Computational Complexity of Certifying Restricted Isometry Property

Abhiram Natarajan and Yi Wu

Purdue University
West Lafayette, IN, USA
`{nataraj2,wu510}@cs.purdue.edu`

—— **Abstract** ——

Given a matrix A with n rows, a number $k < n$, and $0 < \delta < 1$, A is (k, δ)-RIP (Restricted Isometry Property) if, for any vector $x \in \mathbb{R}^n$, with at most k non-zero co-ordinates,

$$(1 - \delta)\|x\|_2 \leq \|Ax\|_2 \leq (1 + \delta)\|x\|_2$$

In other words, a matrix A is (k, δ)-RIP if Ax preserves the length of x when x is a k-sparse vector. In many applications, such as compressed sensing and sparse recovery, it is desirable to construct RIP matrices with a large k and a small δ. It is known that, with high probability, random constructions produce matrices that exhibit RIP. This motivates the problem of certifying whether a randomly generated matrix exhibits RIP with suitable parameters.

In this paper, we prove that it is hard to approximate the RIP parameters of a matrix assuming the SMALL-SET-EXPANSION HYPOTHESIS. Specifically, we prove that for any arbitrarily large constant $C > 0$ and any arbitrarily small constant $0 < \delta < 1$, there exists some k such that given a matrix M, it is SMALL-SET-EXPANSION-HARD to distinguish the following two cases:

- (Highly RIP) M is (k, δ)-RIP.
- (Far away from RIP) M is not $(k/C, 1 - \delta)$-RIP.

Most of the previous results on the topic of hardness of RIP certification only hold for certification when $\delta = o(1)$; i.e, when the matrix exhibits strong RIP. In practice, it is of interest to understand the complexity of certifying a matrix with δ being close to $\sqrt{2} - 1$, as it suffices for many real applications to have matrices with $\delta = \sqrt{2} - 1$. Our hardness result holds for any constant δ. Specifically, our result proves that even if δ is indeed very small, i.e. the matrix is in fact *strongly RIP*, certifying that the matrix exhibits *weak RIP* itself is SMALL-SET-EXPANSION-HARD.

In order to prove the hardness result, we prove a variant of the Cheeger's Inequality for sparse vectors. Although a similar result is already known, our proof technique gives better constants in the inequality which may be useful for other applications. Specifically, let A be the adjacency matrix of a d-regular graph $G(V, E)$, and $L = I - \frac{1}{d}A$ be the normalized Laplacian matrix of G. For any $\eta \leq 1/2$, we show that

$$\lambda_\eta \leq \phi_\eta(G) \leq \sqrt{(2 - \lambda_\eta)\lambda_\eta}$$

where $\lambda_\eta = \min_{\|x\|_0 = \eta n} \frac{x^T L x}{|x|_2^2}$ and $\phi_\eta(G)$ is the minimum edge expansion among all the sets of size at most $\eta|V|$.

It is interesting to note that the relationship between λ_η and $\Phi_\eta(G)$ is different from (and tighter than) the relation between λ and $\phi(G)$ in the regular version of Cheeger's Inequality (which states that $\frac{\lambda}{2} \leq \phi(G) \leq \sqrt{2\lambda}$). We will see that obtaining this tighter relationship between λ_η and $\phi_\eta(G)$ is crucial in proving our hardness result.

1 Introduction

Moore's law has enabled the creation of very robust and effective sensing systems. As a result of the ubiquity of such systems, the amount of data generated by these systems has increased vastly. In fact, in most real applications, there is so much data that sampling at the required rates (called Nyquist rate) becomes impractical due to data storage problems as well as the sheer magnitude of the sampling rate [8]. Signal processing literature shows us that this problem is circumvented by constructing compressible representations of signals. This technique leverages the notion of *sparse approximation* and is called *compressed sensing*.

