when $r = \varepsilon n$) remains an open question. The best known lower bound is $\Omega(n^2/r \log n/r)$ proved for various classes of explicit matrices by Friedman [Fr93], Shokrollahi *et al.* [SSS97], and the author [Lok00]. A slightly weaker bound of $\Omega(n^2/r)$ is also proved for explicit matrices by Pudlák [Pu94] (see also [PV91, KS92]), Razborov [Ra89], and Kashin and Razborov [KR97]. A good candidate for high rigidity over $\mathbb{R}$ seems to be an Hadamard matrix. The best known lower bound on the rigidity of a Hadamard matrix is $\Omega(n^2/r)$ proved in [KR97].

In view of the difficulty of proving strong lower bounds on rigidity, it seems natural to consider restricted versions of the rigidity problem and their applications to computational complexity. Such an approach was first taken by Krause and Waack [KW91]. They use spectral methods to derive lower bounds on a weak form of the rigidity function (which they call "variation rank") and use them to prove lower bounds on certain depth-2 circuits. In this paper, we expand the scope of spectral techniques as well as the range of applications of weak rigidity of matrices over the real and complex numbers. We derive new lower bounds on certain variants of rigidity. We show that these weaker problems still have interesting consequences in complexity theory.

Our first variant of rigidity considers the L_2-norm of changes (as opposed to the *number* of changes) needed to reduce the rank below a certain value. Lower bounds on this norm are used to prove lower bounds on size–depth trade-offs for linear transformations in a model of arithmetic circuits. In a *linear* circuit, each gate computes a linear combination of its inputs. Each output of a linear circuit is

[3] Over an infinite field, "almost all" is to be interpreted as a Zariski open set, i.e., the complement of the solution set of a finite system of algebraic equations; over a finite field it is interpreted in the usual counting sense.

clearly a linear form of its inputs and hence a linear circuit computes a linear transformation. A linear circuit uses bounded coefficients if the coefficients in the linear combination computed by each gate are bounded in absolute value by a constant. We show lower bounds of the form $\Omega(n^{1+\varepsilon/d})$ on the size of linear circuits of depth d with bounded coefficients that compute linear transformations given by explicit classes of matrices. This generalizes a result of Nisan and Wigderson [NW95], where they essentially consider the case $d=2$.

Other weaker forms of rigidity we consider constrain the changes in absolute value and in sign. Such questions are revelant to communication complexity. We prove the conditional result that it would suffice to prove lower bounds on such weaker forms to separate the communication complexity analogs of PH and PSPACE —a long-standing open question in communication complexity [BFS86]. This result strengthens a result of Razborov [Ra89].

Finally, when changes to the entries of a Hadamard matrix are bounded in absolute value *by a constant*, we get asymptotically optimal lower bounds, $\Omega(n(n-r))$, on the number of changes needed to reduce its rank below r.

At the same time, our techniques are general enough to yield a number of known results. These include an alternate proof of Alon's [Al94] lower bound of $\Omega(n^2/r^2)$ on the rigidity of a Hadamard matrix. We also obtain a generalization of a lower bound due to Krause and Waack [KW91] on variation rank as a corollary to one of our results. Another by-product is a lower bound on the rank of a matrix B in terms of its inner product with another matrix A and the spectral norms of A and B. An inequality due to Hoffman and Wielandt [HW53] plays a central role in our proofs. We note that this inequality was also used by Nisan and Wigderson [NW95] and simultaneously and independently in a previous version of the present paper [Lok94]. Our use of the Hoffman–Wielandt inequality, however, differs from its use by Nisan and Wigderson in several respects, leading to our more general results.

As mentioned before, we consider lower bounds on linear circuits over $\mathbb{C}$ with *bounded coefficients* that compute linear transformations. While the restriction of bounded coefficients is a severe one, studying arithmetic complexity in this model has some motivation, as discussed in [NW95]. Most significantly, no superlinear lower bounds are known in the general model for explicity defined matrices[4]. In fact, even for depth 2, with no restrictions on the coefficients, the best known lower bound is only $\Omega(n \log^2 n/\log\log n)$ that follows from results of Pudlák *et al.* [PRS97] and Radhakrishnan and Ta-Shma [RT97]. Second, Morgenstern [Mo73] and Chazelle [Ch94] suggest linear circuits with bounded coefficients as a natural model. Morgenstern observes that natural algorithms like FFT use only small constants and actually proves an optimal $n \log n$ lower bounds under this restriction. Chazelle considers the model with coefficients in $\{+1, 0, -1\}$ in the context of half plane range searching over a group. We note that Chazelle [Ch94] directly relates

[4] Shoup and Smolensky [SS91] give a lower bound of $\Omega(n \log n/\log d)$ on the size of a depth d linear circuit. However, the entries of the matrix they construct grow doubly exponentially with the dimension of the matrix and hence that matrix cannot be said to be explicity given in some natural sense of the phrase

(i.e., without using rigidity) the complexity of a linear circuit with bounded coefficients computing a linear transformation to its spectrum and proves similar $n \log n$ lower bounds in this model.

Our lower bounds on size–depth trade-offs for a linear transformation, given by a complex matrix A, are in general expressed as a function of the spectrum of the transformation. The bounds are interesting when the matrix AA^* has $\Omega(n)$ eigenvalues of value $\Omega(n^\varepsilon)$. Classes of such matrices include the Fourier transform matrix, any Hadamard matrix, and the incidence matrix of a projective plane. For such matrices, our lower bounds take the form $\Omega(n^{1+\varepsilon/2d})$ on the size of a linear circuit of depth d and bounded coefficients. Our result actually has a clean matrix interpretation: we prove a lower bound of $\Omega(n^{1+\varepsilon/2d})$ on the minimum of $\|B_1\|_1 + \cdots + \|B_d\|_1$ over all factorizations $A = B_1 \cdot \cdots \cdot B_d$, where $\|B_1\|$ is the sum of absolute values of the entries of B and A is any of the matrices mentioned above. In particular, this yields as a corollary the lower bounds of $\Omega(n^{1+\delta})$ due to Nisan and Wigderson [NW95] on the *bilinear* formula complexity with bounded coefficients computing the *bilinear* forms given by these classes of explicit matrices.

Next, we turn our attention to Boolean complexity. Babai *et al.* [BFS86] defined analogs of various complexity classes, like PH, PP, $\oplus$P, and PSPACE, in Yao's [Ya79] two-party communication complexity model (denoted by $\mathrm{PH}^{\mathrm{cc}}$, etc.). In this model, the characteristic function of a language L_A on pairs of m-bit strings can be thought of as a $2^m \times 2^m$ Boolean matrix A_n (with 0-1 or ± 1 entries), where $n := 2^m$. Razborov [Ra89] proves that good lower bounds on rigidity *over a finite field* imply strong separation results in communication complexity: For an explicit infinite sequence of (0,1)-matrices $\{A_n\}$ and a finite field $\mathbb{F}$, if $\mathscr{R}_A^{\mathbb{F}}(r) \geqslant n^2/2^{(\log r)^{o(1)}}$ for some $r \geqslant 2^{(\log\log n)^{\omega(1)}}$, then there is an expicit language $L_A \notin \mathrm{PH}^{\mathrm{cc}}$. At present, no explicit languages are known to be outside Σ_2^{cc}. We remark that *lower bounds in this communication complexity model imply lower bounds in Boolean circuit complexity.* An example involving the circuit complexity class ACC will be mentioned in Section 4.

We complement and strengthen the results of Razborov [Ra89] by relating variants of the rigidity problem *over* $\mathbb{R}$ to separation questions in communication complexity. To state our results, let $\mathscr{R}_A(r, \theta)$ denote the number of entries of a real matrix A that must be changed to reduce its rank below r, where the changes are constrained to be bounded in absolute value by θ. Then, we show the following: For an explicit infinte sequence of ± 1 matrices $\{A_n\}$, for some constant $c > 0$ and all constants $c_1, c_2 > 0$, if $\mathscr{R}_A(2^{(\log\log n)^{c_1}}, 2^{(\log\log n)^{c_2}}) \geqslant n^2/2^{(\log\log n)^c}$, then $L_A \notin \mathrm{PH}^{\mathrm{cc}}$. We are able to prove the following lower bound for any Hadamard matrix H: For any constant $c > 0$ and $r \leqslant n/2^{(\log\log n)^c}$, $\mathscr{R}_H(r, 2^{(\log\log n)^c}) = \Omega(n^2/2^{(\log\log n)^c})$. We note that improving this lower bound to insolve separate arbitrary positive constants on the l.h.s and r.h.s., respectively, would give a language in $\mathrm{PSPACE}^{\mathrm{cc}}$–$\mathrm{PH}^{\mathrm{cc}}$.

We note some structural results in communication complexity. The model of *interactive proof systems in communication complexity* can be defined in a natural way analogous to the corresponding notion in the Turing machine model [BaMo88, GMR89]. We consider this model and observe that the well-known result [LFKN92, Sha92] IP = PSPACE holds in the communication complexity world as well. Other simple observations include the validity of Toda's theorem ("PP is at least as hard as PH") in communication complexity and the containment of NC in PSPACE^{cc}. It is worth mentioning that certain simple functions in AC^0, such as equality testing, do not belong to NP^{cc}. Note, here and below, that NC and AC^0 refer to the usual circuit complexity classes.

In this paper we introduce a combinatorial measure on *sets* of Boolean functions called AC^0-dimension. Informally, a set of m-variable functions $\{f_1, ..., f_K\}$ is said to have AC^0-dimension d if there exist m-variable functions $\{g_1, ..., g_d\}$ such that each of the f_i's can be computed as the output of a constant depth, quasi-polynomial in m size circuit whose inputs are the g_j's. In our applications, K will be 2^m and we seek lower bounds larger than $2^{polylog(m)}$ on the AC^0-dimension of explicit sets of Boolean functions. While high rigidity implies large AC^0-dimension, large AC^0-dimension for explicit sets of Boolean functions would already give explicit languages outside PH^{cc}. Moreover, the concept of AC^0-dimension allows us to formulate interesting combinatorial problems which may be easier than rigidity and which would still have consequences to separation questions in communication complexity. In particular, using a simple OR-gate in the definition of AC^0-dimension, we formulate a combinatorial question with implications to separating bounded round interactive proof systems from their unbounded round counterparts in communication complexity.

