2022

Local-Global Results on Discrete Structures

Alexander Lewis Stevens
University of Denver

Local-Global Results on Discrete Structures

Abstract

Local-global arguments, or those which glean global insights from local information, are central ideas in many areas of mathematics and computer science. For instance, in computer science a greedy algorithm makes locally optimal choices that are guaranteed to be consistent with a globally optimal solution. On the mathematical end, global information on Riemannian manifolds is often implied by (local) curvature lower bounds. Discrete notions of graph curvature have recently emerged, allowing ideas pioneered in Riemannian geometry to be extended to the discrete setting. Bakry- Émery curvature has been one such successful notion of curvature. In this thesis we use combinatorial implications of Bakry- Émery curvature on graphs to prove a sort of local discrepancy inequality. This then allows us to derive a number of results regarding the local structure of graphs, dependent only on a curvature lower bound. For instance, it turns out that a curvature lower bound implies a nontrivial lower bound on graph connectivity. We also use these results to consider the curvature of strongly regular graphs, a well studied and important class of graphs. In this regard, we give a partial solution to an open conjecture: all SRGs satisfy the curvature condition $CD(\infty, 2)$. Finally we transition to consider a facility location problem motivated by using Unmanned Aerial Vehicles (UAVs) to guard a border. Here, we find a greedy algorithm, acting on local geometric information, which finds a near optimal placement of base stations for the guarding of UAVs.

Document Type

Dissertation

Degree Name

Ph.D.

Department

Mathematics

First Advisor

Paul K. Horn

Second Advisor

Jing Li

Third Advisor

Mei Yin

Keywords

Bakry Emery curvature, Graph theory, Local-global results, Spectral graph theory

Subject Categories

Computer Sciences | Mathematics | Other Computer Sciences | Other Mathematics | Physical Sciences and Mathematics

Local-global results on discrete structures

A Dissertation

Presented to

the Faculty of the College of Natural Sciences and Mathematics

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Alexander Lewis Stevens

June 2022

Advisor: Dr. Paul Horn

Author: Alexander Lewis Stevens
Title: Local-global results on discrete structures
Advisor: Dr. Paul Horn
Degree Date: June 2022

ABSTRACT

Local-global arguments, or those which glean global insights from local information, are central ideas in many areas of mathematics and computer science. For instance, in computer science a greedy algorithm makes locally optimal choices that are guaranteed to be consistent with a globally optimal solution. On the mathematical end, global information on Riemannian manifolds is often implied by (local) curvature lower bounds. Discrete notions of graph curvature have recently emerged, allowing ideas pioneered in Riemannian geometry to be extended to the discrete setting. Bakry-Émery curvature has been one such successful notion of curvature. In this thesis we use combinatorial implications of Bakry-Émery curvature on graphs to prove a sort of local discrepancy inequality. This then allows us to derive a number of results regarding the local structure of graphs, dependent only on a curvature lower bound. For instance, it turns out that a curvature lower bound implies a nontrivial lower bound on graph connectivity. We also use these results to consider the curvature of strongly regular graphs, a well studied and important class of graphs. In this regard, we give a partial solution to an open conjecture: all SRGs satisfy the curvature condition $CD(\infty, 2)$. Finally we transition to consider a facility location problem motivated by using *Unmanned Aerial Vehicles* (UAVs) to guard a border. Here, we find a greedy algorithm, acting on local geometric information, which finds a near optimal placement of base stations for the guarding of UAVs.

ACKNOWLEDGEMENTS

I would like to thank Michael for keeping me fed watered and sheltered over the past few years. I owe you many dates and flowers and I hope to repay you over the next sixty years. I would also like to thank my family. This thesis is the product of generations of Spencers, Knudsons, Dionnes and Stevens.

I would like to thank the DU mathematics department. Every professor here has affirmed my decision to attend DU. Thank you Dr. Dobrinen and Dr. Pavlov for the independent studies – this was an incredible opportunity to learn from two brilliant mathematicians. Dr. Yin – your kindness and authenticity are inspiring. Susan – without you I would owe DU thousands in late fees and would still be applying for a Masters. Dr. Vojtěchovský – thank you for holding the department together over the past several years. Dr Ormes – I so admire your dedication to the students of DU. I would also like to thank the wonderful mathematicians and computer scientists at UNAM, and UMF.

I would like to thank Dr. Flaherty. I will never forget how much fun we had camping with our E-STEM students. I would like to thank Dr. Lopez. Your passionate for computer science and life is inspiring and I feel so fortunate to call you my friend. Most of all I would like to thank Dr. Paul Horn. I owe you a debt that I can never repay. Thank you so much for helping shape the way I look at mathematics and at life. You are a fantastic human being and I feel so very fortunate to have been your student.

I would also like to thank the many grad students who have been with me for the past six years. In particular Ryan, Danielle, Terry, Adam, Matt, Hwajin, Dylan, Lauren, Wesley, Riley, Vlad, Jesse, Xiao, Cheng-Chih, Mariah, and Casey. I would also like to thank a wonderful mathematician and a dear friend, Bradford Greene. We were attracted to mathematics because we know something. What we know, we can't explain. But we can feel it.

TABLE OF CONTENTS

1 Introduction **1**

1.1 Spectral Graph Theory and Curvature 3

1.2 Strongly Regular Graphs 6

1.3 Intro to Base Station Placement Problem 8

2 Curvature **11**

2.1 Introduction 11

2.2 Preliminaries 13

2.3 Discrepancy Inequalities 20

2.3.1 Local Buser Inequality 26

2.4 Applications and Examples 32

2.4.1 Connectivity 32

2.4.2 Examples 36

2.5 Curvature and Graph Operations 40

2.6 Future Work 43

3 Strongly Regular **44**

3.1 Introduction 44

3.2 Proof of Main Lemma 47

3.3 Curvature of Collinearity Graphs of Partial Geometries 60

3.4 Future Work 64

4 Drones **65**

4.1 Introduction 65

4.1.1 Related work 67

4.2 Terminology and Problem Formulation 68

4.3 Preliminary results 72

4.4 OptSol 74

4.4.1 Time complexity 76

4.5 Discretization 86

4.6 Extending Results 90

4.6.1 The 2 step algorithm for pockets 91

4.6.2 The Algorithm 91

4.7 Acknowledgements 95

LIST OF FIGURES

1.1 Cospectral Graphs 4

2.1 Sharpness example for $\mathring{B}_2(x)$ 36

3.1 Shikrande vs. Rook's 60

4.1 Polygon with Unique Optimal Base Station Configuration 66
4.2 A d path on a polygon. 69
4.3 Solution to MinStation is not on minimal perimeter d-hull 71
4.4 Solution to MinStation restricted to minimal perimeter d-hull can be very bad 71
4.5 Illustration of Sandwich Lemma 73
4.6 Greedy Algorithm bad approximation. 77
4.7 The 2-hop problem 78
4.8 Storing Edge Projections 80
4.9 Visibility and vertex projections 80
4.10 Turning vertices 83
4.11 Salamis Island in the Saronic Gulf 89
4.12 MinStation Variant 91
4.13 Optimal Solution of MinStationVariant 92

Chapter 1: Introduction

A graph G consists of a set of vertices, $V(G)$, and a set of edges $E(G)$ where each edge joins a pair of vertices. As graphs are so versatile, they have proven an important tool for many things, including storing and analysing massive sets of relational data. As graphs being analyzed have grown in both size and complexity it has become impossible to precisely understand all their properties. Instead, it has become increasingly fruitful to focus attention on key graph properties which allow us to understand general underlying behavior. Unfortunately, many of these properties are prohibitively computationally expensive. Uncovering connections between the most enlightening properties and those which are most computationally feasible has been a huge area of research in virtually all areas of applied mathematics, and the study of graphs is no exception.

These sorts of questions lie at the heart of extremal graph theory, which may broadly be described as the study of relationships between graph parameters. For example, a typical theorem in this area may be of the form, 'Graphs which satisfy property P also satisfy property T'. These kinds of extremal arguments allow us to relate parameters we are most concerned with to those which we can most easily compute. A fundamental result in this extremal graph theory is Turan's Theorem.

Theorem 1 (Turan's Theorem)**.** [58] A *clique* is a subset of pairwise adjacent vertices. The clique number of a graph, ω is the size of the largest clique. For any graph $G = (V, E)$ with $|E| \geq \left(1 - \frac{1}{r}\right)\frac{|V|^2}{2}$, we have $\omega \geq r + 1$.

Turan's theorem relates the number of edges – a parameter that is computationally inexpensive to compute – to the clique number, which is known to be NP-complete.

Some graph parameters which are both computationally inexpensive and insightful are found through viewing spectral properties of various graph matrices. Viewing graphs though their associated matrices has allowed research in graph theory to take advantage of a wealth of results and insights from linear algebra and other fields. One such area is differential geometry. As we explore later in this thesis, there are many parallels between the Laplace-Beltrami operator and the graph Laplacian matrix.

In Chapters 2 and 3 we explore Bakry-Émery Graph Curvature, a local graph parameter that has been found to serve as a good discrete analogue of curvature due to relationships between the Laplace Beltrami operator and the graph Laplacian. Investigating discrete notion of curvature has been an attractive venture in part because in the continuous setting it has been wildly successful in relating local properties to global phenomena. In the discrete setting, local conditions are typically computationally inexpensive, making them very attractive. We explore relationships between local and global properties further when considering optimal drone base-station placements in Chapter 4.

As graphs model relational data particularly well, polygons and polygonal lines have been wonderful aids in representing and understanding geographic data. One may see almost immediately why these objects are so useful in representing roads, islands, borders between territories, etc. The formal study of algorithms on these objects falls in the area of computational geometry. A major problem in this field has been finding efficient algorithms to compute the convex hull of a polygon. In our final chapter we explore a variant of this problem and provide a near optimal solution to this problem extremely efficiently. Like our work with curvature, which enabled us to derive global properties of graphs from local measurements, our algorithm uses local information from the polygon to give global insights.

In the remainder of the introduction we survey some important results in these areas and provide some notation and preliminary information.

1.1 Spectral Graph Theory and Curvature

In this thesis, $G = (V, E)$ will denote a simple undirected graph. We denote the degree of the vertex $x \in V(G)$ by $\mathrm{d}(x)$ and for a subset $X \subset V$ we denote by $\mathrm{d}_X(x)$ the number of neighbors of x in X. The minimum degree is denoted by δ. We write $x \sim y$ when $xy \in E(G)$. We only consider graphs which are locally finite; that is, $\mathrm{d}(x) < \infty$ for all $x \in V(G)$ (which encompasses all finite graphs). We will make use of the Landau notations O, o, so that $f(n) = O(g(n))$ means $\dfrac{f(n)}{g(n)}$ is bounded above as $n \to \infty$ and $f(n) = o(g(n))$ means $\dfrac{f(n)}{g(n)}$ tends to 0 as $n \to \infty$.

We begin by introducing the most commonly used graph matrix. The adjacency matrix A of a graph G on n vertices is the square matrix indexed by the vertices of A such that $A_{xy} = 1$ when $x \sim y$, and $A_{xy} = 0$ otherwise. Because A is real and symmetric, A has n real valued eigenvalues. We commonly label them $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and refer to the collection (with multiplicity) of these values as the spectrum.

Many graph properties are directly related to the spectrum of the adjacency matrix. For instance, when G is d-regular it is not difficult to see that $\vec{1}$ is an eigenvector with associated eigenvalue d. One classical result relating the spectrum of A to graph properties is the expander mixing lemma[2].

Theorem 2 (Expander Mixing Lemma)**.** Suppose G is d-regular.
Define $\rho = \max\{|\lambda_2|, |\lambda_n|\}$. For any two sets X and Y, the number of edges between them $e(X, Y)$ satisfies

$$\left|e(X,Y) - \frac{d|X| \cdot |Y|}{|G|}\right| \leq \rho \cdot \sqrt{|X| \cdot |Y|} \tag{1.1}$$

The term on the left of (1.1) compares the number of edges between X and Y to what we would expect had the edges of G been distributed randomly. It is known this bound is asymptotically tight for random d-regular graphs, for which $\rho = O(\sqrt{d})$ with high probability [29]. For this reason, ρ has become a measure of randomness of edge distribution. We explore the spectrum of A further later in this chapter and again in Chapter 3.

While the spectrum of the adjacency matrix is useful, it also has limitations when used to understand geometric properties of general graphs. For example, graph connectivity – a very natural geometric property – is *not* determined by the adjacency spectrum. For a small example, the two graphs in Figure 1.1 have the same adjacency spectrum, while one is connected and the other isn't.

(a) (b)

Figure 1.1: The adjacency spectra of both graphs is $[-2, 0, 0, 0, 2]$. On the other hand, their Laplace spectra are $[0, 0, 2, 2, 4]$ and $[0, 1, 1, 1, 5]$ respectively.

There is a family of graph matrices for which spectra does a particularly good job of capturing properties like connectivity– the graph Laplacians. These matrices have close ties to the Laplace-Beltrami operator, allowing graphs to be viewed from a more geometric perspective. In this thesis, we focus on the combinatorial Laplacian $L = D - A$ where $D = (d_1, d_2, \cdots, d_n)$ is the diagonal degree matrix. Considering the spectra of the two graphs in Figure 1.1, we see the Laplacian spectrum of graph (a) has 0 with multiplicity 2. In fact the Laplacian always has an eigenvalue of 0 with eigenvector $\vec{1}$ and the multiplicity of 0 is exactly the number of connected components of G.

There are several other natural ways to define the graph Laplacian– an incidence matrix definition is used in the proof of the celebrated Kirkhoff's matrix tree theorem [26]. We place an arbitrary orientation on the edges of G and define the $|V| \times |E|$ incidence matrix B as

$$B(v,e) = \begin{cases} 1 & \text{if the edge } e \text{ has initial vertex } v \\ -1 & \text{if the edge } e \text{ has final vertex } v \\ 0 & \text{otherwise.} \end{cases} \tag{1.2}$$

One can confirm that $BB^\intercal = L$. More enlightening perhaps is that for the edge e from v to v' and for all ϕ, $(B^\intercal \phi)(e) = \phi(v) - \phi(v')$. From this we see

$$\vec{\phi}^\intercal L \vec{\phi} = \sum_{v \sim v'} (\phi(v) - \phi(v'))^2 \tag{1.3}$$

and so L is positive semidefinite and as a consequence has non-negative spectrum. We find the quadratic form of (1.3) suits our results in Chapters 2 and 3 particularly well. Notice that the Laplacian is also a 'local' operator, meaning $(L\phi)(v)$ only depends on values of ϕ for $v' \sim v$. From the Laplacian, we define two other local differential operators – the gradient, Γ and iterated gradient Γ_2. We derive them later in Chapter 2. This allows us to define the Bakry-Émery notion of discrete curvature– a graph $G = (V, E)$ satisfies $CD(\infty, K)$ if for all $x \in V$ and all $f : V \to \mathbb{R}$, $(\Gamma_2 f)(x) \geq K(\Gamma f)(x)$.

The local nature of Bakry-Émery curvature makes calculating it computationally inexpensive. Computing local information in this way often allows us to certify graph properties, which could otherwise, be prohibitively difficult to determine. For instance in [45] Matthews and Sumner extended Ore-type results relating minimum degree to longest paths for $K_{1,3}$-free graphs. While the general Hamiltonicity problem is NP-complete, certification of Hamiltonicity through verification of minimum degree and forbidden subgraphs conditions can be considerably less computationally expensive.

Bakry-Émery curvature, and its variants, have proven particularly adept at deriving global consequences from curvature lower bounds. For example, the celebrated work of

Li and Yau which provides an upper bound on the gradient of positive solutions of the heat equation in compact manifolds – was extended to the graph setting by Bauer et al. [6]. From this, consequences such as the volume doubling property and diameter bounds were derived, see [22, 38]. Quite recently, Salez proved the nonexistence of non-negatively curved sparse expanders in both the Ollivier and Bakry-Émery sense, the former being an open problem for a decade [54].

In Chapter 2 we continue to explore these ideas by using combinatorial arguments. We show, for example, that curvature implies a local discrepancy inequality which then provides a non-trivial lower bound on vertex connectivity of connected graphs, dependent only on the degree and a lower bound on Bakry-Émery curvature.

1.2 Strongly Regular Graphs

In Chapter 3 we consider the curvature of strongly regular graphs (SRGs). A graph G is said to be strongly regular with parameters (n, d, λ, μ) if G is d-regular of order n, every two adjacent vertices have exactly λ common neighbors and every two nonadjacent vertices have exactly μ common neighbors.