A formal statement of the central problem of compressed sensing is as follows. Assume the presence of a matrix $\Phi \in \mathbb{R}^{m \times n}$, called the sensing matrix, with $m \ll n$. We are also given a vector $y \in \mathbb{R}^m$, which contains a set of m measurements. We are interested in reconstructing $x \in \mathbb{R}^n$, such that

$$y = \Phi x$$

Given that $m \ll n$, this setting is under-determined. However, under the completely reasonable premise that x is compressible, i.e., it is well approximated by k-sparse representations, the problem of recovering x becomes feasible. In other words, if we restrict ourselves to vectors which have at most k non zero co-efficients, i.e. $\|x\|_0 = k$ and $k \ll n$, we can efficiently search for solutions. In fact, Candes et al. [5, 6, 7] show that it is possible recover a k-sparse x *exactly* if the sensing matrix Φ exhibits the *Restricted Isometry Property* (RIP).

▶ **Definition 1.** A matrix Φ is said to exhibit (k, δ)-RIP iff $\forall x \in \mathbb{R}^n$ with $\|x\|_0 = k$

$$(1 - \delta)\|x\|_2 \leq \|\Phi x\|_2 \leq (1 + \delta)\|x\|_2$$

Please note that k might be referred to as 'order' and δ might be referred to as 'Restricted Isometry Constant (RIC)'. Candes et al. showed that it is possible to reconstruct a k-sparse x very efficiently, by solving the minimization problem

$$\min_{a \in \mathbb{R}^n} \|a\|_1 \qquad \text{subject to} \quad \Phi a = y$$

when we have RIP matrices[1] with $\delta < \sqrt{2} - 1$. Their work has resulted in significant effort towards both deterministic and probabilistic constructions of RIP matrices. Naturally, we want to be able to construct RIP matrices with k as large as possible, but deterministic constructions, such as those presented by Bourgain et al. [4] and DeVore [9], cannot produce RIP matrices of order much greater than $\sqrt{n}$. Deterministic constructions are far from achieving the orders achieved by probabilistic constructions. On the other hand, it has been proven that ± 1 symmetric Bernoulli matrices, or matrices formed by sampling from a Gaussian distribution $\mathcal{N}(0, \frac{1}{n})$, satisfy RIP with $k \in \Theta(n)$ [3, 21] with high probability.

The superiority of random constructions motivates the problem of certifying whether a randomly drawn matrix Φ, from any of the models mentioned above, exhibits RIP with the required parameters. If we find that a randomly drawn matrix is unsuitable to our purposes, we re-generate it and repeat the certification process. Terry Tao posted the following question on his blog [18]: "An alternate approach (to deterministic construction of RIP matrix), and one of interest in its own right, is to work on improving the time it takes to verify that a

---

[1] Henceforth, we shall use the phrase 'RIP matrices' instead of saying 'matrices exhibiting the RIP' every time. It is worth noting that Tao once used the abbreviation 'UUP', which stands for 'Uniform Uncertainty Principle', for what is now commonly known as RIP.

given matrix (possibly one of a special form) obeys the UUP (RIP)." In this paper, we prove that RIP certification is NP-hard to approximate in a *strong* sense assuming the truth of the Small-Set-Expansion Hypothesis.

We now state Small-Set-Expansion Hypothesis, which was proposed by Raghavendra and Steurer [14], and is one of the most important conjectures in complexity theory. In order to present the conjecture, we first define the expansion of a graph.

▶ **Definition 2.** Given a graph d-regular graph $G(V, E)$ with n vertices, we define the expansion of a non-empty set $S \subseteq V$ as

$$\phi_G(S) = \frac{|E(S, V - S)|}{d \cdot \min(|S|, |V - S|)}.$$

where $E(S, V - S)$ denotes the collection of edges of G that have one end in S and the other end in $V - S$. The expansion of the graph G is defined as the minimum expansion among all subset of its vertices:

$$\phi(G) = \min_{S \subseteq V} \phi(S).$$

For any $0 < \delta \leq 1/2$, we also define the minimum expansion among all subsets of size $\leq \delta n$ as

$$\phi_\delta(G) = \min_{\substack{S \subseteq V \\ S \leq \delta n}} \phi(S).$$

The Small Set Expansion conjecture states that:

▶ **Conjecture 3.** *For every $\epsilon > 0$, $\exists\, 0 \leq \delta \leq \frac{1}{2}$, such that it is NP-hard to distinguish between:*
- $\exists\, S \subseteq V$, with $|S| = \delta|V|$, such that $\phi_G(S) \leq \epsilon$
- $\forall S \subseteq V$, with $|S| \leq \delta|V|$, we have $\phi_G(S) \geq 1 - \epsilon$

