Pittsburgh Interdisciplinary Mathematics Review
Volume 2, December 2024, Pages 89–105
https://doi.org/10.5195/pimr.2024.41

# On The Number of Robust Geometric Graphs in a Euclidean Space

Lufei Yang, Colin Yip, Madison Zhao, and Konstantin Tikhomirov

(Communicated by David Philips)

**Abstract**

We say that a graph $\Gamma = (V, E)$ with vertices in a $k$–dimensional Euclidean space is an $\varepsilon$–robust distance graph with threshold $\tau$ if any two vertices $v, w$ in $V$ are adjacent if and only if $\|v - w\|_2 \leq (1+\varepsilon)^{-1}\tau$ and are not adjacent if and only if $\|v - w\|_2 > (1+\varepsilon)\tau$. We show that there are universal constants $C', c', c > 0$ with the following property. For $k \geq C' d \log n$, asymptotically almost every $d$–regular graph on $n$ vertices is isomorphic to a $\frac{c}{\sqrt{d}}$–robust distance graph in $\mathbb{R}^k$, whereas for $k \leq \frac{c' d \log n}{\log d}$, a.a.e $d$–regular graph on $n$ vertices cannot be represented as an $\frac{c}{\sqrt{d}}$–robust distance graph.

## 1. Introduction

In this paper, we adopt the following definition. A graph $\Gamma = (V, E)$ with vertices in $\mathbb{R}^k$ is a **distance graph** (or a **geometric graph**[1]) with a threshold $\tau$ if for any $v, w$ in $V$, the vertices are adjacent if and only if $\|v - w\|_2 \leq \tau$.

Let $N$ be a very large integer, and $\mathbb{R}^N$ be a feature space (for example, images or pieces of text). We further suppose that there exists a map

$$\mathbf{Sim} : \mathbb{R}^N \times \mathbb{R}^N \to \{0, 1\},$$

which for every pair of data points $p_1, p_2 \in \mathbb{R}^N$ returns 1 if $p_1, p_2$ are *similar*, and 0 otherwise. As a concrete example, consider a professional network in which each person's account is represented by a vector in $\mathbb{R}^N$. If for two data points $p_1, p_2 \in \mathbb{R}^N$ we have $\mathbf{Sim}(p_1, p_2) = 1$ then the two persons are considered to have similar professional interests. The network algorithm can use the information for advertisement, suggesting new contacts, etc. As another example, consider a set of documents (for example, academic papers or homework assignments) represented as vectors in $\mathbb{R}^N$. If $\mathbf{Sim}(\mathrm{doc}_1, \mathrm{doc}_2) = 1$ then the documents are considered similar, which may be an indication of plagiarism.

[1] Note that there exist other definitions of geometric graphs in the literature. In particular, $k$ nearest neighbors graphs, planar graphs, dot product graphs are considered to be geometric graphs as well.

Often, the map **Sim** is either not explicitly defined or takes too much time to compute (for example, **Sim** can be an abstraction of an expert human). A standard approach in dealing with large feature spaces is to construct a dimension reduction map $f : \mathbb{R}^N \to \mathbb{R}^k$ (with $k$ much smaller than $N$) and evaluate similarities of data points by comparing the Euclidean distance between the corresponding "reduced" feature vectors with a threshold $\tau$. Let $k$ be an integer much smaller than $N$, and $f : \mathbb{R}^N \to \mathbb{R}^k$ be a linear or non-linear *dimension reduction* map into a "reduced" feature space $\mathbb{R}^k$. For a finite set of data points $D$ in $\mathbb{R}^N$, consider a distance graph $\Gamma$ with vertex set $V := f(D)$ and the edge set

$$E := \Big\{\{f(p_1), f(p_2)\} : \|f(p_1) - f(p_2)\|_2 \le \tau\Big\}.$$

The constructed dimension reduction map is *good* if for *typical* configurations $D$ of data points in $\mathbb{R}^N$ it correctly identifies the similarity network i.e

$$f(p) \text{ and } f(p') \text{ are adjacent in } \Gamma \text{ if and only if } \quad \mathbf{Sim}(p,p') = 1, \quad \text{for all } p,p' \in D. \tag{1}$$

Existence or non-existence of a good quality dimension reduction map $f$ obviously depends on the structure of the typical configurations $D$ of data points, and is necessarily problem-specific. At the same time, there are fundamental *information-theoretic* limits on how good a mapping $f$ can be, determined by the dimensions $N$ and $k$ and the vertex degrees and size of considered similarity networks.

To make the problem mathematically tractable, we have to define an explicit model of a "typical" similarity network in $\mathbb{R}^N$.

*Assumption.* We assume that a similarity network in $\mathbb{R}^N$ has $n$ vertices and is $d$–regular i.e the degree of every vertex is $d$. Further, we suppose that all topologies (isomorphism classes) of $n$–vertex $d$–regular graphs are equally likely to occur in a typical similarity network, i.e the network can be viewed as a uniform random $d$–regular graph.

As a necessary condition for the existence of a good quality reduction map $f$ into $\mathbb{R}^k$ in this setting is that $1 - o(1)$ fraction of $d$–regular graphs $G$ on $n$ vertices have the property that there is a distance graph $\Gamma$ in $\mathbb{R}^k$ isomorphic to $G$. Thus, the information-theoretic problem of the existence of a good mapping can be directly related to the following question:

**Problem 1.1.** *Given parameters $n, d, k$, does the set of all $d$–regular distance graph on $n$ vertices in $\mathbb{R}^k$ comprise $1 - o(1)$ fraction of all $d$–regular graphs on n vertices?*

The last problem has been addressed in the literature; in particular, it follows from a result in [5] that under the assumption $k \ge Cd \log n$ (for a large universal constant $C > 0$), the answer is positive i.e asymptotically almost all $d$–regular graphs are representable as Euclidean distance graphs in $\mathbb{R}^k$. On the other hand, it follows from [9] that for $k \le cd$, the number of Euclidean distance graphs in $\mathbb{R}^k$ is asymptotically infinitely small compared to the total number of labeled $d$–regular graphs on $n$ vertices. Note that there is a significant gap (of order $\log n$) between the two bounds.

