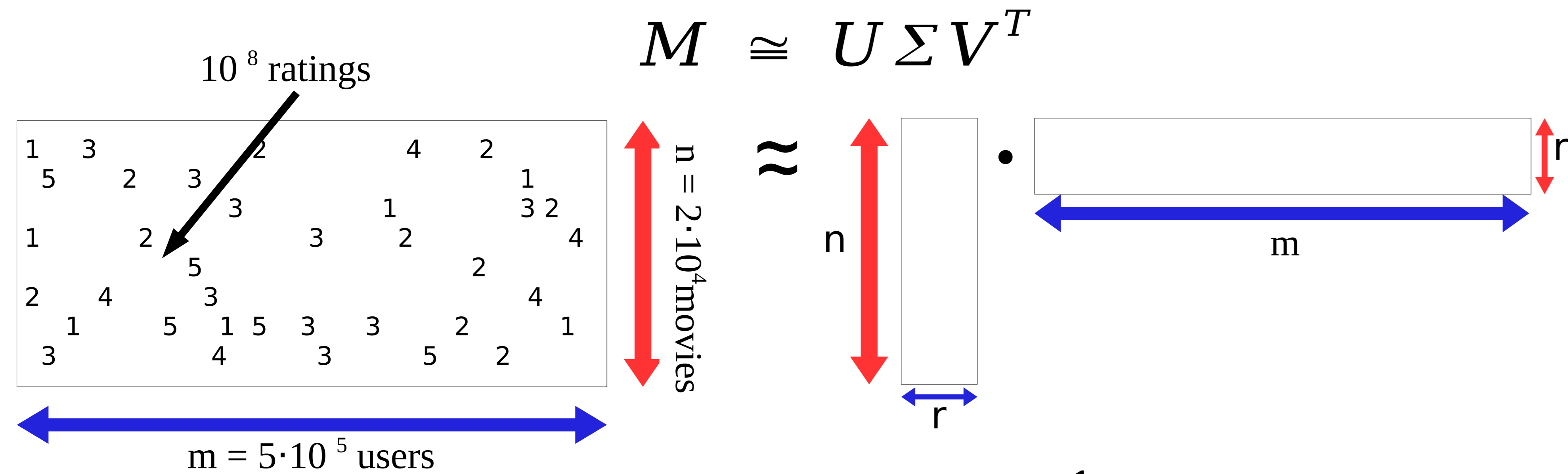


Matrix Completion from a Few Entries

Raghunandan H. Keshavan, Andrea Montanari, Sewoong Oh,
Department of Electrical Engineering, Stanford University

What is Matrix Completion?



Problem : How many revealed entries $|E|$ do we need to get $\frac{1}{mn} \|M - \hat{M}\|_F^2 \leq \delta$?