Recent related work. After a preliminary version of this paper [Lok95] was published, several researchers contributed to substantial progress on the questions and techniques studied in this paper. We review some of these results here.

Razborov and Kashin [KR97] applied spectral techniques to prove lower bounds on the rigidity of Hadamard matrices. In particular, they prove the current best lower bound of $\Omega(n^2/r)$ on the rigidity of a Hadamard matrix, improving the bound of $\Omega(n^2/r^2)$ due to Alon [Al94] (Theorem 2.4(i)). They also improve our lower bound on $\mathscr{R}_A(r, \theta)$ (Theorem 2.4(ii)) by extending the applicable range of the parameter θ.

Pudlák [Pu98] uses determinant-based arguments similar to Morgenstern's [Mo73] to improve our size–depth trade-offs in Section 3 on linear circuits with bounded coefficients. In particular, he obtains a lower bound of $\Omega(dn^{1+1/d})$ for the Hadamard and Fourier transforms improving from $\Omega(n^{1+1/2d})$ presented in this paper. He also proves lower bounds on $\mathscr{R}_A(r, \theta)$ in terms of det A, the determinant of A. In a comment on this result, Razborov [Ra98] explains how to obtain Pudlák's lower bounds on $\mathscr{R}_A(r, \theta)$ using techniques from this paper and [KR97].

Shokrollahi *et al.* [SSS97] prove a lower bound of $\Omega(n^2/r \log n/r)$ on several classes of matrices over infinite and sufficiently large finite fields. In [Lok00], we use their technique to derive the same bound for the Fourier transform matrix. In the same paper, we observe that all the proofs, to our knowledge, of lower bounds on the rigidity of explicit matrices exploit the property that almost all submatrices of the candidate matrices have close to full rank. We note [Lok00] that such techniques are inherently limited in the sense that they cannot be used to prove that $\mathscr{R}_A(\varepsilon n) \geqslant n^{1+\delta}$, for constants $\varepsilon, \delta > 0$, as required in Valiant's [Va77] criterion.

Organization of the paper. In Section 2, we define various forms of the rigidity function and prove our lower bounds on them. Section 3 contains the lower bound proofs on size–depth trade-offs. Applications of matrix rigidity to the PH vs PSPACE question in communication complexity and ACC are given in Section 4. Section 5 includes our observations about some structural results in communication complexity. The concept of AC^0-dimension is discussed in Section 6. Section 7 concludes the paper with some open questions.

2. LOWER BOUNDS ON RIGIDITY

The set of all $n \times n$ complex matrices will be denoted by $\mathbb{C}^{n \times n}$. The superscript in the notation $\mathcal{R}_A^{\mathbb{F}}$ will be omitted when $\mathbb{F}$ is the field $\mathbb{C}$. We give below the formal definitions of the rigidity function and some variants.

DEFINITION 2.1 (Rigidity). For a matrix M, let $wt(M)$ denote the number of nonzero entries in M. Let $A \in \mathbb{C}^{n \times n}$ and $\theta \geqslant 0$.

(i) $\mathcal{R}_A(r) := \min_B \{wt(A-B)\colon \operatorname{rank}(B) \leqslant r\}$.

(ii) $\mathcal{R}_A(r, \theta) := \min_B \{wt(A-B)\colon \operatorname{rank}(B) \leqslant r, \forall i, j\ |a_{i,j} - b_{i,j}| \leqslant \theta\}$.

(iii) $\Delta_A^2(r) := \min_B \{\sum_{i,j} |a_{i,j} - b_{i,j}|^2\colon \operatorname{rank}(B) \leqslant r\}$.

We prove our lower bounds for a generalized Hadamard matrix. Although we state our results for this class of matrices, our proof technique can be adapted to prove lower bounds on rigidity and its variants of *any* matrix in terms of its *spectrum.*

DEFINITION 2.2. An $n \times n$ *complex* matrix H is called a *generalized Hadamard* matrix if (i) $|h_{ij}| = 1$ for all $1 \leqslant i, j \leqslant n$, and (ii) $HH^* = nI_n$, where H^* is the conjugate transpose of H and I_n is the $n \times n$ identity matrix.

Note that when H has only real entries, $h_{ij} = \pm 1$, and we get the usual definition of a Hadamard matrix. Also when $h_{ij} = \zeta^{ij}$, where ζ is a primitive nth root of unity, we get the Fourier transform matrix (character table of the cyclic group). More generally, the character table of any finite abelian group G (DFT matrix for G) is a generalized Hadamard matrix. The character table of an elementary abelian 2-group is called a *Sylvester matrix* and can be recursively defined by

$$H_1 = 1 \qquad \text{and} \qquad H_{2n} = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix}.$$

DEFINITION 2.3. Let $A \in \mathbb{C}^{n \times n}$. Then,

the *Frobenius norm* of A is

$$\|A\|_F := \left(\sum_{i,j} |a_{ij}|^2 \right)^{1/2}.$$

The *spectral norm* of A, $\|A\|_2$, usually, denoted by $\|A\|$, is defined by

$$\|A\| := \max_{x \neq 0} \|Ax\|/\|x\|,$$

where $\|\cdot\|$ on the r.h.s. is the Euclidean vector norm.

The ith *singular value*, $\sigma_i(A)$, is defined by $\sigma_i(A) = \sqrt{\lambda_i(AA^*)}$, $1 \leqslant i \leqslant n$, where $\lambda_i(AA^*)$ denotes the ith largest eigenvalue of AA^*.

The proposition below recalls some standard facts about singular values and their relations to ranks and norms of matrices.

PROPOSITION 2.1. *The following statements hold for any matrix $A \in \mathbb{C}^{n \times n}$.*

(a) *There exist unitary matrices $U, V \in \mathbb{C}^{n \times n}$ such that*

$$U^*AV = \operatorname{diag}(\sigma_1, ..., \sigma_n).$$

(b) *For $i = 1, ..., n$,*

$$\sigma_i(A) = \max_{dim(S) = i} \min_{0 \neq x \in S} \|Ax\|/\|x\|,$$

where S is an i-dimensional subspace of $\mathbb{C}^n$.

(c) $\operatorname{rank}(A) = r$ *if and only if* $\sigma_1(A) \geqslant \cdots \geqslant \sigma_r(A) > \sigma_{r+1}(A) = \cdots = \sigma_n(A) = 0$.

(d) $\|A\|_F^2 = \sigma_1^2(A) + \cdots + \sigma_n^2(A)$.

(e) $\|A\| = \sigma_1(A)$.

(f) *For any submatrices B of a matrix A*, $\operatorname{rank}(B) \geqslant \|B\|_F^2/\|A\|^2$.

Proof. Part (a) is a standard fact and its proof can be found, for instance, in [GV83, Sect. 2.3]. Part (b) follows from the Courtant–Fischer minimax theorem for eigenvalues. Parts (c), (d), and (e) follow from (a) and (b) by oberving that the rank, the Frobenius norm, and the spectral norm are invariant under unitary transformations.

To prove Part (f), let rank (B) be r. From (d) and (c), we have $\|B\|_F^2 = \sigma_1^2(B) + \cdots + \sigma_r^2(B) \leqslant r\sigma_1^2(B) = r\,\|B\|^2$. Since B is a submatrix of A, it is obvious that $\|B\| \leqslant \|A\|$. It follows that $r \geqslant \|B\|_F^2/\|B\|^2 \geqslant \|B\|_F^2/\|A\|^2$. ∎

It is clear from the definition that for a generalized Hadamard matrix H, $\sigma_i(H) = \sqrt{n}$ for all $1 \leqslant i \leqslant n$. Combined with Proposition 2.1(f), this immediately gives

COROLLARY 2.2. *For any $u \times v$ submatrix B of an $n \times n$ generalized Hadamard matrix H, rank$(B) \geqslant uv/n$.*

Remark 2.1. We see that any submatrix with $n^{1+\varepsilon}$ entries of a generalized Hadamard matrix must have rank at least n^{ε}. We remark that Borodin *et al.* [BRS93] proved a lower bound of $\Omega(n^{\varepsilon}/\log n)$ over a broader class of fields for the generalized Fourier transform matrix. They applied it to derive lower bounds on a restricted model of branching programs.

The following inequality of Hoffman and Wielandt [HW53] plays a central role in our proofs.

Lemma 2.3 (Hoffman–Wielandt). *Let A and B be matrices in $\mathbb{C}^{n\times n}$. Then,*

$$\sum_{i=1}^{n} [\sigma_i(A)-\sigma_i(B)]^2 \leqslant \|A-B\|_F^2.$$

Hoffman and Wielandt [HW53] proved their result for eigenvalues of normal matrices using the Birkhoff–von Neumann characterization of doubly stochastic matrices. The theorem for singular values as stated here can be found in [GV83, Sect. 8.3].

We now state our main results in this section.

Theorem 2.4. *Let H be an $n\times n$ generalized Hadamard matrix. Then*

(i) (Alon) $\mathscr{R}_H(r) \geqslant \max\{n^2/(r+1)^2, n-r\}$.

(ii) *For* $\theta \leqslant n/r-1$, $\mathscr{R}_H(r,\theta) \geqslant n^2/(1-1/(\theta+1))/4(\theta+1)$.

(iii) $\Delta_H^2(r) = n(n-r)$.

Remark 2.2. Kashin and Razborov [KR97] improved this result after a preliminary version of this paper [Lok95] was published. In particular, they show that $\mathscr{R}_H(r) = \Omega(n^2/r)$ for $r \leqslant n/2$ improving part (i). They also improve part (ii) by showing that when $\theta \geqslant n/r$, $\mathscr{R}_H(r,\theta) = \Omega(n^2/r\theta^2)$. Their proofs are also based on spectral techniques.

