On a special case of rigidity

Rocco A. Servedio* Emanuele Viola†

November 6, 2012

Abstract

We highlight the special case of Valiant's rigidity problem in which the low-rank matrices are truth-tables of sparse polynomials. We show that progress on this special case entails that Inner Product is not computable by small AC^0 circuits with one layer of parity gates close to the inputs. We then prove that the sign of any $-1/1$ polynomial with $\leq s$ monomials in $2n$ variables disagrees with Inner Product in $\geq \Omega(1/s)$ fraction of inputs, a type of result that seems unknown in the rigidity setting.

Valiant's rigidity problem [Val77] asks to build explicit matrixes that are far in Hamming distance from low-rank matrixes. Valiant proved that if an $N \times N$ matrix M has hamming distance $\geq N^{1+\Omega(1)}$ from any matrix of rank $R = (1-\Omega(1))N$, then the corresponding linear transformation $x \mapsto Mx$ requires circuits of superlogarithmic depth or superlinear size. Exhibiting an explicit such matrix remains a long-standing challenge. Despite significant efforts, the best lower bounds are of the form $(N^2/R)\lg(N/R)$ against matrixes of rank R. The matrix corresponding to the inner product function IP has been conjectured to satisfy better better bounds. We refer the reader to Lokam's survey [Lok09] for more on rigidity.

In this note we highlight a special case of the rigidity problem, and we suggest that attacks should be directed towards it. Recall that an $N \times N$ matrix has rank R if and only if it is the sum of R rank-1 matrixes, i.e., matrixes $u_i v_i^T$, where u_i, v_i are N-entry column vectors. We consider the special case of this problem where the rank-1 matrixes are the truth-tables of *monomials* over the variables $x_1, \ldots, x_n, y_1, \ldots, y_n$, where $N = 2^n$ and the variables range over $\{-1, 1\}$. For example, the truth-table of a monomial $c \prod_{i \in S} x_i \prod_{i \in T} y_i$, where $S, T \subseteq \{1, \ldots, n\}$, is the $N \times N$ matrix whose entry indexed by $(a, b) \in \{-1,1\}^n \times \{-1,1\}^n$ is $c \prod_{i \in S} a_i \prod_{i \in T} b_i$. This matrix can be written as uv^T where the a-th entry of u is $c \prod_{i \in S} a_i$ and the b-th entry of v is $\prod_{i \in T} b_i$. This special case of the rigidity problem is stated without direct reference to rank as follows.

Challenge 0.1 (Sparsity)**.** Exhibit an explicit function $f : \{-1,1\}^n \times \{-1,1\}^n \to \{-1,1\}$ such that for any real polynomial p with $\leq R$ monomials we have

$$\Pr_{x,y \in \{-1,1\}^n}[f(x,y) \neq p(x,y)] \geq \epsilon,$$

*Supported by NSF grant CCF-1115703. Email: `rocco@cs.columbia.edu`

†Supported by NSF grant CCF-0845003. Email: `viola@ccs.neu.edu`

for as large ϵ as possible.

Again, $\epsilon = \Omega(\lg(2^n/R)/R)$ follows from the rigidity bounds.

The concurrent work [RV12] raises a similar challenge for low-degree (as opposed to sparse) polynomials.

Motivation: AC^0 **with parity gates.** Besides hopefully paving the way for the original rigidity question, a motivation for making progress on Challenge 0.1 is that stronger bounds would yield new circuit lower bounds. Let AC^0-$\oplus$ denote the class of AC^0 circuits augmented with a bottom level (right before the input bits) of parity gates. To our knowledge, it is not known whether the Inner Product function IP is computable by poly-size AC^0-$\oplus$ circuits:

Challenge 0.2. Show that IP cannot be computed by poly-size AC^0-$\oplus$ circuits.

Challenge 0.2 seems open even for AC^0-$\oplus$ circuits of depth 4, but it is known to be true for AC^0-$\oplus$ circuits of depth 3, i.e. poly-size DNF-$\oplus$ circuits. Indeed, it follows from Fact 8 in [Jac97] that any function computable by such circuits has 1/poly correlation with parity on some subset of the variables, but it is well-known that IP has exponentially small correlation with parity on any subset of the variables.

Solving Challenge 0.2 is a step towards a more thorough understanding of AC^0 with parity gates. For example, no strong correlation bound is known for this class, see e.g. [SV10]. In fact, this is not even known for AC^0-$\oplus$, and IP is a natural candidate.

Next we formally connect the two challenges.