Algorithm

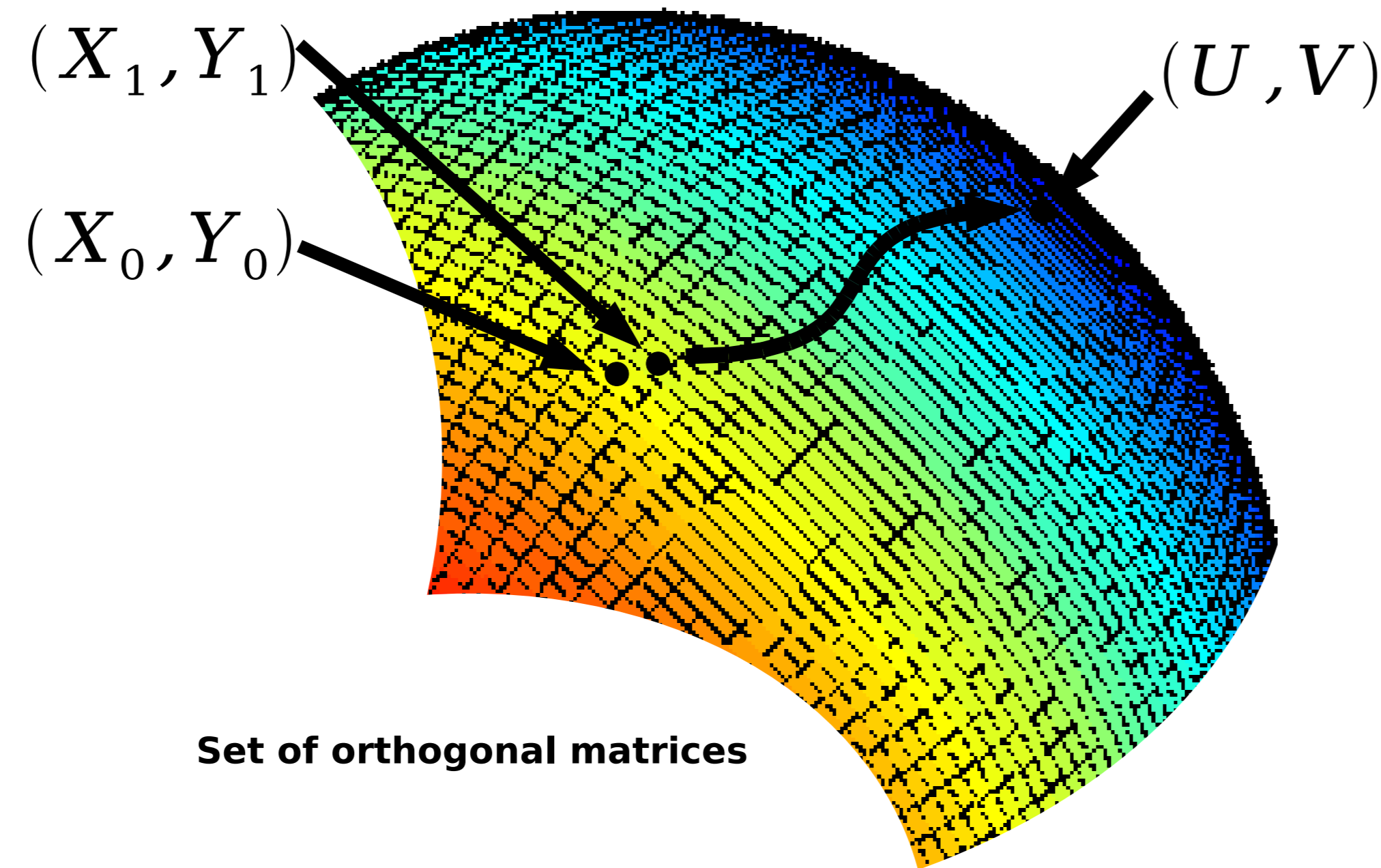
Algorithm [OptSpace]

Trim : Trim M^E to M^E ;

Project : Project M^E onto $\text{Tr}(M^E)$;

