An Approximation Algorithm for Approximation Rank

Troy Lee
Columbia University

Adi Shraibman
Weizmann Institute

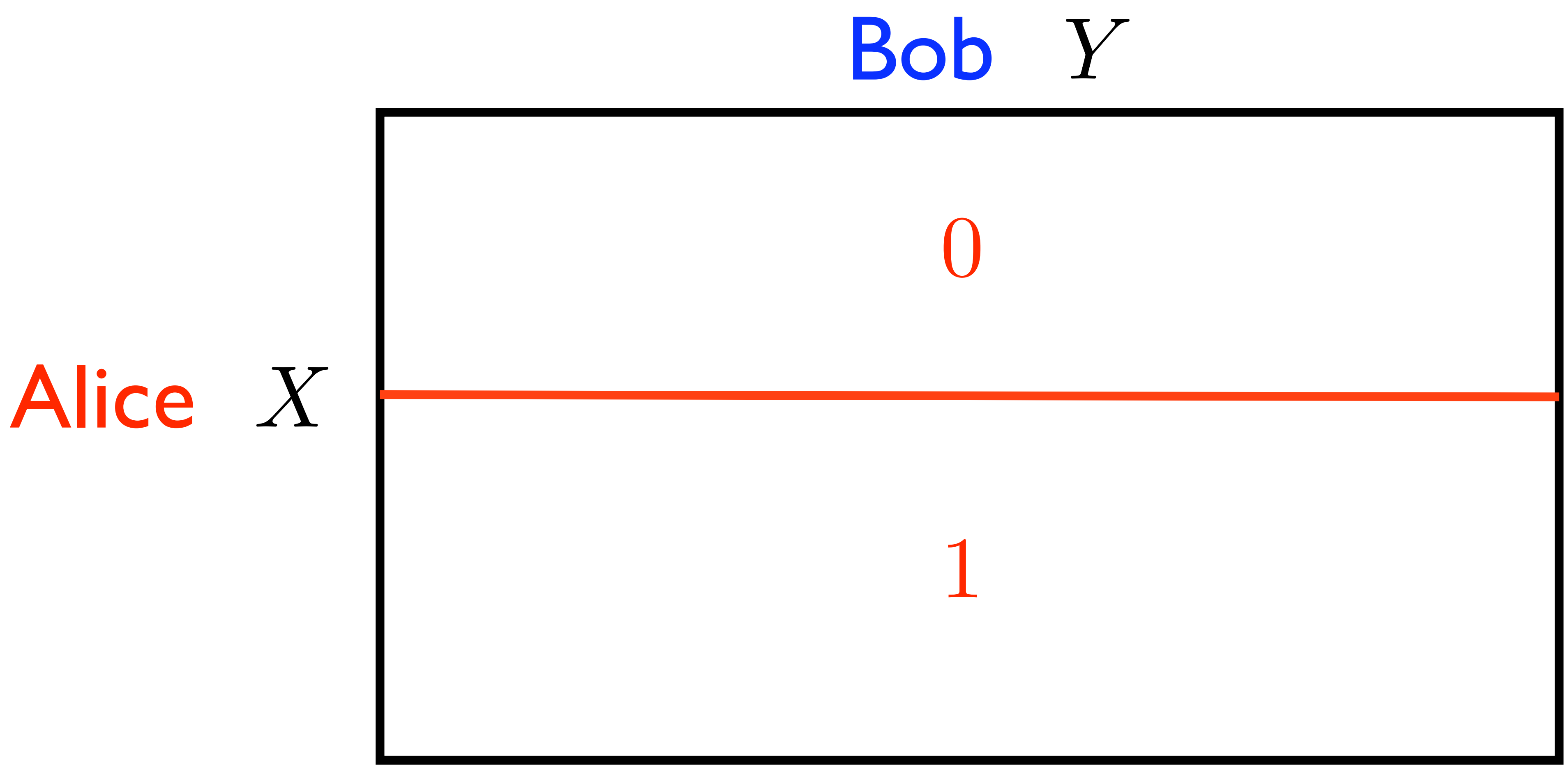
Conventions

Identify a communication function $f : X \times Y \to \{-1, +1\}$ with the associated X-by-Y matrix A(x,y)=f(x,y).

