

Lecture 11: ARV

In which we introduce semi-definite programming and a semi-definite programming relaxation of sparsest cut, and we reduce its analysis to a key lemma that we will prove in the next lecture(s)

1 The Goemans-Linial Relaxation

Recall that, for two undirected graphs G, H , the sparsest cut problem is to optimize

$$\text{sc}(G, H) := \min_{S \subseteq V} \frac{\sum_{\{u,v\} \in E_G} |1_S(u) - 1_S(v)|}{\sum_{\{u,v\} \in E_H} |1_S(u) - 1_S(v)|}$$

and the Leighton-Rao relaxation is obtained by noting that if we define $d(u, v) := |1_S(u) - 1_S(v)|$ then $d(\cdot, \cdot)$ is a semimetric over V , meaning that the following quantity is a relaxation of (G, H) :

$$LR(G, H) = \min_{\substack{d : V \times V \rightarrow \mathbb{R} \\ d \text{ semimetric}}} \frac{\sum_{\{u,v\} \in E_G} d(u, v)}{\sum_{\{u,v\} \in E_H} d(u, v)}$$

If G is r -regular, H is a clique, and $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ are the eigenvalues of the normalized Laplacian of G , then

$$\frac{r}{n} \lambda_2 = \min_{f: V \rightarrow \mathbb{R}} \frac{\sum_{\{u,v\} \in E_G} |f(u) - f(v)|^2}{\sum_{\{u,v\} \in E_{K_n}} |f(u) - f(v)|^2} \quad (1)$$

which is a relaxation of $\text{sc}(G, K_n)$, because, for every S , every u and every v , $|1_S(u) - 1_S(v)| = |1_S(u) - 1_S(v)|^2$.

We note that if we further relax (1) by allowing V to be mapped into a higher dimension space \mathbb{R}^m instead of \mathbb{R} , and we replace $|\cdot - \cdot|$ by $\|\cdot - \cdot\|^2$, the optimum remains the same.

Fact 1

$$\lambda_2 = \inf_{m, F: V \rightarrow \mathbb{R}^m} \frac{\sum_{\{u,v\} \in E_G} \|F(u) - F(v)\|^2}{\sum_{\{u,v\} \in E_{K_n}} \|F(u) - F(v)\|^2}$$

PROOF: For every $F : V \rightarrow \mathbb{R}^m$, if we write $F(v) = (f_1(v), \dots, f_n(v))$, we have

$$\begin{aligned} & \frac{\sum_{\{u,v\} \in E_G} \|F(u) - F(v)\|^2}{\sum_{\{u,v\} \in E_{K_n}} \|F(u) - F(v)\|^2} \\ &= \frac{\sum_i \sum_{\{u,v\} \in E_G} (f_i(u) - f_i(v))^2}{\sum_i \sum_{\{u,v\} \in E_{K_n}} (f_i(u) - f_i(v))^2} \\ &\geq \min_{i=1, \dots, m} \frac{\sum_{\{u,v\} \in E_G} (f_i(u) - f_i(v))^2}{\sum_{\{u,v\} \in E_{K_n}} (f_i(u) - f_i(v))^2} \\ &\geq \lambda_2 \end{aligned}$$

□

The above observations give the following comparison between the Leighton-Rao relaxation and the spectral relaxation: both are obtained by replacing $|1_S(u) - 1_S(v)|$ with a “distance function” $d(u, v)$; in the Leighton-Rao relaxation, $d(u, v)$ is constrained to satisfy the triangle inequality; in the spectral relaxation, $d(u, v)$ is constrained to be the square of the Euclidean distance between $F(u)$ and $F(v)$ for some mapping $F : V \rightarrow \mathbb{R}^m$.

The Arora-Rao-Vazirani relaxation is obtained by *enforcing both conditions*, that is, by considering distance functions $d(u, v)$ that satisfy the triangle inequality *and* can be realized of $\|F(u) - F(v)\|^2$ for some mapping $F : V \rightarrow \mathbb{R}^m$.

Definition 2 A semimetric $d : V \times V \rightarrow \mathbb{R}$ is called of *negative type* if there is a dimension m and a mapping $F : V \rightarrow \mathbb{R}^m$ such that $d(u, v) = \|F(u) - F(v)\|^2$ for every $u, v \in V$.