1.1 Our Main Result

We give a gap preserving reduction from the Small-Set-Expansion problem to the RIP certification problem. More formally, we prove the following theorem:

▶ **Theorem 4.** *For any $0 \leq \delta \leq 1$, and $C \geq 1$, there exists k such that, given a matrix M it is* Small-Set-Expansion-hard[2] *to distinguish between:*

- *(Highly RIP) M is (k, δ)-RIP.*
- *(Far away from RIP) M is not $(k/C, 1 - \delta)$-RIP.*

We claim that our result has a very strong form, which we will justify in more detail a little later. Also, as corollaries, we have that

▶ **Corollary 5.** *Given a matrix M and k, it is* Small-Set-Expansion-hard *to distinguish whether the matrix is (k, δ)-RIP or not $(k, 1 - \delta)$-RIP for any $\delta > 0$*

▶ **Corollary 6.** *Given a fixed δ and matrix M, it is* Small-Set-Expansion-hard *to get a constant approximation for the smallest k such that M exhibits (k, δ)-RIP.*

---

[2] A problem $\mathcal{I}$ is defined to be Small-Set-Expansion-hard if Small-Set-Expansion $\leq_P \mathcal{I}$

1.2   Comparison with Previous Work

Let us go over some previous work on the topic of hardness of RIP certification, and also make a few observations about Theorem 4 to justify our claim that we are proving hardness of RIP certification in a very strong sense. Bandeira et al. [2] prove that the exact decision version of the problem of RIP certification is NP-hard. In other words, they proved that given δ, k, it is NP-hard to certify whether a matrix exhibits (k, δ)-RIP or not. It was later established by Tillmann and Pfetsch [19] that the same problem is also co-NP-hard. Both works reduce from the problem of determining the spark of a matrix, which is known to be NP-hard. It should be mentioned that δ in both results is in $o_n(1)$, where n is the number of rows of the matrix. Also, we must note that the exact decision version of the problem is restrictive.

Results by Koiran and Zouzias (KZ) are the only works we know of on the approximation version of the problem. KZ obtain various inapproximability results by making assumptions on the hidden clique problem [12] and the dense subgraph problem [11]. Most of the results are of the form that, for some k, δ_1, δ_2, (depending on the assumption used), it is hard to distinguish whether a matrix is (k, δ_1)-RIP, or not (k, δ_2)-RIP. In almost all of the cases, δ_1, δ_2 are $\in o_n(1)$, with the exception of one result, which we shall state below:

- No polynomial time algorithm can distinguish matrices that satisfy the $(\Theta(n), \frac{\kappa}{2})$-RIP from matrices that do not satisfy the $(\Theta(n), \kappa)$-RIP

where $\kappa \leq \frac{\sqrt{5}}{3}$ is an unknown constant depending on the correctness of hidden-clique and densest-subgraph conjectures. In practice, it is known that an RIP matrix is useful for many applications as long as $\delta \leq \sqrt{2} - 1$. Clearly, the above theorem does not rule out the existence of an algorithm for deciding whether the RIC of a matrix is $\leq \sqrt{2} - 1$. This is because there is no guarantee that $\kappa \in (\sqrt{2} - 1, 2\sqrt{2} - 2)$. KZ also state "...our hardness results do not rule out the existence of a polynomial-time algorithm distinguishing between matrices with a very small RIP parameter and matrices with a RIP parameter larger than say 0.1...".

Theorem 4 is clearly equipped to make stronger statements than any previous work on inapproximability of RIP parameters. It suggests that certifying RIP for any constant $0 < \delta < 1$ is SMALL-SET-EXPANSION-hard. In addition, even if the matrix indeed exhibits *strong* RIP (small constant δ and very large k), it is still SMALL-SET-EXPANSION-hard to even certify if it exhibits *weak RIP* (with large δ close 1 and small k).