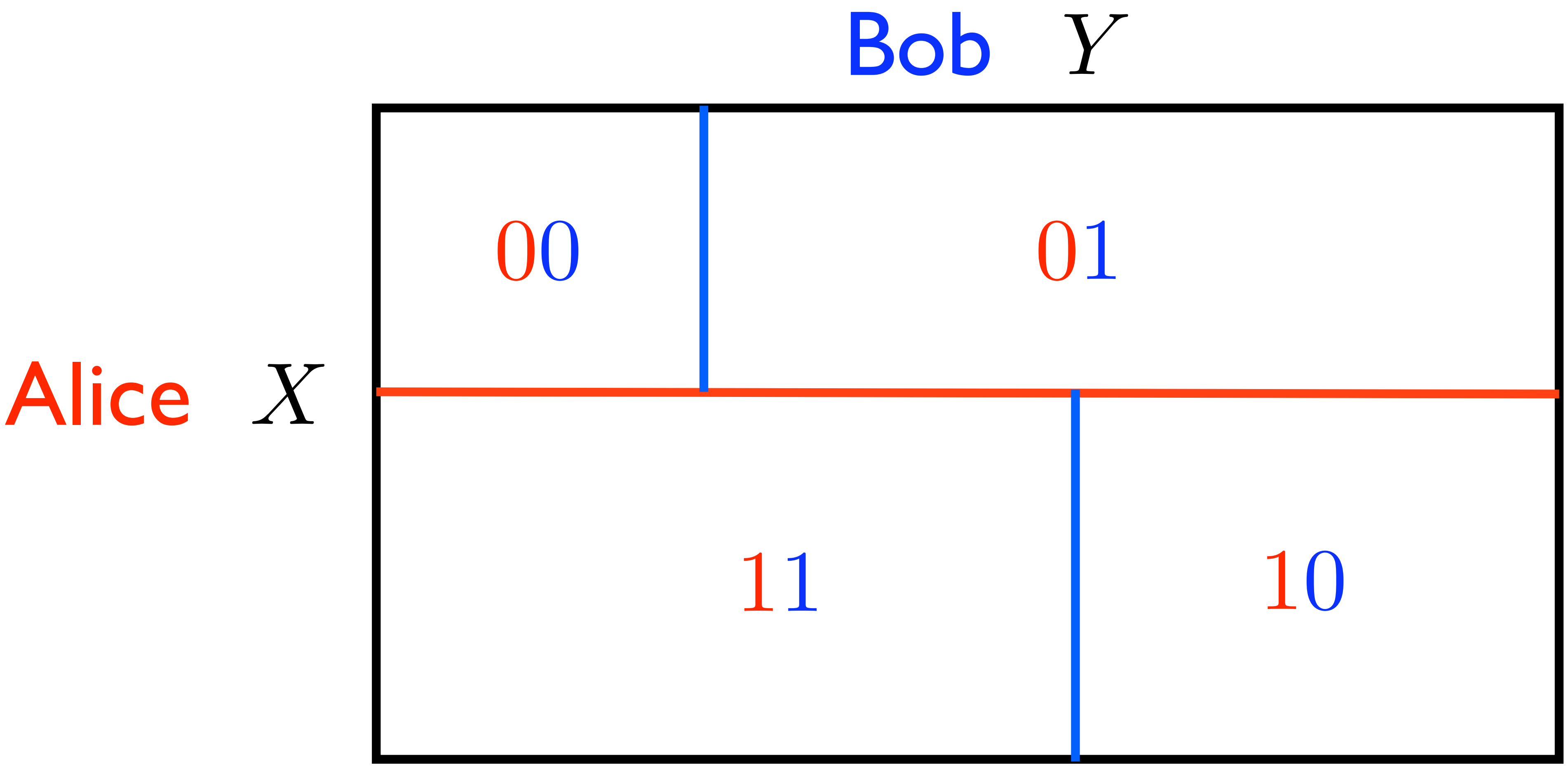
Denote by D(A) the deterministic communication complexity of the sign matrix A.

Denote by $R_\epsilon(A)$ the randomized complexity with error at most ϵ.