With the above definition, we can formulate the Goemans-Linial relaxation as

$$ARV(G, H) := \min_{\substack{d : V \times V \rightarrow \mathbb{R} \\ d \text{ semimetric of negative type}}} \frac{\sum_{\{u,v\} \in E_G} d(u, v)}{\sum_{\{u,v\} \in E_H} d(u, v)} \quad (2)$$

Remark 3 The relaxation (2) was first proposed by Goemans and Linial. Arora, Rao and Vazirani were the first to prove that it achieves an approximation guarantee which is better than the approximation guarantee of the Leighton-Rao relaxation.

We have, by definition,

$$\text{sc}(G, H) \leq ARV(G, H) \leq LR(G, H)$$

and, when H is a clique and G is r -regular,

$$\text{sc}(G, K_n) \leq ARV(G, K_n) \leq \frac{r}{n} \lambda_2(G)$$

and so the approximation results that we have proved for λ_2 and LR apply to ARV . For every graphs G and H :

$$ARV(G, H) \leq O(\log |V|) \cdot \text{sc}(G, H)$$

and for every r -regular graph G

$$\frac{n}{r} ARV(G, K_n) \leq \sqrt{8 \frac{n}{r} \cdot \text{sc}(G, K_n)}$$

Interestingly, known examples of graphs for which LR and λ_2 give poor approximation are complementary. When H is a clique, if G is a cycle, then $\frac{r}{n} \lambda_2$ is a poor approximation of $\text{sc}(G, K_n)$, but $LR(G, K_n)$ is a good approximation of $\text{sc}(G, K_n)$; if G is a constant-degree expander then $LR(G, K_n)$ is a poor approximation of $\text{sc}(G, K_n)$, but $\frac{r}{n} \lambda_2$ is a good approximation.

When Goemans and Linial (separately) proposed to study the relaxation (2), they conjectured that it would always provide a constant-factor approximation of $\text{sc}(G, H)$. Unfortunately, the conjecture turned out to be false, but Arora, Rao and Vazirani were able to prove that (2) does provide a strictly better approximation than the Leighton-Rao relaxation. In the next lectures, we will present parts of the proof of the following results.

Theorem 4 *There is a constant c such that, for every graph $G = (V, E)$,*

$$\text{sc}(G, K_n) \leq c \cdot \sqrt{\log |V|} \cdot ARV(G, K_n)$$

Theorem 5 *There is a constant c such that, for every graphs $G = (V, E_G)$, $H = (V, E_H)$,*

$$\text{sc}(G, H) \leq c \cdot \sqrt{\log |V|} \cdot \log \log |V| \cdot ARV(G, H)$$

Theorem 6 *There is a constant c and an infinite family of graphs $G_n = (V_n, E_n)$ such that*

$$\text{sc}(G_n, K_n) \geq c \cdot \log \log |V_n| \cdot ARV(G_n, K_n)$$

Theorem 7 *There are families of graphs $G_n = (V_n, EG_n)$ and $H_n(V_n, EH_n)$ such that, for every $\epsilon > 0$ and every sufficiently larger n ,*

$$ARV(G_n, K_n) \geq (\log |V|)^{\frac{1}{2}-\epsilon} \cdot \text{sc}(G_n, K_n)$$

2 Polynomial Time Solvability

In this section we show that the Ellipsoid algorithm can compute $ARV(G, H)$ in polynomial time.

Definition 8 *If $C \subseteq \mathbb{R}^m$ is a set, then a separation oracle for C is a procedure that, on input $\mathbf{x} \in \mathbb{R}^m$,*

- *If $\mathbf{x} \in C$, outputs “yes”*
- *If $\mathbf{x} \notin C$, outputs coefficients a_1, \dots, a_m, b such that*

$$\sum_i x_i a_i < b$$

but, for every $\mathbf{z} \in C$,

$$\sum_i z_i a_i \geq b$$

Note that a set can have a separation oracle only if it is convex. Under certain additional mild conditions, if C has a polynomial time computable separation oracle, then the optimization problem

$$\begin{aligned} &\text{minimize} && \sum_i \mathbf{c}^T \mathbf{x} \\ &\text{subject to} && \\ &&& \mathbf{A}\mathbf{x} \geq b \\ &&& \mathbf{x} \in C \end{aligned}$$

is solvable in polynomial time using the Ellipsoid Algorithm.

It remains to see how to put the Arora-Rao-Vazirani relaxation into the above form.