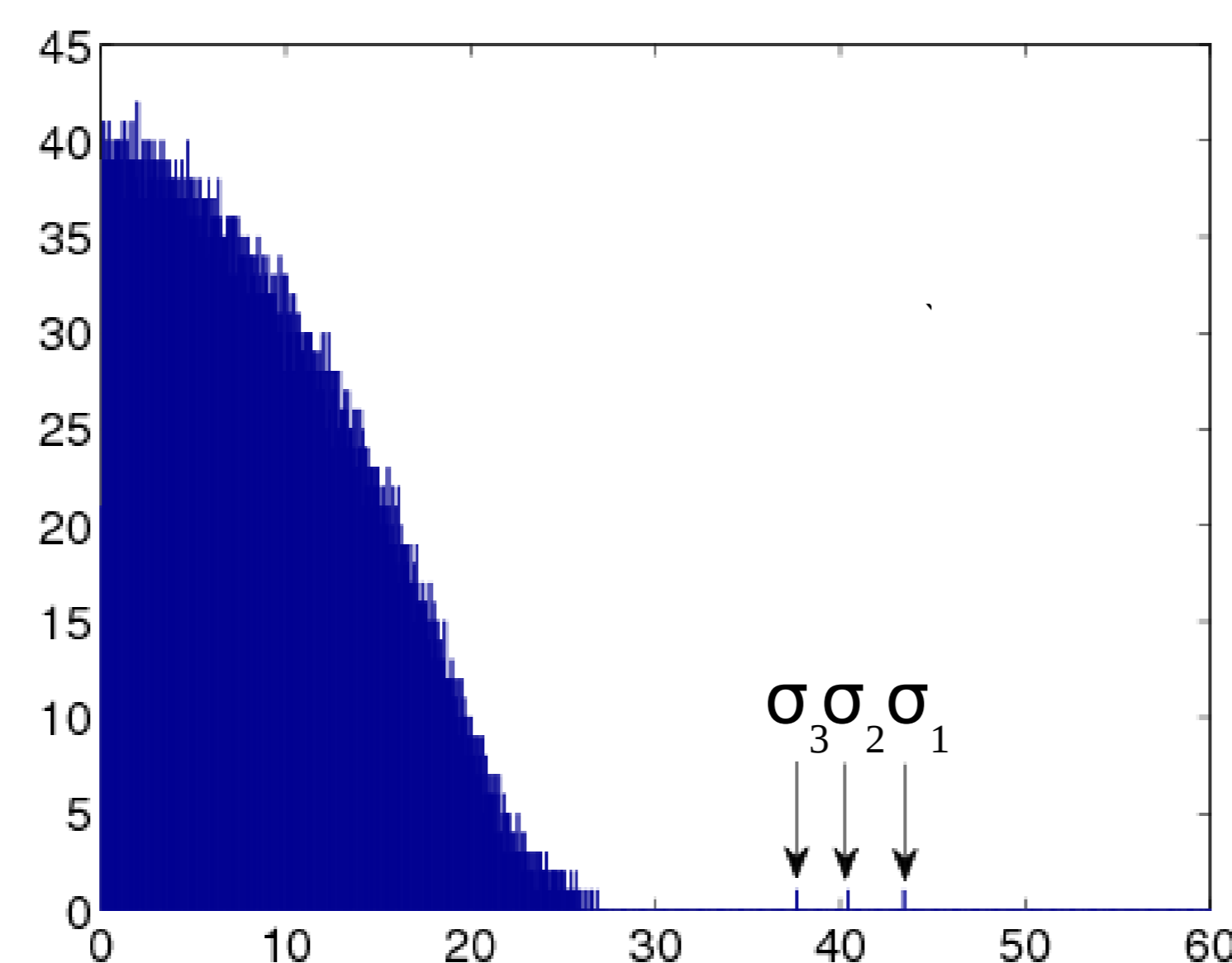
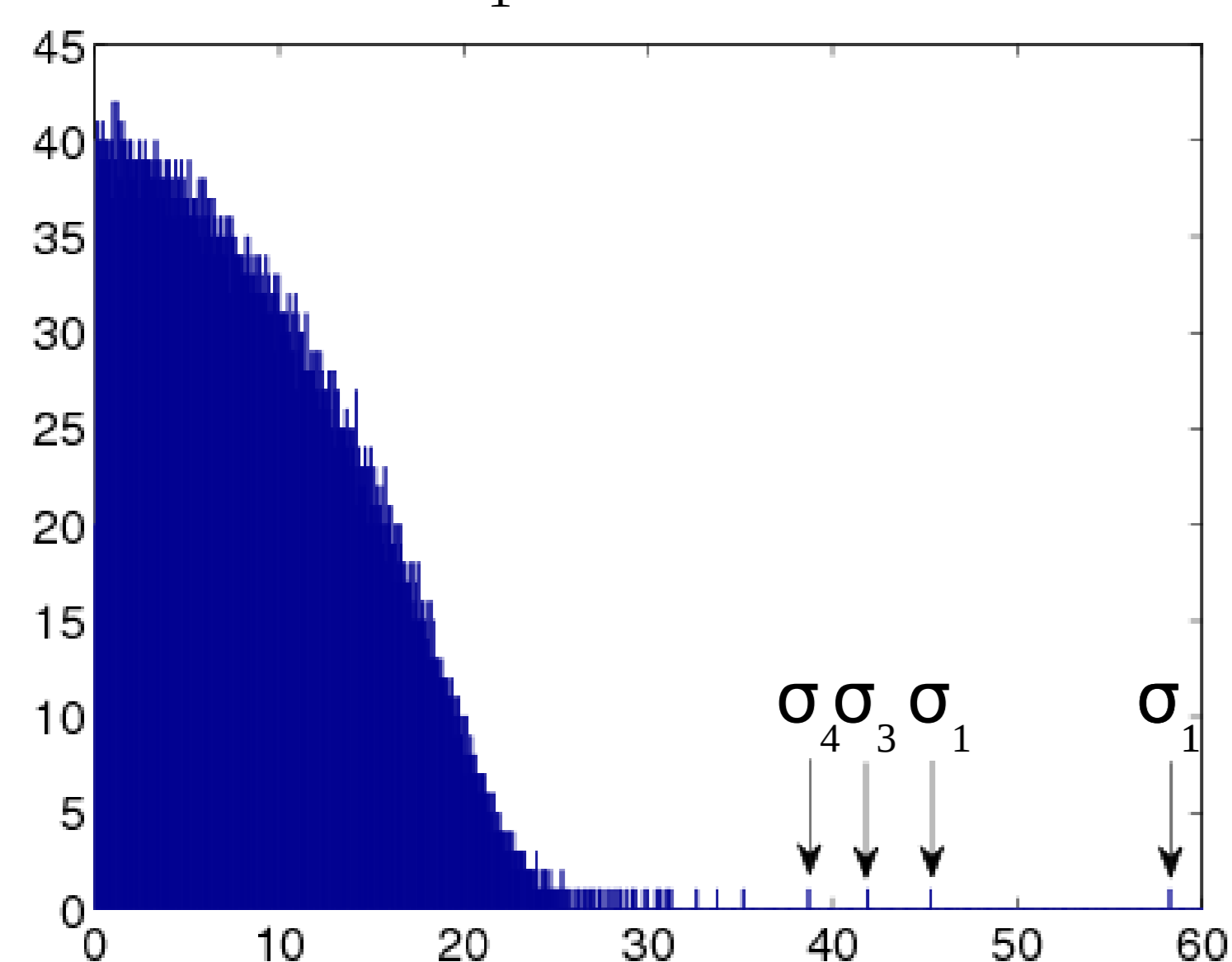
Clean : Minimize Cost $F(X, Y)$,
s.t. X, Y orthogonal.