Immediately from the definition of strongly regular graphs, structural properties begin to emerge. It is easy to see, for instance, that connected strongly regular graphs have diameter at most two and that the complement of a strongly regular graph is strongly regular with parameter set $(n, n-1-d, n-2-(2d-\mu), n-(2d-\lambda))$. Once one starts studying them, more non-trivial relations between the parameter sets emerge. For instance, double counting the edges between the first and second neighborhoods of a vertex reveals that $(n-d-1)\mu = d(d-\lambda-1)$.

More interesting, is that there is an immediate connection between strong regularity and the spectrum of the graph. Supposing M is the adjacency matrix of G, we have $M^2 = dI + \lambda M + \mu(J - I - M)$. To see this, we may view M^2 as a count on the number of paths of length 2. Through dI we see the d paths each vertex has to a neighbor then back to itself. In the second term we see the λ paths a vertex has to its neighbor, and in the third, the μ paths between a vertex and its non-neighbors. Since M is real symmetric, it is diagonalizable and aside from the d eigenvalue, eigenvalues γ of M satisfy $\gamma^2 = d + \lambda\gamma - \mu - \mu\gamma$. Therefore γ can take at most two unique values aside from d. Conversely, it turns out that regular graphs with 3 distinct eigenvalues are strongly regular. It is good here to note that since the sum of eigenvalues with multiplicity of M is $tr(M) = 0$, the smallest eigenvalue is negative. Because the spectrum of SRGs are so well understood, it is often given as a supplementary or alternative parameter set, in which case we say the SRG has spectrum d^1 r^f, and $(-m)^g$, where 1, f, and g represent the multiplicities of the eigenvalues respectively. Understanding these multiplicities has also helped further restrict parameter sets. For example, in [51] Neumaier shows that all SRGs in which $-m$ is not integral have the form $n = 4\mu + 1$, $d = 2\mu$, and $\lambda = \mu - 1$. These are the so called *conference graphs*.

A number of the most commonly encountered SRGs are from highly structured families of graphs, notably Cage and Cayley graphs. Such highly structured graphs are a natural first place to look at curvature – vertex transitivity, for instance, makes computing curvature much simpler as the curvature at all vertices is the same. Beyond this, *generalizations* of these families seem like natural graph classes to look at the curvature of.

One such generalization is the class of Ricci-Flat graphs, introduced by Chung and Yau [14]. Motivated by the structure of the d-dimensional grid $\mathbb{Z}^d$, these graphs satisfy nonnegative curvature conditions for Bakry-Émery curvature and Ollivier curvature – the other most studied notion of discrete curvature.

Ricci-flat graphs are described completely by bijections between neighboring balls of radius one. Ricci-flat graphs are closed under tensor, Cartesian and strong products, making them a good candidate for the prototypical non-negatively curved graph. In [15] Cushing et al. showed that Ricci-flat graphs which satisfy an additional reflexively condition also satisfy $CD(\infty, 2)$. As it turns out, this is the maximal Bakry-Émery curvature attainable for triangle free graphs, a proof of which is provided in Chapter 2.

The strongly regular graphs can be thought of as another such generalization. It was shown by Cushing et al. [15] that despite not generally falling within the family of Ricci-Flat graphs with Reflexivity, SRGs with girth 4 *do* attain this maximal curvature of 2. This suggests perhaps, that the general SRG may also have a strong connection to Bakry-Émery curvature.

Cushing et al., having computed Bakry-Émery curvatures of many smaller SRGs, posited the conjecture that all SRGs satisfy $CD(\infty, 2)$. In Chapter 3 we explore these ideas, show this is largely the case, and that the full result may rest on a longstanding conjecture regarding open feasible parameter sets of SRGs.

1.3 Intro to Base Station Placement Problem

Finally we consider a problem with a more geometric flavor.

Imagine an island modeled as a simple polygon $\mathcal{P}$ with n vertices whose coastline we wish to monitor. We consider the problem of building the minimum number of refueling stations along the boundary of $\mathcal{P}$ in such a way that a drone can follow a polygonal route enclosing the island without running out of fuel. A drone can fly a maximum distance d between consecutive stations and is restricted to move either along the boundary of $\mathcal{P}$ or its exterior (i.e., over the sea). We present an algorithm that, given $\mathcal{P}$, finds the locations for a set of refueling stations whose cardinality is at most the optimal plus one. The time complexity of this algorithm is $O(n^2 + \frac{L}{d}n)$, where L is the length of $\mathcal{P}$. We are able to achieve this time complexity because our algorithm is 'greedy'.

Greedy algorithms which, given an input make a sequence of locally optimal choices in order to try and build a globally optimal solution, are a fundamental construct in computer science. In particular, they are often used to construct good approximate solutions to problems where finding the exact solution is intractable. Of course, making local decisions does not always work – there are many problems where greedy solutions can be arbitrarily bad in relation to the optimal solution. Take, for instance, the knapsack problem.

Problem 1 (Knapsack Problem)**.** Given a set of n items numbered from 1 to n, each with a weight w_i and a value v_i, along with a maximum weight capacity W,

$$\text{Maximize } \sum_{i=1}^{n} v_i x_i$$
$$\text{subject to } \sum_{i=1}^{n} w_i x_i \leq W \text{ and } x_i \in \{0,1\}.$$

The decision problem version of the knapsack problem is known to be NP-complete. For this reason, we would like to find a good approximation algorithm which is easy to implement and computationally inexpensive. A natural idea for an approximation algorithm is to greedily choose the items with the largest value to weight ratio.

Algorithm 1 Greedy Knapsack Algorithm

Sort items in nondecreasing order of $\frac{v_i}{w_i}$
Greedily select items from this order while remaining under maximum weight capacity

Unfortunately, this algorithm can be arbitrarily bad: consider an item with weight 1 and value 2 and another item of weight W and value W.

Nonetheless, for broad classes of problems greedy algorithms do work particularly well. For instance, a wide array of optimization problems are 'submodular' where one is trying to find a set S of a fixed size which maximizes an objective function. We find that a greedy algorithm provides a multiplicative $1/e$ approximation – that is, if the best set S has value $f(S)$, proceeding greedily yields a set S' with $f(S') \geq f(S)/e$. In fact Matroids are characterized by the property that the greedy algorithm correctly solves optimization problem.

For this drone base location problem we find that while a simple greedy algorithm gives an arbitrarily bad bound, a modification of this algorithm works quite well. In Chapter 4 we begin with some related work and the presentation of this algorithm. We then consider a sort of dual to this problem: what if we have some number of base stations which can be built and wish to minimize the maximum distance that a drone is forced to travel between refuelings? We then finish Chapter 4 with an extension of the base station location problem. Instead of enclosing the polygon, we wish to find an optimal set of locations allowing a drone to fly from one point to another point on a polygon.

Chapter 2: Curvature

2.1 Introduction

A fundamental issue with the computational study of large networks is that their sheer size makes many algorithmic approaches unrealistic. Then, instead of computing various graph properties exactly, one often studies simpler properties which 'certify' that various graph properties hold – perhaps not giving the best possible value of a parameter, but enabling a quick guarantee.

One important route to such information is through spectral graph theory. The spectra of various matrices associated with graphs is known to capture various structural and geometric properties. For instance, the well-known Cheeger inequality relates isoperimetry in a graph (that is, the size of normalized cuts) with the first non-trivial eigenvalue of the Laplacian. Kirkhoff's matrix tree theorem [26] relates the spectra of the (combinatorial) Laplacian with the number of spanning trees – which, again, can be thought of as a measure of how well connected a graph is. As another example, the discrepancy inequality commonly known as the expander mixing lemma uses the eigenvalues of the adjacency matrix (for regular graphs) or the normalized Laplacian (for general graphs) to certify the pseudo-random properties of the edge set of a graph. These facts underlie many commonly used graph partitioning, clustering, and drawing algorithms.

One particularly interesting aspect of the spectral graph theory of the Laplacian matrices is the strong analogies between Laplace operators on graphs and the Laplace-Beltrami operator on Riemannian manifolds – perhaps most notably through the Cheeger inequality, but also through inequalities relating eigenvalues and diameters (for instance). As a result of this, recent years have seen researchers try to adapt other concepts on manifolds to the discrete setting. A particular point of interest has been on the development of notions of curvature for graphs. One markedly successful aspect of this work has been the adaptation of the so-called curvature dimension inequality $CD(N, K)$, introduced by Bakry and Émery in a different setting, along with some variants to graphs. Curvature lower bounds in this sense, which is more precisely defined in Chapter 3, have been used to prove bounds on eigenvalues, and diameter, along with being used to study heat flow on graphs and, through this, to establish other geometric properties of networks.

Many of these works have been largely analytic in their character – using curvature to study networks via eigenfunctions and the mixing of random walks (through heat flow). In this work, we take a more directly combinatorial view of curvature and aim to use the definition of curvature to establish a number of purely combinatorial consequences of a curvature lower bound directly.

Our main result is that, just as eigenvalues of a graph certify the pseudo-randomness of the edge set, that a curvature lower bound at a point establishes a local pseudo-randomness – and that the more 'positively curved' a graph looks at a point, the more the edge sets in neighborhoods of that vertex behave (in a precise sense) pseudo-randomly. This result, and some variants, are then used to obtain a number of combinatorial properties of graphs. For instance, a bound on the connectivity of a connected graph is given in terms of a curvature lower bound – interestingly, curvature (being a local property) cannot detect whether a given graph is connected but it *can* detect that a connected graph is well connected.

The remainder of the chapter is organized as follows. In Section 2.2 we introduce the curvature-dimension inequality introduced by Bakry and Émery [5], and give a reformulation of this inequality which is more useful for our purposes. In Section 2.3 we derive our main result, a discrepancy inequality depending on curvature. Several corollaries and variants are also presented to highlight the flexibility of results of this type. In Section 2.4 we present several combinatorial consequences of our result, dealing with the connectivity of a graph and its local neighborhoods, and small subgraph containment. In Section 14 we explore graph operations which don't decrease curvature and in Section 2.6 we consider future work.

2.2 Preliminaries

In this chapter, $G = (V, E)$ will denote a simple graph. We denote the degree of the vertex $x \in V(G)$ by $\mathrm{d}(x)$ and for a subset $X \subset V$ we denote by $\mathrm{d}_X(x)$ the number of neighbors of x in X. The minimum degree is denoted by δ. We write $x \sim y$ when $xy \in E(G)$. We only consider graphs which are locally finite; that is, $\mathrm{d}(x) < \infty$ for all $x \in V(G)$ (which encompass all finite graphs). Given a measure $\mu : V \to \mathbb{R}$, the μ-Laplacian on G is the operator $\Delta : \mathbb{R}^{|V|} \to \mathbb{R}^{|V|}$ defined by

$$\Delta f(x) = \frac{1}{\mu(x)} \sum_{y \sim x} (f(y) - f(x)).$$

In the case $\mu(x) = 1$ for all $x \in V(G)$, we have $\Delta = -L$, where L is the combinatorial graph Laplacian as defined in the introduction. Also of interest is when $\mu(x) = \mathrm{d}(x)$ for all $x \in V(G)$, in which case Δ is the normalized graph Laplacian with a change of sign. Both operators are used in various applications. For the types of combinatorial results that we aim for, we will use the case $\mu = 1$. This corresponds to the combinatorial (standard) graph Laplacian.

The gradient form $\Gamma = \Gamma^{\Delta}$ is defined by

$$\begin{aligned}\Gamma(f, g)(x) &= \frac{1}{2} \left(\Delta(f \cdot g) - f \cdot \Delta(g) - \Delta(f) \cdot g \right)(x) \\ &= \frac{1}{2\mu(x)} \sum_{y \sim x} (f(y) - f(x))(g(y) - g(x))\end{aligned}$$

for all $f, g \in \mathbb{R}^{|V|}$. For brevity's sake we write $\Gamma(f) = \Gamma(f, f)$. The iterated gradient form $\Gamma_2 = \Gamma_2^{\Delta}$ is defined by

$$2\Gamma_2(f, g) = \Delta\Gamma(f, g) - \Gamma(f, \Delta g) - \Gamma(\Delta f, g)$$

for all $f, g \in \mathbb{R}^{|V|}$. Again, we write $\Gamma_2(f) = \Gamma_2(f, f)$.

In their 1985 paper [5], Bakry and Émery demonstrated that in the manifold setting curvature lower bounds may be understood completely through the Laplace-Beltrami operator. They did so through a modification of Bochner's identity, a fundamental identity in Riemannian Geometry. Bakry and Émery suggested using this *consequence* of curvature in the manifold setting as a *definition* in the more general setting of the Markov Semigroup, which encompasses both diffusion and random walks on graphs. This approach has been met with much success see e.g. [6, 23–25, 31, 34–37, 39, 46, 48–50, 56]. This is in no small part due to the close ties between the Laplace-Beltrami Operator and the Laplacian, as defined above. Through use of the differential operators Γ and Γ_2, constructed from the Laplacian, we state the Bakry-Émery notion of discrete curvature below.

Definition 1 (Bakry-Émery Curvature)**.** Let G be a locally finite graph. For $K \in \mathbb{R}$ and $N \in (0, \infty]$, we say that a vertex $x \in V$ satisfies Bakry-Émery's curvature-dimension inequality *CD*(N, K), if for any $f : V \to \mathbb{R}$, we have

$$\Gamma_2(f)(x) \geq \frac{1}{N}(\Delta f(x))^2 + K\Gamma(f)(x),$$

where N is a dimension parameter and K is regarded as a lower Ricci curvature bound at x. A graph G is said to satisfy *CD*(N, K) if every vertex satisfies *CD*(N, K). Since graphs do not have a well-defined dimension, a natural choice simplifying this inequality is to take $n = \infty$. In this thesis the curvature of a vertex (or graph) is defined as the maximum value K for which *CD*(∞, K) holds, for that vertex (or globally).

When considering curvature as a measure of local expansion in the manifold case, we associate large negative curvature with spaces that expand well and large positive curvature to spaces without much local expansion. We find this sentiment echoed in the discrete case. For example, it is known due to [27] that the curvature of the complete graph K_n is $1 + \frac{n}{2}$; it turns out for any d regular graph the curvature is at most $1 + \frac{d+1}{2}$. We provide a general upper bound on curvature in Example 1 and a lower bound for the curvature of the complete graph in Example 2.

The curvature of a d-regular tree, on the other hand, is $2 - d$ [27], the smallest curvature of a d-regular graph. We use Corollary 7 to prove this in in Example 3. The curvature of most graphs, of course, fall somewhere between these two extremes.

In Riemannian geometry, there are numerous results that hold for nonnegatively curved manifolds. If we are to extend these notions to graphs, a natural goal is to determine which classes of graphs have non-negative curvature. In [27], it is proven that the curvature of the hypercube Q_d is 2, regardless of the dimension. Furthermore, all finite abelian Cayley graphs, the discrete torus for example, have non-negative curvature. A notable case is that of the complete bipartite graph $K_{s,t}$, as studied in [17]. When $s = t$, the curvature of $K_{s,t}$ is 2. As the parts become more unbalanced, the curvature decreases, and is in fact negative in many cases.

Often graph curvature is studied on highly symmetric graphs. For vertex transitive graphs, such as Cayley graphs, the curvature of all vertices is the same. For less structured graphs, neighborhood structures differ among vertices and the curvatures of different vertices differs. In these cases, global graph curvature is a lower bound on curvature computed at the vertex level. In this thesis we explore local combinatorial implications of such curvature lower bounds. As we shall see, this gives a way to certify that graphs satisfy certain interesting properties.

As defined above, curvature is formulated via the Laplacian (Δ) and the gradient (Γ), each of which only depend on the the structure of the first neighborhood of vertices. Since the iterated gradient (Γ_2) is a composition of these two operators, the curvature is determined solely by the structure of the balls of radius 2. Therefore, we can translate *CD*(∞, K) into more combinatorial terms, highlighting the contribution of each edge type (that is, the edges within the first neighborhood of x, the edges between the first and second neighborhoods of x, etc.) to the curvature dimension inequality.

The proposition below allows us to begin to rework the CD inequality. For notational purposes, let $N_1(x) = \{y : d(x, y) = 1\}$ be the neighborhood of x and let $N_2(x) = \{z : d(x, z) = 2\}$ be the second neighborhood of x. When x has been fixed, we will often write these simply as N_1 and N_2.

