Electronic Colloquium on Computational Complexity, Report No. 93 (2016)

Bounded Matrix Rigidity and John's Theorem

Cyrus Rashtchian

Department of Computer Science & Engineering
University of Washington, Seattle
cyrash@cs.washington.edu

June 3, 2016

Abstract

Using John's Theorem, we prove a lower bound on the *bounded rigidity* of a sign matrix, defined as the Hamming distance between this matrix and the set of low-rank, real-valued matrices with entries bounded in absolute value. For Hadamard matrices, our asymptotic leading constant is tighter than known results by a factor of two whenever the approximating matrix has sufficiently small entries.

1 Introduction

Motivated to prove circuit lower bounds, Valiant [Val77] introduced rigidity as a complexity measure on matrices. For a matrix $M \in \mathbb{R}^{n\times n}$ and target rank $r \in [n]$ the *rigidity* $\mathcal{R}_M(r)$ is the minimum number of entries that must be changed in M to reduce its rank to r over $\mathbb{R}$. In other words, M has Hamming distance $\mathcal{R}_M(r)$ from the set of rank r real matrices[1]. Valiant showed $\mathcal{R}_M(r) \geqslant (n-r)^2$ when M is a random real matrix, and he proved that sufficiently rigid matrices correspond to functions requiring either super-linear size or super-logarithmic depth linear arithmetic circuits. A family of matrices with $\mathcal{R}_M(r) \geqslant \Omega(n^{1.01})$ for any $r \leqslant n/100$ would suffice. Alas, finding explicit rigid matrices remains an open question, as does proving that an explicit function is not computable by any $O(\log n)$-depth circuit with $O(n)$ gates. The bound $\mathcal{R}_M(r) \geqslant \Omega(\frac{n^2}{r}\log(n/r))$ by [Fri93, SSS97] is the best we know. It becomes trivial when r is linear in n.

The lack of progress in finding explicit rigid matrices led researchers to investigate relaxations such as the requirement that the low-rank approximation matrix have magnitude-bounded entries. This relaxed measure, called *bounded rigidity*, still has applications to separating communication complexity classes [Lok01] and captures the natural question of whether sign matrices suffice for low-rank approximations or real-valued matrices provide a significant advantage. Krause and Waack [KW91] derive depth-two circuit lower bounds using a very related measure. We focus on bounded rigidity and refer the reader to Lokam's survey [Lok09] for more on general matrix rigidity.

The following definition formalizes bounded rigidity. The next section reviews known bounds.

Definition 1.1. *The* bounded rigidity *of a matrix $M = (M_{ij}) \in \mathbb{R}^{n\times n}$ for target rank $r \in [n]$ and entry bound $\theta \in \mathbb{R}_+$ is defined as*

$$\mathcal{R}_M(r,\theta) = \min_C \{|C| \ : \ \mathrm{rk}(M-C) \leqslant r, \quad \max_{ij} |M_{ij} - C_{ij}| \leqslant \theta\},$$

where $|C|$ denotes the number of nonzero entries of the matrix $C \in \mathbb{R}^{n\times n}$ with entries $C = (C_{ij})$.

[1] Although we focus on real matrices, the definition of rigidity and the circuit implications hold over any field.

ISSN 1433-8092

1.1 Prior Results on Bounded Rigidity

While general rigidity lower bounds remain elusive, previous work on bounded rigidity provides asymptotically optimal lower bounds for popular candidate rigid matrices: Hadamard matrices. Defined as any $H \in \{-1,1\}^n$ satisfying $HH^T = nI$, where I is the identity matrix, Hadamard matrices arise frequently. An example is the Walsh (a.k.a. Sylvester) matrix, defined for n a power of two as $W_{x,y} = (-1)^{\sum x_i y_i}$, where we index W by $x, y \in \{0,1\}^n$, and the sum in the exponent is over the reals. The best general rigidity lower bounds for H stand at $\mathcal{R}_H(r) \geqslant \Omega(n^2/r)$ [KR98].