While the above problem appears to be of considerable interest, the aforementioned framework completely avoids any discussion of *robustness* i.e resilience to noise. Assume for a moment that $f$ is a good quality dimension reduction map for a threshold $\tau$, i.e. (1) holds for a typical $n$–point configuration $D$ of data points in $\mathbb{R}^N$. Since (1) imposes no restrictions on how close to $\tau$ the lengths $\|f(p)-f(p')\|_2$ can be, even a small perturbation of $f$ can completely distort the similarity relation on the reduced feature vectors. For that reason, it is desirable that the distances in the distance network $\Gamma$ are "separated" from the threshold value, which motivates the following definition:

**Definition 1.2** ($\varepsilon$–robust distance graphs)**.** *Let $\varepsilon \geq 0$, $\tau > 0$ be parameters. Assume that a geometric graph $\Gamma = (V,E)$ in a Euclidean space has the property that any two vertices $v,w$ in $V$ are adjacent if and only if $\|v-w\|_2 \leq (1+\varepsilon)^{-1}\tau$ and are not adjacent if and only if $\|v-w\|_2 > (1+\varepsilon)\tau$. Then we call the graph $\Gamma$ an $\varepsilon$–**robust distance graph** with the threshold $\tau$.*

The dimension reduction maps $f$ are often defined as functions of data distribution. Since the precise distribution is typically unknown, an actually constructed map $\hat{f}$ is based on training data, and is only an approximation of the desired map $f$. Let $\varepsilon > 0$ be a parameter, and assume that $\hat{f}$ and $f$ satisfy

$$(1+\varepsilon)^{-1}\|\hat{f}(p)-\hat{f}(p')\|_2 \leq \|f(p)-f(p')\|_2 \leq (1+\varepsilon)\|\hat{f}(p)-\hat{f}(p')\|_2 \quad \text{for all } p,p' \in \mathbb{R}^N. \quad (2)$$

Now, if for a data set $D$ the mapping $f$ generates an $\varepsilon$–robust distance graph, i.e

$$\|f(p)-f(p')\|_2 \leq (1+\varepsilon)^{-1}\tau \quad \text{if and only if} \quad \mathbf{Sim}(p,p') = 1;$$
$$\|f(p)-f(p')\|_2 > (1+\varepsilon)\tau \quad \text{if and only if} \quad \mathbf{Sim}(p,p') = 0, \quad p,p' \in D,$$

then, by (2), the image of $D$ under $\hat{f}$ also gives an accurate similarity network. To summarise, the requirement that the distance graph generated by $f$ is $\varepsilon$–robust can make sure that the similarities are correctly identified even when the dimension reduction map is only known up to an $\varepsilon$–distortion.

For $\varepsilon$–robust networks, we can pose an information-theoretic problem similar to Problem 1.1 above:

**Problem 1.3.** *Given parameters $n,d,k,\varepsilon$, does the set of $d$–regular $\varepsilon$–robust distance graphs on $n$ vertices in $\mathbb{R}^k$ comprise $1-o(1)$ fraction of all $d$–regular graphs on $n$ vertices?*

Solving Problem 1.3 is a main goal of this project. The $d$–regular graphs are a popular mathematical model of sparse networks. Although the real networks are often not $d$–regular (i.e some vertices may have more neighbors than others), the model is useful since it combines common features of real-world networks (such as *expansion* properties) with relative ease of rigorous analysis.

The first main result of this paper is the following answer to problem 1.3, presented later as Theorem 6.1:

**Theorem.** *There are universal constants $\tilde{C}, c > 0$ with the following property. Suppose $k \geq \tilde{C} d \log n$. Then the set of $d$–regular $\frac{c}{\sqrt{d}}$–robust geometric graphs on $n$ vertices in $\mathbb{R}^k$ comprise $1 - o(1)$ fraction of all $d$–regular graphs on $n$ vertices.*

The second main result of this paper is a lower bound on $k$ in the setting that almost all $d$-regular graphs on $n$ vertices can be represented as robust graphs in $\mathbb{R}^k$:

**Theorem.** *Suppose $1 - o(1)$ fraction of $d$-regular graphs on $n$ vertices can be represented as $\frac{c}{\sqrt{d}}$-robust geometric graphs in $\mathbb{R}^k$, where $c$ is the small constant from the above theorem. Then $k \geq \frac{c'' d \log n}{\log d}$, for a small universal constant $c'' > 0$.*

It must be noted that to establish the second main result we apply two lemmas — Lemma 7.1 and Lemma 7.2 in Section 7 — without proof. We believe that both lemmas can be proved using established probabilistic tools.

## 2. Definitions

We start by recalling some standard definitions.

### 2.1. Graph-theoretic notions

**Definition 2.1** (Adjacent vertices)**.** *Two vertices $v, w$ in a graph $\Gamma = (V, E)$ are* **adjacent** *if $\{v, w\} \in E$, and not adjacent otherwise.*

**Definition 2.2** (Adjacency matrices)**.** *For a graph $\Gamma = (V, E)$ on $n$ vertices $v_1, v_2, \ldots, v_n$, the* **adjacency matrix** *$A$ of $\Gamma$ is the $n \times n$ matrix with entries*

$$a_{ij} = \begin{cases} 1, & \text{if } v_i \text{ and } v_j \text{ are adjacent;} \\ 0, & \text{otherwise.} \end{cases}$$

**Definition 2.3** (Isomorphism classes of graphs)**.** *Two graphs $\Gamma_1 = (V_1, E_1)$ and $\Gamma_2 = (V_2, E_2)$ are* **isomorphic** *if $|V_1| = |V_2|$, and there is a bijective mapping $\phi : V_1 \to V_2$ such that for any $v, w \in V_1$, $\{v, w\} \in E_1$ if and only if $\{\phi(v), \phi(w)\} \in E_2$. Graphs isomorphism is an equivalence relation, thereby splitting the collection of all graphs into isomorphism classes.*

**Definition 2.4** ($d$–regular graphs)**.** *Let $d \geq 3$. A graph $\Gamma = (V, E)$ is $d$–***regular***, if each vertex $v \in V$ has exactly $d$ neighbors (adjacent vertices), equivalently, the graph is $d$–regular if the adjacency matrix of $\Gamma$ has exactly $d$ ones in every row and column.*

**Definition 2.5** (Random d-regular graphs). *Let $\tilde{G}_{n,d}$ be the set of all d-regular graphs on n vertices. Let G be a random variable such that for any $G' \in \tilde{G}_{n,d}$, $\mathbb{P}(G = G') = \frac{1}{|\tilde{G}_{n,d}|}$: Then G is a* **random d-regular graph**.