Proposition 3. *G satisfies* *CD*(∞, K) *at x if and only if for all $f : V(G) \to \mathbb{R}$,*

$$\sum_{z \in N_2} \sum_{\substack{y \sim z \\ y \in N_1}} \left[\frac{1}{4}(f(z) - f(y))^2 - \frac{1}{2}(f(z) - f(y))(f(y) - f(x)) \right] + \sum_{\substack{y, y' \in N_1 \\ y \sim y'}} (f(y) - f(y'))^2 \geq \left(\frac{2K + \mathrm{d}(x) - 3}{2} \right) \Gamma(f)(x) - \frac{1}{2}\Delta(f(x))^2. \tag{2.1}$$

Proof. Expanding Γ_2 according to the combinatorial definitions of Δ and Γ yields

$$\begin{aligned}
\Gamma_2(f)(x) &= \frac{1}{2}\Delta\Gamma(f) - \Gamma(f, \Delta f) \\
&= \frac{1}{2}\sum_{y \sim x}[\Gamma(f)(y) - \Gamma(f)(x)] - \frac{1}{2}\sum_{y \sim x}(f(y) - f(x))(\Delta f(y) - \Delta f(x)) \\
&= \frac{1}{4}\sum_{y_1 \sim x}\left[\sum_{z \sim y_1}(f(z) - f(y_1))^2 - \sum_{y_2 \sim x}(f(y_2) - f(x))^2\right]
\end{aligned}$$

$$
\begin{aligned}
&- \frac{1}{2} \sum_{y_1 \sim x} (f(y_1) - f(x)) \left[\sum_{z \sim y_1} (f(z) - f(y)) - \sum_{y_2 \sim x} (f(y_2) - f(x)) \right] \\
= &\frac{1}{4} \sum_{y \sim x} \sum_{z \sim y} (f(z) - f(y))^2 - \frac{1}{4} \sum_{y_1 \sim x} \sum_{y_2 \sim x} (f(y_2) - f(x))^2 \\
&- \frac{1}{2} \sum_{y_1 \sim x} \sum_{z \sim y_1} (f(y_1) - f(x))(f(z) - f(y_1)) \\
&+ \frac{1}{2} \sum_{y_1 \sim x} \sum_{y_2 \sim x} (f(y_1) - f(x))(f(y_2) - f(x)) \\
= &\frac{1}{4} \sum_{y \sim x} \sum_{z \sim y} (f(z) - f(y))^2 - \frac{\mathrm{d}(x)}{4} \sum_{y \sim x} (f(y) - f(x))^2 \\
&- \frac{1}{2} \sum_{y_1 \sim x} \sum_{z \sim y_1} (f(y_1) - f(x))(f(z) - f(y_1)) \\
&+ \frac{1}{2} \sum_{y_1 \sim x} \sum_{y_2 \sim x} (f(y_1) - f(x))(f(y_2) - f(x)).
\end{aligned}
$$

To this sum, the edges between N_1 and N_2 contribute

$$
\sum_{z \in N_2} \sum_{\substack{y \sim z \\ y \in N_1}} \left[\frac{1}{4} (f(z) - f(y))^2 - \frac{1}{2} (f(z) - f(y))(f(y) - f(x)) \right]. \tag{2.2}
$$

To the sum, the edges between two vertices in N_1 contribute

$$
\begin{aligned}
&\sum_{\substack{y, y' \in N_1 \\ y \sim y'}} \Big[\frac{1}{2} (f(y) - f(y'))^2 - \frac{1}{2} (f(y) - f(x))(f(y') - f(y)) \\
&\qquad\qquad - \frac{1}{2} (f(y') - f(x))(f(y) - f(y')) \Big] \\
&= \sum_{\substack{y, y' \in N_1(x) \\ y \sim y'}} (f(y) - f(y'))^2.
\end{aligned} \tag{2.3}
$$

Lastly, the edges between x and N_1 contribute

$$
\sum_{y \sim x} \Big[\frac{1}{4} (f(x) - f(y))^2 - \frac{\mathrm{d}(x)}{4} (f(y) - f(x))^2 + \frac{1}{2} (f(y) - f(x))^2
$$

$$+ \frac{1}{2}(f(y) - f(x))\Delta f(x)\Big]$$

$$= \frac{3 - \mathrm{d}(x)}{4} \sum_{y \sim x} (f(x) - f(y))^2 + \frac{1}{2}\Delta f(x) \sum_{y \sim x} (f(y) - f(x)) \tag{2.4}$$

$$= \frac{3 - \mathrm{d}(x)}{2} \Gamma(f)(x) + \frac{1}{2}(\Delta f(x))^2. \tag{2.5}$$

By combining (2.2), (2.3), and (2.5) we see then that $\Gamma_2(f)(x)$ may be rewritten in combinatorial terms as

$$\sum_{z \in N_2} \sum_{\substack{y \sim z \\ y \in N_1}} \left[\frac{1}{4}(f(z) - f(y))^2 - \frac{1}{2}(f(z) - f(y))(f(y) - f(x)) \right]$$

$$+ \sum_{\substack{y, y' \in N_1(x) \\ y \sim y'}} (f(y) - f(y'))^2 + \frac{3 - \mathrm{d}(x)}{2} \Gamma(f)(x) + \frac{1}{2}(\Delta f(x))^2. \tag{2.6}$$

Substituting this rewritten Γ_2 into Definition 1 we see *CD*(∞, K) is satisfied at a vertex x if and only if for all functions f

$$\sum_{z \in N_2} \sum_{\substack{y \sim z \\ y \in N_1}} \left[\frac{1}{4}(f(z) - f(y))^2 - \frac{1}{2}(f(z) - f(y))(f(y) - f(x)) \right]$$

$$+ \sum_{\substack{y, y' \in N_1 \\ y \sim y'}} (f(y) - f(y'))^2$$

$$\geq \left(\frac{2K + \mathrm{d}(x) - 3}{2} \right) \Gamma(f)(x) - \frac{1}{2}\Delta(f(x))^2. \tag{2.7}$$

□

A similar combinatorial reworking of the curvature dimension inequality was established by Klartag, et al. in [27]. While their presentation of the curvature dimension inequality is useful in many applications, ours proved more fruitful for the types of inequalities derived in this thesis.

There are several approaches to calculating curvature of graph families. In this chapter, we relate edge distributions to curvature bounds. If we fix some vertex in a graph of interest and find the $CD(\infty, K)$ inequality is tight for a particular function, this K serves as an upper bound of the vertex and graph curvature. For example if we choose to evaluate the $CD(\infty, K)$ inequality of a function which measures the distance from a vertex x, we are provided with an upper bound of curvature dependent on the edge distribution in the ball of radius two around x. A nice example of this may be found in [27]. In Example 1 we show how we may use the combinatorial interpreitation of Γ_2 and Γ in Proposition 3 to find such bounds.

Example 1 (General Curvature upper bound)**.** Given G satisfying $CD(\infty, K)$ at x where $d = d(x)$,

$$K \leq \frac{d}{2} + \frac{3}{2} - \frac{e\Big(N_1(x), N_2(x)\Big)}{2d}.$$

Proof. By fixing some vertex x with degree d and applying the distance function $f(v) = \text{dist}(v, x)$ we find,

$2\Gamma_2(f)(x) = d^2 + \sum_{v \sim x} \Big(2 - \frac{d + deg(v)}{2}\Big) + \sum_{\Delta(x,v,u)} 1$

and $\Gamma(f)(x) = \frac{1}{2}d$.

We have the CD(∞,K) inequality simplifies to

$$d^2 + \sum_{v \sim x} \Big(2 - \frac{d + deg(v)}{2}\Big) + \sum_{\Delta(x,v,u)} 1 \geq dK.$$

Now let $d_{N_1(x)}(v) :=$ the number of edges from vertex v into $N_1(x)$ and $d_{N_2(x)}(v) :=$ the number of edges from vertex v into $N_2(x)$. We can simplify the inequality above to find

$$d^2 + \sum_{v \sim x} \Big(2 - \frac{d + deg_{N_1(x)}(v) + deg_{N_2(x)}(v) + 1}{2}\Big) + \sum_{v \sim x} \frac{deg_{N_1(x)}(v)}{2} \geq dK$$

$$\frac{d^2}{2} + \frac{3}{2}d - \frac{1}{2}\sum_{v\sim x} deg_{N_2(x)}(v) \geq Kd$$

and so

$$K \leq \frac{d}{2} + \frac{3}{2} - \frac{e\Big(N_1(x), N_2(x)\Big)}{2d}.$$

□

If we hope to find curvature lower bounds we instead find which values of K the *CD*(∞, K) inequality holds for *all* functions on *all* vertices, we explore this more in Chapter 3.

2.3 Discrepancy Inequalities

Spectral graph theory, in particular the spectral theory of Laplace operators on graphs, is particularly useful in giving 'cheap' certifications of graph properties; for instance, despite the question of graph Hamiltonicity being NP-complete, polynomial time checks of regularity and spectral gap conditions may be used to certify Hamiltonicity [28]. One of the most important properties that the spectra of a graph certifies is the pseudo-randomness of the edge set, through what is commonly known as the Expander Mixing Lemma, which we state now in its form for regular graphs.

Lemma 4 (Expander Mixing Lemma)**.** [2] Suppose $G = (V, E)$ is a d-regular n-vertex graph. Denote by λ the second largest eigenvalue in absolute value of the adjacency matrix. Then $\forall S, T \subseteq V$,

$$\left| e(S,T) - \frac{d}{n}|S||T| \right| \leq \lambda\sqrt{|S||T|},$$

where $e(S,T)$ denotes the number of edges between S and T.

This simple proposition implies that graphs with a large spectral gap behave like a random graph with respect to its edge distribution. Having global control on the edge distribution is extremely useful when it occurs, but in some cases it is too much to hope for. For example, many real world graphs display clustering where neighbors of a vertex are more likely to be connected among themselves. Furthermore, as most degrees tend to be small in comparison to the size of the entire network, much of the 'action' in the graph occurs locally.

For these types of graphs, Bakry-Émery Curvature can be particularly useful. The Bakry-Émery Curvature at a vertex x may also be seen as the largest K such that the matrix $\Gamma_2 - K\Gamma$ is positive semidefinite. The order of this matrix depends only on the size of second neighborhood of x, as opposed to the number of vertices in the graph n, and hence for a graph with max degree $\Delta(G)$, the curvature at a point x may be calculated with high accuracy in time complexity $O(\Delta(G)^6 \log(\Delta(G)))$ as documented in [16]. Hence, for a graph of bounded degree, a global curvature lower bound can be computed in linear time in the order of the graph, with the implied constant depending on the maximum degree.

This makes using curvature to certify certain graph properties efficient computationally. Moreover, curvature lower bounds certify similar properties to those of the graph spectrum. For instance, curvature bounds imply bounds on mixing of random walks, on the diameter of a graph, and (through Buser's inequality) regarding cut properties of the graph.

In this section, we show that curvature actually implies a type of 'local' discrepancy inequality, which implies that at vertices with a lower curvature bound the edge distribution within the first and second neighborhood behaves, in a sense, pseudo-randomly. We now present our main result below.

Theorem 5. Suppose G is a graph satisfying *CD*(∞, K) at $x \in V(G)$. Fix $X \subseteq N_1(x)$ with $\overline{X} = N_1(x) \setminus X$. If $|X| = \alpha|N_1(x)|$, then

$$\sum_{z \in N_2} \left(\frac{\left[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right]^2}{\mathrm{d}_{N_1}(z)} \right) \leq$$

$$\frac{3}{4}[\alpha^2 \cdot e(\overline{X}, N_2) + (1-\alpha)^2 \cdot e(X, N_2)] + e(X, \overline{X}) - \left(\frac{2K + \mathrm{d}(x) - 3}{4}\right)\alpha(1-\alpha)\mathrm{d}(x).$$

The sum on the lefthand side of the inequality of Theorem 5 measures how 'randomly' edges are distributed from $N_1(x)$ to $N_2(x)$ in the following sense. If the edges were distributed randomly, and a set $X \subseteq N_1(x)$ were fixed we would expect a proportion close to $|X|/|N_1(x)|$ of the edges to have an endpoint in X – moreover, we would expect this to (roughly) be true for the edges incident to every vertex in $N_2(x)$. The sum on the left hand side exactly measures this deviation from 'random' in the aggregate – note that if the edges were precisely distributed in this way then this discrepancy term would be zero. Of course, this is too much to expect even if the edges in the neighborhood are truly distributed randomly!

Then, what we obtain is an upper bound on this deviation in such a way that the larger the curvature is the more 'random-like' the behavior is. Note that this mimics the theme of the Expander Mixing Lemma – a graph parameter is certifying a random-like behavior of the edge set – only now locally, between the first and second neighborhoods. Hence, this theorem acts as a *local* discrepancy inequality.

Proof. Fix $x \in V(G)$. We proceed by defining an explicit function f and interpreting the curvature dimension inequality, per Proposition 3. Define $f(x) = 0$. Let $X \subseteq N_1(x)$ such that $|X| = \alpha|N_1(x)|$ and let $\overline{X} = N_1(x) \setminus X$. For each $y \in X$, let $f(y) = 1 - \alpha$ and for each $y \in \overline{X}$, let $f(y) = -\alpha$. Note that for this function $\Delta f(x) = 0$ and

$$\begin{aligned}
\Gamma(f)(x) &= \frac{1}{2}[\alpha(1-\alpha)^2 + (1-\alpha)(\alpha)^2]\mathrm{d}(x) \\
&= \frac{1}{2}(\alpha - 2\alpha^2 + \alpha^3 + \alpha^2 - \alpha^3)\mathrm{d}(x) \\
&= \frac{1}{2}(\alpha - \alpha^2)\mathrm{d}(x) \\
&= \frac{1}{2}\alpha(1-\alpha)\mathrm{d}(x).
\end{aligned}$$

Our goal is, for each $z \in N_2(x)$, to select $f(z)$ in order to minimize the left side of the inequality derived from $CD(\infty, K)$. Thus, fix $z \in N_2(x)$. Then

$$\begin{aligned}
&\sum_{\substack{y \sim z \\ y \in N_1}} \left[\frac{1}{4}(f(z)-f(y))^2 - \frac{1}{2}(f(z)-f(y))(f(y)-f(x))\right] \\
&= \mathrm{d}_X(z)\left[\frac{1}{4}(f(z)-(1-\alpha))^2 - \frac{1}{2}(f(z)-(1-\alpha))(1-\alpha)\right] \\
&\qquad\qquad + \mathrm{d}_{\overline{X}}(z)\left[\frac{1}{4}(f(z)+\alpha)^2 + \frac{1}{2}(f(z)+\alpha)\alpha\right] \\
&= \frac{1}{4}\mathrm{d}_{N_1}(z)f(z)^2 + \left[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right] f(z) + \frac{3}{4}[\alpha^2 \mathrm{d}_{\overline{X}}(z) + (1-\alpha)^2 \mathrm{d}_X(z)].
\end{aligned}$$

By taking a formal derivative with respect to $f(z)$, we see that the sum is minimized when

$$f(z) = -\frac{2\left[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right]}{\mathrm{d}_{N_1}(z)}. \tag{2.8}$$

For this selection of $f(z)$,

$$\begin{aligned}
&\sum_{\substack{y \sim z \\ y \in N_1}} \left[\frac{1}{4}(f(z)-f(y))^2 - \frac{1}{2}(f(z)-f(y))(f(y)-f(x))\right] \\
&= -\frac{\left[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right]^2}{\mathrm{d}_{N_1}(z)} + \frac{3}{4}\left[\alpha^2 \mathrm{d}_{\overline{X}}(z) + (1-\alpha)^2 \mathrm{d}_X(z)\right].
\end{aligned}$$

Summing over $z \in N_2$, we find

$$\begin{aligned}
&\sum_{z \in N_2} \sum_{\substack{y \sim z \\ y \in N_1}} \left[\frac{1}{4}(f(z)-f(y))^2 - \frac{1}{2}(f(z)-f(y))(f(y)-f(x))\right] \\
&= \sum_{z \in N_2}\left(-\frac{\left[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right]^2}{\mathrm{d}_{N_1}(z)} + \frac{3}{4}\left[\alpha^2 \mathrm{d}_{\overline{X}}(z) + (1-\alpha)^2 \mathrm{d}_X(z)\right]\right) \\
&= \sum_{z \in N_2}\left(-\frac{\left[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right]^2}{\mathrm{d}_{N_1}(z)}\right) + \frac{3}{4}[\alpha^2 e(\overline{X}, N_2) + (1-\alpha)^2 e(X, N_2)].
\end{aligned}$$