We summarize prior results on bounded rigidity. Theorem 1.2 is due to De Wolf [dW06] and involves quantum arguments. Theorem 1.3 comes from Kashin and Razborov [KR98] and improves a theorem by Lokam [Lok01]. Pudlák [Pud00] proves a similar result.

Theorem 1.2 ([dW06])**.** *Let $H \in \{-1,1\}^{n\times n}$ satisfy $HH^T = nI$. For any $r \in [n]$ and any $\theta \in \mathbb{R}_+$*

$$\mathcal{R}_H(r,\theta) \geqslant \frac{n^2(n-r)}{2\theta n + r(\theta^2 + 2\theta)}.$$

Theorem 1.3 ([KR98])**.** *Let $H \in \{-1,1\}^{n\times n}$ satisfy $HH^T = nI$. For any $r \in [n]$ and any $\theta \in \mathbb{R}_+$*

$$\mathcal{R}_H(r,\theta) \geqslant \frac{n^2(n-r)}{(\theta+1)(2n + r(\theta-1))}.$$

Observe that since $|C| < n^2$ in Definition 1.1, we can assume $\theta \geqslant 1$ without loss of generality.

Although best and simplest for Hadamard matrices, Theorem 1.3 also extends to give bounds for arbitrary matrices in terms of their full set of singular values. All previous results hold for generalized (complex) Hadamard matrices, defined as any $H \in \mathbb{C}^{n\times n}$ having unit entries and satisfying $HH^* = nI$. The Discrete Fourier Transform matrix is one such example.

1.2 Bounded Rigidity as a Function of the Largest Singular Value

We now describe our main theorem. For a rank r matrix $A \in \mathbb{R}^{n\times n}$ agreeing with $M \in \{-1,1\}^{n\times n}$ in most entries, the proof of our main theorem gives upper and lower bound on the matrix inner product $\sum_{u,v} A_{u,v} M_{u,v}$. The upper bound utilizes a consequence of John's Theorem to control vector magnitudes, and the lower bound follows directly from the bounded rigidity definition.

Theorem 1.4. *For any $M \in \{-1,1\}^{n\times n}$ and for any $r \in [n]$ and $\theta \leqslant \frac{n}{\sigma_1(M)\sqrt{r}} - 1$*

$$\mathcal{R}_M(r,\theta) \geqslant \frac{n^2}{\theta+1} - \frac{\theta n\sqrt{r}\cdot\sigma_1(M)}{\theta+1},$$

where $\sigma_1(M) = \max_{x,y\in R^n} \dfrac{\langle x, My\rangle}{\|x\|_2\cdot\|y\|_2}$ denotes the largest singular value of M.

For a Hadamard matrix H, we get the following corollary, since $\sigma_1(H) = \sqrt{n}$.

Corollary 1.5. *Let $H \in \{-1,1\}^{n\times n}$ satisfy $HH^T = nI$. For any $r \in [n]$ and $\theta \leqslant \left(\frac{n}{r}\right)^{1/2} - 1$*

$$\mathcal{R}_H(r,\theta) \geqslant \frac{n^2}{\theta+1} - \frac{\theta n^{3/2}\sqrt{r}}{\theta+1}.$$

Previous theorems already provide asymptotically tight bounds for Hadamard matrices when $\theta = O(1)$. This follows from the simple fact that $\mathcal{R}_M(r, 1) \leqslant n(n-r)$ for any M with unit entries. Our bound is stronger in the regime of $1 < \theta \ll \left(\frac{n}{r}\right)^{1/2}$ because prior bounds scale with n like $\frac{n^2}{(2+o(1))\theta}$, whereas our bound scales like $n^2/(\theta+1)$. Finally, notice that whenever $\theta = \omega(n)$, the general rigidity bound of $\mathcal{R}_H(r) \geqslant \Omega(n^2/r)$ trumps the bounded rigidity results.