1.3   Proof Overview

Let us assume that G is a d-regular graph with adjacency matrix A, and $L = I - \frac{1}{d}A$ is the normalized Laplacian matrix of the graph. it is easy to see that, given $x_S \in \{0, 1\}^n$ as the indicator vector of set S, we have that

$$\phi(S) = \frac{x_S^T L x_S}{x_S^T x_S} = \frac{\|M x_S\|_2^2}{\|x_S\|_2^2}$$

for M satisfying $M^T M = L$. We know that the Laplacian is a quadratic form and thus is positive semi-definite. Thus, L always admits the decomposition $L = M^T M$.

The strategy of the reduction is to take a SMALL-SET-EXPANSION instance and to construct the corresponding M for the RIP certification problem. Our reduction has a similar flavor to the reduction in Koiran and Zouzias [12] (they call their reduction as *Cholesky Reduction*). If there is a small set S with expansion less than ϵ, we know that x_S,

corresponding to this S, is a sparse vector such that $\|Mx_S\|_2 \leq \sqrt{\epsilon}\|x_S\|_2$ and this suggests that M is far from being a RIP matrix. The second case of the proof is more involved. Here, we would like to show that if there exists a k-sparse $x \in \mathbb{R}^n$ such that $x^{\mathrm{T}}Lx$ is bounded away from 1, then we can find a small set S such that $\phi(S)$ is also bounded away from 1.

If we ignore the sparsity constraint, this kind of conversion from a real vector x to a boolean vector x_S is exactly reminiscent of the "hard direction" of Cheeger's Inequality [1, 13, 16]. In this paper, we prove the following generalization of Cheeger's Inequality for sparse vectors, which we use to prove the "hard direction".

▶ **Theorem 7.** *(Sparse Cheeger's Inequality) Let A be the adjacency matrix of a d-regular graph G, and $L = I - \frac{1}{d}A$ be its normalized Laplacian matrix. For any $\delta \leq 1/2$, we have that*

$$\lambda_\delta \leq \phi_\delta(G) \leq \sqrt{(2 - \lambda_\delta)\lambda_\delta}$$

where $\lambda_\delta = \min_{|x|_0 = \delta|V|} \frac{x^{\mathrm{T}}Lx}{x^{\mathrm{T}}x}$.

The above inequality establishes a relationship between the minimum expansion of G on small sets with the minimum value of $\frac{x^{\mathrm{T}}Lx}{x^{\mathrm{T}}x}$ for sparse real vectors x. A similar and independent, but not identical, result is known – Theorem 2.1 in Steurer [17]. We observe that the constants in Theorem 7 are better, and this might find applications elsewhere.

For the purpose of comparison, we also list the original Cheeger's Inequality here:

▶ **Theorem 8.** *(Cheeger's Inequality) Let A be the adjacency matrix of a graph G, and $L = I - \frac{1}{d}A$ be its normalized Laplacian matrix. We have that*

$$\frac{\lambda}{2} \leq \phi(G) \leq \sqrt{2\lambda}$$

where

$$\lambda = \min_{\substack{x \in \mathbb{R}^n \\ Lx \neq \vec{0}}} \frac{\|x^T Lx\|_2}{\|x\|_2^2}$$

is the second smallest eigenvalue of L.

It is interesting to note here that the relationship between λ_δ and $\phi_\delta(G)$ in Theorem 7 is tighter than the relationship between λ and $\phi(G)$ in Theorem 8. It is crucial for our proof that we get $\sqrt{(2 - \lambda_\delta)\lambda_\delta}$ instead of $\sqrt{2\lambda_\delta}$ in Theorem 7. We need to prove that if there exists a δ-sparse vector x^3 such that λ_δ is bounded away from 1, then there is a small set whose expansion is also bounded away from 1. If what we had was only $\phi_\delta(G) \leq \sqrt{2\lambda_\delta}$, the bound becomes trivial even when we know that $\lambda_\delta = 1/2$. It is only because of Theorem 7 that we find that, as long as λ_δ is bounded away from 1, we also have that $\sqrt{(2 - \lambda_\delta)\lambda_\delta}$ is bounded away from 1. This turns out to be exactly what we need to prove.