Note that Theorem 2.4(iii) immediately implies that $\mathscr{R}_H(r,\theta) \geqslant n(n-r)/(\theta+1)^2$. In particular, when the changes are bounded by a constant, we get the asymptotically optimal lower bound $\Omega(n(n-r))$ on the number of changes needed to reduce the rank of H below r.

Theorem 2.4(ii) gives,

Corollary 2.5. *For any constant $c>0$ and $r \leqslant n/2^{(\log\log n)^c}$, $\mathscr{R}_H(r, 2^{(\log\log n)^c})$ $= \Omega(n^2/2^{(\log\log n)^c})$.*

We note that improving this lower bound to involve separate arbitrary positive constants on the l.h.s. and r.h.s., respectively, would give a language outside $\mathrm{PH}^{\mathrm{cc}}$ (see Theorem 4.1). The language corresponding to the function inner product mod 2 is in $\mathrm{PSPACE}^{\mathrm{cc}}$. The associated infinite family of matrices are Hadamard (Sylvester matrices). Thus the improvement mentioned above would give an explicit language in $\mathrm{PSPACE}^{\mathrm{cc}}$–$\mathrm{PH}^{\mathrm{cc}}$.

Proof of Theorem 2.4. Part (i): Clearly, for any r, we need to change at least $n-r$ entries to bring the rank of the full-rank matrix H down to r.

The rest of the proof follows from Corollary 2.2 and a counting argument identical to Alon's [Al94]. We include it for completeness. Suppose there are fewer than $n^2/(r+1)^2$ changes in H. Then, there is an $(r+1)\times n$ submatrix in which there are fewer than $n/(r+1)$ changes. By removing the columns in which a change occurred, we get an $(r+1)\times t$ submatrix in which no change took place, where $t>n-n/(r+1) = nr/(r+1)$. Hence by Corollary 2.2, this submatrix has rank at least $(r+1)$ since $(r+1)\,t/n>r$. This shows that at least $n^2/(r+1)^2$ changes must occur to reduce the rank below $r+1$.

To prove Parts (ii) and (iii) of Theorem 2.4, we will use the Hoffman–Wielandt inequality, Lemma 2.3.

Part (iii): Let B be a matrix of rank r achieving the minimum in Definition 2.1 (iii). Then from Proposition 2.1(c), $\sigma_{r+1}(B)= \cdots =\sigma_n(B)=0$. Thus,

$$\sum_{i=1}^{n} [\sigma_i(H)-\sigma(B)]^2 \geqslant \sum_{i=r+1}^{n} (\sigma_i(H))^2=n(n-r).$$

Using the Hoffman–Wielandt inequality,

$$\Delta_H^2(r)=\|H-B\|_F^2 \geqslant n(n-r).$$

It is easy to see that equality is achieved in the above bound for the matrix $B=U\,\mathrm{diag}(\sigma_1, ..., \sigma_r, 0, ..., 0)\,V^*$, where U and V are from the singular value decomposition of A given in Proposition 2.1.

Part (ii): Let B be a matrix such that $wt(H-B)=\mathscr{R}_H(r,\theta)$. Define ε to be the fraction of entries where b_{ij} differs from h_{ij}. Clearly, $wt(H-B)\geqslant \varepsilon n^2$.

Let us define $B' := B/(\theta+1)$. When H and B agree $|h_{ij}-b'_{ij}|=(1-1/(\theta+1))$, and when they do not $|h_{ij}-b'_{ij}|\leqslant 2$, since $|b'_{ij}|\leqslant 1$. Thus,

$$\|H-B'\|_F^2 \leqslant (1-1/(\theta+1))^2\,(1-\varepsilon)\,n^2+4\varepsilon n^2. \tag{1}$$

Since rank $(B')=\mathrm{rank}(B)=r$, we also have, using Proposition 2.1(c), that

$$\sum_{i=1}^{n} (\sigma_i(H)-\sigma_i(B'))^2 \geqslant \sigma_n^2(H)(n-r)=n(n-r). \tag{2}$$

From inequalities (1), (2) and Theorem 2.3,

$$n(n-r)\leqslant n^2(1-1/(\theta+1))^2+n^2\varepsilon(4-(1-1/\theta+1))^2). \tag{3}$$

This gives

$$\begin{aligned}\mathscr{R}_H(r,\theta)&=wt(H-B)\geqslant \varepsilon n^2\\ &\geqslant \frac{n(n-r)-n^2(1-1/(\theta+1))^2}{4-(1-1/(\theta+1))^2}\\ &\geqslant n^2(1-1/(\theta+1))/(4(\theta+1)),\end{aligned}$$

since $r/n\leqslant 1/(\theta+1)$.

This concludes the proof of Theorem 2.4. ∎

Using the ideas from the previous proof we can give a simpler and slightly more general proof of a theorem of Krause and Waack [KW91] on variation rank. Their result is obtained by setting $\varepsilon = 0$ in the following theorem.

THEOREM 2.6. *Let A be an $n \times n \pm 1$-matrix and B an $n \times n$ real matrix such that*

- $1 \leqslant |b_{ij}| \leqslant \theta$, *and*
- *$sign(a_{ij}) = sign(b_{ij})$ for all i, j* except *an ε-fraction.*

Then, $\operatorname{rank}(B) \geqslant n^2(1-4\varepsilon\theta)/(\theta \cdot \|A\|^2)$.

Proof. As before, let us set $B' := B/\theta$. Since $1/\theta \leqslant |b'_{ij}| \leqslant 1$, when H and B agree in sign, $|h_{ij} - b'_{ij}| \leqslant (1-1/\theta)$, and when they do not $|h_{ij} - b'_{ij}| \leqslant 2$. Hence,

$$\|A - B'\|_F^2 \leqslant (1-1/\theta)^2\,(1-\varepsilon)\,n^2 + 4\varepsilon n^2 \leqslant (1-1/\theta)^2\,n^2 + 4\varepsilon n^2. \tag{4}$$

On the other hand, since $\operatorname{rank}(B') = r$,

$$\begin{aligned}\sum_{i=1}^{n} (\sigma_i(A) - \sigma_i(B'))^2 &\geqslant \sum_{i=r+1}^{n} (\sigma_i(A))^2 \\ &= \sum_{i=1}^{n} (\sigma_i(A))^2 - \sum_{i=1}^{r} (\sigma_i(A))^2 \\ &\geqslant \|A\|_F^2 - r \cdot \|A\|^2.\end{aligned}$$

Using this and (4) in the Hoffman–Wielandt inequality, and noting that $\|A\|_F^2 = n^2$, we get

$$n^2 - r\,\|A\|^2 \leqslant n^2(1-1/\theta)^2 + 4\varepsilon n^2,$$

which gives

$$r \geqslant \frac{n^2}{\theta \cdot \|A\|^2}\,((2-1/\theta) - 4\varepsilon\theta),$$

and the theorem follows since $\theta \geqslant 1$. ∎

This lower bound (with $\varepsilon = 0$) was used by Krause and Waack [KW91] to derive exponential size lower bounds on certain depth-2 circuits. It can also be used to prove a separation result in communication complexity: $\mathrm{PP}^{cc} \neq \mathrm{PSPACE}^{cc}$.

In the notation of Theorem 2.6, a lower bound on $\operatorname{rank}(B)$ when $\varepsilon = 0$ and $\theta = \infty$ (sign-preserving changes of arbitrarily large size) is a fundamental question. It arises in the model of unbounded probabilistic communication complexity defined by Paturi and Simon [PS86]. An equivalent combinatorial problem concerns geometric realizations of set systems [AFR85]. Even proving that for almost all A, $\operatorname{rank}(B) = \Omega(n)$ under sign-preserving changes (of arbitrary size) is a nontrivial result due to Alon *et al.* [AFR85] making use of the Milnor–Thom bound on Betti numbers of real algebraic varieties.

We conclude this section with the following proposition that may be of independent interest. It is a simple consequence of the Hoffman–Wielandt inequality (Lemma 2.3) and generalizes Proposition 2.1(f).

PROPOSITION 2.7. *Let $A, B \in \mathbb{C}^{n \times n}$. Then,*

$$\operatorname{rank}(B) \geqslant \frac{\Re\langle A, B\rangle}{\|A\| \, \|B\|},$$

where $\langle A, B\rangle := Tr(AB^*)$ and $\Re x$ denotes the real part of a complex number x.

Proof. Using Theorem 2.3,

$$\begin{aligned}
\|A-B\|_F^2 &\geqslant \sum_{i=1}^{n} (\sigma_i(A)-\sigma_i(B))^2 \\
&= \|A\|_F^2 + \|B\|_F^2 - 2\sum_{i=1}^{n} \sigma_i(A)\,\sigma_i(B), \\
&\qquad \text{using Proposition 2.1(d)} \\
&\geqslant \|A\|_F^2 + \|B\|_F^2 - 2\operatorname{rank}(B)\,\|A\|\,\|B\|, \\
&\qquad \text{using Proposition 2.1(c) and (e).}
\end{aligned}$$

Observe that for any matrix M, $\|M\|_F^2 = Tr(MM^*)$.
Using this in the last inequality above, we get

$$\begin{aligned}
2\operatorname{rank}(B)\,\|A\|\,\|B\| &\geqslant \|A\|_F^2 + \|B\|_F^2 - \|A-B\|_F^2 \\
&= Tr(AB^*) + Tr(BA^*) \\
&= 2\Re Tr(AB^*),
\end{aligned}$$

and the proposition is proved. ∎

3. SIZE–DEPTH TRADE-OFFS FOR LINEAR TRANSFORMATIONS

For a matrix A over a field $\mathbb{F}$, let ℓ_A and b_A denote the linear transformation and bilinear form, respectively, defined by A, i.e., $\ell_A(x) := Ax$ and $b_A(x, y) := y^T A x$.