As a result, satisfying *CD*(∞, K) implies that

$$-\sum_{z\in N_2}\left(\frac{\left[\alpha \mathrm{d}_{\overline{X}}(z)-(1-\alpha)\mathrm{d}_X(z)\right]^2}{\mathrm{d}_{N_1}(z)}\right)+\frac{3}{4}[\alpha^2 e(\overline{X},N_2)+(1-\alpha)^2 e(X,N_2)]+e(X,\overline{X})\geq \left(\frac{2K+\mathrm{d}(x)-3}{4}\right)\alpha(1-\alpha)\mathrm{d}(x).$$

Solving for this first term yields

$$\sum_{z\in N_2}\left(\frac{\left[\alpha \mathrm{d}_{\overline{X}}(z)-(1-\alpha)\mathrm{d}_X(z)\right]^2}{\mathrm{d}_{N_1}(z)}\right)\leq \frac{3}{4}[\alpha^2 e(\overline{X},N_2)+(1-\alpha)^2 e(X,N_2)]+e(X,\overline{X})-\left(\frac{2K+\mathrm{d}(x)-3}{4}\right)\alpha(1-\alpha)\mathrm{d}(x).$$

□

Corollary 6. Suppose G is a graph satisfying *CD*(∞, K) at $x\in V(G)$. Let $X\subseteq N_1(x)$ with $\overline{X}=N_1(x)\setminus X$. If $|X|=\alpha|N_1(x)|$, then

$$\frac{\left(\sum_{z\in N_2}|\alpha \mathrm{d}_{\overline{X}}(z)-(1-\alpha)\mathrm{d}_X(z)|\right)^2}{e(N_1,N_2)}\leq \frac{3}{4}\left[\alpha^2 e(\overline{X},N_2)+(1-\alpha)^2 e(X,N_2)\right]+e(X,\overline{X})-\left(\frac{2K+\mathrm{d}(x)-3}{4}\right)\alpha(1-\alpha)\mathrm{d}(x).$$

Proof. We use Cauchy-Schwarz to simplify the sum

$$\sum_{z\in N_2}\frac{\left[\alpha \mathrm{d}_{\overline{X}}(z)+(1-\alpha)\mathrm{d}_X(z)\right]^2}{\mathrm{d}_{N_1}(z)}.$$

Using Cauchy-Schwarz in the form of

$$\sum a_i^2\geq \frac{\left(\sum a_i b_i\right)^2}{\sum b_i^2},$$

with

$$a_i = \frac{\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)}{\sqrt{\mathrm{d}_{N_1}(z)}} \text{ and } b_i = \sqrt{\mathrm{d}_{N_1}(z)}$$

we obtain

$$\begin{aligned}\sum_{z\in N_2} \frac{\left[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right]^2}{\mathrm{d}_{N_1}(z)} &\geq \frac{\left(\sum_{z\in N_2}\left[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right]\right)^2}{\sum_{z\in N_2}\mathrm{d}_{N_1}(z)} \\ &= \frac{\left(\sum_{z\in N_2}\left[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right]\right)^2}{e(N_1,N_2)}.\end{aligned}$$

As a result, satisfying $CD(\infty, K)$ implies that

$$\begin{aligned}-\frac{\left(\sum\left|\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right|\right)^2}{e(N_1,N_2)} + \frac{3}{4}\left[\alpha^2 e(\overline{X},N_2) + (1-\alpha)^2 e(X,N_2)\right] + e(X,\overline{X}) \geq \\ \left(\frac{2K + \mathrm{d}(x) - 3}{4}\right)\alpha(1-\alpha)\mathrm{d}(x).\end{aligned}$$

□

This corollary gives us a much cleaner version of the above theorem. However, in some situations, this corollary gives away too much in its use of Cauchy-Schwarz. While we typically think of applying these results to bound the discrepancy of a graph whose curvature is known, they can also be applied to bound the curvature in cases where the local discrepancy is understood. For example, our theorem gives a sharp bound on curvature for the graph $\mathbb{Z}^d$, while the corollary only gives us an asymptotic upper bound of d on the curvature. We will more fully explore similar examples later in Section 2.4.

We highlight a few special instances of these results.

Corollary 7. Suppose G is a graph satisfying *CD*(∞, K) at $x \in V(G)$. For any $X \subseteq N(x)$ with $|X| = \frac{1}{2}|N(x)|$, if $\overline{X} = N(x) \setminus X$, then

$$\frac{\left(\sum_{z\in N_2} |\mathrm{d}_{\overline{X}}(z) - \mathrm{d}_X(z)|\right)^2}{e(N_1, N_2)} \leq \frac{3}{4}e(N_1, N_2) + 4e(X, \overline{X}) - \frac{2K + \mathrm{d}(x) - 3}{4}\mathrm{d}(x).$$

Proof. Taking $\alpha = \frac{1}{2}$ in the above theorem directly gives the corollary. □

Corollary 8. Suppose G is a triangle-free graph satisfying *CD*(∞, K) at $x \in V(G)$. Let $X \subseteq N_1(x)$ with $\overline{X} = N_1(x) \setminus X$. If $|X| = \alpha|N_1(x)|$, then

$$\frac{\left(\sum_{z\in N_2} |\alpha\mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)|\right)^2}{e(N_1, N_2)} \leq$$
$$\frac{3}{4}\left[\alpha^2 e(\overline{X}, N_2) + (1-\alpha)^2 e(X, N_2)\right] - \left(\frac{2K + \mathrm{d}(x) - 3}{4}\right)\alpha(1-\alpha)\mathrm{d}(x).$$

Proof. If $e(X, \overline{X}) > 0$, then there exist adjacent vertices $y, y' \in N_1$. This edge then creates a triangle with x. Since G is triangle-free, then, $e(X, \overline{X}) = 0$. □

2.3.1 Local Buser Inequality. The isoperimtric problem in a graph – how sparse is the sparsest normalized cut? – is well-known to be closely related to the spectrum of the normalized Laplace operator via the Cheeger and Buser inequalities. It concerns the isoperimetric constant

$$\Phi(G) = \min_{S\subseteq V(G)} \frac{e(S, \overline{S})}{\min(\mathrm{vol}(S), \mathrm{vol}(\overline{S}))}.$$

The 'standard' form of the Cheeger inequality for graphs states that

$$\Phi(G)^2/2 \leq \lambda_2(\mathcal{L}) \leq 2\Phi(G).$$

Here the lower bound is the analogue of Cheeger's inequality from Riemannian geometry, while the upper bound is roughly an analogue of Buser's inequality. We remark here that the upper bound can be improved under a curvature assumption; such a result depending on curvature is the original Buser's inequality from manifolds, and is known for graphs satisfying $CD(\infty, K)$ [27].

The simple upper bound $\lambda_2(\mathcal{L}) \leq 2\Phi(G)$ on the first non-trivial eigenvalue of the Laplacian proceeds by fixing a set S minimizing the isoperimetric constant $\dfrac{e(S,\overline{S})}{\text{vol}(S)}$. Then one explicitly defines a vector φ based on this cut and computes via the Rayleigh quotient that

$$\lambda_2(\mathcal{L}) \leq \frac{e(S,\overline{S})\text{vol}(V)}{\text{vol}(S)\text{vol}(\overline{S})} \leq 2\Phi(G).$$

This inequality is tremendously useful in practice, as it implies that for all sets S

$$e(S,\overline{S}) \geq \lambda_2(\mathcal{L})\frac{\text{vol}(S)\text{vol}(\overline{S})}{\text{vol}(G)},$$

which when restricted to the d-regular case states

$$e(S,\overline{S}) \geq \frac{d\lambda_2(\mathcal{L})}{n}|S||\overline{S}|. \tag{2.9}$$

We remark that local variations of Cheeger's inequality have been the basis of several successful local graph partitioning algorithms. While Cheeger's inequality is a statement about all subsets of G, of which there are exponentially many, most proofs of Cheeger's inequality (and its variants) follow by considering a much smaller number of cuts defined by some process – eg. via (classically) an eigenvector sweep, or a sweep of the distribution of a random walk after some number of steps. The underlying ideas have, then, been used to develop a number of fast algorithms for finding sparse cuts in a graph. For example, one such algorithm given by Spielman and Teng in 2004 [57] considers subsets based off of a rapid mixing result for random walks given by Lovász and Simonovits [42, 43]. Through a modified Cheeger's inequality, this graph partitioning algorithm runs in time proportional to the size of the output – independent of the size of G. This approach and others like it have been particularly useful in solving problems involving massive networks.

It turns out that, just as curvature provides a 'local' version of a discrepency inequality it also provides a local version of the Buser inequality in the form of (2.9). We explore this in a variant of Theorem 5 and in Section 2.4 we use it to prove two connectivity results, one local and the other global. Both proofs rely on the connectivity of the punctured ball of radius two, which we define as $\mathring{B}_2(x)$, the subgraph of G containing all vertices with distance 1 or 2 from x, and all edges between these vertices, except those edges between vertices of distance 2 from x. Theorem 5 and the corrolaries that followed, show how a curvature lower bound ensures that most vertices have a close to proportionate edge count between bipartitions of $N_1(x)$. We use this to speak to edge counts between sets in a variant of our main result, Theorem 10. This allows us to show, for example, Corrolary 11: Suppose G satisfies $\boldsymbol{CD}(\infty, K)$ at x, and x has degree d, for all partitions of $\mathring{B}_2(x)$ into S and $\overline{S}$,

$$e(S, \overline{S}) \geq \frac{2K + d - 3}{4d} |N(x) \cap S| \cdot |N(x) \cap \overline{S}|.$$

Notice the similarity between this inequality and the Cheeger inequality on regular graphs, equation (2.9). To prove Theorem 10 we first present the following lemma which gives us a reworking of the discrepancy term of Theorem 2.3.

Lemma 9. Suppose $x \in V(G)$ and let $X, \overline{X}$ partition $N_1(x)$ and $A, \overline{A}$ partition $N_2(x)$. Let $\alpha = \dfrac{|X|}{|N_1(x)|}$. Then

$$\sum_{z \in N_2(x)} \frac{\left[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right]^2}{\mathrm{d}_{N_1}(z)}$$
$$= \Big[\alpha^2 e(\overline{X}, N_2) + (1-\alpha)^2 e(X, N_2)\Big] - e(X, \overline{A}) - e(\overline{X}, A) + \sum_{z \in \overline{A}} \frac{\mathrm{d}_X(z)^2}{\mathrm{d}_{N_1}(z)} + \sum_{z \in A} \frac{\mathrm{d}_{\overline{X}}(z)^2}{\mathrm{d}_{N_1}(z)}.$$

Proof. Equality follows from a series of algebraic manipulations. First, we split the sum over $N_2(x)$ of the discrepancy term into a sum over A and a sum over $\overline{A}$,

$$\sum_{z\in\overline{A}} \frac{\left[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right]^2}{\mathrm{d}_{N_1}(z)} + \sum_{z\in A} \frac{\left[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right]^2}{\mathrm{d}_{N_1}(z)}.$$

The squares are rewritten in terms of edges to N_1 to find

$$\sum_{z\in\overline{A}} \frac{\left[\alpha \mathrm{d}_{N_1}(z) - \mathrm{d}_X(z)\right]^2}{\mathrm{d}_{N_1}(z)} + \sum_{z\in A} \frac{\left[(1-\alpha)\mathrm{d}_{N_1}(z) - \mathrm{d}_{\overline{X}}(z)\right]^2}{\mathrm{d}_{N_1}(z)}.$$

We expand the squares to find

$$\sum_{z\in\overline{A}} \alpha^2 \mathrm{d}_{N_1}(z) - 2\alpha \mathrm{d}_X(z) + \frac{\mathrm{d}_X(z)^2}{\mathrm{d}_{N_1}(z)} + \sum_{z\in A} (1-\alpha)^2 \mathrm{d}_{N_1}(z) - 2(1-\alpha)\mathrm{d}_{\overline{X}}(z) + \frac{\mathrm{d}_{\overline{X}}(z)^2}{\mathrm{d}_{N_1}(z)}.$$

Summing over z in A and $\overline{A}$ yields

$$\begin{aligned} &\alpha^2 e(\overline{X}, \overline{A}) + \alpha^2 e(X, \overline{A}) - 2\alpha e(X, \overline{A}) + \sum_{z\in\overline{A}} \frac{\mathrm{d}_X(z)^2}{\mathrm{d}_{N_1}(z)} \\ &\quad + (1-\alpha)^2 e(X, A) + (1-\alpha)^2 e(\overline{X}, A) - 2(1-\alpha) e(\overline{X}, A) + \sum_{z\in A} \frac{\mathrm{d}_{\overline{X}}(z)^2}{\mathrm{d}_{N_1}(z)} \end{aligned}$$

which can be seen as

$$\alpha^2 e(\overline{X}, N_2) - \alpha^2 e(\overline{X}, A) + \alpha^2 e(X, \overline{A}) - 2\alpha e(X, \overline{A}) + \sum_{z \in \overline{A}} \frac{\mathrm{d}_X(z)^2}{\mathrm{d}_{N_1}(z)}$$
$$+ (1-\alpha)^2 e(X, N_2) - (1-\alpha)^2 e(X, \overline{A})$$
$$+ (1-\alpha)^2 e(\overline{X}, A) - 2(1-\alpha) e(\overline{X}, A) + \sum_{z \in A} \frac{\mathrm{d}_{\overline{X}}(z)^2}{\mathrm{d}_{N_1}(z)}.$$

The discrepancy term is then rewritten as to more easily compare it against other terms in the *CD*(∞, K) inequality

$$\begin{aligned}
&\Big[\alpha^2 e(\overline{X}, N_2) + (1-\alpha)^2 e(X, N_2)\Big] - 2\alpha e(X, \overline{A}) - (1-\alpha)^2 e(X, \overline{A}) + \alpha^2 e(X, \overline{A}) \\
&\quad - 2(1-\alpha) e(\overline{X}, A) - \alpha^2 e(\overline{X}, A) + (1-\alpha)^2 e(\overline{X}, A) + \sum_{z \in \overline{A}} \frac{\mathrm{d}_X(z)^2}{\mathrm{d}_{N_1}(z)} + \sum_{z \in A} \frac{\mathrm{d}_{\overline{X}}(z)^2}{\mathrm{d}_{N_1}(z)} \\
&\quad = \Big[\alpha^2 e(\overline{X}, N_2) + (1-\alpha)^2 e(X, N_2)\Big] - \big[2\alpha + (1-\alpha)^2 - \alpha^2\big] e(X, \overline{A}) \\
&\qquad - \Big[2(1-\alpha) + \alpha^2 - (1-\alpha)^2\Big] e(\overline{X}, A) + \sum_{z \in \overline{A}} \frac{\mathrm{d}_X(z)^2}{\mathrm{d}_{N_1}(z)} + \sum_{z \in A} \frac{\mathrm{d}_{\overline{X}}(z)^2}{\mathrm{d}_{N_1}(z)} \\
&\quad = \Big[\alpha^2 e(\overline{X}, N_2) + (1-\alpha)^2 e(X, N_2)\Big] - e(X, \overline{A}) - e(\overline{X}, A) \\
&\qquad\qquad + \sum_{z \in \overline{A}} \frac{\mathrm{d}_X(z)^2}{\mathrm{d}_{N_1}(z)} + \sum_{z \in A} \frac{\mathrm{d}_{\overline{X}}(z)^2}{\mathrm{d}_{N_1}(z)}.
\end{aligned}$$

□

In Lemma 9, we see that the discrepancy term may be rewritten to provide information on the distribution of edges between all bipartitions of $\mathring{B}_2(x)$. We may use Lemma 9 to relate this distribution to graph curvature in Theorem 10.