The proof of the sparse Cheeger's Inequality bears resemblance to the proof of the classical Cheeger's Inequality (e.g., see [20]). The proof makes it necessary to strengthen the analysis for the sparse vector case so as to obtain a tighter relationship between λ_δ and $\Phi_\delta(G)$.

One final thing to notice is that our hardness result amplifies the dependence on the order k. We show that it is hard to distinguish $(k, 1 - \delta)$-RIP from $(k/C, \delta)$-RIP, where C is any arbitrary constant. To this end, we need to use an equivalent statement of the SMALL-SET-EXPANSION HYPOTHESIS by Raghavendra et al. [15] which gives us a stronger starting point for the hardness reduction.

---

[3] $\|x\|_0 \leq \delta n$

1.4 Organization

The paper is organized as follows. In Section 2, we prove the sparse Cheeger's Inequality, namely Theorem 7. We then use this theorem to prove our main result, i.e Theorem 4 which is presented in Section 3.

2 Sparse Cheeger's Inequality

Below we state the proof of Theorem 7.

Proof. Assuming that $|V| = n$, let us first prove the left side of the inequality, analogous to what is commonly called the *easy direction*. Choose $x_S \in \{0,1\}^n$, as a bit vector representation of a set S of size at most $\delta|V|$. We then easily get

$$\phi_\delta(G) = \frac{x_S^T L x_S}{\|x_S\|^2} \leq \lambda_\delta$$

Next, we prove the right hand side of inequality in Theorem 7. Assume we are given any $x \in \mathbb{R}^n$ such that

$$\frac{x^T L x}{\|x\|_2^2} = \lambda_\delta$$

We shall prove that, using x, we will be able to construct some set S such that $\phi(S) \leq \sqrt{(2 - \lambda_\delta)\lambda_\delta}$. This will complete the proof because $\phi_\delta(G)$ is the minimum value of $\phi(S)$ over all S with $|S| \leq \delta n$.

Let us use x_i to indicate the i-th coordinate of x, by the property of the Laplacian matrix of a graph, we know that

$$\frac{x^T L x}{\|x\|_2^2} = \frac{\sum_{1 \leq i,j \leq n} (x_i - x_j)^2 A_{ij}}{2d \sum_{i=1}^n |x_i|^2} \tag{1}$$

Without loss of generality, we can assume that all coordinates x_i have nonnegative value because changing x_i to $|x_i|$ does not increase $\sum_{1 \leq i,j \leq n} (x_i - x_j)^2 A_{ij}$ and $\sum |x_i|^2$ remains the same. Also, we can scale x such that $0 \leq x_1 \leq x_2 \ldots \leq x_n = 1$ because $\frac{x^T L x}{\|x\|_2^2}$ does not change when we scale x.

Consider a distribution P with density $f(x) = 2x$ on the interval $(0,1)$. It is easy to verify this is a valid density function as

$$\int_0^1 2x \cdot \mathrm{d}x = 1.$$

Now consider the following randomized construction of set S from x.

1. Choose t in $(0,1)$ according to P
2. Set S to be $S_t = \{i \mid x_i \geq t\}$

Given the fact that x is δn sparse, $|S_t| \leq \delta n$ for any $0 < t < 1$. We can easily see that for any $0 \leq a \leq b \leq 1$, we have

$$\Pr(t \in [a,b]) = \int_a^b 2x dx = b^2 - a^2.$$

Therefore,

$$\Pr(x_i \in S_t) = \Pr(x_i \geq t) = x_i^2$$

which implies that

$$\mathbf{E}_t\left[|S_t|\right] = \sum_{i=1}^{n} x_i^2$$

Also we know that for any i, j, there is an edge between vertex i and vertex j only if t falls between x_i and x_j. Therefore,

$$\mathbf{E}_t\left[|E(S_t, V - S_t)|\right] = \frac{1}{2}\sum_{i,j} A_{ij}|x_i^2 - x_j^2| = \frac{1}{2}\sum_{1 \le i,j \le n} A_{ij}|x_i - x_j||x_i + x_j|$$

$$\le \sqrt{\frac{\sum_{1 \le i,j \le n} A_{ij}(x_i + x_j)^2}{2}} \cdot \sqrt{\frac{\sum_{1 \le i,j \le n} A_{ij}(x_i - x_j)^2}{2}}$$