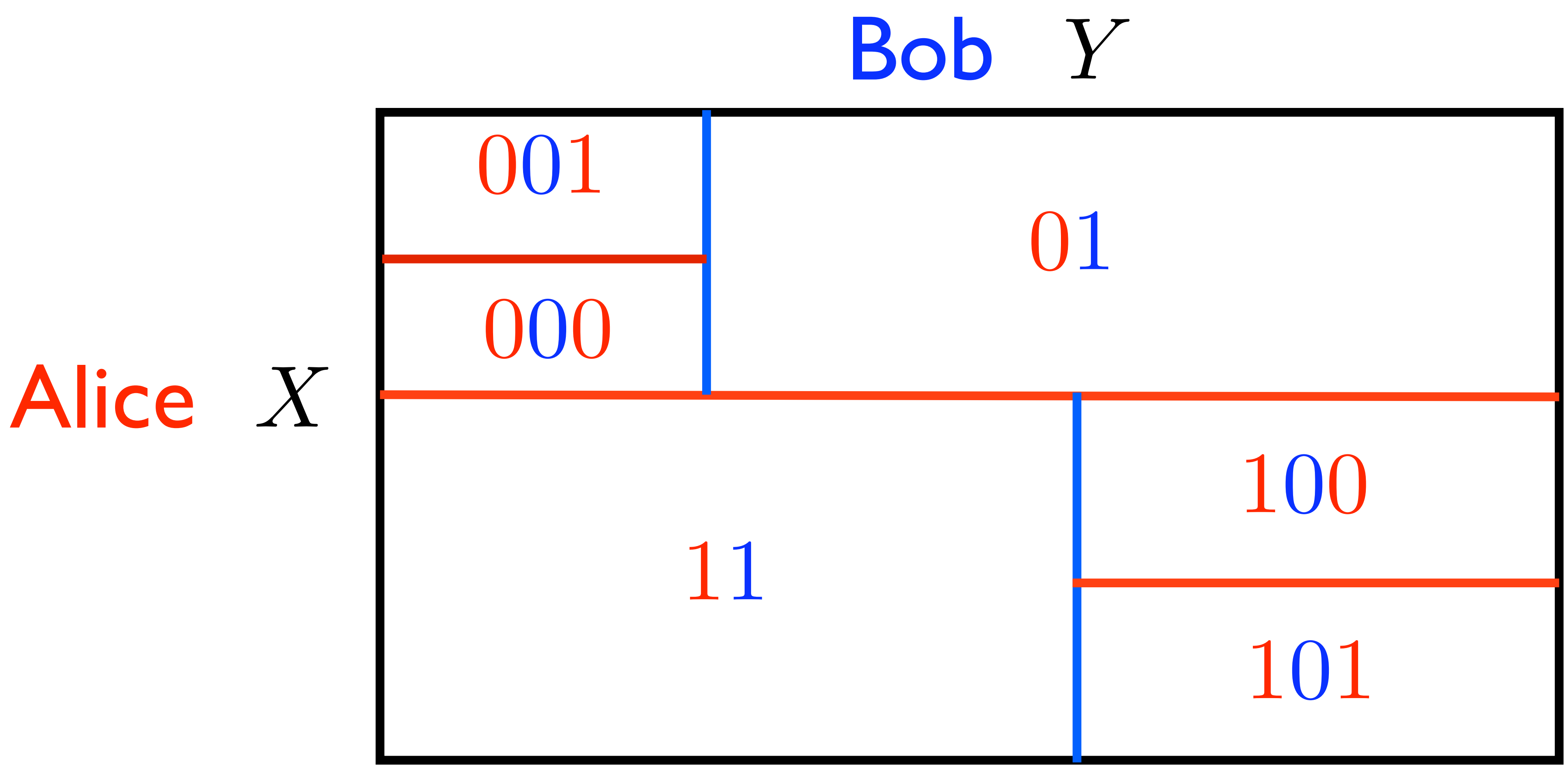
How a protocol partitions a matrix



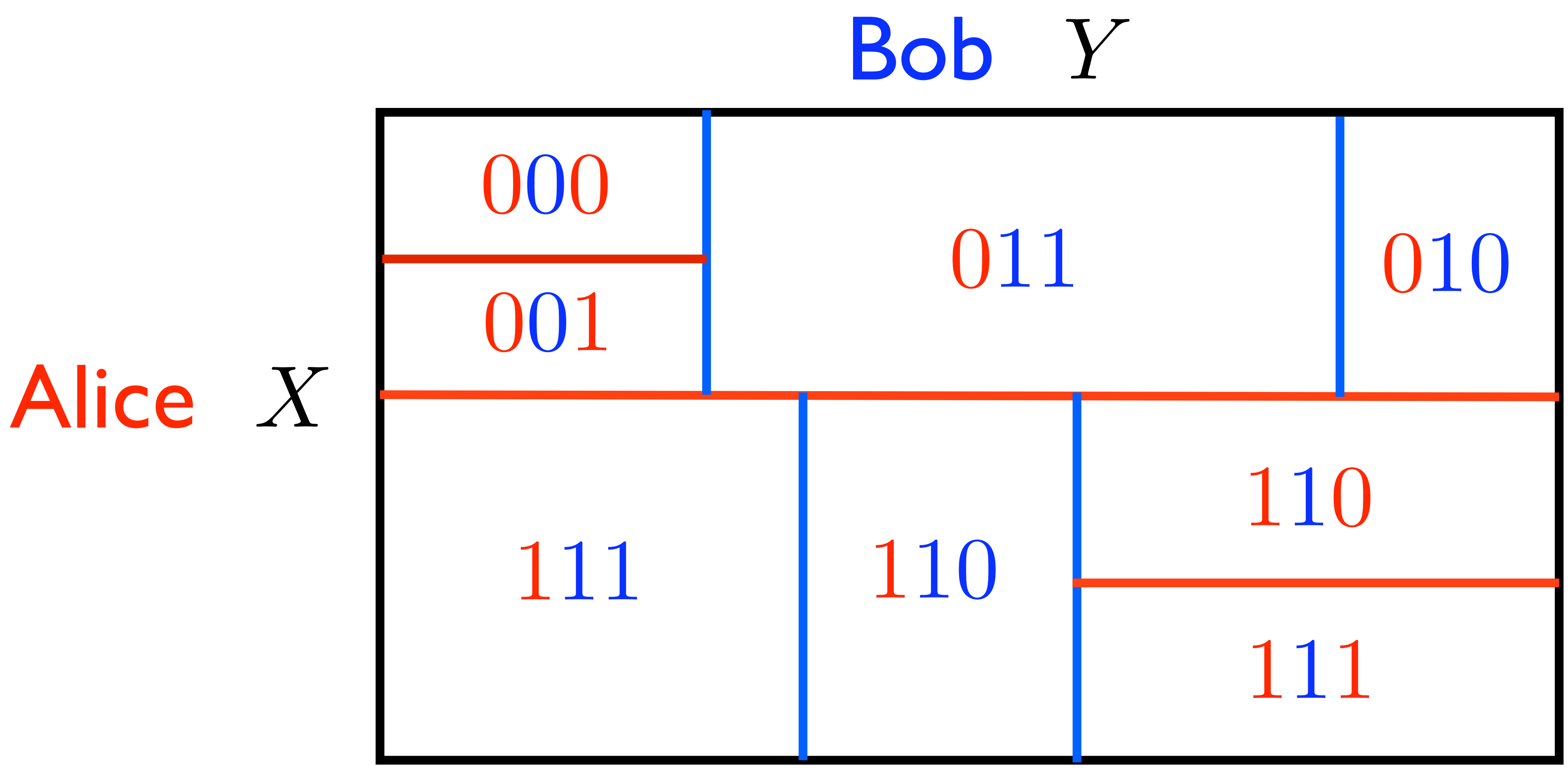
How a protocol partitions a matrix



How a protocol partitions a matrix



How a protocol partitions a matrix



Log rank bound

Log rank bound

A c bit protocol partitions the communication matrix into at most 2^c many rectangles.