Theorem 10. Suppose G is a graph satisfying *CD*(∞, K) at $x \in V(G)$. Let $X, \overline{X}$ partition $N_1(x)$ and $A, \overline{A}$ partition $N_2(x)$. It follows that

$$e(X, \overline{A}) + e(\overline{X}, A) + e(X, \overline{X}) \geq \left(\frac{2K + \mathrm{d}(x) - 3}{4}\right) \frac{|X| \cdot |\overline{X}|}{\mathrm{d}(x)}.$$

Proof. Using Theorem 5 , let $\alpha = \dfrac{|X|}{\mathrm{d}(x)}$. We have the inequality,

$$\sum_{z \in N_2(x)} \left(-\frac{\left[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)\right]^2}{\mathrm{d}_{N_1}(z)} \right) + \frac{3}{4}\left[\alpha^2 e(\overline{X}, N_2) + (1-\alpha)^2 e(X, N_2)\right]$$
$$+ e(X, \overline{X}) \geq \left(\frac{2K + \mathrm{d}(x) - 3}{4}\right) \alpha(1-\alpha)\mathrm{d}(x).$$

Through Lemma 9 we use the partition of $N_2(x)$ into A and $\overline{A}$ and substitute this into the inequality to find

$$-\left[\alpha^2 e(\overline{X}, N_2) + (1-\alpha)^2 e(X, N_2)\right] + e(X, \overline{A}) + e(\overline{X}, A) - \sum_{z \in \overline{A}} \frac{\mathrm{d}_X(z)^2}{\mathrm{d}_{N_1}(z)} - \sum_{z \in A} \frac{\mathrm{d}_{\overline{X}}(z)^2}{\mathrm{d}_{N_1}(z)}$$
$$+ \frac{3}{4}\left[\alpha^2 e(\overline{X}, N_2) + (1-\alpha)^2 e(X, N_2)\right] + e(X, \overline{X}) \geq \left(\frac{2K + \mathrm{d}(x) - 3}{4}\right) \alpha(1-\alpha)\mathrm{d}(x)$$

which simplifies to

$$e(X, \overline{X}) + e(X, \overline{A}) + e(\overline{X}, A)$$
$$- \frac{1}{4}\left[\alpha^2 e(\overline{X}, N_2) + (1-\alpha)^2 e(X, N_2)\right] - \sum_{z \in \overline{A}} \frac{\mathrm{d}_X(z)^2}{\mathrm{d}_{N_1}(z)} - \sum_{z \in A} \frac{\mathrm{d}_{\overline{X}}(z)^2}{\mathrm{d}_{N_1}(z)}$$
$$\geq \left(\frac{2K + \mathrm{d}(x) - 3}{4}\right) \alpha(1-\alpha)\mathrm{d}(x).$$

Most notably,

$$e(X,\overline{A}) + e(\overline{X},A) + e(X,\overline{X}) \geq \left(\frac{2K + \mathrm{d}(x) - 3}{4}\right)\alpha(1-\alpha)\mathrm{d}(x).$$

With the substitution $\alpha = \frac{|X|}{\mathrm{d}(x)}$ and $1-\alpha = \frac{|\overline{X}|}{\mathrm{d}(x)}$ we have

$$e(X,\overline{A}) + e(\overline{X},A) + e(X,\overline{X}) \geq \left(\frac{2K + \mathrm{d}(x) - 3}{4}\right) \cdot \frac{|X| \cdot |\overline{X}|}{\mathrm{d}(x)}.$$

□

This immediately gives us the following corollary.

Corollary 11. Suppose G is a graph satisfying $CD(\infty, K)$ at $x \in V(G)$. Let the degree of x be d. For all partitions of $\mathring{B}_2(x)$ into S and $\overline{S}$,

$$e(S,\overline{S}) \geq \frac{2K + d - 3}{4d}|N(x) \cap S| \cdot |N(x) \cap \overline{S}|.$$

Proof. Given some partition of $\mathring{B}_2(x)$ into S and $\overline{S}$ we use Theorem 10 and let $X = S \cap N_1(x)$, $A = S \cap N_2(x)$, $\overline{X} = \overline{S} \cap N_1(x)$ and $\overline{A} = \overline{S} \cap N_2(x)$. □

2.4 Applications and Examples

2.4.1 Connectivity. Thus far we have seen several ways in which curvature of a graph can be used to certify similar properties to spectral properties, only locally. Eigenvalues of the Laplacian are well known to certify notions of connectivity of a graph in various forms: through Fielder's Theorem[18], Kirchhoff's matrix tree theorem [26], and Cheeger's inequality, for example.

In this section we present two results relating the curvature to connectivity. One, like before, is a 'local' connectivity result. The other, however, actually provides a truly global connectivity bound. We highlight here that a curvature bound alone cannot imply that a graph is connected. If one takes two disjoint copies of a single graph, curvature can never see the existence of the other copy – a lower curvature bound for one copy implies a lower curvature bound for both. None the less, a curvature lower bound *does* imply a strong bound on the connectivity of a connected graph. We note that a lower bound on connectivity for positively curved graphs can also be obtained by combining results of Chung, Lin and Yau [13] and Fielder [18]. However, we find a result which applies even when the curvature lower bound for a graph is not too negative.

We say G is l-connected if after the removal of any $l-1$ vertices G remains connected. We define $\kappa(G)$ as the largest l for which G is l-connected. We see through Corollary 11 that even moderate curvature conditions ensure balls of radius 2 to be well stitched together, we exploit the manner in which these balls overlap to provide a global connectivity bound.

Theorem 12 (Global Connectivity)**.** Suppose G is a connected graph which satisfies *CD*(∞, K) with minimum degree δ. Then

$$\kappa(G) \geq \frac{2K+\delta+5}{8}.$$

Remark: Unlike most theorems in this thesis, Theorem 12 requires the graph to satisfy *CD*(∞, K) at *all* $x \in V(G)$.

Proof. Suppose $G=(V,E)$ is a connected graph with vertex connectivity $\kappa(G)$. Then there exists $U=\{x_1,\ldots,x_{\kappa(G)}\} \subseteq V$ such that $G \setminus U$ is not connected. Label the components of $V \setminus U$ as $A_1, A_2, \cdots, A_k$. Let $S=A_1$ and $\overline{S}=\cup_{i=2}^{k} A_i$. Choose U_S and $U_{\overline{S}}$ to partition U in such a way as to minimize $e(S \cup U_S, \overline{S} \cup U_{\overline{S}})$. Notice that as a consequence of the minimization of $e(S \cup U_S, \overline{S} \cup U_{\overline{S}})$, for all $x_i \in U_S$, it is the case that $\mathrm{d}_{S \cup U_S}(x_i) \geq \mathrm{d}_{\overline{S} \cup U_{\overline{S}}}(x_i)$ and for $x_i \in U_{\overline{S}}$, it is the case that $\mathrm{d}_{\overline{S} \cup U_{\overline{S}}}(x_i) \geq \mathrm{d}_{S \cup U_S}(x_i)$. Fix $x \in U$ which maximizes $\min\{\mathrm{d}_S(x_i), \mathrm{d}_{\overline{S}}(x_i)\}$. Without loss of generality, assume $x \in U_S$. We now apply Corollary 11 with $X = N_1(x) \cap (S \cup U_S)$, $\overline{X} = N_1(x) \cap (\overline{S} \cup U_{\overline{S}})$, $A = N_2(x) \cap (S \cup U_S)$, and $\overline{A} = N_2(x) \cap (\overline{S} \cup U_{\overline{S}})$. From this we see,

$$e(N_{1,2}(x) \cap (S \cup U_S), N_{1,2}(x) \cap (\overline{S} \cup U_{\overline{S}})) \geq \left(\frac{2K + \mathrm{d}(x) - 3}{4}\right) \frac{|X| \cdot |\overline{X}|}{\mathrm{d}(x)}.$$

Notice that every edge passing from $S \cup U_S$ to $\overline{S} \cup U_{\overline{S}}$ is adjacent to a member of U, and by our choice of x, each member of U is incident to at most $|\overline{X}|$ of such edges. This allows us to bound the number of edges in $N_{1,2}(x)$ which pass from $S \cup U_S$ to $\overline{S} \cup U_{\overline{S}}$ above by $(\kappa(G) - 1)|\overline{X}|$,

$$(\kappa(G) - 1)|\overline{X}| \geq \left(\frac{2K + \mathrm{d}(x) - 3}{4}\right) \frac{|X| \cdot |\overline{X}|}{\mathrm{d}(x)}.$$

Recalling that $\frac{|X|}{\mathrm{d}(x)} \geq \frac{1}{2}$, we see

$$(\kappa(G) - 1) \geq \left(\frac{2K + \mathrm{d}(x) - 3}{8}\right).$$

Therefore,

$$\kappa(G) \geq \frac{2K + \mathrm{d}(x) + 5}{8}.$$

□

We remark that such a curvature based lower bound can, in theory at least, provide an efficient way of bounding the connectivity of a graph. Classical algorithms for determining vertex connectivity run in $O(|V|^3|E|)$ time ($O(|V| \cdot |E|)$ for determining the number of vertex disjoint paths between any fixed pair of vertices). Meanwhile, as briefly discussed above in Section 2.3, for a bounded degree graph the curvature at a given point can be computed in constant (depending on the degree bound) time and hence a global curvature bound can be computed in linear time. For details see [16]. A reasonably accurate curvature calculation at vertex requires a binary search on potential curvatures, taking $\log(\Delta)$ checks, and each check requires an eigenvalue calculation of a matrix of size $O(\Delta^2)$ (serving as an upper bound on the size of the second neighborhood of x) which takes at most $O(\Delta^6)$ time. Therefore, a curvature computation of graph with maximum degree Δ takes at most time $O(\Delta^6 \log(\Delta)|V(G)|)$. In practice, we expect to see degrees much smaller than the size of the network, making a curvature calculation in linear in V potentially more efficient than a more classical computation.

Finally, we present a slight improvement of a result in [17]. In [17], Cushing et al. showed that, with very few exceptions, if the curvature at x is positive then $\mathring{B}_2(x)$ is connected. Through Corollary 11 we show that with more moderate curvature assumptions, allowing for negative curvature, this remains true.

Theorem 13 (Local Connectivity)**.** Suppose G is a graph satisfying *CD*(∞, K) at $x \in V(G)$ and $2K + \mathrm{d}(x) > 3$, then $\mathring{B}_2(x)$ is connected.

Proof. Let $S, \overline{S}$ be nonempty and partition $\mathring{B}_2(x)$. Suppose $N_1 \subseteq \overline{S}$. Since S is nonempty it contains a member of N_2 and we see $e(S, \overline{S}) > 0$. A symmetric argument covers the case when $N_1 \subset S$. Supposing that $N_1 \cap S$ and $N_1 \cap \overline{S}$ are both nonempty, we use Theorem 10 letting $X = N_1 \cap S$, $A = N_2 \cap S$, $\overline{X} = N_1 \cap \overline{S}$ and $\overline{A} = N_2 \cap \overline{S}$ to find $e(S, \overline{S}) \geq \left(\frac{2K + \mathrm{d}(x) - 3}{4}\right) \cdot \frac{|X| \cdot |\overline{X}|}{\mathrm{d}(x)}$ which by assumption is greater than 0. Therefore, for all $S, \overline{S}$ nonempty and partitioning $\mathring{B}_2(x)$, $e(S, \overline{S}) > 0$ and therefore $\mathring{B}_2(x)$ is connected. □

Sharpness Example for Theorem 13: The two graphs in Figure 2.1 show that Theorem 13 is sometimes sharp. The graph on the right has curvature -1.5 and $d(x) = 6$ so $2K + d(x) = 3$ and we see that $\mathring{B}_2(x)$ is disconnected. On the other hand, the graph on the left has curvature $-.5$, $d(x) = 4$, we also have $2K + d(x) = 3$ but in this case $\mathring{B}_2(x)$ is connected. Following the string of inequalities leading to Corollary 11, we observe that the inequality in Theorem 13 only needs to be made strict when $N_2(x)$ is empty.

Figure 2.1: In the graph on the left, $\mathring{B}_2(x)$ is connected and $N_2(x)$ is nonempty, while in the graph on the right $\mathring{B}_2(x)$ is disconnected and $N_2(x)$ is empty.

2.4.2 Examples. As promised, we now provide the lower bound of K_n using Proposition 3.

Example 2 (K_n satisfies $CD(\infty, \frac{n}{2}+1)$).

Proof. By Proposition 3 we have K_n satisfies $CD(\infty, \frac{n}{2}+1)$ at x provided for all f,

$$\sum_{z\in N_2}\sum_{\substack{y\sim z\\ y\in N_1}}\left[\frac{1}{4}(f(z)-f(y))^2-\frac{1}{2}(f(z)-f(y))(f(y)-f(x))\right]$$
$$+\sum_{\substack{y,y'\in N_1\\ y\sim y'}}(f(y)-f(y'))^2$$
$$\geq\left(\frac{2K+\mathrm{d}(x)-3}{2}\right)\Gamma(f)(x)-\frac{1}{2}\Delta(f(x))^2.$$

Notice that since N_2 is empty this may be simplified to

$$\sum_{\substack{y,y'\in N_1\\ y\sim y'}}(f(y)-f(y'))^2\geq\left(\frac{2K+\mathrm{d}(x)-3}{2}\right)\Gamma(f)(x)-\frac{1}{2}\Delta(f(x))^2. \tag{2.10}$$

We may simplify the lefthand side of (2.10).

$$\begin{aligned}\sum_{\substack{y,y'\in N_1\\ y\sim y'}}(f(y)-f(y'))^2 &= -2\cdot\sum_{\substack{y,y'\in N_1\\ y\sim y'}} f(y)f(y')+(n-2)\sum_y f(y)^2\\ &= -\left(\sum_{y\in N_1} f(y)\right)^2+(n-1)\sum_y f(y)^2\\ &= -\Delta(f(x))^2+2(n-1)\Gamma(f)(x). \qquad (2.11)\end{aligned}$$

With the substitution of $\frac{n}{2}+1$ for K, and $n-1$ for $d(x)$, the righthand side of (2.10) becomes

$$(n-1)\Gamma(f)(x)-\frac{1}{2}\Delta(f(x))^2. \qquad (2.12)$$

Since (2.11) is nonnegative and twice (2.12), we see (2.10) holds. □

We now present a few examples to illustrate that the inequality presented in Theorem 5 can actually be tight. This implies, in turn, that the functions considered actually witness the limiting curvature for some family of graphs. This, in turn, means that the discrepancy properties of graph neighborhoods can, at times, imply sharp upper bounds on the curvature for graph families.

Example 3 (Regular Trees)**.**

We find that Corollary 7 gives a sharp upper bound on graph curvature of regular trees. Suppose that G is a d-regular tree. In any tree, every vertex in N_2 must be adjacent to exactly one vertex in N_1, as any vertex $z\in N_2$ with $\mathrm{d}_{N_1}(z)$ would form a 4-cycle with x and its neighbors in N_1. Thus, we have that

$$\sum_{z\in N_2}|\mathrm{d}_{\overline{X}}(z)-\mathrm{d}_X(z)|=\sum_{z\in N_2}1=d(d-1).$$

Furthermore, every tree is triangle-free, which implies that $e(X, \overline{X}) = 0$. Our theorem, in this case, states that

$$d(d-1) \leq \frac{3}{4}d(d-1) - \frac{2K + d - 3}{4}d.$$

Solving for K here yields that $K \leq 2 - d$, and as shown in [27], the curvature of a tree is exactly $K = 2 - d$. Therefore, our theorem gives a sharp upper bound on the curvature of a d-regular tree.

Example 4 ($\mathbb{Z}^d$)**.**

In this example, we find Theorem 5 provides a sharp upper bound on the curvature of $\mathbb{Z}^d$. Consider the graph $\mathbb{Z}^d$. Let $x = (0, \ldots, 0)$. Then $N_1(x) = \{(y_1, \ldots, y_d) : \sum |y_i| = 1\}$ and $N_2(x) = \{(z_1, \ldots, z_d) : \sum |z_i| = 2\}$. Define $X \subseteq N_1(x)$ to be the points $(y_1, \ldots, y_d)$ where $y_i = 1$ for some $i \in \{1, \ldots, d\}$ and $y_j = 0$ for all $j \neq i$. Then $\overline{X} \subseteq N_1(x)$ is the set of points $(y_1, \ldots, y_d)$ where $y_i = -1$ for some $i \in \{1, \ldots, d\}$ and $y_j = 0$ for all $j \neq i$. Note here that $\alpha = \frac{1}{2}$, as N_1 is split evenly according to this partition.

Let $z \in N_2(x)$. We will analyze how each possible z contributes to the sum on the left side.