The last inequality in the above sequence of steps is due to the Cauchy-Schwarz Inequality. We then calculate ratio between $\mathbf{E}_t\left[|E(S_t, V - S_t)|\right]$ and $d \cdot \mathbf{E}_t\left[|S_t|\right]$, which is

$$\frac{\mathbf{E}_t\left[|E(S_t, V - S_t)|\right]}{d \cdot \mathbf{E}\left[|S_t|\right]} \le \sqrt{\frac{\sum_{1 \le i,j \le n} A_{ij}(x_i + x_j)^2}{2d\sum_{i=1}^n x_i^2}}\sqrt{\frac{\sum_{1 \le i,j \le n} A_{ij}(x_i - x_j)^2}{2d\sum_{i=1}^n x_i^2}}.$$

By Equation (1), we know that

$$\frac{\sum_{1 \le i,j \le n} A_{ij}(x_i - x_j)^2}{2d\sum_{i=1}^n x_i^2} = \frac{x^T L x}{|x|^2} = \lambda_\delta$$

We also know that

$$\frac{\sum_{1 \le i,j \le n} A_{ij}(x_i + x_j)^2}{2d\sum_{i=1}^n x_i^2} + \frac{\sum_{1 \le i,j \le n} A_{ij}(x_i - x_j)^2}{2d\sum_{i=1}^n x_i^2} = \frac{\sum_{1 \le i,j \le n} A_{ij}(2x_i^2 + 2x_j^2)}{2d\sum_{i=1}^n x_i^2} = 2$$

which implies that

$$\frac{\sum_{1 \le i,j \le n} A_{ij}(x_i + x_j)^2}{2d\sum_{i=1}^n x_i^2} = 2 - \lambda_\delta$$

This suggests that

$$\frac{\mathbf{E}_t[|E(S_t, V - S_t)|]}{d \cdot \mathbf{E}_t[|S_t|]} \le \sqrt{\lambda_\delta(2 - \lambda_\delta)}$$

or equivalently

$$\mathbf{E}_t\left[|E(S_t, V - S_t)| - \sqrt{\lambda_\delta(2 - \lambda_\delta)} \cdot d \cdot |S_t|\right] \le 0$$

Therefore, there must exist some $t \in (0, 1)$ such that

$$|E(S_t, V - S_t)| - \sqrt{\lambda_\delta(2 - \lambda_\delta)} \cdot d \cdot |S_t| \le 0$$

or in other words

$$\phi(S_t) = \frac{|E(S, V - S)|}{d \cdot |S_t|} \le \sqrt{\lambda_\delta(2 - \lambda_\delta)}$$

Therefore, if we choose the best $t \in (0, 1)$, we know that

$$\min_t \phi(S_t) \le \sqrt{\lambda_\delta(2 - \lambda_\delta)}$$

This finishes the proof for the right hand side of the inequality in Theorem 7. ◀

3 Proof for the Hardness of Certifying RIP

3.1 Equivalent Variant of the Small-Set-Expansion Hypothesis

The starting point is the following Theorem 9 from [15], which states that a strengthened form of Small-Set-Expansion Hypothesis is equivalent to the original Small-Set-Expansion Hypothesis.

▶ **Theorem 9.** *Given a d-regular graph $G(V, E)$, for all constant integer $q \in \mathbb{N}$, and any constant $\gamma, \epsilon \in (0, 1)$, it is* Small-Set-Expansion-hard *to distinguish the following two cases:*

- *there are q disjoint sets $S_1, S_2, \ldots, S_q \subseteq V$ of size $\frac{n}{q}$ such that $\phi(S) \leq \epsilon + o(\epsilon)$.*
- *for any $0 < \beta < 1$, every set of $S \leq \beta n, S \subseteq V$ has expansion at least $1 - \frac{T_{1-\epsilon}(\beta)}{\beta} - \gamma/\beta$*