Recall that a matrix $X \in \mathbb{R}^{n \times n}$ is *positive semidefinite* if all its eigenvalues are nonnegative. We will use the set of all $n \times n$ positive semidefinite matrices as our set C (thinking of them as n^2 -dimensional vectors). If we think of two matrices $M, M' \in \mathbb{R}^{n \times n}$ as n^2 -dimensional vectors, then their “inner product” is

$$M \bullet M' := \sum_{i,j} M_{i,j} \cdot M'_{i,j}$$

Lemma 9 *The set of $n \times n$ positive semidefinite matrices has a separation oracle computable in time polynomial in n .*

PROOF: Given a symmetric matrix X , its smallest eigenvalue is

$$\min_{\mathbf{z} \in \mathbb{R}^n, \|\mathbf{z}\|=1} \mathbf{z}^T X \mathbf{z}$$

the vector achieving the minimum is a corresponding eigenvector, and both the smallest eigenvalue and the corresponding eigenvector can be computed in polynomial time.

If we find that the smallest eigenvalue of X is non-negative, then we answer “yes.” Otherwise, if \mathbf{z} is an eigenvector of the smallest eigenvalue we output the matrix $A = \mathbf{z}^T \mathbf{z}$. We see that we have

$$A \bullet X = \mathbf{z}^T X \mathbf{z} < 0$$

but that, for every positive semidefinite matrix M , we have

$$A \bullet M = \mathbf{z}^T M \mathbf{z} \geq 0$$

□

This implies that any optimization problem of the following form can be solved in polynomial time

$$\begin{aligned} & \text{minimize} && C \bullet X \\ & \text{subject to} && A^1 \bullet X \geq b_1 \\ & && \dots \\ & && A^m \bullet X \geq b_m \\ & && X \succeq 0 \end{aligned} \tag{3}$$

where C, A^1, \dots, A^m are square matrices of coefficients, b_1, \dots, b_m are scalars, and X is a square matrix of variables. An optimization problem like the one above is called a *semidefinite program*.

It remains to see how to cast the Arora-Rao-Vazirani relaxation as a semidefinite program.

Lemma 10 *For a symmetric matrix $M \in \mathbb{R}^{n \times n}$, the following properties are equivalent:*

1. M is positive semidefinite;

2. there are vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$ such that, for all i, j , $M_{i,j} = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$;

3. for every vector $\mathbf{z} \in \mathbb{R}^n$, $\mathbf{z}^T M \mathbf{z} \geq 0$

PROOF: That (1) and (3) are equivalent follows from the characterization of the smallest eigenvalue of M as the minimum of $\mathbf{z}^T M \mathbf{z}$ over all unit vectors \mathbf{z} .

To see that (2) \Rightarrow (3), suppose that vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ exist as asserted in (2), and let X be the matrix whose columns are the vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$, so that $X^T \cdot X = M$. Take any vector \mathbf{z} , and see that

$$\mathbf{z}^T M \mathbf{z} = \mathbf{z}^T X^T X \mathbf{z} = \|X \mathbf{z}\|^2 \geq 0$$

Finally, to see that (1) \Rightarrow (2), let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of M with multiplicities, and let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a corresponding orthonormal set of eigenvectors. Then

$$M = \sum_i \lambda_k \mathbf{v}_k \mathbf{v}_k^T$$

that is,

$$M_{i,j} = \sum_k \lambda_k v_k(i) v_k(j) = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$$

if we define $\mathbf{x}_1, \dots, \mathbf{x}_n$ as the vectors such that $x_i(k) := \sqrt{\lambda_k} v_k(i)$. \square

This means that the generic semidefinite program (4) can be rewritten as an optimization problem in which the variables are the vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ as in part (2) of the above lemma.

$$\begin{aligned} & \text{minimize} && \sum_{i,j} C_{i,j} \langle \mathbf{x}_i, \mathbf{x}_j \rangle \\ & \text{subject to} && \sum_{i,j} A_{i,j}^1 \langle \mathbf{x}_i, \mathbf{x}_j \rangle \geq b_1 \\ & && \dots \\ & && \sum_{i,j} A_{i,j}^m \langle \mathbf{x}_i, \mathbf{x}_j \rangle \geq b_m \\ & && \mathbf{x}_i \in \mathbb{R}^m \quad \forall i \in \{1, \dots, n\} \end{aligned} \tag{4}$$

where the dimension m is itself a variable (although one could fix it, without loss of generality, to be equal to n). In this view, a semidefinite program is an optimization problem in which we wish to select n vectors such that their pairwise inner products satisfy certain linear inequalities, while optimizing a cost function that is linear in their pairwise inner product.