- If z contains two 1s, then $|\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)| = 1$ and $\mathrm{d}_{N_1}(z) = 2$. Thus, the contribution of z to the sum is $\frac{1}{2}$. There are $\binom{d}{2}$ such vertices.
- If z contains two -1s, then $|\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)| = 1$ and $\mathrm{d}_{N_1}(z) = 2$. Thus, the contribution of z to the sum is $\frac{1}{2}$. There are again $\binom{d}{2}$ such vertices.
- If z contains a 2, then $|\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)| = \frac{1}{2}$ and $\mathrm{d}_{N_1}(z) = 1$. Thus, the contribution of z to the sum is $\frac{1}{4}$. There are d such vertices.
- If z contains a -2, then $|\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)| = \frac{1}{2}$ and $\mathrm{d}_{N_1}(z) = 1$. Thus, the contribution of z to the sum is $\frac{1}{4}$. There are d such vertices.

- If z contains a 1 and a -1, then $|\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)| = 0$ and $\mathrm{d}_{N_1}(z) = 2$. Thus, the contribution of z to the sum is 0. There are $2\binom{d}{2}$ such vertices.

Thus, the sum over all z yields

$$\sum_{z \in N_2} \left(\frac{[\alpha \mathrm{d}_{\overline{X}}(z) - (1-\alpha)\mathrm{d}_X(z)]}{\mathrm{d}_{N_1}(z)} \right) = \frac{1}{2}\binom{d}{2} + \frac{1}{2}\binom{d}{2} + \frac{1}{4}d + \frac{1}{4}d + 0 \cdot 2\binom{d}{2} = \frac{1}{2}d^2.$$

On the right side, $\alpha = \frac{1}{2}$. Thus,

$$\frac{3}{4}[\alpha^2 \cdot e(\overline{X}, N_2) + (1-\alpha)^2 e(X, N_2)] = \frac{3}{16} e(N_1, N_2).$$

To compute $e(N_1, N_2)$, every vertex in N_2 with two nonzero coordinates has two neighbors in N_1 and there are $4\binom{d}{2}$ of these vertices. Also, every vertex in N_2 with one nonzero coordinate has one neighbor in N_1 and there are $2d$ of these vertices. Thus, $e(N_1, N_2) = 8\binom{d}{2} + 2d = 4d^2 - 2d$. Since $\mathbb{Z}^d$ is triangle-free, we have that $e(X, \overline{X}) = 0$. Finally, the curvature term yields $\frac{d(2K + 2d - 3)}{8}$.

Therefore, our discrepancy inequality yields that

$$\frac{1}{2}d^2 \leq \frac{3}{16}(4d^2 - 2d) - \frac{d(2K + 2d - 3)}{8}.$$

Solving for K yields that $K \leq 0$, which again is a tight upper bound as [27] gives that $K = 0$.

We remark here that while we have chosen to consider combinatorial properties as they relate to the infinite curvature dimensional inequality, the functions we consider in our main result and the corollaries that follow effectively ignore the dimension term in the CD inequality. As a result, these results hold curvature when normalized to any dimension. We leave the reader to confirm that the function used in 2.3 satisfies $\Delta f = 0$.

2.5 Curvature and Graph Operations

Another way to better understand how this curvature on graphs works is to explore how curvature is affected by various graph operations. As we have seen before, the curvature of the disjoint union of graphs is calculated by considering the graphs separately. Ricci-flat graphs are closed under tensor, Cartesian and strong products. In the paper [17] the authors propose the following two conjectures, both of which we confirm below.

Conjecture 14. Let $G = (V, E)$ be a graph and $x \in V$ be a vertex. Let $G' = (V', E')$ be the graph obtained from G by one of the two operations:

- Delete a leaf in $S_2(x)$ and its incident edge.
- Delete $z \in S_2(x)$ and its incident edges $\{\{y, z\} \in E : y \in S_1(x)\}$; Adding a new edge between every two of $\{y \in S_1(x) : \{y, z\} \in E\}$.

Then we have for any $N \in (0, \infty]$ and $K \in \mathbb{R}$

If G satisfies *CD*(N, K) at x then so does G'.

We are only interested in the $N = \infty$ case, which we have proofs of below.

We will start with the following lemma.

Lemma 15. Fix some $z \in N_2(x)$ and let $Y_z = \{y : x \sim y \sim z\}$. If for all $y, y' \in Y_z$, $y \sim y'$, it will be the case that for any function f on Y_z, and for optimized values of $f(z)$,

$$\sum_{y \sim y' \in Y_z} (f(y) - f(y'))^2 \geq \sum_{y \in Y_z} \left(\frac{1}{4}(f(z) - f(y))^2 - \frac{1}{2}(f(z) - f(y))f(y) \right)$$

Proof. Fix some $z \in N_2(x)$ and let $Y_z = \{y : x \sim y \sim z\}$. Let $r = |Y_z|$.
By Cauchy-Schwarz we can say $\dfrac{(\sum_{y \in Y_z} f(y))^2}{r} \leq \sum_{y \in Y_z} f(y)^2$.

We have shown that the optimal value of $f(z) = \frac{2\sum_{y\in Y_z} f(y)}{r}$.

Notice also that

$$\sum_{y\sim y'\in Y_z} (f(y) - f(y'))^2 = \sum_{y\sim y'\in Y_z} f(y)^2 + f(y')^2 - 2f(y)f(y'). \tag{2.13}$$

Under the hypothesis that for $y, y' \in Y_z$, $y \sim y'$ then the above may be written as

$$\begin{aligned}
&= \sum_{y\in Y_z} \left[(rf(y)^2 - f(y) \sum_{y'\in Y_z} f(y') \right] \\
&= \sum_{y\in Y_z} rf(y)^2 - (\sum_{y'\in Y_z} f(y'))^2 \\
&\geq (r-1) \sum_{y\in Y_z} f(y)^2 + \frac{3}{4} \sum_{y\in Y_z} f(y)^2 - \left(\sum_{y'\in Y_z} f(y') \right)^2 \\
&\geq (r-1) \sum_{y\in Y_z} f(y)^2 + \frac{3}{4} \sum_{y\in Y_z} f(y)^2 - \frac{(r-1)(\sum_{y'\in Y_z} f(y'))^2}{r} - \frac{(\sum_{y'\in Y_z} f(y'))^2}{r}
\end{aligned} \tag{2.14}$$

Using Cauchy-Schwarz this may be simplified to

$$(2.14) \geq \frac{3}{4} \sum_{y\in Y_z} f(y)^2 - \frac{(\sum_{y'\in Y_z} f(y'))^2}{r} \tag{2.15}$$

$$\geq \frac{3}{4} \sum_{y\in Y_z} f(y)^2 - \frac{2(\sum_{y'\in Y_z} f(y'))^2}{r} + \frac{(\sum_{y'\in Y_z} f(y'))^2}{r} \tag{2.16}$$

$$\geq \frac{3}{4} \sum_{y\in Y_z} f(y)^2 - \frac{2(\sum_{y'\in Y_z} f(y'))^2}{r} + \frac{1}{4} r \left(\frac{2\sum_{y'\in Y_z} f(y')}{r} \right)^2 \tag{2.17}$$

$$\geq \sum_{y\in Y_z} \left(\frac{3}{4} f(y)^2 - \frac{2\sum_{y'\in Y_z} f(y')}{r} f(y) + \frac{1}{4} \left(\frac{2\sum_{y'\in Y_z} f(y')}{r} \right)^2 \right) \tag{2.18}$$

with the assumption that $f(z) = \dfrac{2\sum_{y\in Y_z} f(y)}{r}$ this becomes.

$$\geq \sum_{y\in Y_z}\left(\frac{3}{4}f(y)^2 - f(z)f(y) + \frac{1}{4}f(z)^2\right) \tag{2.19}$$
$$\geq \sum_{y\in Y_z}\left(\frac{1}{4}(f(z)-f(y))^2 - \frac{1}{2}(f(z)-f(y))f(y)\right). \tag{2.20}$$

Therefore,

$$\sum_{y\sim y'\in Y_z}(f(y)-f(y'))^2 \geq \sum_{y\in Y_z}\left(\frac{1}{4}(f(z)-f(y))^2 - \frac{1}{2}(f(z)-f(y))f(y)\right)$$

□

Theorem 16 (Proof of Conjecture 14)**.**

Proof. Let G, z and G' be as in Conjecture 14. Let Y_z be as in the Lemma above. WLOG let $f(x) = 0$, let $f : N_1(x) \to \mathbb{R}$ and choose optimized values of $f(z)$. Let K be the largest value such that G satisfies *CD*(N, K). We have

$$\sum_{z\in N_2(x)}\sum_{\substack{y\sim z\\ y\in N_1(x)}}\left[\frac{1}{4}(f(z)-f(y))^2 - \frac{1}{2}(f(z)-f(y))(f(y)-f(x))\right]$$
$$+\sum_{\substack{y,y'\in N_1(x)\\ y\sim y'}}(f(y)-f(y'))^2$$
$$\geq \frac{1}{N}(\Delta f(x))^2 + \left(\frac{2K+\deg(x)-3}{2}\right)\Gamma(f)(x) - \frac{1}{2}\Delta(f(x))^2.$$

Letting K' be the largest value such that G' satisfies *CD*(N, K), we have

$$\sum_{z\in N_2(x)}\sum_{\substack{y\sim z\\ y\in N_1(x)}}\left[\frac{1}{4}(f(z)-f(y))^2 - \frac{1}{2}(f(z)-f(y))(f(y)-f(x))\right]$$

$$
\begin{aligned}
&- \sum_{y \in Y_z} \left(\frac{1}{4}(f(z) - f(y))^2 - \frac{1}{2}(f(z) - f(y))f(y) \right) \\
&+ \sum_{y \sim y' \in Y_z} (f(y) - f(y'))^2 + \sum_{\substack{y,y' \in N_1(x) \\ y \sim y'}} (f(y) - f(y'))^2 \\
&\geq \frac{1}{N}(\Delta f(x))^2 + \left(\frac{2K' + \deg(x) - 3}{2} \right) \Gamma(f)(x) - \frac{1}{2}\Delta(f(x))^2.
\end{aligned}
$$

By Lemma 3

$$
\begin{aligned}
\sum_{z \in N_2(x)} \sum_{\substack{y \sim z \\ y \in N_1(x)}} &\left[\frac{1}{4}(f(z) - f(y))^2 - \frac{1}{2}(f(z) - f(y))(f(y) - f(x)) \right] \\
&- \sum_{y \in Y_z} \left(\frac{1}{4}(f(z) - f(y))^2 - \frac{1}{2}(f(z) - f(y))f(y) \right) \\
&+ \sum_{y \sim y' \in Y_z} (f(y) - f(y'))^2 + \sum_{\substack{y,y' \in N_1(x) \\ y \sim y'}} (f(y) - f(y'))^2 \\
&\geq \frac{1}{N}(\Delta f(x))^2 + \left(\frac{2K + \deg(x) - 3}{2} \right) \Gamma(f)(x) - \frac{1}{2}\Delta(f(x))^2.
\end{aligned}
$$

Since our choice of f was arbitrary, and K is the largest value for which the above inequality holds for arbitrary f, $K \leq K'$. Note also that if G satisfies *CD*(N, K) it also satisfies larger values of N and smaller values of K. With this we see that whenever G satisfies *CD*(N, K) at x, so does G'. □

2.6 Future Work

Moving forward, we hope that these discrepancy-type bounds may be used to certify other global graph properties, Hamiltonicity for instance. Local Ore-type conditions have had great success in this regard. We hope that through this local combinatorial lens we can uncover structural implications of curvature which may then be used to guarantee the existence of long cycles and paths.

Chapter 3: Strongly Regular

3.1 Introduction

A graph G is said to be strongly regular with parameters (n, d, λ, μ) if G is d-regular of order n and, so that every two adjacent vertices have exactly λ common neighbors and every two nonadjacent vertices have exactly μ common neighbors.

One can see from the definition that structural properties of strongly regular graphs quickly present themselves. It is easy to see, for instance, that connected strongly regular graphs have diameter at most two and that the complement of a strongly regular graph is strongly regular with parameter sets $(n, n-1-d, n-2-(2d-\mu), n-(2d-\lambda))$. Once one starts studying them, more non-trivial relations between the parameter sets emerge. For instance, double counting the edges between the first and second neighborhoods of a vertex reveals that $(n-d-1)\mu = d(d-\lambda-1)$.

More interesting, is that there is an immediate connection between strong regularity and the spectrum of the graph. Supposing M is the adjacency matrix of G, we have $M^2 = dI + \lambda M + \mu(J - I - M)$. To see this, we may view M^2 as a count on the number of paths of length 2. Through dI we see the d paths each vertex has to a neighbor then back to itself. In the second term we see the λ paths a vertex has to its neighbor, and in the third, the μ paths between a vertex and its non-neighbors. Since M is real symmetric, it is diagonalizable and aside from the d eigenvalue, eigenvalues γ of M satisfy $\gamma^2 = d + \lambda\gamma - \mu - \mu\gamma$. Therefore γ can take at most two unique values aside from d. Conversely, it turns out that regular graphs with 3 distinct eigenvalues are strongly regular. It is good here to note that since the sum of eigenvalues with multiplicity of M is $tr(M) = 0$, the smallest eigenvalue is negative. Because the spectrum of SRGs are so well understood, it is often given as a supplementary or alternative parameter set, in which case we say the SRG has spectrum d^1 r^f, and $(-m)^g$, where 1, f, and g represent the multiplicities of the eigenvalues respectively. Understanding these multiplicities has also helped further restrict parameter sets. For example, in [51] Neumaier shows that all SRGs in which $-m$ is not integral have the form $n = 4\mu + 1$, $d = 2\mu$, and $\lambda = \mu - 1$. These are the so called *conference graphs*.

Because the spectrum is so well understood, algebraic techniques work particularly well. Due to the close ties between graph spectra and curvature conditions, it is natural then to ask what the curvatures of these graphs can be. If G is an SRG of girth 5, then G is a Moore graph in which case balls of radius 2 are trees. As seen in Chapter 2, the curvature of these graphs is well understood. When the girth is 4 (or equivalently $\lambda = 0$ and $\mu > 1$), Cushing et al [15] showed that such graphs satisfy *CD*$(\infty, 2)$. It remains to be seen however, if there is a general lower bound on the curvature of SRGs with girth 3. We do not even know if they satisfy *CD*$(\infty, 0)$. Through calculating curvatures of many known SRGs, Cushing conjectured the following

Conjecture 17 (Cushing SRG Conjecture)**.** All SRGs with girth 3 satisfy *CD*$(\infty, 2)$.

However, proving that all *known* SRGs satisfy *CD*$(\infty, 2)$ and all *possible* SRGs satisfy *CD*$(\infty, 2)$ are two different problems.

Determining for which parameter sets strongly regular graphs exist is a major open question. To this end, a number of less obvious conditions have been uncovered which serve to further restrict feasible parameter sets. While some sets may have thousands of graphs, in many cases we see the parameters alone are enough to certify a number of graph properties. In this chapter we use the the parameters to understand these SRGs well enough to give nontrivial curvature bounds. This makes it important for us to understand what the feasible parameters of SRGs are. In this chapter we use established bounds, and some that we uncover, to play parameters off each other.

In one particularly impressive result regarding the classification of feasible parameter sets of SRGs, Brouwer and Maldeghen combined several known powerful relations between the parameters and their structural implications to establish the following theorem, which they credit largely to Neumaier.

Theorem 18 (Claw Bound)**.** (Theorem 8.6.3 [10]) Let G be a primitive strongly regular graph with integral eigenvalues k (with multiplicity 1) $r = n - m$ (with multiplicity f) and $s = -m$ (with multiplicity g). Let $f(m, \mu) = \frac{1}{2}m(m-1)(\mu+1) + m - 1$.
Then
(i) (Bruck [11]) If $\mu = m(m-1)$ and $n > f(m, \mu)$ then Γ is the collinearity graph of a partial geometry $pg(K, R, T)$ with $T = R - 1$, that is, is a Latin square graph $LS_m(n)$.
(ii) (Bose [9]) If $\mu = m^2$ and $n > f(m, \mu)$ then Γ is the collinearity graph of a partial geometry $pg(K, R, T)$ with $T = R$, that is, the block graph of a $2-(mn+m-n, m, 1)$ design.
(iii) ('Claw bound', Neumaier [51]) If $\mu \neq m(m-1)$ and $\mu \neq m^2$ then $n \leq f(m, \mu)$.
In other words: If $r + 1 > \frac{1}{2}s(s+1)(\mu+1)$ then $\mu = s(s+1)$ or $\mu = s^2$.

Using this bound, we show that the truth of Cushing's conjecture may rest on the existence of a hypothetical family of SRGs, yet to be found. Specifically, we find the following.