Log rank bound

A c bit protocol partitions the communication matrix into at most 2^c many rectangles.

If the protocol is correct, each rectangle will be monochromatic, and so rank one.

Log rank bound

A c bit protocol partitions the communication matrix into at most 2^c many rectangles.

If the protocol is correct, each rectangle will be monochromatic, and so rank one.

By subadditivity of rank, this gives

$$D(A) \geq \log \operatorname{rank}(A)$$

Log rank bound

A c bit protocol partitions the communication matrix into at most 2^c many rectangles.

If the protocol is correct, each rectangle will be monochromatic, and so rank one.

By subadditivity of rank, this gives

$$D(A) \geq \log \operatorname{rank}(A)$$

One of the greatest open problems in communication complexity, the log rank conjecture of Lovasz and Saks, states that this bound is polynomially tight

Approximation rank

Approximation rank

For randomized (and quantum) models, one can show that a lower bound is given by *approximation rank*.

[Krause96, Buhrman and de Wolf 01]

Approximation rank

For randomized (and quantum) models, one can show that a lower bound is given by *approximation rank*.
[Krause96, Buhrman and de Wolf 01]

Given a target matrix A, approximation rank looks at the minimal rank matrix entrywise close to A:

$$
\begin{aligned}
\mathrm{rank}_\epsilon(A) \quad &= \quad \min_X \mathrm{rank}(X) \\
& \quad\quad |X(i,j) - A(i,j)| \leq \epsilon \text{ for all } i,j.
\end{aligned}
$$

Approximation rank

Approximation rank

In analogy with the log rank conjecture, can also conjecture that log approximation rank is polynomially tight for randomized communication complexity.

Approximation rank

In analogy with the log rank conjecture, can also conjecture that log approximation rank is polynomially tight for randomized communication complexity.

We do not know if approximation rank is NP-hard to compute, but similar rank minimization problems are NP-hard.

Approximation rank

In analogy with the log rank conjecture, can also conjecture that log approximation rank is polynomially tight for randomized communication complexity.

We do not know if approximation rank is NP-hard to compute, but similar rank minimization problems are NP-hard.

While approximation rank gives a lower bound on quantum communication complexity, not known to work for model with entanglement.

Main result

Main result

We show a semidefinite programming quantity which is polynomially related to approximation rank for constant $0 < \varepsilon < 1/2$

Main result

We show a semidefinite programming quantity which is polynomially related to approximation rank for constant $0 < \varepsilon < 1/2$

Moreover, this quantity is known to be a lower bound even in the quantum model with entanglement.

Convex relaxation

Convex relaxation

Part of what makes rank minimization problems hard is that rank is not a convex function.

Convex relaxation

Part of what makes rank minimization problems hard is that rank is not a convex function.

As is common in approximation algorithms, we look at a convex relaxation of the problem.

Convex relaxation

Part of what makes rank minimization problems hard is that rank is not a convex function.

As is common in approximation algorithms, we look at a convex relaxation of the problem.

For this, we need to introduce some matrix norms.

Matrix norms

- Define the i[th] singular value as

$$\sigma_i(A) = \sqrt{\lambda_i(AA^t)}$$

- Denote

$$\begin{aligned} \|A\|_1 &= \sum_i \sigma_i(A) \\ \|A\|_\infty &= \sigma_1(A) \\ \|A\|_2 &= \sqrt{\sum_i \sigma_i^2(A)} \end{aligned}$$

Matrix norms

- Define the i^{th} singular value as

$$\sigma_i(A) = \sqrt{\lambda_i(AA^t)}$$

- Denote

$$\|A\|_1 = \sum_i \sigma_i(A) \quad \text{“Trace Norm”}$$

$$\|A\|_\infty = \sigma_1(A)$$

$$\|A\|_2 = \sqrt{\sum_i \sigma_i^2(A)}$$

Matrix norms

- Define the i[th] singular value as

$$\sigma_i(A) = \sqrt{\lambda_i(AA^t)}$$

- Denote

$$\|A\|_1 = \sum_i \sigma_i(A) \quad \text{"Trace Norm"}$$

$$\|A\|_\infty = \sigma_1(A) \quad \text{"Spectral Norm"}$$

$$\|A\|_2 = \sqrt{\sum_i \sigma_i^2(A)}$$

Matrix norms

- Define the i[th] singular value as

$$\sigma_i(A) = \sqrt{\lambda_i(AA^t)}$$

- Denote

$$\begin{aligned} \|A\|_1 &= \sum_i \sigma_i(A) && \text{“Trace Norm”} \\ \|A\|_\infty &= \sigma_1(A) && \text{“Spectral Norm”} \\ \|A\|_2 &= \sqrt{\sum_i \sigma_i^2(A)} && \text{“Frobenius Norm”} \end{aligned}$$

Matrix norms

- Define the i[th] singular value as

$$\sigma_i(A) = \sqrt{\lambda_i(AA^t)}$$

- Denote

$$\begin{aligned} \|A\|_1 &= \sum_i \sigma_i(A) && \text{"Trace Norm"} \\ \|A\|_\infty &= \sigma_1(A) && \text{"Spectral Norm"} \\ \|A\|_2 &= \sqrt{\sum_i \sigma_i^2(A)} && \text{"Frobenius Norm"} \\ &= \sqrt{\sum_{i,j} A(i,j)^2} \end{aligned}$$

Trace norm method

Replace the rank objective function by the trace norm.