The square of the Euclidean distance between two vectors is a linear function of inner products

$$\|\mathbf{x} - \mathbf{y}\|^2 = \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle - 2\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle$$

and so, in a semidefinite program, we can include expressions that are linear in the pairwise squared distances (or squared norms) of the vectors. The ARV relaxation can be written as follows

$$\begin{aligned} & \text{minimize} && \sum_{\{u,v\} \in E_G} \|\mathbf{x}_u - \mathbf{x}_v\|^2 \\ & \text{subject to} && \sum_{\{u,v\} \in E_H} \|\mathbf{x}_u - \mathbf{x}_v\|^2 = 1 \\ & && \|\mathbf{x}_u - \mathbf{x}_v\|^2 \leq \|\mathbf{x}_u - \mathbf{x}_w\|^2 + \|\mathbf{x}_w - \mathbf{x}_v\|^2 \quad \forall u, v, w \in V \\ & && \mathbf{x}_u \in \mathbb{R}^d \quad \forall u \in V \end{aligned}$$

and so it is a semidefinite program, and it can be solved in polynomial time.

Remark 11 *Our discussion of polynomial time solvability glossed over important issues about numerical precision. To run the Ellipsoid Algorithm one needs, besides the separation oracle, to be given a ball that is entirely contained in the set of feasible solutions and a ball that entirely contains the set of feasible solutions, and the running time of the algorithm is polynomial in the size of the input, polylogarithmic in the ratio of the volumes of the two balls, and polylogarithmic in the desired amount of precision. At the end, one doesn't get an optimal solution, which might not have a finite-precision exact representation, but an approximation within the desired precision. The algorithm is able to tolerate a bounded amount of imprecision in the separation oracle, which is an important feature because we do not have exact algorithms to compute eigenvalues and eigenvectors (the entries in the eigenvector might not have a finite-precision representation).*

The Ellipsoid algorithm is typically not a practical algorithm. Algorithms based on the *interior point* method have been adapted to semidefinite programming, and run both in worst-case polynomial time and in reasonable time in practice.

Arora and Kale have developed an $\tilde{O}((|V| + |E|)^2 / \epsilon^{O(1)})$ time algorithm to solve the ARV relaxation within a multiplicative error $(1 + \epsilon)$. The dependence on the error is worse than that of generic algorithms, which achieve polylogarithmic dependency, but this is not a problem in this application, because we are going to lose an $O(\sqrt{\log |V|})$ factor in the rounding, so an extra constant factor coming from an approximate solution of the relaxation is a low-order consideration.

3 Rounding when H is a clique

Given the equivalence between the sparsest cut problem and the “ $L1$ relaxation” of sparsest cut, it will be enough to prove the following result.

Theorem 12 (Rounding of ARV) *Let $G = (V, E)$ be a graph, and $\{\mathbf{x}_v\}_{v \in V}$ be a feasible solution of the relaxation $ARV(G, K_n)$.*

Then there is a mapping $f : V \rightarrow \mathbb{R}$ such that

$$\frac{\sum_{\{u,v\} \in E} |f(u) - f(v)|}{\sum_{\{u,v\}} |f(u) - f(v)|} \leq O(\sqrt{\log |V|}) \cdot \frac{\sum_{\{u,v\} \in E} \|\mathbf{x}_u - \mathbf{x}_v\|^2}{\sum_{\{u,v\}} \|\mathbf{x}_u - \mathbf{x}_v\|^2}$$

As in the rounding of the Leighton-Rao relaxation via Bourgain's theorem, we will identify a set $S \subseteq V$, and define

$$f_S(v) := \min_{s \in S} \|\mathbf{x}_s - \mathbf{x}_v\|^2 \tag{5}$$

Recall that, as we saw in the proof of Bourgain's embedding theorem, no matter how we choose the set S we have

$$|f_S(u) - f_S(v)| \leq \|\mathbf{x}_u - \mathbf{x}_v\|^2 \tag{6}$$

where we are not using any facts about $\|\cdot - \cdot\|^2$ other than the fact that, for solutions of the ARV relaxation, it is a distance function that obeys the triangle inequality.