Remark 1. Kashin and Razborov [KR98] and Lokam [Lok01] state their results for Hadamard matrices, but their arguments, which both use the Hoffman-Wielandt inequality [HW53], provide a general relationship between singular values and bounded rigidity. Letting $\sigma_i(M)$ denote the ith largest singular value of $M \in \{-1,1\}^{n\times n}$, their implicit bound for any parameter $\beta \in (0, 1]$ is

$$\mathcal{R}_M(r,\theta) \geqslant \frac{n^2 - \sum_{i=1}^{r} \sigma_i(M)^2 - (1-\beta)^2 n^2}{(1+\beta\theta)^2 - (1-\beta)^2}.$$

Hadamard matrices have $\sigma_i(H) = \sqrt{n}$ for $i \in [n]$. Our bound only depends on the largest singular value, but the asymptotic dependence on n never surpasses that of previous results. Indeed, we can replace $\sum_{i=1}^{r} \sigma_i(M)^2$ with $n\sqrt{r}\sigma_1(M)$ in the above bound, since $\sum_{i=1}^{n} \sigma_i(M)^2 = n^2$ for any $n \times n$ sign matrix. Although we do not improve the scaling with n, we believe that our alternate proof technique is interesting. Kashin and Razborov [KR98] set $\beta = r/n$ to achieve Theorem 1.3. We observe that setting $\beta = (r/n)^{1/2}$ leads to an improved bound in the regime of $1 < \theta \ll \left(\frac{n}{r}\right)^{1/2}$. With this choice of β, their bound achieves asymptotic scaling of $n^2/(\theta+1)$, matching our bound.

The largest singular value $\sigma_1(M)$ also makes an appearance in a related result by Lokam [Lok01], which improves a theorem of Krause and Waack [KW91]. It generalizes the above theorems to hold for sign representations at the cost of an added assumption on the approximating matrix. Define the sign function $\mathsf{sgn} : \mathbb{R} \to \{-1, 1\}$ as $\mathsf{sgn}(x) = 1$ if $x \geqslant 0$ and $\mathsf{sgn}(x) = -1$ if $x < 0$.

Theorem 1.6 ([Lok01])**.** *Consider $M \in \{-1,1\}^{n\times n}$. If $A \in \mathbb{R}^{n\times n}$ has rank r and $1 \leqslant |A_{ij}| \leqslant \theta$ for all $i, j \in [n]$, then $\mathsf{sgn}(A_{ij}) \neq \mathsf{sgn}(M_{ij})$ for at least $\frac{1}{4}\left(n^2/\theta - r\sigma_1^2(M)\right)$ pairs $(i, j) \in [n] \times [n]$.*

The proof of our main theorem straightforwardly extends to prove the following related theorem.

Theorem 1.7. *Consider $M \in \{-1,1\}^{n\times n}$. If $A \in \mathbb{R}^{n\times n}$ has rank r and $1 \leqslant |A_{ij}| \leqslant \theta$ for all $i, j \in [n]$, then $\mathsf{sgn}(A_{ij}) \neq \mathsf{sgn}(M_{ij})$ for at least $n^2/(\theta+1) - n\sqrt{r}\sigma_1(M)\theta/(\theta+1)$ pairs $(i, j) \in [n]\times[n]$.*

2 John's Theorem and Matrix Factorization

We will use John's Theorem to find a matrix factorization with vectors bounded in 2-norm. For more about John's Theorem, see Ball's survey [Bal97]. We only need the following consequence.

Theorem 2.1 (John's Theorem [Joh48])**.** *For any full-dimensional, symmetric convex set $K \subseteq \mathbb{R}^r$ there exists an invertible linear map T such that $\mathcal{B}_r \subseteq T(K) \subseteq \sqrt{r} \cdot \mathcal{B}_r$, where $\mathcal{B}_r$ denotes the r-dimensional ℓ_2-unit ball and $\sqrt{r} \cdot \mathcal{B}_r = \{x \in \mathbb{R}^r \mid \|x\|_2 \leqslant \sqrt{r}\}$.*

The following corollary encapsulates our use of John's Theorem. Essentially the same result appears in Rothvoß' [Rot14] in the context of communication complexity. Our result differs in that we explicitly incorporate the maximum magnitude of any matrix entry.