As rank equals the number of nonzero singular values, we have

$$\sum_i \sigma_i(A) \leq \sqrt{\operatorname{rank}(A)}\ \|A\|_2$$

For a M-by-N sign matrix this gives

$$\operatorname{rank}(A) \geq \frac{\|A\|_1^2}{MN}$$

Example: Trace Norm

Let H be a N-by-N Hadamard matrix.

Example: Trace Norm

Let H be a N-by-N Hadamard matrix.

Then $HH^t = NI$ and so $\|H\|_1 = N^{3/2}$

Example: Trace Norm

Let H be a N-by-N Hadamard matrix.

Then $HH^t = NI$ and so $\|H\|_1 = N^{3/2}$

Thus trace norm method gives a bound on the rank of

$$\frac{\|H\|_1^2}{\|H\|_2^2} = \frac{N^3}{N^2} = N$$

Trace norm method: drawback

Trace norm bound is not monotone. Consider

$$\begin{pmatrix} H_N & 1_N \\ 1_N & 1_N \end{pmatrix}$$

Trace norm method: drawback

Trace norm bound is not monotone. Consider

$$\begin{pmatrix} H_N & 1_N \\ 1_N & 1_N \end{pmatrix}$$

Trace norm is at most $N^{3/2} + 3N$

Trace norm method: drawback

Trace norm bound is not monotone. Consider

$$\begin{pmatrix} H_N & 1_N \\ 1_N & 1_N \end{pmatrix}$$

Trace norm is at most $N^{3/2} + 3N$

Trace norm method gives

$$\frac{(N^{3/2} + 3N)^2}{4N^2} = \frac{N}{4} + O(\sqrt{N})$$

Trace norm method: drawback

Trace norm bound is not monotone. Consider

$$\begin{pmatrix} H_N & 1_N \\ 1_N & 1_N \end{pmatrix}$$

Trace norm is at most $N^{3/2} + 3N$

Trace norm method gives

$$\frac{(N^{3/2} + 3N)^2}{4N^2} = \frac{N}{4} + O(\sqrt{N})$$

Worse bound than on Hadamard submatrix!

A fix

We can fix this by considering

$$\max_{\substack{u,v \\ \|u\|=\|v\|=1}} \|A \circ uv^t\|_1$$

As this entrywise product does not increase the rank we still have

$$\mathrm{rank}(A) \geq \left(\frac{\|A \circ uv^t\|_1}{\|A \circ uv^t\|_2} \right)^2$$

For a sign matrix A, this simplifies nicely:

$$\mathrm{rank}(A) \geq \left(\|A \circ uv^t\|_1 \right)^2$$

We have arrived

At the γ_2 norm, introduced in communication complexity by [LMSS07, LS07].

$$\gamma_2(A) = \max_{\substack{u,v \\ \|u\|=\|v\|=1}} \|A \circ uv^t\|_1$$

We have arrived

At the γ_2 norm, introduced in communication complexity by [LMSS07, LS07].

$$\gamma_2(A) = \max_{\substack{u,v \\ \|u\|=\|v\|=1}} \|A \circ uv^t\|_1$$

Known by a variety of names in matrix analysis:

Hadamard/Schur product operator/trace norm

We have arrived

At the γ_2 norm, introduced in communication complexity by [LMSS07, LS07].

$$\gamma_2(A) = \max_{\substack{u,v \\ \|u\|=\|v\|=1}} \|A \circ uv^t\|_1$$

Known by a variety of names in matrix analysis:
Hadamard/Schur product operator/trace norm

Has many nice properties. For this talk, we use the fact that it can be written as a semidefinite program and computed to high accuracy in polynomial time.

Application to approximation rank

Application to approximation rank

Use γ_2 norm as surrogate in rank minimization problem

$$\begin{aligned} \gamma_2^{\epsilon}(A) &= \min_B \gamma_2(B) \\ & |A(i,j) - B(i,j)| \leq \epsilon \text{ for all } i,j. \end{aligned}$$