This means that, in order to prove the theorem, we just have to find a set $S \subseteq V$ such that

$$\sum_{u,v} |f_S(u) - f_S(v)| \geq \frac{1}{O(\sqrt{\log |V|})} \cdot \sum_{u,v} \|\mathbf{x}_u - \mathbf{x}_v\|^2 \tag{7}$$

and this is a considerable simplification because the above expression is completely independent of the graph! The remaining problem is purely one about geometry.

Recall that if we have a set of vectors $\{\mathbf{x}_v\}_{v \in V}$ such that the distance function $d(u, v) := \|\mathbf{x}_u - \mathbf{x}_v\|^2$ satisfies the triangle inequality, then we say that $d(\cdot, \cdot)$ is a (semi-)metric of *negative type*.

After these preliminaries observations, our goal is to prove the following theorem.

Theorem 13 (Rounding of ARV – Revisited) *If $d(\cdot, \cdot)$ is a semimetric of negative type over a set V , then there is a set S such that if we define*

$$f_S(v) := \min_{s \in S} \{d(s, v)\}$$

we have

$$\sum_{u,v} |f_S(u) - f_S(v)| \geq \frac{1}{O(\sqrt{\log |V|})} \cdot \sum_{u,v} d(u, v)$$

Furthermore, the set S can be found in randomized polynomial time with high probability given a set of vector $\{\mathbf{x}_v\}_{v \in V}$ such that $d(u, v) = \|\mathbf{x}_u - \mathbf{x}_v\|^2$.

Since the statement is scale-invariant, we can restrict ourselves, with no loss of generality, to the case $\sum_{u,v} d(u, v) = |V|^2$.

Remark 14 Let us discuss some intuition before continuing with the proof.

As our experience proving Bourgain's embedding theorem shows us, it is rather difficult to pick sets such that $|f_S(u) - f_S(v)|$ is not much smaller than $d(u, v)$. Here we have a somewhat simpler case to solve because we are not trying to preserve all distances, but only the average pairwise distance. A simple observation is that if we find a set S which contains $\Omega(|V|)$ elements and such that $\Omega(|V|)$ elements of V are at distance $\Omega(\delta)$ from S , then we immediately get $\sum_{u,v} |f_S(u) - f_S(v)| \geq \Omega(\delta|V|^2)$, because there will be $\Omega(|V|^2)$ pairs u, v such that $f_S(u) = 0$ and $f_S(v) \geq \delta$. In particular, if we could find such a set with $\delta = 1/O(\sqrt{\log |V|})$ then we would be done. Unfortunately this is too much to ask for in general, because we always have $|f_S(u) - f_S(v)| \leq d(u, v)$, which means that if we want $\sum_{u,v} |f_S(u) - f_S(v)|$ to have $\Omega(|V|^2)$ noticeably large terms we must also have that $d(u, v)$ is noticeably large for $\Omega(|V|^2)$ pairs of points, which is not always true.

There is, however, the following argument, which goes back to Leighton and Rao: either there are $\Omega(|V|)$ points concentrated in a ball whose radius is a quarter (say) of the average pairwise distance, and then we can use that ball to get an ℓ_1 mapping with only *constant* error; or there are $\Omega(|V|)$ points in a ball of radius twice the average pairwise distance, such that the pairwise distances of the points in the ball account for a constant fraction of all pairwise distances. In particular, the sum of pairwise distances includes $\Omega(|V|^2)$ terms which are $\Omega(1)$.

After we do this reduction and some scaling, we are left with the task of proving the following theorem: suppose we are given an n -point negative type metric in which the points are contained in a ball of radius 1 and are such that the sum of pairwise distances is $\Omega(n^2)$; then there is a subset S of size $\Omega(n)$ such that there are $\Omega(n)$ points whose distance from the set is $1/O(\sqrt{\log n})$. This theorem is the main result of the Arora-Rao-Vazirani paper. (Strictly speaking, this form of the theorem was proved later by Lee – Arora, Rao and Vazirani had a slightly weaker formulation.)

We begin by considering the case in which a constant fraction of the points are concentrated in a small ball.