### 2.2. Gaussian distributions

**Definition 2.6** (Gaussian Variables). *A random variable X is said to be Gaussian (or normally distributed) if it has a probability density function given by:*

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

*where $\mu$ is the mean and $\sigma^2$ is the variance of X.*

**Definition 2.7** (Covariance Matrix). *The covariance matrix $\Sigma$ is a symmetric matrix that encodes the variance of each component of a random vector and the covariance between each pair of components. Mathematically, it is defined as:*

$$\Sigma = \mathbb{E}[(\mathbf{X}-\mu)(\mathbf{X}-\mu)^T]$$

*where $\mathbf{X}$ is a random vector and $\mu$ is the mean vector. A diagonal covariance matrix $\Sigma = diag(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$ indicates that the components of $\mathbf{X}$ are uncorrelated, with variances $\sigma_i^2$.*

**Definition 2.8** (Gaussian Vectors). *A random vector $\mathbf{X} \in \mathbb{R}^n$ is said to be multivariate Gaussian (or normally distributed) if it has a probability density function given by:*

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

*where $\boldsymbol{\mu}$ is the mean vector and $\Sigma$ is the covariance matrix.*

**Definition 2.9** (Gaussian Matrices). *A random matrix $\mathbf{X} \in \mathbb{R}^{m\times n}$ is said to have a matrix normal distribution if its vectorized form* vec($\mathbf{X}$) *follows a multivariate normal distribution. Specifically, $\mathbf{X} \sim \mathcal{MN}(\mathbf{M}, \Sigma_r, \Sigma_c)$ if:*

$$\text{vec}(\mathbf{X}) \sim \mathcal{N}(\text{vec}(\mathbf{M}), \Sigma_c \otimes \Sigma_r)$$

*where $\mathbf{M}$ is the mean matrix, $\Sigma_r$ is the row covariance matrix, $\Sigma_c$ is the column covariance matrix, and $\otimes$ denotes the Kronecker product.*

## 3. Use $\sqrt{A + C\sqrt{d}\,\mathrm{Id}}$ to embed G into $\mathbb{R}^n$

Consider $\tilde{G}_{n,d}$ to be the set of labeled d-regular graphs on $\{1, 2, \ldots, n\}$. We denote the eigenvalues of an adjacency matrix of these d-regular graphs as $\lambda_1, \lambda_2, \ldots, \lambda_n$, where

$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$. First, we want to think about the second-largest eigenvalue under absolute value.

**Theorem 3.1** ([1, 4, 10]). *Let $G$ be a uniform random graph from $\tilde{G}_{n,d}$. Denote $\lambda_i(G)$ as the eigenvalues for the adjacency matrix of $G$. Then there is a universal constant $C > 0$ such that*

$$\mathbb{P}(\max(|\lambda_2(G)|, |\lambda_3(G)|, ..., |\lambda_n(G)|) \leq C\sqrt{d}) = 1 - o(1).$$

**Lemma 3.2.** *(Spectrum Shifting) Suppose we have matrix $A$. Denote the corresponding eigenvalues of $A$ as $\lambda_1, ..., \lambda_k$, where $Av_k = \lambda_k v_k$. Then, for matrix $A + C\,\mathrm{Id}_k$, we have eigenvalues being shifted, meaning for each $v_k$, $(A + C\,\mathrm{Id}_k)v_k = Av_k + Cv_k = \lambda_k v_k + Cv_k = (\lambda_k + C)v_k$. Therefore, $\lambda_k + C$ are the eigenvalues for matrix $A + C\mathrm{Id}_k$.*

**Corollary 3.3.** *Let $G$ be a uniform random graph from $\tilde{G}_{n,d}$. Denote $A$ be the corresponding adjacency matrix. If $C$ is sufficiently large, then with high probability, $A + C\sqrt{d}$ is a positive semi-definite matrix.*

**Definition 3.4** (Square Root of a Positive-Semi Definite Matrix). *Consider a positive-semi definite matrix $A$. By the definition of a positive-semi definite matrix, we know that $A = U\Lambda U^{-1}$ under eigendecomposition where $\Lambda$ is a diagonal matrix with non-negative eigenvalues on its diagonal and $U$ is a square matrix with the ith column is the eigenvector corresponding eigenvalue $\Lambda_{ii}$. Then, $A^{\frac{1}{2}} = U\Lambda^{\frac{1}{2}}U^{-1}$ and $\Lambda^{\frac{1}{2}}$ is just taking the square root of the eigenvalues on the diagonal.*

Since we are considering d-regular graph $G$, the corresponding matrix $A$ should be a random symmetric matrix of $G$. With the previous theorem, $A + C\sqrt{d}$ is a positive semi-definite matrix with $1 - o(1)$ probability. At the same time, the eigenvalues for the new matrix should be $\lambda_1 + C\sqrt{d}, ..., \lambda_n + C\sqrt{d}$ and $\geq 0$ with high probabilities.

By shifting the matrix by $C\sqrt{d}\,\mathrm{Id}_n$, we have a high probability that the matrix will be positive semi-definite and the eigenvalues are shifted together to preserve the relationships. At the same time this preserves the structure of the graph as much as possible. In this case, the ratio of distance between non-adjacent vertices over the distance between adjacent vertices will be $1 + \frac{C}{\sqrt{d}}$. So the choice of $\varepsilon$ will be of $\Theta(\frac{1}{\sqrt{d}})$. Later in Section 5, the choice of $k$ depends on $\varepsilon$, $k \geq \frac{C\log n}{\varepsilon^2} = \Theta(d\log n)$, having a lower bound with larger $\varepsilon$.

Considering the special properties of the positive-semi definite matrix, we naturally can consider $(A + C\sqrt{d})^{\frac{1}{2}}$.