A *linear circuit* is a directed acyclic graph whose inputs are labeled by elements of $\{x_1, ..., x_n\}$ and edges are labeled by nonzero scalars from the field $\mathbb{F}$. Each internal node (a linear gate) of the circuit computes a linear combination of its inputs; the coefficients of the linear combination are given by the scalars on the input wires to the gate. Hence every gate computes a linear form in the input vector x. A circuit is said to compute the linear forms $\{f_1, ..., f_m\}$ if each f_i is computed at some internal node of the circuit. Given an $m \times n$ matrix A, the linear transformation ℓ_A naturally defines a set of m linear forms in x. We say a circuit computes the linear transformation ℓ_A if it computes the corresponding linear forms. The *size* of a linear circuit is the number of wires in it. The *depth* of a circuit is the length of the longest path from an input to an output. Let $C^{[d]}(\ell_A)$ denote the minimum size of a depth d linear circuit (with n inputs and m outputs) computing ℓ_A.

A *bilinear formula* for b_A is defined by t pairs of vectors p_i, q_i, $1 \leqslant i \leqslant t$, for some t, such that

$$b_A(x, y) = \sum_{i=1}^{t} y^T q_i p_i^T x.$$

The *size* of this bilinear formula is defined to be

$$\sum_{i=1}^{t} (wt(p_i) + wt(q_i)).$$

Recall that $wt(p)$ denotes the number of nonzero entries of the vector p. Such a formula is naturally represented by a depth-3 tree T where the root of T is an unbounded fan-in addition gate, the next level has multiplication gates of fan-in 2, and the bottom level has linear gates. The pair of inputs to the i-th multiplication gate compute linear forms $p_i^T x$ and $y^T q_i$. The non-zero coefficients of these linear forms appear as the scalars on the input wires of the bottom level gates. The size of the bilinear formula is then the number of leaves of this tree. Let $L(b_A)$ denote the minimum size of a bilinear formula computing b_A.

DEFINITION 3.1. For a matrix A over a fixed field $\mathbb{F}$, we define $w_d(A)$ by

$$w_d(A) := \min \left\{ \sum_{i=1}^{d} wt(B_i) : A = B_1 \cdot \cdots \cdot B_d \right\},$$

where B_i are matrices of arbitrary dimensions over $\mathbb{F}$.

The next lemma is implicit in [Pu94, Sect. 3]. We include its proof for completeness.

LEMMA 3.1. *For any matrix A, $w_d(A) \geqslant C^{[d]}(\ell_A) \geqslant w_d(A)/d$.*

Proof. Let C be a depth d circuit computing ℓ_A, where A is an $m \times n$ matrix. At the expense of at most a factor of d, we can assume that the circuit C is leveled; i.e., for $k = 0, \ldots, d-1$, wires go from level k only to level $k+1$. Let t_k be the number of nodes on level k. Thus $t_0 = n$ and $t_d = m$. For $1 \leqslant k \leqslant d$ define the $t_{k-1} \times t_k$ matrix B_{d-k+1} by setting its (i, j)th entry, $\beta_{ij}^{(d-k+1)}$, to be the scalar on the wire connecting the ith node on the kth level to the jth node on the $(k-1)$st level ($\beta_{ij}^{(d-k+1)} = 0$ if there is no such edge). Let z_k be the length-t_k vector computed by nodes at level k. Then, it is easy to see that $z_k = B_{d-k+1} z_{k-1}$. So, we must have $Ax = z_d = B_1 \cdot \cdots \cdot B_d x$, since z_d is the output vector of C and $z_0 = x$ is its input vector. Furthermore, the number of wires between levels k and $k-1$ is equal to the number of nonzero entries in B_{d-k+1}. Thus $A = B_1 \cdot \cdots \cdot B_d$ and the complexity of the circuit C is $\sum_{k=1}^{d} wt(B_k)$.

Conversely, given a decomposition $A = B_1 \cdot \cdots \cdot B_d$, we can construct a leveled circuit of depth d and number of wires $\sum_{k=1}^{d} wt(B_k)$. ∎

COROLLARY 3.2. $L(b_A) = \Theta(C^{[2]}(\ell_A))$.

Proof. From Lemma 3.1 for $d=2$, it follows that $C^{[2]}(\ell_A) = \Theta(w_2(A))$. Nisan and Wigderson [NW95, Eqs. 1 and 2] show that $L(b_A) = w_2(A)$. ∎

We will prove lower bounds, for explicit matrices A over the field $\mathbb{C}$ of complex numbers, on the complexity of linear circuits for ℓ_A when the *scalars on the wires are bounded in absolute value by a constant*. For full generality, we allow multiple wires out of one gate into another. We may asume that the scalars on the wires are bounded in absolute value by 1. Modifications to the calculations when the scalars are bounded by an arbitrary constant are straighforward. We will use the subscript 1 to denote these restricted complexities: $C_1^{[d]}(\ell_A)$ denotes the minimum size of a depth d linear circuit computing ℓ_A with coefficients of absolute value at most 1, and $L_1(b_A)$ denotes the minimum size of a bilinear formula computing b_A with coefficients of absolute value at most 1.

In fact, our lower bounds apply to the L_1-norm of linear circuits: for a linear circuit C, let $\|C\|_1$ denote the sum of absolute values of the scalars on the wires of C. For a matrix A, let us define $\|C^{[d]}(\ell_A)\|_1$ to be the minimum L_1-norm, $\|C\|_1$, of a linear circuit C of depth d that computes ℓ_A. Clearly,

PROPOSITION 3.3. *For any complex matrix A, $C_1^{[d]}(\ell_A) \geqslant \|C^{[d]}(\ell_A)\|_1$.*

The following lemma uses ideas from [Pu94] and [Va77]. We remark that this lemma and the next theorem are proved using the L_2-norm of changes as in Definition 2.1(iii). The connection to the spectrum of A is made explicit below in Theorem 3.5 using the Hoffman–Wielandt inequality (Lemma 2.3).

LEMMA 3.4. *For any $r \geqslant 1$,*

$$\|C^{[d]}(\ell_A)\|_1 \geqslant r \cdot \left(\frac{\Delta_A^2(r)}{n}\right)^{1/2d},$$

where $\Delta_A^2(r) := \min_B \{\sum_{i,j} |a_{ij} - b_{ij}|^2 : \operatorname{rank}(B) \leqslant r\}$.

Proof. Let S be the L_1-norm of a circuit C computing ℓ_A. Call a node g of C special if the sum of absolute values of scalars on the outgoing edges of g is at least S/r. There are at most r special nodes.

Form the matrix B by setting $B(i, j) =$ sum of the products along the paths from input node j to output node i that go through at least one special node. Then, $rank(B) \leqslant r$ since it can be written as the sum of at most r rank-1 matrices, one for each special node. Indeed, let $g_1, ..., g_r$ be the special nodes. For $k = 1, ..., r$, let $z_k = q_k^T x$ be the linear form computed at g_k. Let Q be the $r \times n$ matrix with rows q_k, $1 \leqslant k \leqslant r$. Define P to be the $n \times r$ matrix of the linear transformation computed by the partial circuit of C with $g_1, ..., g_r$ as its inputs obtained by retaining a path from an output to a special node iff it contains no other special node in its interior. Let p_k, for $1 \leqslant k \leqslant r$, be the kth column of P. Then, it is easy to show that $B = PQ = p_1 q_1^T + \cdots + p_r q_r^T$. Since each summand is a rank-1 matrix, it follows that B has rank at most r.

Now remove these special nodes and let $K = A - B$ be the matrix corresponding to the linear transformation computed by the remaining circuit C'.

We will now estimate the L_1-norm of any column k_j of the matrix K. Expand the subcircuit of C' from the input j to the set of all outputs into a tree. For notational convenience, let us define the weight of a path in this tree to be the product of the scalars appearing on the edges of the path. Define the weight of the tree to be the sum of absolute values of the weights of all the paths from the root (input node j) to the leaves of this tree (output nodes of C', possible repeated). Clearly $\|k_j\|_1$ is at most the weight of the tree. The tree has depth at most d and contains only non-special nodes. Hence the sum of the absolute values of scalars on the outgoing edges of any node is at most S/r. By induction on d, it is easily seen that the weight of the tree is at most $(S/r)^d$.

Thus, every column k_j of K has L_1-norm bounded by $(S/r)^d$. Hence $\|K\|_F^2 = \sum_{j=1}^n \|k_j\|_2^2 \leqslant n(S/r)^{2d}$, since $\|k_j\|_2 \leqslant \|k_j\|_1$. We therefore have

$$\Delta_A^2(r) \leqslant \|A - B\|_F^2 \leqslant n(S/r)^{2d},$$

and the lemma follows by solving for S. ∎

THEOREM 3.5. *For any complex matrix A and any constant ε, $0 < \varepsilon < 1$, $C_1^{[d]}(\ell_A) = \Omega(\sum_{\varepsilon n < j \leqslant n} (\sigma_j(A))^{1/d})$.*

Proof. From Hoffman–Wielandt inequality (see proof of Theorem 2.4(iii)), we know that for integer r, $0 \leqslant r \leqslant n$,

$$\Delta_A^2(r) \geqslant \sum_{j=r+1}^{n} (\sigma_j(A))^2.$$

Using this in Lemma 3.4, we get

$$C_1^{[d]}(\ell_A) \geqslant r\left(\frac{1}{n}\sum_{j=r+1}^{n} (\sigma_j(A))^2\right)^{1/2d}.$$

Using Hölder's inequality,

$$(\sigma_{r+1}^2 + \cdots + \sigma_n^2)^{1/2d} \geqslant (\sigma_{r+1}^{1/d} + \cdots + \sigma_n^{1/d})/(n-r)^{1-1/2d}.$$

Setting $r = \varepsilon n$ and the inequality in Proposition 3.3, we get the result. ∎

Using Corollary 3.2 and Theorem 3.5 for depth-2, we get

COROLLARY 3.6. *For any constant ε, $0 < \varepsilon < 1$, $L_1^b(b_A) = \Omega(\sum_{\varepsilon n < j \leqslant n} \sqrt{\sigma_j(A)})$.* ∎

COROLLARY 3.7 [NW95, Therorem 12]. (i) *If A is a generalized Hadamard matrix, then $L_1^b(b_A) = \Omega(n^{5/4})$.*

(ii) *If A is the incidence matrix of a projective plane, then $L_1^b(b_A) = \Omega(n^{9/8})$.* ∎

Remark 3.1. Using Theorem 3.5 and an analog of Lemma 3.1 (with L_1-norm in place of weight *wt*), we get a lower bound of $\Omega(\sum_{\varepsilon n < j \leqslant n} (\sigma_j(A))^{1/d})$ on the sum $\|B_1\|_1 + \cdots + \|B_d\|_1$ for any factorization $A = B_1 \cdot \cdots \cdot B_d$, where $\|B\|_1$ is the sum of absolute values of the entries of B.