$$F(X, Y) = \min \sum_{i, j \in E} ((XSY^T)_{ij} - M_{ij})^2$$



SVD : $\text{Tr}(M^E) = \frac{mn}{|E|} \sum_{i=1}^r x_i \sigma_i y_i^T$

Problem : $\sigma_1 = \Omega(\log(n)/\log \log(n))$



Histogram of singular values of a partially revealed random rank 3 matrix before(left) and after(right) trimming

Main Results

Theorem (Keshavan, Montanari, Oh, 2009 [1])

Assume $r = O(1)$, and let M be an $n \times n$ matrix satisfying (μ_0, μ_1) -incoherence with $\sigma_1(M)/\sigma_r(M) = O(1)$. If $|E| \geq C'n \log n$, then OPTSPACE returns, whp., the matrix M .

Theorem (Keshavan, Montanari, Oh, 2009 [2])

Let $N = M + Z$ with M as above and Z any $n \times n$ matrix. If $|E| \geq C'n \log n$, then (under appropriate technical conditions) OPTSPACE with input N^E returns \hat{M} such that whp.,

$$\frac{1}{\sqrt{mn}} \|M - \hat{M}\|_F \leq C \frac{n\sqrt{\alpha r}}{|E|} \|Z^E\|_2$$

Implementation

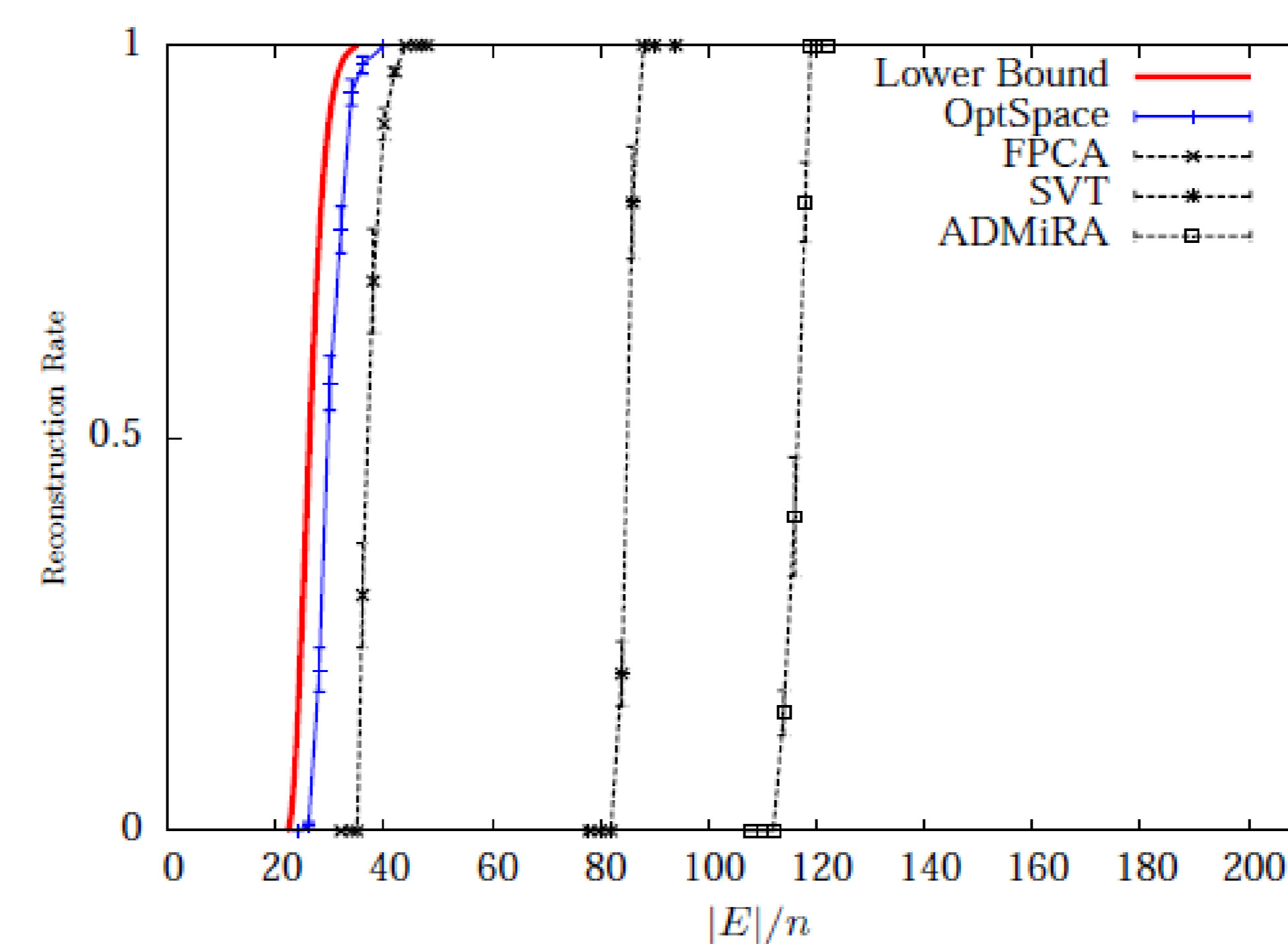


Figure 1: Reconstruction Rate vs. No. of samples revealed ($m = n = 1000, r = 10$)