**Theorem 3.5** (See [5] for a related theorem). *Let $G$ be a d-regular random graph with $n$ vertices. Then with the probability of $1 - o(1)$, the following statement holds. Let*

$$\{z_i\}_{i=1}^n = \left\{\sqrt{A + C\sqrt{d}\,\mathrm{Id}_n e_i}\right\}_{i=1}^n,$$

*where $e_i$ is the i-th standard basis vector in $\mathbb{R}^n$ and A is the adjacency matrix for G. Consider the geometric graph $\Gamma = (V_\Gamma, E_\Gamma)$ where:*

$$V_\Gamma = \{z_1, \dots, z_n\},$$
$$E_\Gamma = \Big\{\{z_i, z_j\} : \|z_i - z_j\|_2 \leq \tau\Big\}.$$

*Take $\tau :- \sqrt{2}C^{1/2}d^{1/4} - \frac{\sqrt{2}}{4C^{1/2}d^{1/4}}$. Then, $\Gamma$ is well-defined and isomorphic to G.*

*Proof.* For $i \neq j$:

$$\begin{aligned}\langle z_i, z_j\rangle &= e^T \sqrt{A + C\sqrt{d}\,\mathrm{Id}_n}^T \sqrt{A + C\sqrt{d}\,\mathrm{Id}_n} e_j \\ &= e^T (A + C\sqrt{d}\,\mathrm{Id}_n) e_j \\ &\quad \text{(since the matrix is positive-semi definite symmetric matrix)} \\ &= e_i^T A e_j + C\sqrt{d} e_i^T e_j \\ &= e_i^T A e_j \\ &= A_{ij} \qquad \text{(which is the } i,j\text{th entry of A)}\end{aligned}$$

For $i = j$:

$$\langle z_i, z_i\rangle = e_i^T A e_i + C\sqrt{d} e_i^T e_i = C\sqrt{d} = \|z_i\|_2^2.$$

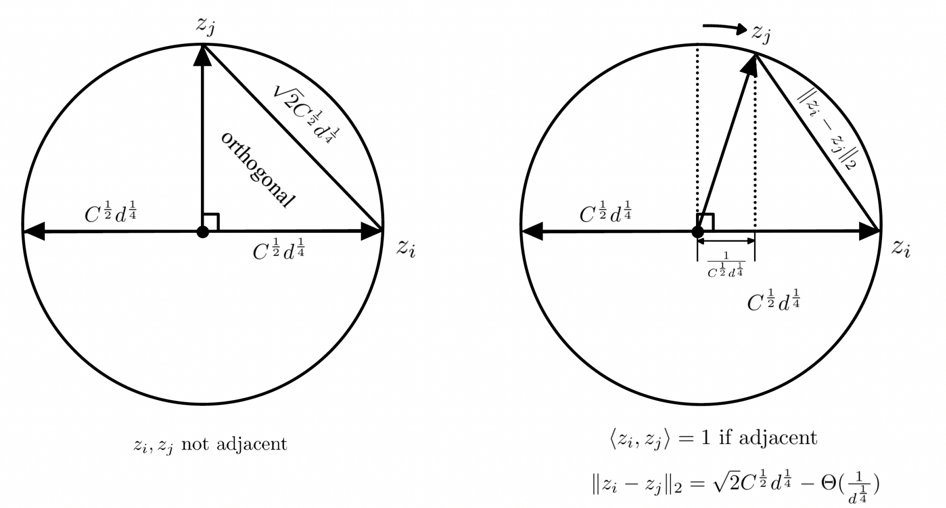


Figure 1: Visualizing how we construct an isomorphic geometric graph in the proof.

**Recall:** We obtain points $z_1, z_2, ..., z_n$ such that (See in Figure 1)

- $\|z\|_i = C^{1/2}d^{1/4}\ \forall i \in [n]$
- $z_i, z_j$ are orthogonal if i is not adjacent to j
- $\langle z_i, z_j\rangle = 1$ if i is adjacent to j

**Choosing threshold $\tau$:** Next, we want to choose a threshold $\tau$ to prove that the resulting geometric graph is isomorphic to $G$.

Consider the graph $\Gamma = (V_\Gamma, E_\Gamma)$ where:

$$V_\Gamma = \{z_1, \dots, z_n\},$$

$$E_\Gamma = \Big\{\{z_i, z_j\} : \|z_i - z_j\|_2 \leq \tau\Big\}.$$

We want to choose $\tau$ such that $\Gamma$ and $G$ are isomorphic under the map $i \mapsto z_i$.

If i is not adjacent to j, then by Pythagorean Theorem $\|z_i - z_j\|_2 = \sqrt{2}C^{1/2}d^{1/4}$. If i is adjacent to j, then $\langle z_i, z_j\rangle = 1$. Then

$$\begin{aligned} \|z_i - z_j\|_2^2 &= \langle z_j - z_i, z_i - z_j\rangle \\ &= \langle z_i, z_i\rangle + \langle z_j, z_j\rangle - 2\langle z_i, z_j\rangle \\ &= 2C\sqrt{d} - 2 \\ &\Rightarrow \|z_i - z_j\| = \sqrt{2}\sqrt{C\sqrt{d} - 1} \end{aligned} \tag{3}$$

Consider the geometric graph $\Gamma = (V_\Gamma, E_\Gamma)$, with $V_\Gamma = \{z_i\}_{i\in[n]}$ and $E_\Gamma = \{(z_i, z_j) \mid \|z_i - z_j\|_2 \leq \tau\}$. Take $\tau :- \sqrt{2}C^{1/2}d^{1/4} - \frac{\sqrt{2}}{4C^{1/2}d^{1/4}}$. This is roughly the mean of $\sqrt{2}C^{1/2}d^{1/4}$ and $\sqrt{2}\sqrt{C\sqrt{d}-1}$ (we used the first order Taylor approximation to estimate $\sqrt{2}\sqrt{C\sqrt{d}-1}$).

Thus, $\Gamma$ is a geometric graph in $\mathbb{R}^n$ isomorphic to $G$. □

# 4. Gaussian Concentration

## 4.1. Rotational Invariance

Consider a Gaussian vector $g = (g_1, g_2, \dots, g_n)$ where each component $g_i$ is independently and identically distributed according to the standard normal distribution $N(0,1)$. In this case, $g$ has a multivariate normal distribution $N(0, \mathrm{Id}_n)$, where $\mathrm{Id}_n$ is the $n$-dimensional identity matrix serving as the covariance matrix.

If $M$ is any orthogonal matrix, then the transformed vector $Mg$ also follows a standard Gaussian distribution. This can be shown through the properties of orthogonal matrices and Gaussian distributions:

- **Preservation of Dot Product**: Orthogonal matrices preserve dot products, meaning that $M^T M = \mathrm{Id}_n$. This preservation implies that the transformed vector $Mg$ retains the same length and statistical properties as $g$.

- **Covariance of Transformed Vector**: The covariance matrix of the transformed vector $Mg$ can be calculated as:

$$\mathrm{Cov}(Mg) = M\mathrm{Cov}(g)M^T = M\mathrm{Id}_n M^T = \mathrm{Id}_n$$

This shows that the covariance matrix of $Mg$ remains the identity matrix, confirming that $Mg$ is still a standard Gaussian vector.