Remark 3.2. Lemma 3.4 and Theorem 3.5 are improved by Pudlák [Pu98]. In particular, he shows a lower bound on the L_2-norm of linear circuits (cf. the definitions before Proposition 3.3),

$$\|C^{[d]}(\ell_A)\|_2^2 \geqslant dn\,|\det A|^{2/dn},$$

where $\det A$ denotes the determinant of the matrix A. This implies a lower bound of $C_1^{[d]}(\ell_H) \geqslant dn^{1+1/d}$, where H is a generalized Hadamard matrix. Accordingly, the bounds in Corollary 3.7(i) and (ii) are improved to $\Omega(n^{3/2})$ and $\Omega(n^{5/4})$, respectively.

Pudlák also shows the following lower bound on $\mathcal{R}_A(r, \theta)$ (cf. Definition 2.1(ii)),

$$\mathcal{R}_A(r, \theta) \geqslant (n-r)\left(\frac{|\det A|}{r^{r/2}}\right)^{2/(n-r)} \theta^{-O(1)},$$

where A has entries of absolute value at most θ, $\theta \geqslant 1$, and $r \leqslant n/2$. Razborov [Ra98] explains how to prove this bound using techniques similar to ours.

4. APPROXIMATING COMMUNICATION MATRICES

It has been a long-standing open question to separate the communication complexity analogs of PH and PSPACE [BFS86]. In this section, we relate this question to weak rigidity. We also mention another simple question that relates weak rigidity to complexity of Boolean circuits with modular gates (this model is used to define the circuit complexity class ACC). This is a slight modification of a question described by Pudlák and Rödl [PR94].

Taking a complexity theoretic view of Yao's [Ya79] model of communication complexity, Babai *et al.* [BFS86] defined analogs of various Turing machine complexity classes. To define communication complexity classes, we consider languages consisting of pairs of strings (x, y) such that $|x| = |y|$. Denote by Σ^{2*} the universe $\{(x, y): x, y \in \{0, 1\}^*$ and $|x| = |y|\}$. For a language $L \subseteq \Sigma^{2*}$, we denote its characteristic function on pairs of strings of length m by L_n, where $n := 2^m$. L_n is naturally represented as an $n \times n$ matrix with 0-1 or ± 1 entries (with -1 for *true* and $+1$ for *false*). Conversely, if $A = \{A_n\}$ is an infinite sequence of ± 1-matrices (where A_n is $n \times n$), then we can associate a language L_A with A and talk about its communication complexity. L_A is not necessarily unique (since the n's may be different from powers of two), but for the purposes of lower bounds we will fix one such language and refer to it as *the* language L_A corresponding to A.

We recall the following definitions from [BFS86]:

DEFINITION 4.1. Let the nonnegative integers $l_1(m), ..., l_k(m)$ satisfy the inequality $l(m) := \sum_{i=1}^{k} l_i(m) \leqslant (\log m)^c$ for a fixed constant $c \geqslant 0$.

A language L is in Σ_k^{cc} if, for some choice of $l_i(m)$, there exist Boolean functions φ, ψ: $\{0, 1\}^{m+l(m)} \to \{0, 1\}$ such that $(x, y) \in L_n$ if

$$\exists u_1\ \forall u_2 ... Q_k u_k(\varphi(x, u) \diamondsuit \psi(y, u)),$$

where $|u_i| = l_i(m)$, $u = u_1 ... u_k$, Q_k is $\forall$ for k even and is $\exists$ for k odd and $\diamondsuit$ stands for $\vee$ if k is even and for $\wedge$ if k is odd.

DEFINITION 4.2. By allowing a bounded number of alternating quantifiers in Definition 4.1, we get an analog of the polynomial time hierarchy: $\mathrm{PH}^{\mathrm{cc}} = \bigcup_{k \geqslant 0} \Sigma_k^{cc}$.

DEFINITION 4.3. By allowing an unbounded, but no more than polylog(m), number of alternating quantifiers in Definition 4.1, we get an analog of PSPACE: $\mathrm{PSPACE}^{\mathrm{cc}} = \bigcup_{c>0} \bigcup_{k \leqslant (\log m)^c} \Sigma_k^{cc}$.

THEOREM 4.1. *Let $\{A_n\}$ be an infinite sequence of ± 1-matrices and L_A be the associated language. For some constant $c > 0$ and all constants $c_1, c_2 > 0$, if $\mathcal{R}_A(2^{(\log\log n)^{c_1}}, 2^{(\log\log n)^{c_2}}) \geqslant n^2/2^{(\log\log n)^c}$, then $L_A \notin \mathrm{PH}^{\mathrm{cc}}$.*

This theorem is proved in Section 4.1 using Tarui's [Ta93] low-degree polynomials (over integers) that approximate AC^0 circuits.

In the case of an ACC circuit, we use the results of Beigel and Tarui [BT91] and Green *et al.* [GKT92] that reduce an ACC circuit to a depth-two circuit with a MidBit gate at the top and polylog fan-in AND gates at the bottom. A MidBit gate over w inputs $x_1, ..., x_w$ outputs the value of the $\lfloor(\log w)/2\rfloor$th bit of the binary representation of the number $\sum_{i=1}^{w} x_i$. Using this depth-two circuit in the proof of Theorem 4.1 gives us a rigidity question, which we state below, with consequences to separating ACC.

DEFINITION 4.4. Fix disjoint subsets S and T of $\mathbb{R}$. A matrix B is said to (S, T)-represent a ± 1-matrix A if for all x and y, $B(x, y) \in S$ if $A(x, y) = +1$ and $B(x, y) \in T$ if $A(x, y) = -1$.

DEFINITION 4.5. $\rho_A(S, T) := \min\{\mathrm{rank}(B)$: B (S, T)-represents $A\}$.

Remark 4.1. The following remarks indicate the significance of obtaining lower bounds on this function.

1. When S and T are the set of positive and negative *integers* respectively, bounded in absolute value by θ, this becomes the definition of *variation rank* [KW91]. In this case, Theorem 2.6 gives a lower bound of $\rho_A(S, T) \geqslant n^2(2 - 1/\theta)/(\theta \cdot \|A\|^2)$, slightly improving the bound of [KW91]. We note that this bound when applied to the Sylvester matrix can be used to give an alternative proof of a result in [HR90] that $\oplus\mathrm{P}^{\mathrm{cc}} \nsubseteq \mathrm{PP}^{\mathrm{cc}}$ (see [Lok94]).

2. When S and T are positive and negative *reals*, respectively, then proving that $\rho_{A_n}(S, T) \geqslant 2^{(\log\log n)^{\omega(1)}}$ for an explicit family $\{A_n\}$ would yield an explicit language outside UPP^{cc} (cf. Discussion after Theorem 2.6).

In the following theorem, the sets S_c and T_c can be explicity described using the MidBit function. We will omit these details.

THEOREM 4.2. *Let $\{A_n\}$ be an infinite sequence of ± 1-matrices and L_A be the associated language. For all $c > 0$, there exist (explicity defined) partitions $S_c \cup T_c$ of the integers $\{-2^{(\log\log n)^c}, ..., +2^{(\log\log n)^c}\}$ such that the following holds: If for all constants $c > 0$, $\rho_{A_n}(S_c, T_c) \geqslant 2^{(\log\log n)^{\omega(1)}}$, then $L_A \notin \mathrm{ACC}$.* ∎

It seems plausible that there may be explicit matrices for which the "if" part of the theorem is true for any nontrivial partition $S \cup T$ of integers in the given range.

4.1. Proof of Theorem 4.1

Let L be a language in PH^{cc} and let A_n be its $n \times n \pm 1$-matrix, where $n := 2^m$. The theorem will follow from the following: for all $c > 0$, there exist $c_1, c_2 > 0$, and integer matrices $\{B_n\}$, where B_n is $n \times n$, such that

(i) $\forall(x, y),\ 1 \leqslant |B_n(x, y)| \leqslant 2^{(\log\log n)^{c_1}}$,

(ii) $\mathrm{rank}(B_n) \leqslant 2^{(\log\log n)^{c_2}}$, and

(iii) $wt(A_n - B_n) \leqslant n^2/2^{(\log\log n)^c}$.

For simplicity of notation, let $L \in \Sigma_k^{cc}$ where k is odd. In Definition 4.1 of Σ_k^{cc}, for any fixed sequence of moves $u = u_1, ..., u_k$, φ is a function of x and ψ is a function of y. Define $f_u(\cdot) \equiv \varphi(\cdot, u)$ and similarly $g_u(\cdot) \equiv \psi(\cdot, u)$. Replacing $\exists$ move by an OR-gate and $\forall$ move by an AND-gate, we see that L has a Σ_k^{cc} protocol iff it can be expressed as the output of an {AND, OR} circuit C of depth k and size $2^{polylog(m)}$ where the inputs of C are $f_u(x) \wedge g_u(y)$ for $1 \leqslant u \leqslant 2^{polylog(m)}$. Hence, for all $(x, y) \in \{0, 1\}^m \times \{0, 1\}^m$,

$$L(x, y) = C(f_1(x) \wedge g_1(y), ..., f_t(y) \wedge g_t(y)), \tag{5}$$

where $t \leqslant 2^{polylog(m)}$ is the number of possible u's.