Application to approximation rank

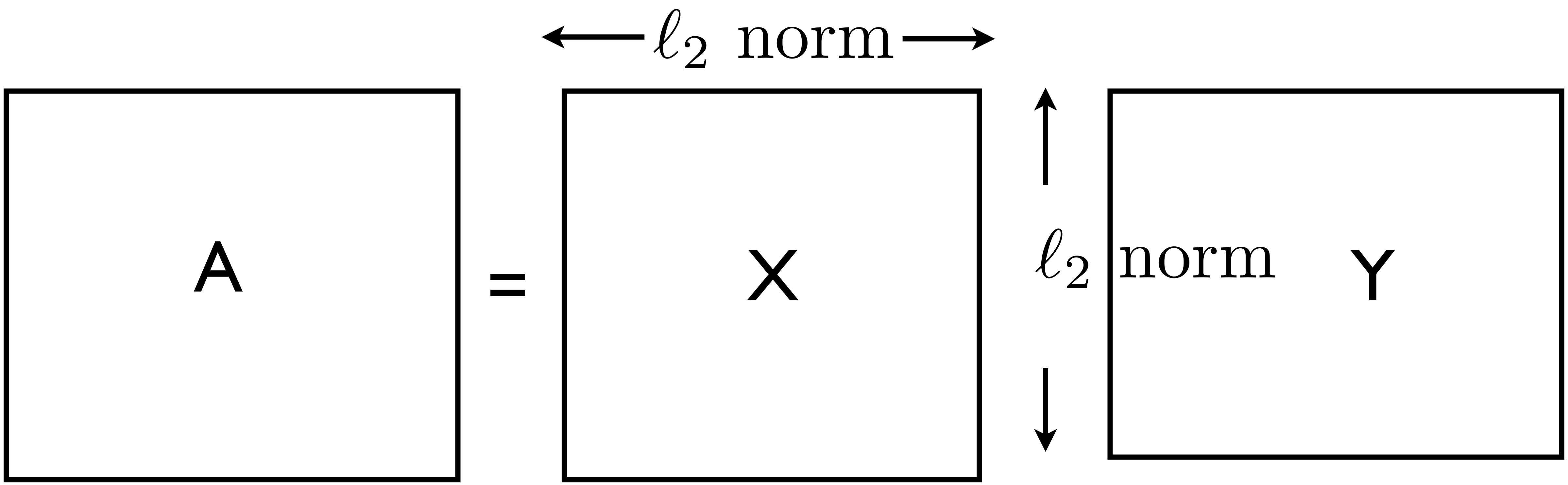
Use γ_2 norm as surrogate in rank minimization problem

$$\begin{aligned} \gamma_2^\epsilon(A) &= \min_B \gamma_2(B) \\ & |A(i,j) - B(i,j)| \leq \epsilon \text{ for all } i,j. \end{aligned}$$

As argued above,

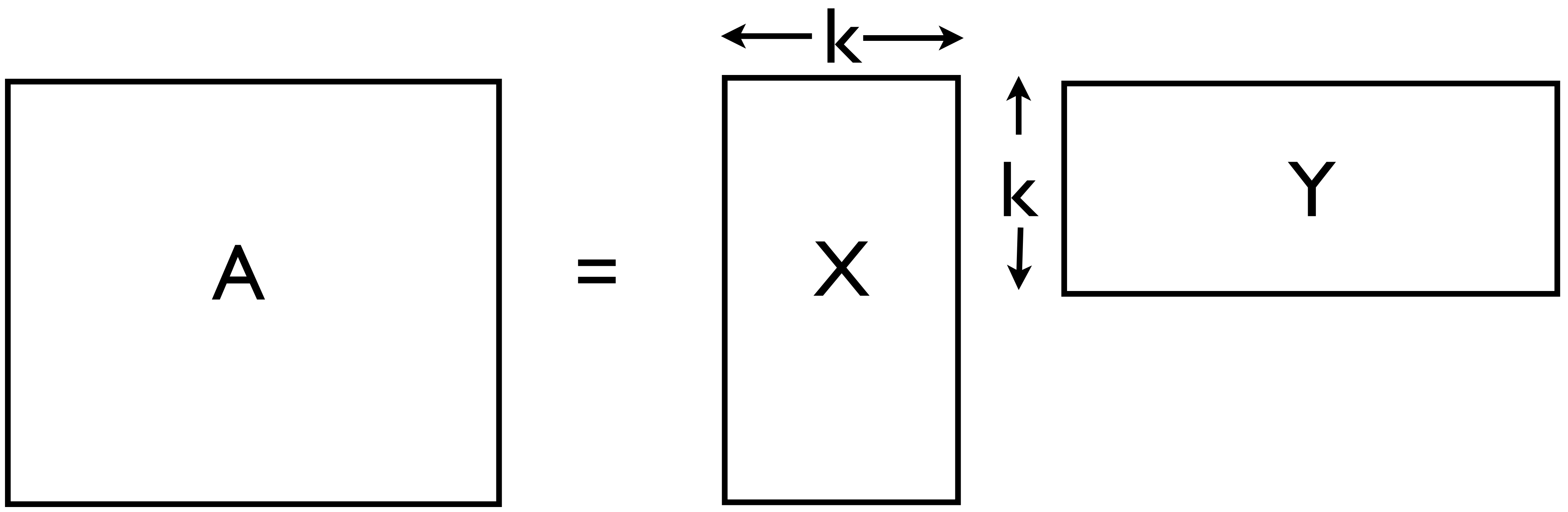
$$\frac{\gamma_2^\epsilon(A)^2}{(1+\epsilon)^2} \leq \mathrm{rank}_\epsilon(A).$$