### 4.2. Tail bound by Gaussian Concentration Inequality

**Definition 4.1.** *A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be Lipschitz continuous if there exists a constant $L \geq 0$ (called the Lipschitz constant) such that for all points $x, y$ in the domain of $f$, the following inequality holds:*

$$|f(x) - f(y)| \leq L\|x - y\|_2,$$

*where $\|x - y\|_2$ denotes the Euclidean distance between $x$ and $y$.*

**Theorem 4.2** (See, for example, [7])**.** *The Gaussian concentration inequality states that for a Lipschitz function $f$ with Lipschitz constant $L$, and a standard Gaussian vector $g$, the following inequality holds:*

$$\mathbb{P}(|f(g) - \mathbb{E}[f(g)]| \geq t) \leq 2\exp\left(-\frac{t^2}{2L^2}\right)$$

*This inequality is particularly useful for functions like the Euclidean norm, which are inherently 1-Lipschitz.*

### 4.3. Applying the Inequality to the Euclidean Norm

For the Euclidean norm of a standard Gaussian vector $g$, since it is a 1-Lipschitz function (i.e., $L = 1$), we can apply the Gaussian concentration inequality to obtain

$$\mathbb{P}\left(|\|g\|_2 - \mathbb{E}[\|g\|_2]| \geq t\right) \leq 2\exp\left(-\frac{t^2}{2}\right), \quad t > 0.$$

This inequality tells us how the norm of a Gaussian vector deviates from its expected value, providing insights into the spread of the norm values around their mean, which is approximately $\sqrt{n}$.

### 4.4. Variance of the Euclidean Norm of a Gaussian Vector

Consider a Gaussian vector $G = (G_1, G_2, \dots, G_n)$, where each component $G_i \sim N(0,1)$ independently. The Euclidean norm of $G$ is given by

$$\|G\|_2 = \sqrt{G_1^2 + G_2^2 + \cdots + G_n^2}.$$

The expected value of $\|G\|_2$ can be approximated by

$$\mathbb{E}[\|G\|_2] = \sqrt{n}.$$

The variance of $\|G\|_2$ can be derived from the properties of the chi-squared distribution:

$$\text{Var}(\|G\|_2) = \text{Var}(\sqrt{\chi_n^2}) \approx \frac{1}{2}.$$

Given that $\|G\|_2^2$ follows a chi-squared distribution with $n$ degrees of freedom:

$$\mathbb{E}[\|G\|_2^2] = n \quad \text{and} \quad \text{Var}(\|G\|_2^2) = 2n.$$

Thus, the variance of the Euclidean norm for large $n$ is:

$$\text{Var}(\|G\|_2) \approx \frac{1}{2}.$$

## 5. The Johnson–Lindenstrauss Lemma

We introduce The Johnson–Lindenstrauss (J-L) lemma for dimensionality reduction. It asserts that a set of points in a high-dimensional space can be embedded into a much lower-dimensional space such that the pairwise distances between the points are approximately preserved. This helps us handling high-dimensional data, where working directly in the original space may be computationally infeasible or inefficient.

Consider a set of points $X_1, X_2, \dots, X_n$ in a high-dimensional space $\mathbb{R}^N$. These points $X_i$'s represent the original data in the high-dimensional feature space. The goal is to find a new set of points $y_1, y_2, \dots, y_n$ in a much lower-dimensional space $\mathbb{R}^k$ (with $k \ll N$) such that the distances between the points are approximately preserved. The $X_i$'s define the relationships and distances in the original space, which we aim to maintain in the reduced space.

Mathematically, for all $i, j \leq n$, we want:

$$\|X_i - X_j\|_2 \approx \|y_i - y_j\|_2$$

This means that the distance between any pair of points $X_i$ and $X_j$ in the original space should be approximately equal to the distance between the corresponding points $y_i$

and $y_j$ in the reduced space. The J-L lemma provides the theoretical guarantee that such an embedding is possible with high probability, given certain conditions on the dimensionality $k$.

The key idea is to use a random projection matrix to map the high-dimensional points $X_i$'s into the lower-dimensional space, thereby creating the points $y_i$'s. This random projection is achieved using a matrix with entries drawn from a Gaussian distribution, ensuring that the pairwise distances are preserved up to a small error.

**Theorem 5.1** (The JL Lemma [2, 3, 6, 8]). *Let $X_1, X_2, \ldots, X_n \in \mathbb{R}^N$ be points in a feature space $\mathbb{R}^N$. For any $\varepsilon \in (0,1)$ and $k \geq \frac{C \log n}{\varepsilon^2}$, there exists a set of points $y_1, \ldots, y_n$ in $\mathbb{R}^k$ such that for all $i, j$ we have*

$$(1-\varepsilon)\|X_i - X_j\|_2 \leq \|y_i - y_j\|_2 \leq (1+\varepsilon)\|X_i - X_j\|_2.$$

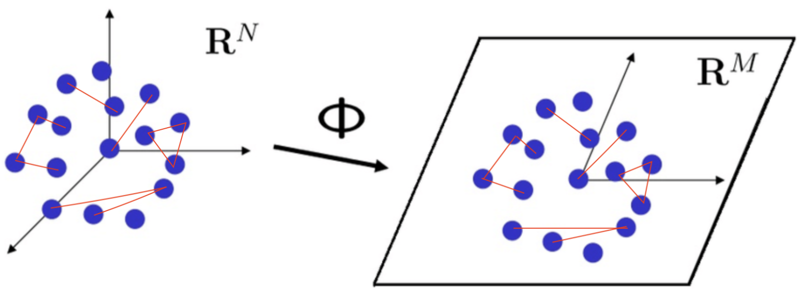


Figure 2: Johnson-Lindenstrauss lemma's illustration: shows how a geometric graph in a high-dimensional space $\mathbb{R}^3$ can be projected to a lower-dimensional space $\mathbb{R}^2$ without changing the topology.

*Proof.* Let $G$ be a $K \times N$ standard Gaussian matrix (all entries are independent $N(0,1)$).