Considering f_i as the characteristic function of a subset U_i of rows and g_i as that of a subset V_i of columns of the $\{0, 1\}^m \times \{0, 1\}^m$ matrix, we observe that $f_i(x) \wedge g_i(y)$ is a *rectangle* $U_i \times V_i$ in the matrix. We will denote this rectangle by R_i and identify it with the corresponding $n \times n$ (0, 1)-matrix of rank 1:

$$R_i(x, y) = \begin{cases} 1 & \text{if } f_i(x) \wedge g_i(y) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

From Eq. (5), it follows that L is in Σ_k^{cc} iff its matrix is expressible by an AC^0 circuit (of quasipolynomial size) acting on a set of rank 1 matrices.

We now use the fact that an AC^0 circuit is well approximated by a low-degree polynomial over $\mathbb{Z}$. Tarui [Ta93] constructs such polynomials.

THEOREM 4.3 (Tarui). *Let C be an* AC^0 *circuit of size* $2^{polylog(t)}$ *and let* $\phi_1, \ldots, \phi_t\colon \{0,1\}^s \to \{0,1\}$ *be arbitrary Boolean functions. Fix* $0 < \delta = 2^{-(\log t)^{c'}}$, *for some constant* $c' \geqslant 0$. *Then there exist constants* $c_1', c_2' \geqslant 0$, *and a polynomial* $\Phi(\phi_1, \ldots, \phi_t)$ *such that*

- Low degree: *the degree of* Φ *in* $\phi_1, \ldots, \phi_t$ *is at most* $(\log t)^{c_2'}$.
- Small error: *The fraction of inputs* $x \in \{0,1\}^s$ *where* $C(\phi_1, \ldots, \phi_t)(x) \neq \Phi(\phi_1, \ldots, \phi_t)(x)$ *is at most* δ.
- Small norm: *The sum of the absolute values of the coefficients of* Φ *is at most* $2^{(\log t)^{c_1'}}$.
- Boolean guarantee: *When* Φ *differs from* C, *the value of* $\Phi(\phi_1, \ldots, \phi_t)(x)$ *is* $\geqslant 2$.

Let L_n be the (0,1)-matrix describing L at input length n; i.e., $L_n = (J_n - A_n)/2$ where J_n is the $n \times n$ all 1's matrix.

From Eq. (5), L_n is computed by an AC^0 circuit $C(z_1, \ldots, z_t)$ of size $2^{polylog(m)}$ where $z_i = f_i(x) \wedge g_i(y) = f_i(x)\, g_i(y)$ since f_i, g_i are $\{0,1\}$ functions. Using Theorem 4.3 for C, there is a $d \leqslant polylog(t)$ such that $L(x, y) = \Phi(x, y)$, except for an ε fraction of $(x, y) \in \{0,1\}^m \times \{0,1\}^m$, where

$$\begin{aligned}\Phi(x, y) &= \sum_{S \subseteq [t],\, |S| \leqslant d} \alpha_S \prod_{i \in S} z_i \\ &= \sum_{S \subseteq [t],\, |S| \leqslant d} \alpha_S \prod_{i \in S} f_i(x)\, g_i(y) \\ &= \sum_{S \subseteq [t],\, |S| \leqslant d} \alpha_S\, f_S(x)\, g_S(y).\end{aligned}$$

Here $f_S(x) = \prod_{i \in S} f_i(x)$ and similarly g_S.

Returning to our matrix interpretation, $f_S(x)\, g_S(y)$ is a (0, 1)-matrix R_S of rank 1, and then, as a matrix, Φ is of rank at most $\sum_{i \leqslant d} \binom{t}{i} \leqslant 2^{polylog(t)}$. L and Φ agree on all but an ε fraction of the entries. Furthermore, by Theorem 4.3, the entries of Φ are all nonnegative integers and > 1 if $L(x, y) \neq \Phi(x, y)$. Let us now define a matrix B_n:

$$B_n := J_n - 2\Phi = J_n - 2 \cdot \sum_{S \subseteq [t],\, |S| \leqslant d} \alpha_S R_S.$$

Clearly,

$$\begin{aligned}\operatorname{rank}(B_n) &\leqslant 1 + \operatorname{rank}(\Phi) \\ &\leqslant 2^{polylog(t)} \\ &\leqslant 2^{polylog(m)},\end{aligned}$$

thus proving (i). Entries of B_n are bounded in absolute value by $2^{polylog(m)}$ and hence (ii) is true. Moreover, B_n differs from A_n in at most a $2^{-polylog(m)}$ fraction of entries.

Thus (iii) follows. (In fact, since Φ is at least 2 on the error points, B_n can only switch the signs of $+1$'s in A_n.) ∎

5. SOME STRUCTURAL RESULTS IN COMMUNICATION COMPLEXITY

In this section, we mention some results from structural complexity, analogs of which remain valid in the communication complexity model. Their proofs involve essentially no new ideas and follow by adaptation of the techniques used in the Turing machine model. We also point out a simple connection between a customary circuit complexity class and a communication complexity class, namely that $\mathrm{NC} \subseteq \mathrm{PSPACE}^{cc}$. It is worth mentioning that certain simple functions in AC^0, such as equality testing, do not belong to NP^{cc}.

First, we observe that Toda's theorem ("PP is as hard as the Polynomial-time Hierarchy") continues to hold in the communication complexity world as well. This can be proved by essentially translating Toda's proof [To91] (cf. [BF91]).

THEOREM 5.1 (Toda's thereom in communication complexity). $\mathrm{PH}^{cc} \subseteq \mathrm{P(PP)}^{cc}$.

One can also naturally consider the notion of interactive proof systems [BaMo88, GMR89] in the communication complexity model. An *interactive proof system in communication complexity* consists of an infinitely powerful, omniscient prover P and the two players Alice and Bob, Alice holding x and Bob holding y, $|x| = |y| = m$. The power knows both x and y and tries to convince Alice and Bob that (x, y) is in L (i.e., $L(x, y) = 1$). Alice and Bob can query the prover to verify the "proof" given by the prover. To form the query string as well as to process the response from the prover, Alice and Bob are allowed to execute a *randomized protocol*. Furthermore, the coin tosses of Alice and Bob are visible to the prover (*public coins model*) and the response from the prover is visible to both Alice and Bob. Thus we can think of the entire communication taking place on a backboard visible to everybody. In an interactive proof system we require that the total number of bits ever written on the blackboard must be bounded by polylog(m). A typical *round* in the protocol consists of

- Alice and Bob execute a randomized protocol of length polylog(m) to agree on a query string and present it to the prover.
- The Prover responds with an answer string.

DEFINITION 5.1. A language $L \subseteq \Sigma^{2*}$ is in IP^{cc} if for all $(x, y) \in \{0, 1\}^n \times \{0, 1\}^n$,

$$(x, y) \in L \Rightarrow \exists \mathrm{P} \;:\; \Pr[\text{Alice and Bob Acept P's proof}] \geqslant 2/3, \text{ and}$$

$$(x, y) \notin L \Rightarrow \forall \mathrm{P} \;:\; \Pr[\text{Alice and Bob Acept P's proof}] \leqslant 1/3.$$

Here the probability is taken over the coin tosses of Alice and Bob.

By adapting the techniques of [LFKN92, Sha92, She92], it is easy to prove that $\mathrm{PSPACE}^{cc} \subseteq \mathrm{IP}^{cc}$. To prove the other direction, that $\mathrm{IP}^{cc} \subseteq \mathrm{PSPACE}^{cc}$, we note that the standard argument that evaluates the game tree between the prover and the verifier in polynomial space does not directly apply in our context since the notion of space-bounded computation does not make sense in the communication complexity model. However, Lautemann's theorem that $\mathrm{BPP} \subseteq \Sigma_2 \cap \Pi_2$ holds in the communication complexity model, as oberved in [BFS86]. Now replacing the randomized moves of Alice and Bob in an interactive protocol using Lautemann's result, we get a sequence of $\exists$ and $\forall$ moves followed by a refereeing protocol. From the definition of $\mathrm{PSPACE}^{\mathrm{cc}}$ (Definitions 4.1 and 4.3), we then conclude that $\mathrm{IP}^{\mathrm{cc}} \subseteq \mathrm{PSPACE}^{\mathrm{cc}}$. Therefore, we have the following analog of a well-known result [LFKN92, Sha92] in turing machine complexity:

THEOREM 5.2 (IP = PSPACE in communication complexity). $\mathrm{IP}^{\mathrm{cc}} = \mathrm{PSPACE}^{\mathrm{cc}}$. ∎

In Section 6, we will refer to bounded round interactive proof systems in communication complexity. The collapse theorem from [Ba85] (cf. [BaMo88]) shows that a constant number of moves can be replaced by just two moves: a randomized (Arthur's) move followed by an existential one (Merlin's move)—the complexity class given by these two moves is denoted by AM. By a straightforward translation of this result into communication complexity, we will denote the class of languages accepted by bounded round interactive proof systems in communication complexity by $\mathrm{AM}^{\mathrm{cc}}$.

We also point out a connection between parallel complexity and communication complexity. For this purpose w.l.o.g., let us consider languages consisting of even-length strings only and treat a language L as a sequence of Boolean functions $\{f_{2m}\}_{m>0}$. By arbitrarily partitioning the variables into two equal pieces, $x, y \in \{0, 1\}^m$, we can naturally talk about the communication complexity of $\{f_{2m}\}$ when we give x to Alice and y to Bob.

PROPOSITION 5.3. $\mathrm{NC} \subseteq \mathrm{PSPACE}^{\mathrm{cc}}$.