γ_2 norm as factorization

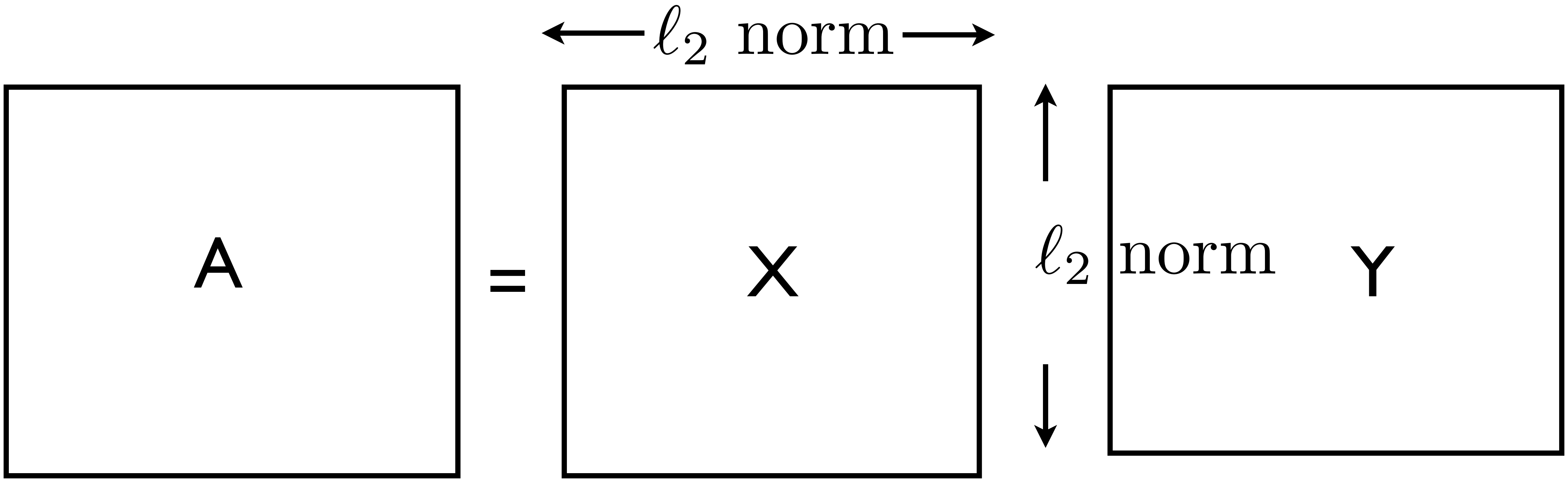


$$\gamma_2(A) = \min_{\substack{X,Y \\ XY=A}} \|X\|_r \|Y\|_c$$

Rank as factorization



γ_2 norm as factorization

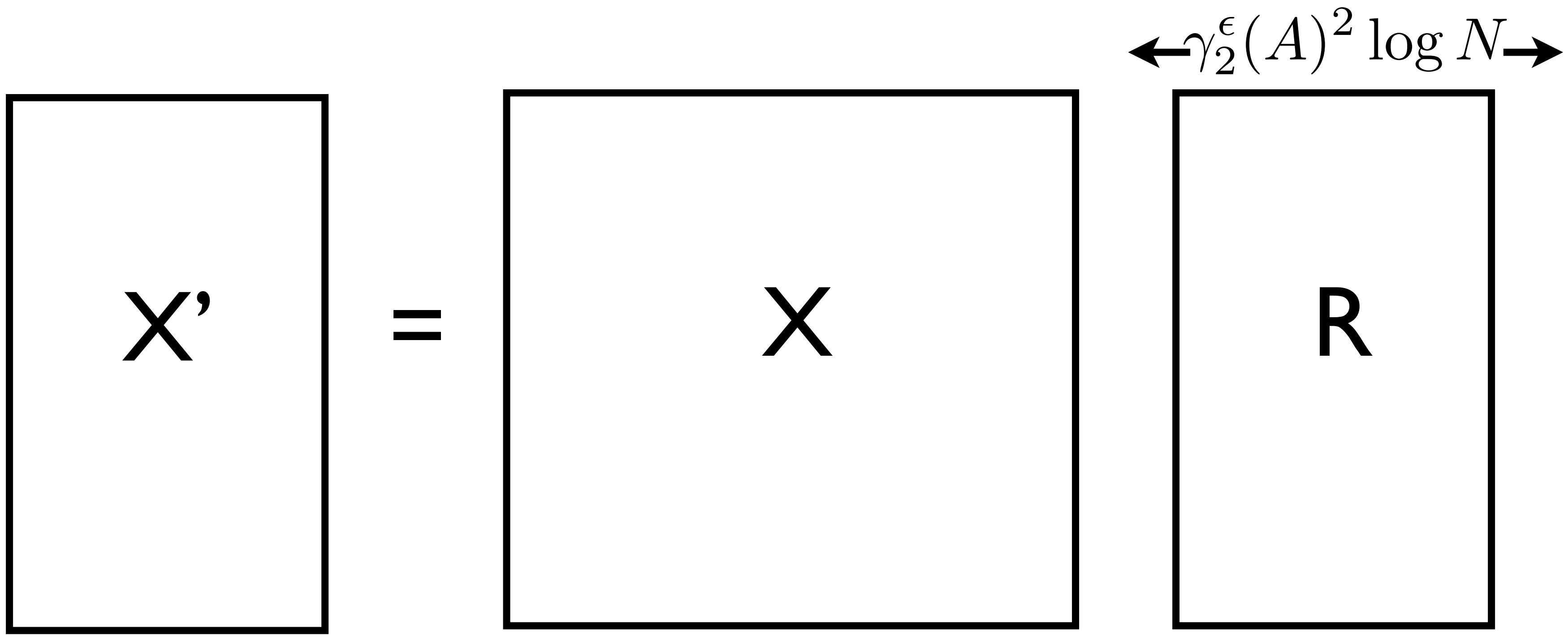


$$\gamma_2(A) = \min_{\substack{X,Y \\ XY=A}} \|X\|_r \|Y\|_c$$

Now shrink the rows

Take X,Y optimal factorization realizing $\gamma_2^{\epsilon}(A)$.