Corollary 2.2. *For any rank r matrix $A = (A_{ij}) \in \mathbb{R}^{n\times n}$ with $\max_{ij} |A_{ij}| \leqslant \theta$ there exists vectors $u_1, \dots, u_n, v_1, \dots, v_n \in \mathbb{R}^r$ such that $\langle u_i, v_j\rangle = A_{ij}$ and $\|u_i\|_2 \leqslant \sqrt{r}$ and $\|v_j\|_2 \leqslant \theta$ for $i, j \in [n]$.*

Proof. Consider any r-dimensional factorization $A_{ij} = \langle x_i, y_j\rangle$. Define K as the convex hull of the set $\{\pm x_i \mid i \in [n]\}$. Use the mapping T from John's Theorem to define vectors $u_i = Tx_i$ and $v_j = (T^{-1})^T y_j$ for $i, j \in [n]$. These new vectors satisfy the first condition because $\langle u_i, v_j\rangle = \langle x_i, y_j\rangle$. They satisfy the second condition because $u_i = Tx_i \in \sqrt{r}\cdot\mathcal{B}_r$ for all $i \in [n]$. For the third condition, for any vector v_j consider a vector $z \in T(K)$ maximizing $|\langle z, v_j\rangle|$. The inclusion $\mathcal{B}_r \subseteq T(K)$ implies

$$|\langle z, v_j\rangle| \geqslant \|v_j\|_2.$$

Since $T(K)$ is the convex hull of the vectors $\{\pm u_i\}$, we know that z can be expressed as a convex combination of $\{\pm u_i\}$. By the triangle inequality, we conclude $|\langle z, v_j\rangle| \leqslant \max_i |\langle u_i, v_j\rangle| \leqslant \theta$. □

3 Proof of Theorem 1.4

Let C be a minimizer for the set defined in $\mathcal{R}_M(r, \theta)$ and denote $A = M - C$. Let $U, V \subseteq \mathbb{R}^r$ be the sets of vectors that Corollary 2.2 guarantees. Index A as $A_{u,v} = \langle u, v\rangle = \sum_{k=1}^r u_k v_k$ and extend this indexing to M. Note that we have switched notation so $u, v \in \mathbb{R}^r$ are vectors (and indices) and $u_k, v_k \in \mathbb{R}$ are vector elements. We will upper and lower bound the quantity $\sum_{u,v}\langle u, v\rangle M_{uv}$.

For the upper bound, we use the definition of $\sigma_1(M)$, the Cauchy-Schwarz inequality, and the 2-norm bounds on $u \in U$ and $v \in V$ to compute

$$\begin{aligned}
\sum_{u\in U, v\in V} \langle u, v\rangle M_{uv} &= \sum_{k=1}^r \sum_{u\in U, v\in V} M_{uv} u_k v_k \\
&\leqslant \sum_{k=1}^r \sigma_1(M) \sqrt{\sum_{u\in U} u_k^2} \sqrt{\sum_{v\in V} v_k^2} \\
&\leqslant \sigma_1(M) \cdot \sqrt{\sum_{k=1}^r \sum_{u\in U} u_k^2} \cdot \sqrt{\sum_{k=1}^r \sum_{v\in V} v_k^2} \\
&= \sigma_1(M) \cdot \sqrt{\sum_{u\in U} \|u\|_2^2} \cdot \sqrt{\sum_{v\in V} \|v\|_2^2} \;\leqslant\; \sigma_1(M) \cdot n\theta\sqrt{r}.
\end{aligned}$$

For the lower bound, we use that $|A_{u,v}| \leqslant \theta$ from the bounded rigidity definition and compute

$$\sum_{u,v} A_{u,v} M_{u,v} \;\geqslant\; \big(n^2 - |C|\big) - \theta|C| \;=\; n^2 - (\theta + 1)|C|.$$

Rearranging gives the theorem.