Theorem 19 (Main Result)**.** All but a finite number of strongly regular graphs with smallest eigenvalue $-m$ satisfy *CD*$(\infty, 2)$.
Furthermore, if there exists an SRG with $\lambda > 0$ and $\mu = 1$, it does not satisfy *CD*$(\infty, 2)$.

This chapter is outlined as follows. We show that all other collinearity graphs of partial geometries of type i and ii in Theorem 18 satisfy *CD*$(\infty, 2)$. We do this first in Section 3.2 by considering those SRGs which satisfy the conditions or our main lemma, Lemma 20.

Lemma 20 (Main Lemma)**.** Strongly Regular Graphs with $\mu \geq \text{Max}\{2, \frac{1}{3}\lambda\}$ satisfy *CD*$(\infty, 2)$.

We then address the remaining SRGs in Section 3.3 by considering the relationships between parameters of the underlying partial geometries and finish with Section 3.4 where we consider future work in this area.

3.2 Proof of Main Lemma

We begin this section by considering the curvature of SRGs with $\mu = 1$. We then rework the CD inequality and consider SRGs with $\lambda = 1$. We then use an averaging argument to compute a curvature lower bounds for SRGs with $\lambda \geq 2$ and $\mu \geq 2\lambda$. We combine these results to show all SRGs with $\mu \geq \frac{1}{3}\lambda$ and $\mu \geq 2$ satisfy *CD*(∞, K).

For the remainder of this chapter we consider an SRG with parameters (n, d, λ, μ) and fix our attention to some vertex x. We refer to the first neighborhood of x as N_1 or N and the second neighborhood of x as N_2. The *local graph* is the subgraph of the SRG induced by the vertices of N. We also fix an arbitrary ordering of the vertices in N which we label, $y_1, \cdots, y_d$. We find it convenient to sometimes think of f as a function and other times, to think of f as a vector, $\vec{f}$ in which case $\vec{f_i} = f(y_i)$. In context, which we are using should be clear.

Theorem 21. Strongly regular graphs with $\lambda > 0$ and $\mu = 1$ do not satisfy *CD*$(\infty, 2)$.

Proof. Let G be an SRG with $\mu = 1$ and $\lambda > 0$. Fix some $x \in G$ and consider the local graph. Since $\mu = 1$ and x is a mutual neighbor of all vertices in the local graph, there are no paths of length two in the local graph, and the local graph is a disjoint union of cliques. Notice that if $\lambda + 1 = d$, then $\mu = 0$. It then becomes clear that $\mathring{B}_2(x)$ is disconnected, and by Theorem 13, *CD*$(\infty, 2)$ is not satisfied. □

It is interesting to note that it is not known if SRGs with $\lambda > 0$ and $\mu = 1$ exist. The smallest open parameter set is $(400, 21, 2, 1)$. If one were known to exist, then the Conjecture 17, that all SRGs with girth 3 satisfy $CD(\infty, 2)$, would be false.

Lemma 22. Strongly Regular Graphs with parameter set (n, d, λ, μ) satisfy $CD(\infty, K)$ at a vertex x if and only if for all functions, $f : N \to \mathbb{R}$,

$$\sum_{i,j} \left[\frac{|z \in N_2 : z \sim y_i, y_j|}{\mu} + \mathbb{1}_{y_i \sim y_j} + \frac{\lambda}{4d} - \frac{1}{2} \right] (f(y_i) - f(y_j))^2 + \frac{\lambda}{4d}(\Delta f)^2 + \frac{2K-4}{2}\Gamma f \geq 0 \tag{3.1}$$

Proof. There are several ways in which we are able to simplify the combinatorial CD inequality when dealing with the SRG case. We know that $f(z)$ is minimized at the value $\frac{2}{\mu} \sum_{x \sim y \sim z} f(y)$ which will make these values much easier to compare to one another. Recall also that $(\Delta f)^2 = (\sum f(y))^2$ and $\Gamma f = \frac{1}{2} \sum f(y)^2$. We use the equality

$$(\sum_{i=1}^{n} a_i)^2 = n \sum_{i=1}^{n} a_i - \sum_{i,j} (a_i - a_j)^2$$

in two ways. In one, we find

$$(\Delta f)^2 = 2d\Gamma f - \sum_{i,j} (f(y_i) - f(y_j))^2. \tag{3.2}$$

This allows us to relate some of our most common terms to one another. Also, it allows us to simplify the previously stubborn $\sum_{z \in N_2} f(z)^2$ as

$$\sum_{z \in N_2} f(z)^2 = \sum_{z \in N_2} \left(\frac{2 \sum_{x \sim y \sim z} f(y)}{\mu} \right)^2$$

$$=\sum_{z\in N_2}\frac{4}{\mu^2}\left[\mu\sum_{x\sim y\sim z}f(y_i)^2-\sum_{x\sim\{y,w\}\sim z}(f(y)-f(w))^2\right]$$

$$=\frac{8(d-\lambda-1)}{\mu}\Gamma f-\frac{4}{\mu^2}\sum_{i,j}|z\in N_2:z\sim y_i,y_j|(f(y_i)-f(y_j))^2. \tag{3.3}$$

Typically we expect $\sum_{z\in N_2} f(z)^2$ to depend heavily on the structure of edges between N_1 and N_2. Interestingly enough, the only term in (3.3) which isn't concerned with N_1, $|z\in N_2 : z\sim y_i,y_j|$, is completely determined by the structure of the local graph. A consequence of this is that curvature of an SRG at some point x is completely determined by the structure of the local graph at x. We note here that Cushing et al. made this observation in [17] (Theorem 11.1).

We begin reworking the CD inequality by considering the first term of (2.1) in Proposition 3

$$\sum_{z\in N_2}\sum_{y_i\sim z}\left[\frac{1}{4}(f(z)-f(y_i))^2-\frac{1}{2}(f(z)-f(y_i))f(y_i)\right]$$

$$=\sum_{z\in N_2}\sum_{y_i\sim z}\left[\frac{1}{4}(f(z))^2-f(y_i)f(z)+\frac{3}{4}f(y_i)^2\right]$$

$$=\sum_{z\in N_2}\sum_{y_i\sim z}\left[\frac{1}{4}(f(z))^2-f(y_i)f(z)\right]+\frac{3(d-\lambda-1)}{4}\sum_{y_i\in N}[f(y_i)^2]$$

$$=-\frac{\mu}{4}\sum_{z\in N_2}[(f(z))^2]+\frac{3(d-\lambda-1)}{4}\sum_{y_i\in N}[f(y_i)^2]$$

$$=-\frac{\mu}{4}\sum_{z\in N_2}[(f(z))^2]+\frac{3(d-\lambda-1)}{2}\Gamma(f). \tag{3.4}$$

With the substitution for $\sum f(z)^2$ in (3.3), this gives us

$$(3.4)=$$

$$-\frac{\mu}{4}\left[\frac{8(d-\lambda-1)}{\mu}\Gamma f-\frac{4}{\mu^2}\sum_{i,j}|z\in N_2:z\sim y_i,y_j|(f(y_i)-f(y_j))^2\right]$$

$$+\frac{3(d-\lambda-1)}{2}\Gamma(f)$$

$$=\sum_{i,j}\frac{|z\in N_2 : z\sim y_i,y_j|}{\mu}(f(y_i)-f(y_j))^2-\frac{(d-\lambda-1)}{2}\Gamma(f). \tag{3.5}$$

Using (3.2) we find the equality

$$\frac{-2d+\lambda}{2}\Gamma f=\frac{-2d+\lambda}{4d}\left((\Delta f)^2+\sum_{i,j}(f(y_i)-f(y_j))^2\right). \tag{3.6}$$

By adding the right side of the equality (3.6) and subtracting the left, we rewrite (3.5) as

$$\sum_{i,j}\left[\frac{|z\in N_2 : z\sim y_i,y_j|}{\mu}+\frac{-2d+\lambda}{4d}\right](f(y_i)-f(y_j))^2$$
$$+\frac{-2d+\lambda}{4d}(\Delta f)^2+\frac{(d+1)}{2}\Gamma(f). \tag{3.7}$$

We also rewrite the second term of (2.1)

$$\sum_{\substack{y,y'\in N_1\\ y\sim y'}}(f(y)-f(y'))^2=\sum_{i,j}\mathbb{1}_{y_i\sim y_j}(f(y_i)-f(y_j))^2. \tag{3.8}$$

Through the substitution of (3.7) and (3.8) we find (2.1) is equivalent to

$$(3.7)+\sum_{i,j}\mathbb{1}_{y_i\sim y_j}(f(y_i)-f(y_j))^2\geq\left(\frac{2K+d-3}{2}\right)\Gamma(f)(x)-\frac{1}{2}(\Delta f)^2. \tag{3.9}$$

Through simplification of (3.9) we find (2.1) is equivalent to

$$\sum_{i,j}\left[\frac{|z\in N_2 : z\sim y_i,y_j|}{\mu}+\mathbb{1}_{y_i\sim y_j}+\frac{\lambda}{4d}-\frac{1}{2}\right](f(y_i)-f(y_j))^2$$
$$+\frac{\lambda}{4d}(\Delta f)^2-\frac{2K-4}{2}\Gamma f\geq 0 \tag{3.10}$$

and through Proposition 3, an SRG satisfies *CD*(∞, K) if and only if (3.10) is satisfied. □

Corollary 23. Strongly Regular Graphs with parameter set (n, d, λ, μ) satisfy *CD*$(\infty, 2)$ at a vertex x if and only if for all functions, $f : N_1 \to \mathbb{R}$

$$\sum_{i,j\in N_1}\left[\alpha_{i,j}(f(y_i)-f(y_j))^2\right]+\frac{\lambda}{4d}(\Delta f)^2\geq 0 \tag{3.11}$$

where $\alpha_{i,j}=\dfrac{|z\in N_2 : y_i,y_j\sim z|}{\mu}+\mathbb{1}_{\{y_i\sim y_j\}}+\dfrac{\lambda}{4d}-\dfrac{1}{2}$. Note that if for all $i\neq j$, $\alpha_{i,j}\geq 0$, then (3.11) is satisfied.

Proof. This follows immediately from Lemma 22. □

Lemma 24. If G is an SRG with $\lambda = 1$ or $\lambda = 0$, and $\mu \geq 2$ then G satisfies *CD*$(\infty, 2)$.

Proof. First let G be an SRG with $\lambda = 1$ and $\mu \geq 2$. Fix some $x \in G$. Notice that the edges within the local graph form a perfect matching. If $y_i \not\sim y_j \in N_1$ then $|z \in N_2 : z \sim y_i, y_j| = \mu - 1$. In this case we have $\alpha_{i,j} \geq \left[\dfrac{\mu-1}{\mu}+\dfrac{\lambda}{4d}-\dfrac{1}{2}\right] \geq 0$. By Lemma 22 we have that G satisfies *CD*$(\infty, 2)$ at x and since our choice of x was arbitrary, G satisfies *CD*$(\infty, 2)$. Similarly, we let G be an SRG with $\lambda = 0$ and $\mu \geq 2$. Fix some $x \in G$. Notice that the local graph has no edges. Then for all $y_i, y_j \in N_1$, $y_i \not\sim y_j$ and $|z \in N_2 : z \sim y_i, y_j| = \mu - 1$. In this case we have $\alpha_{i,j} \geq \left[\dfrac{\mu-1}{\mu}+\dfrac{\lambda}{4d}-\dfrac{1}{2}\right] \geq 0$. By Lemma 22 we have that G satisfies *CD*$(\infty, 2)$ at x and since our choice of x was arbitrary, G satisfies *CD*$(\infty, 2)$. □

We see that Lemma 24 gives an alternate proof of the result in [15] that all SRGs with girth 4 and $\mu \geq 2$ satisfy *CD*$(\infty, 2)$.

We have now covered the cases where λ or μ is 1. Next we redistribute the $\alpha_{i,j}$ coefficients from adjacent vertices to paths of length 2. We then show that these new coefficients are non-negative provided $\lambda \geq 2$ and $\mu \geq 2\lambda$.

Lemma 25. If G is an SRG with $\lambda \geq 2$ and $\mu \geq 2\lambda$, then G satisfies *CD*$(2, \infty)$.

Proof. Let G be an SRG with $\lambda \geq 2$ and $\mu \geq 2\lambda$. By Corollary 23 it suffices to show

$$\sum_{i,j} \left[\frac{|z \in N_2 : y_i, y_j \sim z|}{\mu} + \mathbb{1}_{y_i \sim y_j} + \frac{\lambda}{4d} - \frac{1}{2} \right] (f(y_i) - f(y_j))^2 \geq 0. \tag{3.12}$$

We start with the observation that every edge in the local graph is a part of $2(\lambda - 1)$ paths of length 2 in the local graph. This observation, along with the triangle inequality gives us the rewritten term

$$\begin{aligned}
\sum_{i,j} \mathbb{1}_{y_i \sim y_j} (y_i - y_j)^2 &= \sum_{i,j} \left(\sum_{y \in N \sim y_i, y_j} \left(\frac{(y - y_i)^2 + (y - y_j)^2}{2(\lambda - 1)} \right) \right) \\
&\geq \sum_{i,j} \left(\sum_{y \in N \sim y_i, y_j} \left(\frac{(\frac{y_i + y_j}{2} - y_i)^2 + (\frac{y_i + y_j}{2} - y_j)^2}{2(\lambda - 1)} \right) \right) \\
&= \sum_{i,j} \left(\sum_{y \in N \sim y_i, y_j} \left(\frac{2(\frac{y_i - y_j}{2})^2}{2(\lambda - 1)} \right) \right) \\
&= \sum_{i,j} \left(\frac{|y \in N_1 : y_i, y_j \sim y|}{4(\lambda - 1)} \right) (y_i - y_j)^2. \qquad (3.13)
\end{aligned}$$

Thus by averaging (3.13) with its original form, we have

$$\sum_{i,j} \mathbb{1}_{y_i \sim y_j} (y_i - y_j)^2 \geq \sum_{i,j} \frac{1}{2} \cdot \mathbb{1}_{y_i \sim y_j} + \frac{|y \in N_1 : y_i, y_j \sim z|}{8(\lambda - 1)} (y_i - y_j)^2. \tag{3.14}$$

Though the substitution of (3.14) into (3.12) we find

$$\sum_{i,j}\left[\frac{|z:y_i,y_j\sim z|}{\mu}+\frac{1}{2}\cdot\mathbb{1}_{y_i\sim y_j}+\frac{|y\in N_1(x):y_i,y_j\sim z|}{8(\lambda-1)}+\frac{\lambda}{4d}-\frac{1}{2}\right](f(y_i)-f(y_j))^2$$
$$\geq 0. \tag{3.15}$$

Therefore the satisfaction of (3.15) implies the satisfaction of (3.12).

We now show that in the parameter regime we consider, all coefficients of $(f(y_i)-f(y_j))^2$ in (3.15) are non-negative. Note that if $y_i\sim y_j$, then clearly this coefficient is non-negative, as the contribution from the indicator is enough to make the term positive.

Supposing $y_i\not\sim y_j$, $|z\in N_2:y_i,y_j\sim z|+|y\in N_1:y_i,y_j\sim y|=\mu-1$. When $\mu>8(\lambda-1)$ the sum $\frac{|z:y_i,y_j\sim z|}{\mu}+\frac{|y:y_i,y_j\sim y|}{8(\lambda-1)}$ is minimized when $|z:y_i,y_j\sim z|=\mu-1$ and since $\frac{\mu-1}{\mu}+\frac{|y:y_i,y_j\sim y|}{8(\lambda-1)}\geq\frac{1}{2}$ and the coefficients are non-negative. When $\mu\leq 8(\lambda-1)$, the sum $\frac{|z:y_i,y_j\sim z|}{\mu}+\frac{|y:y_i,y_j\sim y|}{8(\lambda-1)}$ is minimized when $|z:y_i,y_j\sim z|=\mu-(\lambda+1)$ and $|y:y_i,y_j\sim y|=\lambda$, in which case $\frac{|z:y_i,y_j\sim z|}{\mu}+\frac{|y:y_i,y_j\sim y|}{8(\lambda-1)}\geq\frac{\mu-(\lambda+1)}{\mu}+\frac{\lambda}{8(\lambda-1)}$. Now observe that for $\mu\geq 2\lambda$ and $\lambda\geq 2$,

$$\frac{\mu-(\lambda+1)}{\mu}+\frac{\lambda}{8(\lambda-1)}\geq\frac{\lambda-1}{2\lambda}+\frac{\lambda}{8(\lambda-1)}=\frac{1}{2}+\frac{1}{8}-\frac{1}{2\lambda}+\frac{1}{8(\lambda-1)}\geq\frac{1}{2}.$$

Thus, with these parameters, all of the coefficients of the $(f(y_i)-f(y_j))^2$ terms in (3.15) are non-negative, and hence the sum is non-negative and hence G satisfies *CD*$(\infty,2)$.