Consider points $GX_1, GX_2, \ldots, GX_n \in \mathbb{R}^K$. Fix $i \neq j$, and consider

$$\|G(X_i - X_j)\|_2.$$

Note that $G(X_i - X_j)$ is a $k$-dimensional random vector. By properties of Gaussian matrices, if $y_1, y_2, \ldots, y_n$ are jointly Gaussian, then any linear combination $\sum_{i=1}^n a_i y_i$ is a Gaussian variable for any choice of numbers $a_1, \ldots, a_n$.

$$\mathbb{E}[G(X_i - X_j)] = \mathbb{E}[G](X_i - X_j) = \vec{0}.$$

The covariance matrix $\Sigma$ of $V = G(X_i - X_j)$ is:

$$\Sigma = \mathbb{E}[(G(X_i - X_j))(G(X_i - X_j))^T] = \|X_i - X_j\|_2^2 \, \mathrm{Id}_k,$$

where $\mathrm{Id}_k$ is the $k \times k$ identity matrix.

Using concentration inequalities for Gaussian distributions:

$$\mathbb{P}\left(\left|\|G(X_i - X_j)\|_2 - \sqrt{K}\|X_i - X_j\|_2\right| \geq t\right) \leq 2\exp\left(-\frac{t^2}{2\|X_i - X_j\|_2^2}\right), \quad t \geq 0.$$

Let $t = \varepsilon\sqrt{K}\|X_i - X_j\|_2$:

$$\mathbb{P}\left(\left|\|G(X_i - X_j)\|_2 - \sqrt{K}\|X_i - X_j\|_2\right| \geq \varepsilon\sqrt{K}\|X_i - X_j\|_2\right) \leq 2\exp\left(-\frac{\varepsilon^2 K}{2}\right).$$

Choosing $K \geq \frac{10\log n}{\varepsilon^2}$ ensures:

$$\mathbb{P}\left(\left|\|G(X_i - X_j)\|_2 - \sqrt{K}\|X_i - X_j\|_2\right| \geq \varepsilon\sqrt{K}\|X_i - X_j\|_2\right) \leq \frac{1}{n^3}.$$

Applying the union bound over all pairs $(i, j)$:

$$\begin{aligned}\mathbb{P}\left(\exists(i,j) : \left|\|G(X_i - X_j)\|_2 - \sqrt{K}\|X_i - X_j\|_2\right| \geq \varepsilon\sqrt{K}\|X_i - X_j\|_2\right) &\leq \binom{n}{2} \cdot \frac{1}{n^3} \\ &\leq \frac{n^2}{2} \cdot \frac{1}{n^3} = \frac{1}{2n}.\end{aligned}$$

Therefore, with high probability:

$$(1 - \varepsilon)\sqrt{K}\|X_i - X_j\|_2 \leq \|G(X_i - X_j)\|_2 \leq (1 + \varepsilon)\sqrt{K}\|X_i - X_j\|_2.$$

Let $y_i = \frac{GX_i}{\sqrt{K}}$ for $i = 1, 2, \ldots, n$. Then:

$$(1 - \varepsilon)\|X_i - X_j\|_2 \leq \|y_i - y_j\|_2 \leq (1 + \varepsilon)\|X_i - X_j\|_2.$$

□

## 6. Embedding for robust graphs: a sufficient condition on $k$

**Theorem 6.1.** *Let $G = (V, E)$ be a random d-regular graph on n vertices. Let $k \geq \tilde{C} d \log n$, where $\tilde{C}$ is a large universal constant. With probability 1-o(1) there exists a $\frac{c}{\sqrt{d}}$-robust geometric graph in $\mathbb{R}^k$ isomorphic to G, where c is a small universal constant.*

**Remark 6.2.** By definition of an $\tilde{\varepsilon}$-robust graph proving the lemma requires that we show the following: There exists a geometric graph $\Gamma' = \left(V'_\Gamma, E'_\Gamma\right)$, $V'_\Gamma = \{v'_1, v'_2, \ldots, v'_n\}$ in $\mathbb{R}^k$ and a distance threshold $\tau$ such that $\forall i, j \in V$:

1. i and j are adjacent $\Leftrightarrow \|v'_i - v'_j\|_2 \leq \tau\left(1 + \frac{c}{\sqrt{d}}\right)^{-1}$.

2. i and j are not adjacent $\Leftrightarrow \|v_i' - v_j'\|_2 > \tau\left(1 + \frac{c}{\sqrt{d}}\right)$.

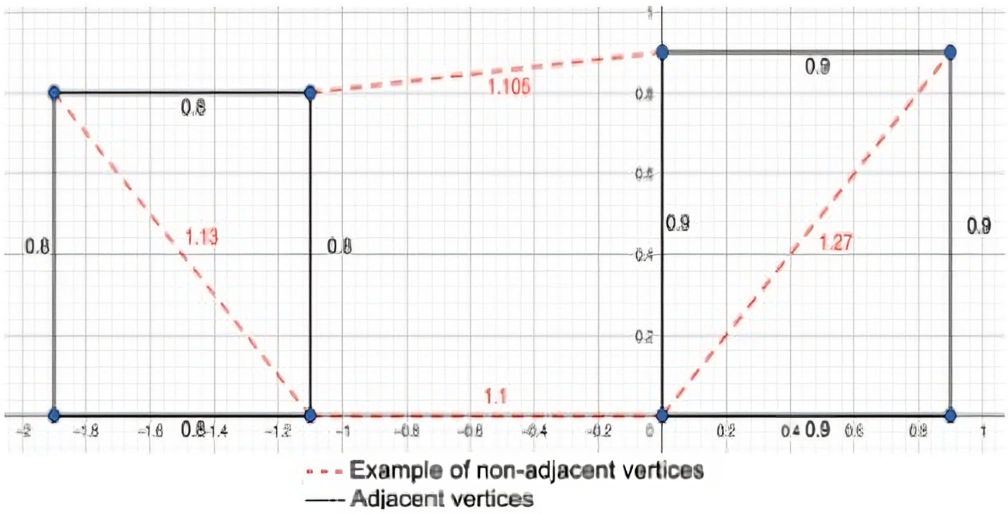


Figure 3: An embedding of a graph $G$ as a 0.1-robust distance graph in $\mathbb{R}^2$ ($\tau = 1$) and the distance between any two adjacent vertices is at most 0.9. The distance between any two non-adjacent vertices is at least 1.1.