4 Conclusion

We introduced a new connection between convex geometry and complexity theory and presented a slightly improved lower bound on the bounded rigidity of sign matrices.

Other theorems, such as Forster's sign rank lower bound [For02] and the discrepancy lower bound via Lindsey's Lemma [Juk12], hold for all $M \in \{-1, 1\}^{n\times n}$ with $\sigma_1(M) = O(\sqrt{n})$. It would be interesting to generalize our result or these theorems to arbitrary real matrices with bounded entries. For example, Codenotti asks as Problem 6 in his survey [Cod99] whether examining singular values suffices to prove better results on the general matrix rigidity of matrices with bounded entries.

5 Acknowledgements

We thank Thomas Rothvoß for sharing his use of John's Theorem, thus inspiring this application.

References

[Bal97] Keith Ball. An Elementary Introduction to Modern Convex Geometry. *Flavors of Geometry*, 31:1–58, 1997.

[Cod99] Bruno Codenotti. Matrix Rigidity, 1999.

[dW06] Ronald de Wolf. Lower Bounds on Matrix Rigidity via a Quantum Argument. ICALP, 2006.

[For02] Jürgen Forster. A Linear Lower Bound on the Unbounded Error Probabilistic Communication Complexity. *J. Comput. Syst. Sci.*, 65(4), December 2002.

[Fri93] Joel Friedman. A Note on Matrix Rigidity. *Combinatorica*, 13(2):235–239, 1993.

[HW53] Alan J Hoffman and Helmut W Wielandt. The Variation of the Spectrum of a Normal Matrix. *Duke Math. J*, 20(1):37–39, 1953.

[Joh48] Fritz John. Extremum Problems with Inequalities as Subsidiary Conditions. In K. O. Friedrichs, O. E. Neugebauer, and J. J. Stoker, editors, *Studies and Essays: Courant Anniversary Volume*, pages 187–204. Wiley-Interscience, New York, 1948.

[Juk12] Stasys Jukna. *Boolean Function Complexity - Advances and Frontiers*, volume 27 of *Algorithms and combinatorics*. Springer, 2012.

[KR98] B.S. Kashin and A.A. Razborov. Improved Lower Bounds on the Rigidity of Hadamard Matrices. *Mathematical Notes*, 1998.

[KW91] Matthias Krause and Stephan Waack. Variation Ranks of Communication Matrices and Lower Bounds for Depth Two Circuits Having Symmetric Gates with Unbounded Fan-in. In *FOCS*, 1991.

[Lok01] Satyanarayana V. Lokam. Spectral Methods for Matrix Rigidity with Applications to Size/Depth Trade-offs and Communication Complexity. *Journal of Computer and System Sciences*, 63(3):449 – 473, 2001.

[Lok09] Satyanarayana V. Lokam. Complexity Lower Bounds using Linear Algebra. *Foundations and Trends in Theoretical Computer Science*, 4(1-2):1–155, 2009.

[Pud00] Pavel Pudlák. A Note on the Use of Determinant for Proving Lower Bounds on the Size of Linear Circuits. *Inf. Process. Lett.*, 74(5-6):197–201, June 2000.

[Rot14] Thomas Rothvoß. A Direct Proof for Lovett's Bound on the Communication Complexity of Low Rank Matrices. *arXiv preprint arXiv:1409.6366*, 2014.

[SSS97] Mohammad Amin Shokrollahi, Daniel A Spielman, and Volker Stemann. A Remark on Matrix Rigidity. *Information Processing Letters*, 64(6):283–285, 1997.

[Val77] Leslie Valiant. Graph-theoretic Arguments in Low-level Complexity. *Mathematical Foundations of Computer Science*, pages 162–176, 1977.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de