Proof. Let L be described by $\{f_{2m}\}$. Then there is a circuit C_m of depth polylog(m) and size $m^{O(1)}$ (in fact, size $\leqslant 2^{polylog(m)}$ suffices) computing f_m. W.l.o.g. we assume C_m consists of AND–OR gates (of fan-in 2) only with literals (variables and their negations) appearing at its input nodes. Let $x, y \in \{0, 1\}^m$ denote the halves given to Alice and Bob. We describe a $\mathrm{PSPACE}^{\mathrm{cc}}$ protocol for $f_m(x, y)$. The $\exists$ player picks the OR gates and the $\forall$ player picks the AND gates of C_m. Then it is easy to write a predicate

$$\exists u_1 \forall u_2 \ldots Q_k u_k (\varphi(x, u) \diamond \psi(y, u))$$

with $k \leqslant$ polylog(m), which is true iff the circuit C_m evaluates to 1. Here u_i specifies a wire feeding into a gate and the sequence $u_1 \ldots u_k$ defines a path from the output to a bottom gate of C_m. The functions $\varphi(x, u)$ and $\psi(y, u)$ specify the input literals to the bottom gate. Each of them is a simple function of at most two literals from x and y, respectively. From Definitions 4.1 and 4.3, it is easy to see that this is a $\mathrm{PSPACE}^{\mathrm{cc}}$ protocol for f_{2m}. ∎

When Alice and Bob are computationally limited to be in NC, the functions $\varphi(x, u)$ and $\psi(y, u)$ are computable by NC circuits for any fixed u. Then it is straightforward to convert a PSPACE^{cc} protocol into an NC circuit. This implies that when the power of Alice and Bob is restricted to NC, the power of the model IP^{cc} reduces to NC.

6. AC^0-DIMENSION

In this section, we introduce a complexity measure, called AC^0-dimension, of a *set of Boolean functions*. One motivation to consider AC^0-dimension is to obtain sufficient conditions for separating PH^{cc} from PSPACE^{cc} involving notions other than matrix rigidity. Since we do not have strong lower bounds on matrix rigidity, one might try attacking the lower bound questions on AC^0-dimension using combinatorial tools different from the ones used in matrix rigidity, such as the switching lemma [Ha86]. A further advantage of using AC^0-dimension is our ability to formulate interesting combinatorial questions potentially simpler than matrix rigidity that would still have consequences to separation questions in communication complexity. We illustrate this approach in Lemma 6.3 below.

In defining AC^0-dimension, we consider circuits whose inputs are *arbitrary* Boolean functions (rather than literals as is usual). First, let us consider a fixed input size m. Let $\mathscr{F} := \{f_1, ..., f_K\}$ be a set of m-variable Boolean functions. We want the smallest set of arbitrary m-variable Boolean functions $\mathscr{G} := \{g_1, ..., g_D\}$ such that *each* f_i, $1 \leqslant i \leqslant K$, can be computed as the output of a circuit C_i where the inputs to C_i are selected from the set $\{g_1, ..., g_D\}$, and C_i has small complexity. Since the g_i's can be the f_i's themselves, trivially $D \leqslant K$. We say $\mathscr{G}$ *generates* $\mathscr{F}$ *in size* s *and depth* d if each circuit C_i is of size at most s and depth at most d (and unbounded fan-in).

The notion of AC^0-dimension is actually defined for an infinite sequence of *sets* of functions, one for each input length m. Just as we informally use the term "complexity of *a* function f" when we really mean the complexity (as a function of m) of an infinite sequence of functions $\{f_m\}$, we may also informally use the term "AC^0-dimension of a set of functions" to really refer to the dimension (as a function of m) of an infinite sequence of sets of functions. Hence, we often use the symbol $\mathscr{F}$ to refer to the infinite sequence $\{\mathscr{F}_m\}$, where $\mathscr{F}_m := \{f_1^m, ..., f_{K(m)}^m\}$ is a set of m-varaible Boolean functions. As an example, consider the set $\mathscr{P}$ of parity functions: $\mathscr{P}$ defines the infinite sequence $\{\mathscr{P}_m\}$, where $\mathscr{P}_m$ is the set of all 2^m parity functions of m variables, namely $\mathscr{P}_m = \{\bigoplus_{i \in S} x_i \mid S \subseteq [m]\}$.

DEFINITION 6.1. Let $\mathscr{F} := \{\mathscr{F}_m\}$ and $\mathscr{G} := \{\mathscr{G}\}$ be (infinite sequences of) sets of functions, where $\mathscr{F}_m := \{f_1^m, ..., f_{K(m)}^m\}$ and $\mathscr{G}_m := \{g_1^m, ..., g_{D(m)}^m\}$. We say $\mathscr{G}$ generates $\mathscr{F}$ via AC^0-combinations if there exist constants $c, d > 0$ such that for all m, $\mathscr{G}_m$ generates $\mathscr{F}_m$ in size $2^{(\log m)^c}$ and depth d. In other words,

$$f_i^m \equiv C_i^m(g_i^m, ..., g_{D(m)}^m), \qquad 1 \leqslant i \leqslant K(m),$$

where C_i^m is a circuit of size $2^{(\log m)^c}$ and depth d.

DEFINITION 6.2. We say $\mathscr{F} \in \mathrm{AC}^0\text{-DIM}[D(m)]$ if there exists a set $\mathscr{G} = \{\mathscr{G}_m\}$, with $\mathscr{G}_m$ of size $D(m)$, that generates $\mathscr{F}$ via AC^0-combinations.

DEFINITION 6.3.

$$\mathrm{AC}^0\text{-DIM}[qP] = \bigcup_{c \geqslant 0} \mathrm{AC}^0\text{-DIM}[2^{(\log m)^c}],$$

where qP is intended for quasipolynomial.

Notation 1. Given an infinite sequence of matrices $A = \{A_m\}$, where $A_m \in \{+1, -1\}^{2^m \times 2^m}$, we will use the corresponding symbol $\mathscr{A}$ to denote the infinite sequence of sets of functions $\{\mathscr{A}_m\}$, where $\mathscr{A}_m$ is the set of Boolean functions corresponding to the rows of A_m.

The proof of Theorem 4.1 shows that high rigidity implies large AC^0-dimension:

LEMMA 6.1. *Let A and $\mathscr{A}$ be given by Notation* 1. *If $\mathscr{A} \in \mathrm{AC}^0\text{-DIM}[qP]$ then, for every $c > 0$ there exist constants $c_1, c_2 > 0$ such that $\mathscr{R}_A(2^{(\log m)^{c_1}}, 2^{(\log m)^{c_2}}) \leqslant 2^{2m}/2^{(\log m)^c}$.*

On the other hand, from Definitions 4.1 and 4.2 it follows that

LEMMA 6.2. *Let $L \subseteq \Sigma^{2*}$ and let A_m be its $2^m \times 2^m$ ± 1-matrix. If $L \in \mathrm{PH}^{\mathrm{cc}}$, then $\mathscr{A} \in \mathrm{AC}^0\text{-DIM}[qP]$, where $\mathscr{A}$ is defined from A as in Notation* 1.

An interesting special case occurs when the circuit C_i^m in Definition 6.1 is a simple OR gate. In this case, we will speak of OR dimension and use the notation OR-DIM.

DEFINITION 6.4. Let $A = \{A_m\}$ and $B = \{B_m\}$ be infinite sequences of ± 1-matrices, where A_m and B_m are of size $2^m \times 2^m$. Let $\varepsilon = \varepsilon(m) > 0$. Then B is said to be ε-close to A if for all m, B_m and A_m differ on at most an ε-fraction of their entries (they can only differ in signs). Now define

$$\tilde{A}_\varepsilon := \{B\colon B \text{ is } \varepsilon\text{-close to } A\}.$$

Finally, define

$$\widetilde{\mathscr{A}}_\varepsilon := \{\mathscr{B}\colon B \in \tilde{A}_\varepsilon\},$$

where, as per Notation 1, $\mathscr{B}$ denotes the infinite sequence of sets of functions given by B. In this case, we will also say $\mathscr{B}$ is ε-close to $\mathscr{A}$.

We have the following lemma:

LEMMA 6.3. *Let $L \subseteq \Sigma^{2*}$ and let A_m be its $2^m \times 2^m$ ± 1-matrix. For all $\varepsilon := \varepsilon(m) \leqslant 2^{\mathrm{polylog}(m)}$, if $\widetilde{\mathscr{A}}_\varepsilon \cap \mathrm{OR\text{-}DIM}[qP] = \varnothing$, i.e., for any $\mathscr{B}$ that is ε-close to $\mathscr{A}$, it holds that $\mathscr{B} \notin \mathrm{OR\text{-}DIM}[qP]$, then $L \notin \mathrm{AM}^{\mathrm{cc}}$.*

In particular, if $L \in \mathrm{IP}^{\mathrm{cc}}$, this will show that $\mathrm{AM}^{\mathrm{cc}} \neq \mathrm{IP}^{\mathrm{cc}}$.

7. OPEN PROBLEMS

The single major open question in this area is to prove better lower bounds on the rigidity of an explicit matrix. However, we will state below two simpler open questions whose solutions would improve existing lower bounds.

Question 1. Give an explicit infinite family of matrices $\{A_n\}$ such that $w_2(A_n) = \Omega(n^{1+\varepsilon})$ for a constant $\varepsilon > 0$, where $w_2(A_n)$ is defined by $w_2(A) := \min\{wt(B) + wt(C) : A = BC\}$.

Generalized Hadamard matrices (see Definition 2.2) seem to be a good class of candidates for A_n in this question. A lower bound of $\mathscr{R}_A(\varepsilon n) = n^{1+\delta}$, for some constants $\varepsilon, \delta > 0$, would imply a lower bound of $\Omega(n^{1+\delta/2})$ on $w_2(A)$.

From Lemma 3.1, $w_2(A)$ is essentially the depth-2 complexity of ℓ_A by linear circuits (with unrestricted coefficients). Currently, the best known lower bound on $w_2(A)$ for an explicit matrix is $\Omega(n \log^2 n / \log \log n)$ [PRS97, RT97].