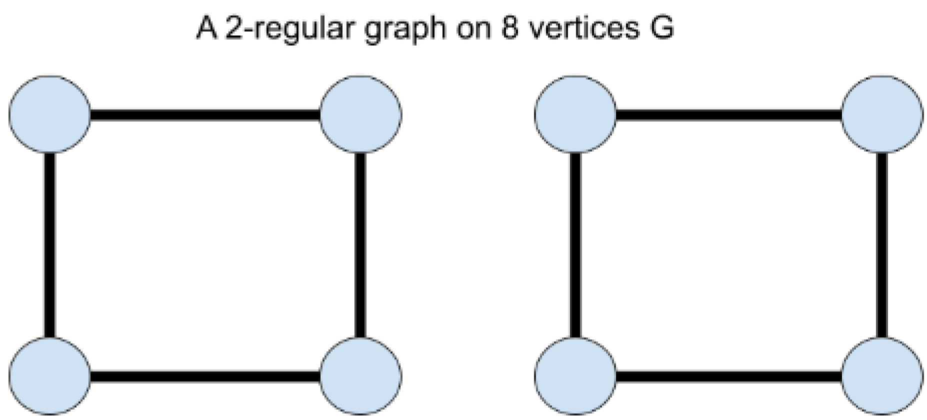


Figure 4

*Proof.* Let $G$ be a random $d$-regular graph on $n$ vertices. Consider the geometric graph in $\mathbb{R}^k$ $\Gamma' = \left(V_\Gamma', E_\Gamma'\right)$ where the vertices $V_\Gamma' = \{v_1', v_2', \ldots, v_n'\}$ are constructed by first embedding G into $\mathbb{R}^n$ (as in Section 3) and then embedding the resulting geometric graph into $\mathbb{R}^k$ using the Johnson-Lindenstrauss Lemma (This process is shown in Section 5). Use $\tau = \sqrt{2}\tilde{C}^{\frac{1}{2}} d^{\frac{1}{4}} - \frac{\sqrt{2}}{4\tilde{C}^{\frac{1}{2}} d^{\frac{1}{4}}}$ as the distance threshold.

To apply the Johnson–Lindenstrauss Lemma with a distortion $\varepsilon$ it is required that

$$k \geq \frac{C \log(n)}{\varepsilon^2}.$$

Let $\varepsilon = \frac{c}{\sqrt{d}}$, for some small constant $c$. Then assuming that $\tilde{C}$ from Lemma 6.1 is sufficiently large, the required inequality is satisfied.

First, let us check that property 1 from Remark 6.2 is satisfied. Let $i$ and $j$ be adjacent vertices. Now, by the Johnson–Lindenstrauss Lemma:

$$\begin{aligned}
\|v_i' - v_j'\|_2 &\le \left(1 + \frac{c}{\sqrt{d}}\right)\left(\sqrt{2}\tilde{C}^{\frac{1}{2}} d^{\frac{1}{4}} - \frac{\sqrt{2}}{2\tilde{C}^{\frac{1}{2}} d^{\frac{1}{4}}}\right) \\
&\le \left(1 + \frac{c}{\sqrt{d}}\right)^{-1}\left(1 + \frac{c}{\sqrt{d}}\right)^{2} \sqrt{2}\left(\tilde{C}^{\frac{1}{2}} d^{\frac{1}{4}} - \frac{1}{2\tilde{C}^{\frac{1}{2}} d^{\frac{1}{4}}}\right) \\
&\le \left(1 + \frac{c}{\sqrt{d}}\right)^{-1}\left(\sqrt{2}\tilde{C}^{\frac{1}{2}} d^{\frac{1}{4}} + \frac{8\sqrt{2}\tilde{C}c - 2\sqrt{2}}{4\tilde{C}^{\frac{1}{2}} d^{\frac{1}{4}}} + \frac{\sqrt{2}c\left(\tilde{C}c - 1\right)}{\tilde{C}^{\frac{1}{2}} d^{\frac{3}{4}}}\right) \\
&= \left(1 + \frac{c}{\sqrt{d}}\right)^{-1}\left(\sqrt{2}\tilde{C}^{\frac{1}{2}} d^{\frac{1}{4}} - \frac{\sqrt{2}\left(2 - 8\tilde{C}c\right)}{4\tilde{C} d^{\frac{1}{4}}} + \frac{\sqrt{2}c\left(\tilde{C}c - 1\right)}{\tilde{C}^{\frac{1}{2}} d^{\frac{3}{4}}}\right) \\
&\le \left(1 + \frac{c}{\sqrt{d}}\right)^{-1} \tau \qquad \text{(provided c is small enough)}
\end{aligned}$$

Thus, if $i$ and $j$ are adjacent, then

$$\|v_i' - v_j'\|_2 \le \left(1 + \frac{c}{\sqrt{d}}\right)^{-1} \tau \tag{4}$$

On the other hand, suppose $i$ and $j$ are not adjacent. Again by the Johnson–Lindenstrauss Lemma:

$$\|v_i' - v_j'\|_2 \ge \left(1 + \frac{c}{\sqrt{d}}\right)^{-1} \sqrt{2}\tilde{C}^{\frac{1}{2}} d^{\frac{1}{4}}$$

Also:

$$\tau\left(1 + \frac{c}{\sqrt{d}}\right)^{2} = \left(1 + \frac{c}{\sqrt{d}}\right)^{2} \sqrt{2}\left(\tilde{C}^{\frac{1}{2}} d^{\frac{14}{-}} \frac{\sqrt{2}}{4\tilde{C}^{\frac{1}{2}} d^{\frac{1}{4}}}\right) < \sqrt{2}\tilde{C}^{\frac{1}{2}} d^{\frac{1}{4}}$$

Combining the previous two inequalities, we get: if $i$ and $j$ are not adjacent, then

$$\|v_i' - v_j'\|_2 > \tau\left(1 + \frac{c}{\sqrt{d}}\right) \tag{5}$$

Statements 1 and 2 from Remark 6.2 follow from inequalities (4) and (5). □

## 7. Embedding for robust graphs: a necessary condition on $k$

For the proof of Theorem 7.3 below, we will need the following two lemmas which we state without proof. We believe that the proof of Lemma 7.1 can be obtained using properties of expansion graphs. We believe that the proof of Lemma 7.2 can be obtained using counting arguments with the union bound.