Definition 15 (Ball) For a point $z \in V$ and a radius $r > 0$, the ball of radius r and center z is the set

$$B(z, r) := \{v : d(z, v) \leq r\}$$

Lemma 16 For every vertex z , if we define $S := B(z, 1/4)$, then

$$\sum_{u,v} |f_S(u) - f_S(v)| \geq \frac{|S|}{2|V|} \sum_{u,v} d(u,v)$$

PROOF: Our first calculation is to show that the typical value of $f_S(u)$ is rather large. We note that for every two vertices u and v , if we call a a closest vertex in S to u , and b a closest vertex to v in S , we have

$$\begin{aligned} d(u,v) &\leq d(u,a) + d(a,z) + d(z,b) + d(b,v) \\ &\leq f_S(u) + f_S(v) + \frac{1}{2} \end{aligned}$$

and so

$$|V|^2 = \sum_{u,v} d(u,v) \leq 2|V| \cdot \sum_v f_S(v) + \frac{|V|^2}{2}$$

that is,

$$\sum_v f_S(v) \geq \frac{|V|}{2}$$

Now we can get a lower bound to the sum of ℓ_1 distances given by the embedding $f_S(\cdot)$.

$$\begin{aligned} &\sum_{u,v} |f_S(u) - f_S(v)| \\ &\geq \sum_{u \in S, v \in V} |f_S(v)| \\ &= |S| \sum_v f_S(v) \\ &\geq \frac{1}{2} |S| \cdot |V| \end{aligned}$$

□

This means that if there is a vertex z such that $|B(z, 1/4)| = \Omega(|V|)$, or even $|B(z, 1/4)| = \Omega(|V|/\sqrt{\log |V|})$, then we are done.

Otherwise, we will find a set of $\Omega(|V|)$ vertices such that their average pairwise distances are within a constant factor of their maximum pairwise distances, and then we

will work on finding an embedding for such a set of points. (The condition that the average distance is a constant fraction of the maximal distance will be very helpful in subsequent calculations.)

Lemma 17 *Suppose that for every vertex z we have $|B(z, 1/4)| \leq |V|/4$. Then there is a vertex w such that, if we set $S = B(w, 2)$, we have*

- $|S| \geq \frac{1}{2} \cdot |V|$
- $\sum_{u,v \in S} d(u, v) \geq \frac{1}{8} |S|^2$

PROOF: Let w be a vertex that maximizes $|B(w, 2)|$; then $|B(w, 2)| \geq |V|/2$, because if we had $|B(u, 2)| < |V|/2$ for every vertex u , then we would have

$$\sum_{u,v} d(u, v) > \sum_u 2 \cdot (|V - B(u, 2)|) > |V|^2$$

Regarding the sum of pairwise distances of elements of S , we have

$$\sum_{u,v \in S} d(u, v) > \sum_{u \in S} \frac{1}{4} (|S - B(u, 1/4)|) \geq |S| \cdot \frac{1}{4} \cdot \frac{|S|}{2}$$

□

The proof of the main theorem now reduces to proving the following geometric fact.

Lemma 18 (ARV Main Lemma) *Let d be a negative-type metric over a set V such that the points are contained in a unit ball and have constant average distance, that is,*

- *there is a vertex z such that $d(v, z) \leq 1$ for every $v \in V$*
- $\sum_{u,v \in V} d(u, v) \geq c \cdot |V|^2$

Then there are sets $S, T \subseteq V$ such that

- $|S|, |T| \geq \Omega(|V|)$;
- *for every $u \in S$ and every $v \in S$, $d(u, v) \geq 1/O(\sqrt{\log |V|})$*

where the multiplicative factors hidden in the $O(\cdot)$ and $\Omega(\cdot)$ notations depend only on c .

Indeed, applying the ARV Main Lemma to $\frac{1}{2}d(u, v)$ tells us that there are subsets S, T of $B(z, 2)$, both of size $\Omega(|B(z, 2)|) = \Omega(n)$ such that $d(u, v) \geq 1/O(\sqrt{\log n})$ for every $u \in S$ and $v \in T$. If we consider the Frechet embedding f_S , we have

$$\begin{aligned}
\sum_{\{u,v\}} |f_S(u) - f_S(v)| &\geq \sum_{u \in S, v \in T} |f_S(u) - f_S(v)| \\
&\geq |S| \cdot |T| \cdot \frac{1}{O(\sqrt{\log n})} \\
&\geq n^2 \cdot \frac{1}{O(\sqrt{\log n})} \\
&= \frac{1}{O(\sqrt{\log n})} \cdot \sum_{\{u,v\}} d(u, v)
\end{aligned}$$

It remains to prove the ARV Main Lemma.