Obtain matrices X',Y' by randomly projecting columnspace down to dimension about $\gamma_2^{\epsilon}(A)^2 \log N$



Johnson-Lindenstrauss Lemma

Now X'Y' will be a matrix of the desired rank.

Furthermore, by the Johnson-Lindenstrauss lemma and a union bound, the inner products between all rows of X' and columns of Y' will approximately equal those between X and Y.

Thus X'Y' will still be entrywise close to A.

Error-reduction

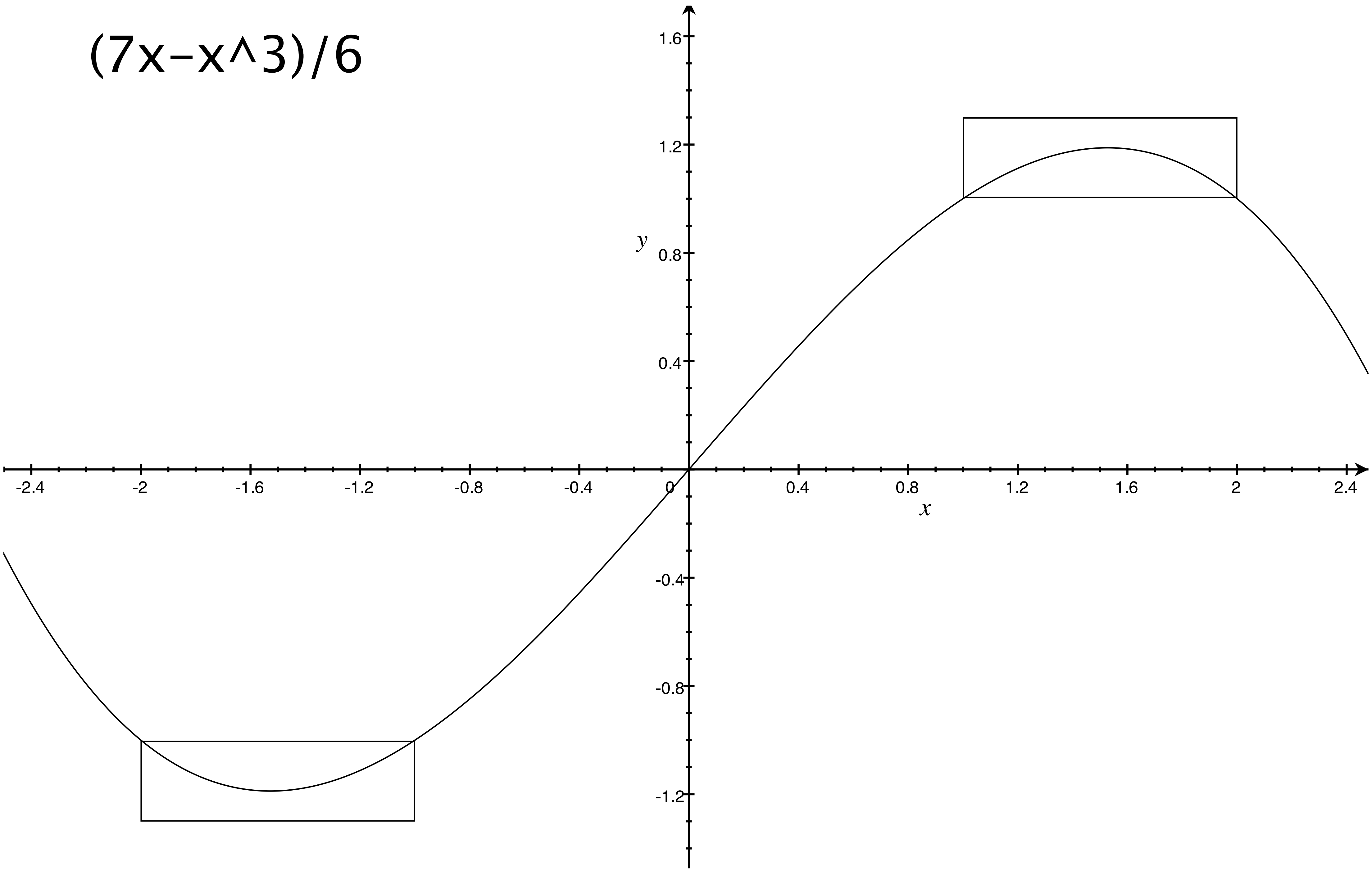
Can argue that if we started with an ε approximation, now will have 2ε approximation.

Can fix this by applying low degree polynomial approximation of the sign function entrywise to the matrix.

Applying a degree d polynomial blows up rank by at most a power of d.

Error-reduction

(7x−x^3)/6



Final result

For any M-by-N sign matrix A and constant $0 < \varepsilon < 1$

$$\frac{\gamma_2^\epsilon(A)^2}{(1+\epsilon)^2} \leq \mathrm{rank}_\epsilon(A) = O\left(\gamma_2^\epsilon(A)^2 \log(MN)\right)^3$$

Implies that the log approximation rank conjecture is essentially equivalent to the existence of a constant c such that

$$R_\epsilon(A) \leq (\log \gamma_2^\epsilon(A))^c + O(\log n).$$

Open questions

- What is the complexity of the real vector inner product function?
- Does the log rank conjecture imply the log approximation rank conjecture?
- Approximation algorithm for the limiting case of sign rank?