Here $T_{1-\epsilon}$ is related to the Gaussian Stability function, which is defined by Khot et al. in [10]. We will use the following upper bound that was presented in [10]:

$$\frac{T_{1-\epsilon/2}(\beta)}{\beta} \leq \beta^{\epsilon/4} \tag{2}$$

for any β, ϵ. By putting $\beta = (\epsilon)^{4/\epsilon}, \alpha = \beta/C, q = 1/\alpha, \gamma = \epsilon^{4/\epsilon+2}$ in inequality (2), we have that

$$\frac{T_{1-\epsilon}(\beta)}{\beta} + \gamma/\beta \leq \epsilon + o(\epsilon)$$

and the following hardness statement of Small-Set-Expansion:

▶ **Theorem 10.** *For any $0 < \epsilon < 1$, and an arbitrarily large constant C, there exists some $k < n$ (functionally dependent on ϵ), for which it is* Small-Set-Expansion-hard *to distinguish the following two cases in a d-regular graph $G(V, E)$:*

- *there is a set $S \subseteq V$ of size k/C such that $\phi(S) \leq O(\epsilon)$*
- *every set $S \subseteq V$ of size less than k has expansion at least $1 - O(\epsilon)$*

3.2 Hardness Reduction

We shall make a gap preserving reduction from the Small-Set-Expansion hardness of Theorem 10. Given any d-regular graph with adjacency matrix A, will consider matrix M such that $M^{\mathrm{T}}M = I - \frac{1}{d}A$ for the RIP certification problem. Also without loss of generality, we can only prove for δ that is sufficiently small constant as if Theorem 4 holds for some $\delta = \delta_0$, it also holds for all $\delta \geq \delta_0$. In order to prove Theorem 4, it suffices to prove the following Lemma 11.

▶ **Lemma 11.** *Let $\delta = \epsilon^{0.4}$ for a sufficiently small constant ϵ. Then:*
1. *If there is a set S of size at most k/C and $\phi_G(S) \leq O(\epsilon)$, then the matrix is M not $(k/C, 1 - \delta)$-RIP*
2. *If for every set S of size at most k, $\phi_G(S) \geq 1 - O(\epsilon)$, then M is (k, δ)-RIP*

The proof of Lemma 11 would complete the proof of Theorem 4.

Proof. Given any d-regular graph with adjacency matrix A, let $x_S \in \{0, 1\}^n$ be the indicator vector of a subset S. We know the number of edges that leave S is equal to $d \cdot |S| - x_S^{\mathrm{T}} A x_S = x_S^{\mathrm{T}}(d \cdot I - A)x_S$. Therefore, we have

$$\phi_G(S) = \frac{x_S^{\mathrm{T}}(d \cdot I - A)x_S}{d|S|} = \frac{x_S^{\mathrm{T}}(d \cdot I - A)x_S}{d\|x_S\|_2^2} = \frac{x_S^{\mathrm{T}}(I - A/d)x_S}{\|x_S\|_2^2} = \frac{x_S^{\mathrm{T}} M^{\mathrm{T}} M x_S}{\|x_S\|_2^2}$$

Let us prove the first claim. We know that when there is a set $S \subseteq V$ of size less than $\frac{k}{C}$ that has expansion less than $O(\epsilon)$. Let us denote $x_S \in \{0, 1\}^n$ as the indicator of set S, then

$$\frac{x_S^{\mathrm{T}} \cdot M^{\mathrm{T}} M(x_S)}{\|x_S\|^2} \leq O(\epsilon)$$

which implies that

$$\|Mx\|_2 \leq O(\sqrt{\epsilon})\|x\|_2$$

Since x_S is k/C-sparse, after applying M, its length is only $O(\sqrt{\epsilon})$ times $\|x\|_2$. Now, given that we know $\delta = \epsilon^{0.4}$, for sufficiently small ϵ, we have that M is not $(k/C, \delta)$-RIP.