Claim 0.3. Suppose that IP on $2n$ variables has AC^0-$\oplus$ circuits of polynomial size. Then for any b there exists c and a polynomial $p(x, y)$ with $\leq 2^{\lg^c n}$ monomials such that

$$\Pr_{x,y}[p(x, y) \neq \mathrm{IP}(x, y)] \leq 2^{\lg^b n}.$$

Proof. Let C be a depth-$(d+1)$ AC^0-$\oplus$ circuit that computes IP over $2n$ input bits $x_1, \ldots, x_n$, $y_1, \ldots, y_n$. Let $N = \mathrm{poly}(n)$ denote the number of parity gates at the leaves. Let C' be the depth-d AC^0 circuit obtained by replacing the i-th parity gate by a fresh input variable z_i (so C' is a circuit over N input bits $z_1, ..., z_N$).

Let D be the distribution over $\{-1, 1\}^N$ induced by drawing a uniform random input x from $\{-1, 1\}^n$ and setting $z_i =$ the value of the i-th parity gate on x (the draw from D is the string $z \in \{0, 1\}^N$). Let $\epsilon := 1/2^{\lg^c n}$. Lemma 5.1 and Corollary 5.2 of [ABFR94] tell us that there is a polynomial $p(z_1, \ldots, z_N)$ of degree $(O(\lg(n))^{2d}$ that computes $C'(z)$ for a $(1 - \epsilon)$ fraction of all inputs drawn from D. Since p has degree $(O(\lg n))^{2d}$ it must have $\leq n^{(O(\lg n))^{2d}}$ monomials. Now let $q(x_1, ..., x_n, y_1, \ldots, y_n)$ be the polynomial obtained by substituting in the i-th parity (monomial) for z_i in p. q has no more monomials than p, and q computes IP on $(1 - \epsilon)$ fraction of all inputs drawn from $\{-1, 1\}^n$. $\square$

We note that for Valiant's connection to lower bounds, we need rank $R = \Omega(N)$, whereas for sparsity much smaller rank $R = \mathrm{poly} \lg N$ suffices. In both cases we need to go beyond error $1/R$.

Sign-rank. The *sign-rank* of a $-1, 1$ matrix M is the minimum rank of a matrix that agrees in sign with M in every entry. Forster proved [For02] that the $N \times N$ matrix corresponding to IP has sign-rank $\geq \sqrt{N}$.

For sparsity, we can prove a stronger type of bound where we also allow errors. As far as we know such a result is not known for sign-rank. Perhaps this gives hope that progress on Challenge 0.1 may be within reach.

Theorem 0.4. Let p be a polynomial in n variables with $\leq s$ monomials. Consider the inner-product function $\mathrm{IP}(x, y)$ where $|x| = |y| = n/2$. Then

$$\Pr_{x,y}[\mathrm{sign}(p(x, y)) \neq \mathrm{IP}(x, y)] \geq (1 - s/2^{n/2}) \cdot (1/s) = \Omega(1/s).$$

The proof of Theorem 0.4 relies on the following lemma.

Lemma 0.5. Let p be a $-1/1$ polynomial on n variables with $\leq s$ not monomials and not containing the monomial (parity) $t(x)$. Then $\mathrm{sign}(p(x))$ disagrees with $t(x)$ on at least $2^n/s$ points.

Proof of Theorem 0.4 assuming Lemma 0.5. Let p be a polynomial with $\leq s$ monomials over variables x, y where $|x| = |y| = n/2$. A uniform random choice of y reduces IP to parity over a uniform random subset of variables $x_1, \ldots, x_{n/2}$. But fixing y does not change the set of monomials of p in x (it merely changes the sign of the coefficients). So with probability $\geq 1 - s/2^{n/2}$ a uniform random choice of y reduces to the setting of Lemma 0.5, in which p is reduced to a polynomial with $\leq s$ monomials over $n/2$ x-variables and IP is reduced to a parity over x-variables not contained in p. Hence the overall error probability over a random choice of both x and y is $\geq (1 - s/2^{n/2}) \cdot (1/s)$. $\square$

Before proving Lemma 0.5 in the next section we remark that it is essentially tight: for $s = 2^k - 1$, there is a polynomial p of sparsity s that does not contain the monomial t but computes t exactly on all but $2^n/(s + 1)$ inputs. We show next a construction for $t = 1$, i.e. the parity on 0 variables, so p is not allowed to have a constant term. (Given such a construction p then $p \cdot t$ is a construction for any monomial t.)

For sparsity $s = 1$ we take $p = x_1$ and the error is $1/2$ (p is wrong exactly when $x_1 = -1$); for sparsity $s = 3$ we take $p = x_1 + x_2 \cdot (1 - x_1)$ and the error is $1/4$ (p is wrong exactly when $x_1 = -1, x_2 = -1$); for sparsity $s = 7$ we take $p = x_1 + x_2(1 - x_1) + x_3(1 - x_1)(1 - x_2)$ and the error is $1/8$ (p is wrong exactly when $x_1 = -1, x_2 = -1, x_3 = -1$); and so on.