The following open question seems interesting in the context of OR dimension:

Let A_n be an $n \times n$ (0, 1)-matrix and let $\tilde{A}_{n,\varepsilon}$ be an $n \times n$ (0, 1)-matrix differing from A_n on at most an ε-fraction of entries. Let $\mathscr{F}_{n,\varepsilon}$ be the set system defined by $\tilde{A}_{n,\varepsilon}$, i.e., the rows of this matrix define the characteristic vectors of n subsets of a universe of size n.

Question 2. Find an explicit infinite family $\{A_n\}$ of (0, 1)-matrices such that the following holds: for any constant $c > 0$ and any set system $\mathscr{F}_{n,\varepsilon}$ obtained as above from A_n and $\varepsilon \leqslant 2^{(\log \log n)^c}$, every family $\mathscr{G} = \{G_1, ..., G_{d(n)}\}$, $G_i \subseteq [n]$, *that generates* $\mathscr{F}_{n,\varepsilon}$ *via unions*, i.e., for every $F \in \mathscr{F}_{n,\varepsilon}$, $F = \bigcup_{i \in S_F} G_i$ for some $S_F \subseteq [d(n)]$ must have size $d(n) = 2^{(\log \log n)^{\omega(1)}}$.

From Lemma 6.1, a lower bound of $\mathscr{R}_A(r) \geqslant n^2 / 2^{(\log \log n)^{o(1)}}$ for $r \geqslant 2^{(\log \log n)^{\omega(1)}}$ would answer Question 2. By Lemma 6.3, a solution to Question 2 would give an explicit language outside bounded round interactive proof systems, $\mathrm{AM}^{\mathrm{cc}}$, in communication complexity. In particular, if H_n is the $n \times n$ Sylvester matrix, where $n := 2^m$, and $A_n := (J_n - H_n)/2$, then a solution to Question 2 for $\{A_n\}$ would separate bounded round interactive proof systems in communication complexity ($\mathrm{AM}^{\mathrm{cc}}$) from their unbounded round counterparts ($\mathrm{IP}^{\mathrm{cc}}$).

ACKNOWLEDGMENTS

I am indebted to Laci Babai and Kati Friedl for numerous insightful discussions throughout this work. I thank Lance Fortnow, Ketan Mulmuley, Janos Simon, Pradyut Shah, and Dieter van Melkebeek for several helpful comments on earlier versions. I am grateful to Noga Alon, Noam Nisan, Pavel Pudlák, and Sasha Razborov for sending me their manuscripts. I thank two anonymous referees for several suggestions that improved the presentation of the paper.

REFERENCES

[Al94] N. Alon, On the rigidity of an Hadamard matrix, manuscript, 1994.

[AFR85] N. Alon, P. Frankl, and V. Rödl, Geometric realizations of set systems and probabilistic communication complexity, *in* "26th IEEE FOCS (1985)," pp. 277–280.

[Ba85] L. Babai, Trading group theory for randomness, *in* "17th ACM STOC (1985)," pp. 421–429.

[BaMo88] L. Babai and S. Moran, Arthur-Merlin games: A randomized proof system, and a hierarchy of complexity classes, *J. Comput. System Sci.* **36** (1988), 254–276.

[BCS97] P. Bürgisser, M. Clausen, and M. A. Shokrollahi, Algebraic complexity theory, *in* "Grundlehren der mathematischen Wissenschaften," Vol. 315, Springer-Verlag, New York, 1997.

[BF91] L. Babai and L. Fortnow, Arithmetization: A new method in structural complexity theory, *Comput. Complexity* **1** (1991), 41–66.

[BFS86] L. Babai, P. Frankl, and J. Simon, Complexity classes in communication complexity theory, *in* "27th IEEE FOCS (1986)," pp. 337–347.

[BoMu75] A. Borodin and I. Munro, "The Computational Complexity of Algebraic and Numeric Problems," American Elsevier, New York, 1975.

[BRS93] A. Borodin, A. Razborov, and R. Smolensky, On lower bounds for read-k-times branching programs, *Comput. Complexity* **3** (1993), 1–18.

[BT91] R. Beigel and T. Tarui, On ACC, *in* "32nd IEEE FOCS," 1991, pp. 783–792.

[Ch94] B. Chazelle, A spectral approch to lower bounds, *in* "35th IEEE FOCS (1994)," pp. 674–682.

[Fr93] J. Friedman, A note on matrix rigidity, *Combinatorica* **13** (1993), 235–239.

[GKT92] F. Green, J. Köbler, and J. Torán, The power of the middle bit, *in* "Proc. 7th IEEE Structure in Complexity Theory, 1992," pp. 111–117.

[GMR89] S. Goldwasser, S. Micali, and C. Rackoff, The knowledge complexity of interactive proof systems, *SIAM J. Comput.* **18** (1989), 186–208.

[GV83] G. H. Golub and C. F. Van Loan, "Matrix Computations," Johns Hopkins Univ. Press, 1983.

[Ha86] J. Håstad, "Computational Limitations of Small Depth Circuits," ACM Doctoral Dissertation Awards, M.I.T. Press, Cambridge, MA, 1986.

[HR90] B. Halstenberg and R. Reischuk, Relations between communication complexity classes, *J. Comput. System Sci.* **41** (1990), 402–429.

[HW53] A. J. Hoffman and H. W. Wielandt, The variation of the spectrum of a normal matrix, *Duke Math. J.* **20** (1953), 37–39.

[KR97] B. S. Kashin and A. A. Razborov, Improved lower bounds on the rigidity of Hadamard matrices, manuscript, Dec. 1997.

[KW91] M. Krause and S. Waack, Variation ranks of communication matrices and lower bounds for depth-two circuits having symmetric gates with unbounded fan-in, *in* "Proc. 32nd IEEE FOCS, 1991," pp. 777–782.

[KS92] P. Kimmel and A. Settle, "Reducing the Rank of Lower Triangular All-Ones Matrix," Univ. of Chicago Tech. Report CS 92–21, Nov. 1992.

[LFKN92] C. Lund, L. Fortnow, H. Karloff, and N. Nisan, Algebraic methods in interactive proof systems, *J. Assoc. Comput. Mach.* **39** (1992), 859–868.

[Lok94] S. V. Lokam, Matrix rigidity and communication complexity, manuscript, Nov. 28, 1994.

[Lok95] S. V. Lokam, Spectral methods for matrix rigidity with applications to size-depth tradeoffs and communication complexity, *in* "Proc. 36th IEEE symp. Foundations of Comp. Sci. (FOCS), 1995," pp. 6–15.

[Lok00] S. V. Lokam, On the rigidity of Vandermonde matrices, *Theoret. Comput. Sci.* **237** (2000), 477–483. Preliminary version presented at the DIMACS-DIMATIA workshop on Arithmetic Circuits and Algebraic Methods, June 2–4, 1999.

[Mo73] J. Morgenstern, Note on a lower bound of the linear complexity of the fast Fourier transform, *J. Assoc. Comput. Mach.* **20** (1973), 305–306.

[NW95] N. Nisan and A. Wigderson, On the complexity of bilinear forms, *in* "ACM STOC 1995," pp. 723–732.

[Pu94] P. Pudlák, Large communication in constant depth circuits, *Combinatorica* **14** (1994), 203–216.

[Pu98] P. Pudlák, A note on the use of determinant for proving lower bounds on the size of linear circuits, *Electron. Colloquium Comput. Complexity (ECCC)*, Report 42, July 2, 1998.

[PR94] P. Pudlák and V. Rödl, Some combinatorial-algebraic problems from complexity theory, *Discrete Math.* **136** (1994), 253–279.

[PRS97] P. Pudlák, V. Rödl, and J. Sgall, Boolean circuits, tensor ranks and communication complexity, *SIAM J. Comput.* **26** (1997), 605–633.

[PS86] R. Paturi and J. Simon, Probabilistic communication complexity, *J. Comput. System Sci.* **33** (1986), 106–123.

[PV91] P. Pudlák and Z. Vavřín, Computation of rigidity of order n^2/r for one simple matrix, *Comment. Math. Univ. Carolinae* **32** (1991), 213–218.

[Ra89] A. A. Razborov, On rigid matrices, manuscript, 1989. [In Russian]

[Ra98] A. A. Razborov, Comment published in *Electron. Colloquium Comput. Complexity (ECCC)*, Report 42, July 1998 (cf. [Pu98]).

[RT97] J. Radhakrishnan and A. Ta-Shma, Bounds on Depth-2 superconcentrators, *in* "Proc. 38th IEEE Symp. Foundations of Computer Science (FOCS), 1997," pp. 585–594.

[She92] A. Shen, IP = PSPACE, a simplified proof, *J. Assoc. Comput. Mach.* **39** (1992), 878–880.

[Sha92] A. Shamir, IP = PSPACE, *J. Assoc. Comput. Mach.* **39** (1992), 869–877.

[SS91] V. Shoup and R. Smolensky, Lower bounds for polynomial evaluation and interpolation, *in* "Proc. IEEE FOCS (1991)," pp. 378–383.

[SSS97] M. A. Shokrollahi, D. A. Spielman, and V. Stemann, A remark on matrix rigidity, *Inform. Process. Lett.* **64** (1997), 283–285.

[Ta93] J. Tarui, Randomized polynomials, threshold circuits and polynomial hierarchy, *Theoret. Comput. Sci.* **113** (1993), 167–183.

[To91] S. Toda, PP is as hard as the polynomial hierarchy, *SIAM J. Comp.* **20** (1991), 865–877.

[Va77] L. Valiant, Graph-theoretic arguments in low-level complexity, *in* "Proc. 6th Math. Foundations of Comp. Sci.," Lecture Notes in computer Science, Vol. 53, pp. 162–176, Springer-Verlag, New York, 1977.

[Ya79] A. C.-C. Yao, Some complexity questions related to distributive computing, *in* "Proc. ACM STOC 1979," pp. 209–213.

[Ya90] A. C.-C. Yao, On ACC and threshold circuits, *in* "Proc. of 31st IEEE FOCS, 1990," pp. 619–627.