We shall prove the second claim of Lemma 11 by contradiction. Suppose there exists some k-sparse vector $x \in \mathbb{R}^n$ such that

$$\|Mx\|_2 \leq (1 - \delta)\|x\|_2$$

We know then that

$$x^{\mathrm{T}} M^{\mathrm{T}} M x \leq (1 - \delta)^2 \|x\|_2,$$

which implies that $\lambda_\delta(G) \leq 1 - 2\delta + \delta^2$. Now, by Theorem 7, we have that there must exist a set such that the expansion is at most

$$\sqrt{\lambda_\delta(G)(2 - \lambda_\delta(G))} \leq \sqrt{1 - (2\delta - \delta^2)^2} = 1 - \Theta(\epsilon^{0.8}),$$

which contradicts the fact that all sets S of size less than k must have expansion at least $1 - \epsilon$.   ◀

4   Conclusion and Open Problems

In this paper, we establish that certifying RIP of a matrix (even approximately) is SMALL-SET-EXPANSION-hard in a strong sense. Although the SMALL-SET-EXPANSION problem is a conjecture, our work helps cement the place of RIP certification relative to other problems in regard to their hardness. In general, whenever we reduce from a known problem to a new problem, it increases the importance of the original problem.

One possible immediate open problems is to prove NP-hardness of RIP certification by reducing from known canonical problems. This would be interesting and important because the correctness of SMALL-SET-EXPANSION HYPOTHESIS is uncertain. Another interesting direction to pursue could be to prove that RIP certification is hard even when the matrix satisfies certain natural properties such as coherence.

──── **References** ────

1 Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.

2 Afonso S. Bandeira, Edgar Dobriban, Dustin G. Mixon, and William F. Sawin. Certifying the restricted isometry property is hard. *IEEE Transactions on Information Theory*, 59(6):3448–3450, 2013.

3 Richard Baraniuk, Mark Davenport, Ronald DeVore, and Michael Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3):253–263, 2008.

4 Jean Bourgain, Stephen J Dilworth, Kevin Ford, Sergei V Konyagin, and Denka Kutzarova. Breaking the k^2 barrier for explicit rip matrices. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, pages 637–644. ACM, 2011.

5 Emmanuel J Candes. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathematique*, 346(9):589–592, 2008.

6 Emmanuel J Candes, Justin K Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on pure and applied mathematics*, 59(8):1207–1223, 2006.

7 Emmanuel J Candes and Terence Tao. Decoding by linear programming. *Information Theory, IEEE Transactions on*, 51(12):4203–4215, 2005.

8 Mark A Davenport, Marco F Duarte, Yonina C Eldar, and Gitta Kutyniok. Introduction to compressed sensing. In *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.

9 Ronald A DeVore. Deterministic constructions of compressed sensing matrices. *Journal of Complexity*, 23(4):918–925, 2007.

10 Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for max-cut and other 2-variable csps? *SIAM Journal on Computing*, 37(1):319–357, 2007.

11 Pascal Koiran and Anastasios Zouzias. On the certification of the restricted isometry property. *CoRR*, abs/1103.4984, 2011.

12 Pascal Koiran and Anastasios Zouzias. Hidden cliques and the certification of the restricted isometry property. *CoRR*, abs/1211.0665, 2012.

13 George Pólya and Gabor Szego. *Isoperimetric inequalities in mathematical physics*. Princeton University Press, 1951.

14 Prasad Raghavendra and David Steurer. Graph expansion and the unique games conjecture. In *Proceedings of the 42nd ACM symposium on Theory of computing*, pages 755–764. ACM, 2010.

15 Prasad Raghavendra, David Steurer, and Madhur Tulsiani. Reductions between expansion problems. In *Computational Complexity (CCC), 2012 IEEE 27th Annual Conference on*, pages 64–73. IEEE, 2012.

16 Alistair Sinclair and Mark Jerrum. Approximate counting, uniform generation and rapidly mixing markov chains. *Information and Computation*, 82(1):93–133, 1989.

17 David Steurer. Subexponential algorithms for d-to-1 two-prover games and for certifying almost perfect expansion. *Available at the author's website*, 1:2–1, 2010.

18 Terence Tao. Open question: deterministic uup matrices, July 2007.

19 Andreas M. Tillmann and Marc E. Pfetsch. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing. *IEEE Transactions on Information Theory*, 60(2):1248–1259, 2014.

20 Luca Trevisan. Cs359g lecture 4: Spectral partitioning. CS359G Lecture 4: Spectral Partitioning.

21 Roman Vershynin. Introduction to the non-asymptotic analysis of random matrices. *arXiv preprint arXiv:1011.3027*, 2010.