0.1 Proof of Lemma 0.5

First, our polynomials are multi-linear without loss of generality. Recall that such a polynomial p in n variables is syntactically zero if and only if $p(x) = 0$ for every $x \in \{-1, 1\}^n$.[Sch80, Zip79] The proof is by contradiction, so we suppose that the conclusion does not hold, i.e. $\mathrm{sign}(p(x))$ disagrees with $t(x)$ on fewer than $2^n/s$ points. ($p(x) = 0$ counts as a disagreement; alternatively, we can assume that $p(x) \neq 0$ for every x without loss of generality.) We show below how to construct a non-zero polynomial g such that $g(x) = 0$ on the few ($< 2^n/s$) disagreement points, and moreover the monomials of $p \cdot g^2$ still do not contain $t(x)$. Given such a g we observe that the polynomial $p \cdot g^2$ is non-zero and always agrees in sign with t, but on the other hand $E[p \cdot g^2 \cdot t] = 0$. This is a contradiction.

The construction of g. We identify monomials with elements of $\{0,1\}^n$ in the obvious way. Note that product of monomials corresponds to bit-wise addition mod 2. Let B be the set of monomials of p, so $s = |B|$. Let t be a monomial not present in B. We construct a set M of size $|M| \geq 2^n/|B|$ such that $t \notin M + M + B$, where $S + T := \{s + t : s \in S, t \in T\}$.

Then we define g to be a polynomial with the monomials in M. We set the coefficients of the monomials in M so that $g(x) = 0$ for $|M| - 1$ inputs x, and still have g be a non-zero polynomial. This is possible because we have a homogeneous system of $|M| - 1$ equations in $|M|$ variables.

The condition $t \notin M + M + B$ translates to the condition that $p \cdot g^2$ does not contain the monomial t.

The construction of M. Call a pair (M, G) *good* if for every $g \in G$, $2(M \bigcup g) + B$ does not contain t. For simplicity here and below we write g for the set $\{g\}$.

The next two claims allow us to construct a pair (M, G) that is good and where $|M| \geq 2^n/|B|$, as desired.

Claim 0.6. $(\emptyset, \{0,1\}^n)$ is good.

Proof. In this case $2(M \bigcup g) + B = g + g + B = B$, which does not contain t by assumption. □

Claim 0.7. If (M, G) is good then for any $g \in G, (M \bigcup g, G \setminus (B + t + g))$ is also good.

Proof. Suppose by contradiction that there is $g' \in G \setminus (B+t+g)$ such that $t \in 2(M \bigcup g \bigcup g') + B$.

Recall $t \notin 2(M \bigcup g) + B$, and $t \notin 2(M \bigcup g') + B$, because both g and g' are in G, and (M, G) is good.

Hence $t \in 2(g \bigcup g') + B$.

Recall again that $t \notin B$ by assumption.

Hence $t \in g + g' + B$, but this contradicts the choice of g'. □

We remark that the proof of Lemma 0.5 in this section may be viewed as a generalization of an argument from [ABFR94]. In the latter the polynomial p has degree d, so B's elements are just strings in $\{0,1\}^n$ of weight $\leq d$, and one defines M to be the set of all strings of weight less than $(n - d)/2$. Our proof employs a slightly more involved greedy construction.

References

[ABFR94] James Aspnes, Richard Beigel, Merrick Furst, and Steven Rudich. The expressive power of voting polynomials. *Combinatorica*, 14(2):135–148, 1994.

[For02] Jürgen Forster. A linear lower bound on the unbounded error probabilistic communication complexity. *J. of Computer and System Sciences*, 65(4):612–625, 2002.

[Jac97] Jeffrey C. Jackson. An efficient membership-query alogithm for learning DNF with respect to the uniform distribution. *J. of Computer and System Sciences*, 55(3):414–440, 1997. 35th Symposium on Foundations of Computer Science (Santa Fe, NM, 1994).

[Lok09] Satyanarayana V. Lokam. Complexity lower bounds using linear algebra. *Foundations and Trends in Theoretical Computer Science*, 4(1-2):1–155, 2009.

[RV12] Alexander Razborov and Emanuele Viola. Real advantage. 2012.

[Sch80] Jacob T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. of the ACM*, 27(4):701–717, 1980.

[SV10] Ronen Shaltiel and Emanuele Viola. Hardness amplification proofs require majority. *SIAM J. on Computing*, 39(7):3122–3154, 2010.

[Val77] Leslie G. Valiant. Graph-theoretic arguments in low-level complexity. In *6th Symposium on Mathematical Foundations of Computer Science*, volume 53 of *Lecture Notes in Computer Science*, pages 162–176. Springer, 1977.

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Symbolic and Algebraic Computation (EUROSAM)*, volume 72 of *Lecture Notes in Computer Science*, pages 216–226. Springer, 1979.