**Lemma 7.1.** *For every $\delta > 0$ there exists $R = R(\delta)$ such that the following holds for $1 - o(1)$ fraction of $d$-regular graphs $G$ on $n$ vertices. If $\Gamma$ is a distance graph in $\mathbb{R}^k$ isomorphic to $G$ then there is a sphere $u + R \cdot B_2^k$ comprising $1 - \delta$ fraction of vertices of $\Gamma$.*

**Lemma 7.2.** *Let $\delta > 0$ be a small constant. For every $d$-regular graph $G$ on $n$ vertices, let $H_G$ be an induced subgraph of $G$ on $n - \delta n$ vertices, and let $T$ be a subset of $\tilde{G}_{n,d}$ of size at least $\frac{1}{2}|\tilde{G}_{n,d}|$. Then the number of distinct isomorphism classes of $H_G$ with $G \in T$ is at least $n^{c'dn}$, for a universal constant $c'$.*

The next theorem is the second main result of the note:

**Theorem 7.3.** *Suppose $1 - o(1)$ fraction of $d$-regular graphs on $n$ vertices can be represented as $\frac{c}{\sqrt{d}}$-robust geometric graphs in $\mathbb{R}^k$, where $c$ is the small constant from Section 5. Then $k \geq \frac{c'' d \log n}{\log d}$, for some sufficiently small $c'' > 0$.*

*Proof.* Assume toward a contradiction $k < \frac{c' dn \log(n)}{\log \frac{4R\sqrt{d}}{c}}$. Suppose $1 - o(1)$ fraction of $d$-regular graphs on $n$ vertices can be represented as $\frac{c}{\sqrt{d}}$-robust geometric graphs in $\mathbb{R}^k$. Let $G$ be a $d$-regular graph on $n$ vertices, with $\Gamma$ an isomorphic geometric graph in $\mathbb{R}^k$. Let $\delta$ be a small constant and $u + R \cdot B_2^k$ be a sphere containing $1 - \delta$ fraction of vertices of $\Gamma$ (the existence of such sphere follows from Lemma 7.1). Let $\Gamma_G$ be the induced subgraph of $\Gamma$ containing all of the vertices that are contained in $u + R \cdot B_2^k$.

Construct a maximal packing of $u + R \cdot B_2^k$ by translates of $\frac{c}{4\sqrt{d}} B_2^k$ (we can do this by adding non-overlapping translates of $\frac{c}{4\sqrt{d}} B_2^k$ into $u + R \cdot B_2^k$ until it is no longer possible to add any more). Let $\Gamma'_G$ be the distance graph in $\mathbb{R}^k$ formed by shifting the vertices of $\Gamma_G$ to the nearest center of a sphere in the packing. This does not change the topology of the graph, as each vertex shifts by at most distance $\frac{c}{2\sqrt{d}}$ and $\Gamma$ is $\frac{c}{\sqrt{d}}$-robust. We refer to Figure 5 for illustration.

By volumetric argument the number of spheres in the packing is at most:

$$\left( \frac{4R\sqrt{d}}{c} \right)^k . \tag{6}$$

Thus, the number of possible isomorphism classes of $\Gamma'_G$ is at most:

$$\left( \frac{4R\sqrt{d}}{c} \right)^{k(n - \delta n)} \leq \left( \frac{4R\sqrt{d}}{c} \right)^{\frac{c' d \log n}{\log \frac{4R\sqrt{d}}{c}} (n - \delta n)} < n^{c'dn}. \tag{7}$$

On the other hand, by Lemma 7.2 the number of possible isomorphism classes of $\Gamma'_G$ is at least $n^{c'dn}$. This is a contradiction, so we conclude that $k \geq \frac{c'' d \log n}{\log \frac{4R\sqrt{d}}{c}}$.

□

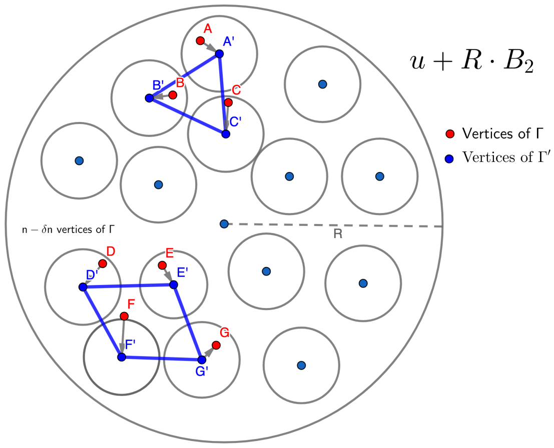


Figure 5: The vertices of Γ are "shifted" to the nearest center of a sphere in the packing to form Γ′. Γ and Γ′ have the same topology.

## References

[1] A. Z. Broder et al., Optimal construction of edge-disjoint paths in random graphs, SIAM J. Comput. **28** (1999), no. 2, 541–573. MR1634360

[2] M. Belkin and P. Niyogi, Dimensionality Reduction: A Short Tutorial, 2003, 1–10.

[3] A. Blum, J. Hopcroft and R. Kannan, *Foundations of Data Science*, Cambridge University Press, 2020, Chapter 13, 412–416. MR4088282

[4] N. A. Cook, L. M. Goldstein and T. L. Johnson, Size biased couplings and the spectral gap for random regular graphs, Ann. Probab. **46** (2018), no. 1, 72–125. MR3758727

[5] P. Frankl and H. Maehara, The Johnson-Lindenstrauss lemma and the sphericity of some graphs, J. Combin. Theory Ser. B **44** (1988), no. 3, 355–362. MR0941443

[6] J. A. Lee and M. Verleysen, *Nonlinear Dimensionality Reduction*, Springer, 2007, Chapter 6, 109–114. MR2328563

[7] V. D. Milman and G. Schechtman, *Asymptotic theory of finite-dimensional normed spaces*, Lecture Notes in Mathematics, 1200, Springer, Berlin, 1986. MR0856576

[8] R. Motwani and P. Raghavan, *Randomized Algorithms*, Cambridge University Press, 1995, Chapter 7, 215–217. MR1344451

[9] J. Reiterman, V. Rödl and E. Pelantová, Embeddings of graphs in Euclidean spaces, Discrete Comput. Geom. **4** (1989), no. 4, 349–364. MR0996768

[10] K. E. Tikhomirov and P. Youssef, The spectral gap of dense random regular graphs, Ann. Probab. **47** (2019), no. 1, 362–419. MR3909972

[11] Johnson-Lindenstrauss lemma's illustration. Available at: https://www.researchgate.net/figure/Johnson-Lindenstrauss-lemmas-illustration_fig3_341576089. Accessed on [insert access date].

Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213
*E-mail addresses*:
lufeiy@andrew.cmu.edu
cdyip@andrew.cmu.edu
mingjinz@andrew.cmu.edu
ktikhomi@andrew.cmu.edu


*Pittsburgh Interdiscip. Math. Rev.* is published by the University Library System of the University of Pittsburgh.