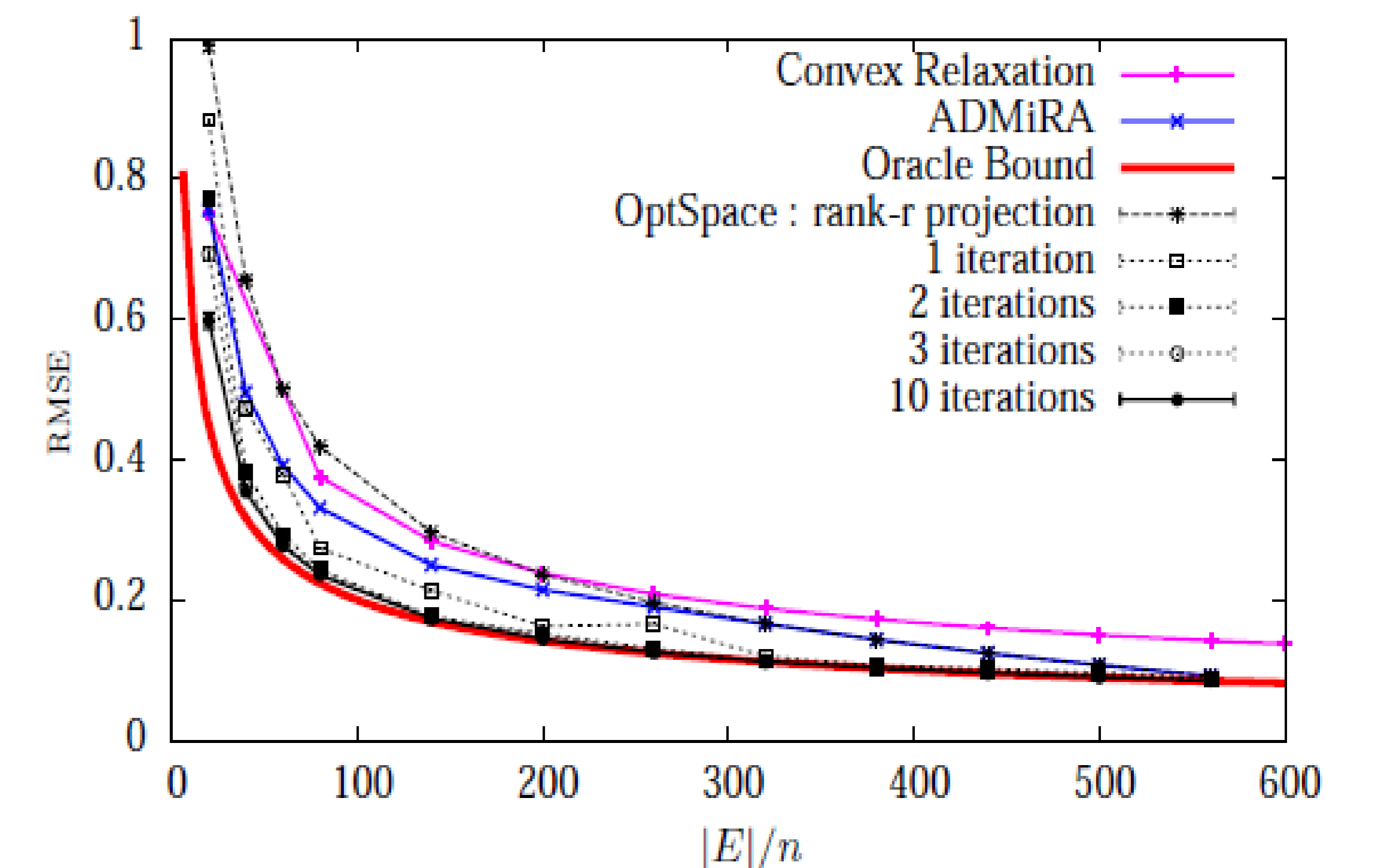


Figure 2: RMSE vs. Number of samples revealed ($m = n = 600, r = 2$)

References

- [1] R.H.Keshavan, A. Montanari, and S. Oh, *Matrix Completion from a few entries*, arXiv:0901.3150, January 2009.
- [2] R.H.Keshavan, A. Montanari, and S. Oh, *Matrix Completion from noisy entries*, arXiv:0906.2027, June 2009.