□

Now that we have taken care of the cases where λ or μ is 1, and when $\mu\geq 2\lambda$. We will now use the spectrum of the local graph to show *CD*$(\infty,2)$ is satisfied when $\mu\geq\max\{\frac{1}{3}\lambda,2\}$. We begin by defining some relevant graph matrices and their row sums.

For a strongly regular graph with parameter set (n, d, λ, μ) and a fixed x, we define $\alpha_{i,j} = \dfrac{|z \in N_2 : y_i, y_j \sim z|}{\mu} + \mathbb{1}_{\{y_i \sim y_j\}} + \dfrac{\lambda}{4d} - \dfrac{1}{2}$ if $i \neq j$ and $\alpha_{i,i} = \dfrac{\lambda}{4} - \dfrac{1}{2}$. Notice that $\sum_i \alpha_{i,j}$ is independent of j, we will call this sum $A = \dfrac{(d-\lambda-1)(\mu-1)}{\mu} + \lambda + \dfrac{\lambda}{4} - \dfrac{d}{2}$. We define $\mathcal{A}$ as the square d matrix with $\mathcal{A}_{i,j} = \alpha_{i,j}$. We remind the reader that a restricted eigenvalue, is one which is not associated $\vec{1}$. We would also like to remind the reader that because $\mathcal{A}$ is real valued and symmetric, it has d orthonormal eigenvectors. This result, attributed to Cauchy, was later generalized by Von Neumann and is known as the Spectral Theorem.

Lemma 26. If there exists a function $f : N \to \mathbb{R}$ which violates the following inequality,

$$\sum_{i,j \in N_1(x)} \left[\alpha_{i,j}(f(y_i) - f(y_j))^2\right] - \left[\frac{(2K-4)}{2}\right] \Gamma f \geq 0 \tag{3.16}$$

where $K = 2$, then there exists a restricted eigenvalue of $\mathcal{A}$ larger than A.

Proof. Note that if a function fails (3.16), it can be scaled so that $\sum_i f(y_i)^2 = 1$. For such a function f, the left hand side (specialized at $K = 2$) is

$$\begin{aligned}
&\sum_{i,j \in N_1(x)} \alpha_{i,j}(f(y_i) - f(y_j))^2 \\
&\quad = 2 \sum_{i,j \in N_1(x)} \left[\alpha_{i,j} f(y_i)^2\right] - 2 \sum_{i,j \in N_1(x)} \left[\alpha_{i,j} f(y_i) f(y_j)\right] \\
&\quad = 2 \sum_{i \in N_1(x)} \left[A f(y_i)^2\right] - 2 \sum_{i,j \in N_1(x)} \left[\alpha_{i,j} f(y_i) f(y_j)\right]
\end{aligned} \tag{3.17}$$

where we used that $\mathcal{A}$ has a constant row and column sum of A. Thinking of f as a vector, we have

$$(3.17) = 2A - 2\vec{f}^{\intercal} \mathcal{A} \vec{f} \tag{3.18}$$

and satisfying (3.16) is hence equivalent to minimizing (3.18) and showing that this minimum is non-negative.

$\mathcal{A}$ is real symmetric, and hence admits an orthonormal basis of eigenvectors. It is precisely such a vector that minimizes (3.18). Further observe that, though $\vec{1}$ is an eigenvector of $\mathcal{A}$, the constant function does satisfy (3.16). Thus if (3.16) is violated it must be must be for the eigenfunction corresponding to a maximal restricted eigenvalue of $\mathcal{A}$, γ. For this function,

$$(3.17) = A - \gamma$$

and we see if (3.16) is not satisfied, then $\gamma > A$. □

From Lemma 26 we find the left-hand side of 3.16 is minimized at eigenvectors perpendicular to $\vec{1}$. Since $\frac{\lambda}{4d}(\Delta f)^2$ is also minimized at eigenvectors perpendicular to $\vec{1}$, (3.1) is also minimized at these eigenvectors and by Lemma 22 we have the following proposition.

Proposition 27. Strongly Regular Graphs with parameter set (n, d, λ, μ) satisfy *CD*$(\infty, 2)$ at a vertex x if and only if there exists some maximal eigenvalue γ of $\mathcal{A}$ such that $\gamma > A$.

We now relate the eigenvectors of $\mathcal{A}$ to the eigenvectors of the adjacency matrix of the local graph. Using this same labeling of the local graph used in the construction of $\mathcal{A}$, let $\bar{A}$ be the adjacency matrix of the local graph. Let γ be the largest restricted eigenvalue of $\mathcal{A}$. Notice that because the local graph is λ-regular, it has eigenvector $\vec{1}$ with associated eigenvalue λ. In the next lemma, we show the relationship between γ and an eigenvector of $\bar{A}$.

Lemma 28. Suppose γ is the largest restricted eigenvalue of $\mathcal{A}$. Suppose $\lambda = \gamma_1 \geq \cdots \geq \gamma_k$ are the eigenvalues of $\bar{\mathcal{A}}$, then

$$\gamma = \max_{i>1} \left[-\frac{1}{\mu}\gamma_i^2 + \frac{(\lambda-\mu)}{\mu}\gamma_i + \frac{\lambda-\mu+1}{\mu} \right]. \tag{3.19}$$

Proof. We remind the reader that

$$\mathcal{A}_{i,j} = \alpha_{i,j} \tag{3.20}$$

$$= \frac{|z \in N_2 : y_i, y_j \sim z|}{\mu} + \mathbb{1}_{\{y_i \sim y_j\}} + \frac{\lambda}{4d} - \frac{1}{2}. \tag{3.21}$$

Since vertices y_i, y_j have $\bar{\mathcal{A}}^2_{i,j}$ mutual neighbors in the local graph and one mutual neighbor x, all other mutual neighbors are in N_2. This allows us to rewrite $\mathcal{A}$,

$$\mathcal{A} = \frac{\mu-1}{\mu}(J - \bar{\mathcal{A}} - I) + \frac{\lambda-1}{\mu}\bar{\mathcal{A}} - \frac{1}{\mu}(\bar{\mathcal{A}}^2 - \lambda I) + \bar{\mathcal{A}} + \left(\frac{\lambda}{4d} - \frac{1}{2}\right) J. \tag{3.22}$$

This expression of $\mathcal{A}$ in terms of $\bar{\mathcal{A}}$ and J ensures that any restricted eigenvector of $\mathcal{A}$ is a restricted eigenvector of $\bar{\mathcal{A}}$ and vice versa.

Notice that $\vec{1}$ is an eigenvector of $\bar{\mathcal{A}}$. Let $\vec{\phi}_i \perp \vec{1}$ be another eigenvector of $\bar{\mathcal{A}}$ with associated eigenvalue γ_i. Then $J\vec{y} = \vec{0}$. From (3.22) we have

$$\begin{aligned}
\mathcal{A}\vec{\phi}_i &= \left[\bar{\mathcal{A}} + \frac{\mu-1}{\mu}(J - \bar{\mathcal{A}} - I) + \frac{\lambda-1}{\mu}\bar{\mathcal{A}} - \frac{1}{\mu}(\bar{\mathcal{A}}^2 - \lambda I) + \left(\frac{\lambda}{4d} - \frac{1}{2}\right) J \right] \vec{\phi}_i \\
&= \left[\gamma_i + \frac{\mu-1}{\mu}(0 - \gamma_i - 1) + \frac{\lambda-1}{\mu}\gamma_i - \frac{1}{\mu}\left(\gamma_i^2 - \lambda\right) \right] \vec{\phi}_i \\
&= \left[-\frac{1}{\mu}\gamma_i^2 + \frac{(\lambda-1)-(\mu-1)}{\mu}\gamma_i + \frac{-(\mu-1)+\lambda}{\mu} \right] \vec{\phi}_i \\
&= \left[-\frac{1}{\mu}\gamma_i^2 + \frac{(\lambda-\mu)}{\mu}\gamma_i + \frac{\lambda-\mu+1}{\mu} \right] \vec{\phi}_i.
\end{aligned}$$

where $\left[-\frac{1}{\mu}\gamma_i^2 + \frac{(\lambda-\mu)}{\mu}\gamma_i + \frac{\lambda-\mu+1}{\mu}\right]$ is a restricted eigenvalue of $\mathcal{A}$ coming from a restricted eigenvalue γ_i of $\bar{\mathcal{A}}$. □

Maximizing the quadratic in (3.19) allows us, in turn, a bound on γ solely in terms of the parameters of the SRG. This allows us to drop the dependence of the spectrum of the local graph in the following corollary.

Corollary 29.

$$A - \gamma \geq d(\frac{\mu-2}{2\mu}) + \frac{\lambda}{4} - \frac{(\lambda-\mu)^2}{4\mu}. \tag{3.23}$$

Proof. By treating γ_i as a formal variable in (3.19), we see γ is maximized when some $\gamma_i = \frac{\lambda-\mu}{2}$. Since γ arises from such a γ_i this implies that

$$\gamma \leq \left[\frac{(\lambda-\mu)^2}{4\mu} + \frac{\lambda-\mu+1}{\mu}\right]$$

as desired. Recall $A = \frac{(d-\lambda-1)(\mu-1)}{\mu} + \lambda + \frac{\lambda}{4} - \frac{d}{2}$. Therefore

$$\begin{aligned} A - \gamma \geq & \frac{(d-\lambda-1)(\mu-1)}{\mu} + \lambda + \frac{\lambda}{4} - \frac{d}{2} - \left(\frac{(\lambda-\mu)^2}{4\mu} + \frac{\lambda-\mu+1}{\mu}\right) \\ =& d(\frac{\mu-1}{\mu} - \frac{1}{2}) + \frac{\lambda-\mu+1}{\mu} + \frac{\lambda}{4} - \left(\frac{(\lambda-\mu)^2}{4\mu} + \frac{\lambda-\mu+1}{\mu}\right) \\ =& d(\frac{\mu-1}{\mu} - \frac{1}{2}) + \frac{\lambda}{4} - \frac{(\lambda-\mu)^2}{4\mu} \\ =& d(\frac{\mu-2}{2\mu}) + \frac{\lambda}{4} - \frac{(\lambda-\mu)^2}{4\mu}. \end{aligned}$$

□

Lemma 30. Fix a $\mu \geq 2$ and let

$$C_\mu = \min\left\{\frac{\lambda}{\mu} : \text{ An SRG with parameters } \lambda \text{ and } \mu \text{ does not satisfy } CD(\infty, 2)\right\},$$

taking $C_\mu = \infty$ if no such graphs exist. Then $C_2 \geq 3$ and $C_\mu \geq 4$ for $\mu \geq 3$.

Proof. Per Lemma 26 and Corollary 29 the inequality

$$d(\frac{\mu - 2}{2\mu}) + \frac{\lambda}{4} - \frac{(\lambda - \mu)^2}{4\mu} \geq 0$$

implies $CD(\infty, 2)$. As K_{d+1} satisfies $CD(\infty, 2)$ we know that in any SRG not satisfying $CD(\infty, 2)$ the local graph has two nonadjacent vertices in which case $d \geq 2\lambda - \mu$.

Hence it must be the case that if a graph fails this inequality, with $\lambda = x \cdot \mu$ and using our lower bound on d we must have

$$(x - 1/2)(\mu - 2) + \frac{x\mu}{4} - \frac{\mu(x-1)^2}{4} < 0.$$

The roots of this polynomial are

$$\frac{7\mu - 8 \pm \sqrt{37\mu^2 - 96\mu + 64}}{2\mu}.$$

Note that

$$\frac{7\mu - 8 - \sqrt{37\mu^2 - 96\mu + 64}}{2\mu} < \frac{1}{2}$$

for $\mu \geq 2$ (and actually, the limit as $\mu \to \infty$ increases to $\dfrac{(7-\sqrt{37})}{2} \approx 0.458...$). Recall Lemma 25 implies that $C_\mu \geq \dfrac{1}{2}$ therefore all SRGs with x less than this smaller root satisfy $CD(\infty, 2)$.

Thus the only way that G fails to satisfy $CD(\infty, 2)$ is if

$$x > \frac{7\mu - 8 + \sqrt{37\mu^2 - 96\mu + 64}}{2\mu}.$$

For $\mu \geq 4$, this is already at least 4, and for $\mu = 3$, note that this is larger than $11/3$ and hence $C_3 \geq \frac{12}{3} = 4$. For $\mu = 2$, this root is $\frac{3+\sqrt{5}}{2} > \frac{5}{2}$, and hence $C_2 \geq \frac{6}{2} = 3$

□

We define the conference graph to be an SRG with parameters, $n, d = \frac{(n-1)}{2}, \lambda = \frac{(n-5)}{4}$, and $\mu = \frac{(n-1)}{4}$. As mentioned earlier, these graphs are known to be the only class of SRG without integral eigenvalues. In the next corollary, we show that conference graphs with girth 3 satisfy $CD(\infty, 2)$.

Corollary 31. Conference graphs with girth 3 satisfy $CD(\infty, 2)$.

Proof. Notice for a conference graph, with girth 3, $n \geq 9$ so $\mu \geq 2$ and $\frac{\lambda}{\mu} = \frac{n-5}{n-1} \leq 1$. By Lemma 30 we see $CD(\infty, 2)$ is satisfied. □

Unfortunately, there are SRGs with parameter sets which are not covered by Lemma 30. For example, The line graphs $L(K_{n,n})$, $n \geq 2$, with parameters $(n^2, 2(n-1), n-2, 2)$. We consider many of these graphs in the next section.

It is also good to note that the parameters do not determine the local graph, or its spectrum. For example, the Shrikhande graph and 4×4 Rook's graph both have parameters $(16, 6, 2, 2)$, but the local graph of the Shrikhande graph and 4×4 Rook's graph (see Figure 3.1) are not isomorphic – this is reflected in the curvatures of these two graphs. The Shrikhande graph has curvature 2 (that is, it satisfies $CD(\infty, 2)$ but not $CD(\infty, K)$ for any $K > 2$), while the Rook's graph has curvature 3 [17].

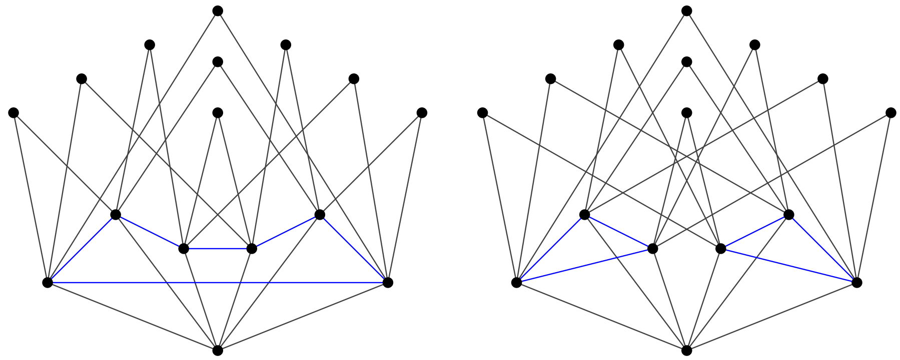

Figure 3.1: On the left is the first and second neighborhood of the Shikrande Graph and on the right, that of the Rook's graph. The local graph of x is highlighted in blue.

3.3 Curvature of Collinearity Graphs of Partial Geometries

In this section, we show that the infinite families of graphs with smallest eigenvalue $-m$ which are not covered by Lemma 20 satisfy $CD(\infty, 2)$. By the claw bound, these are all collinearity graphs of partial geometries. We begin by defining SRGs of this form.

Definition 2. A partial geometry $pg(K, R, T)$ is a partial linear space $(X, \mathcal{L})$ such that each line has K points, each point is on R lines, and given a point x outside a line L, there are precisely T lines on x meeting L.

It is known that the collinearity graph of a partial geometry is strongly regular with parameters $n = K + K(K-1)(R-1)/T$, $d = R(K-1)$, $\lambda = (R-1)(T-1) + K - 2$, and $\mu = RT$.
We define the set of exceptional graphs as described in the claw bound to be $\mathcal{L}$. We note here that when $R = 1$, G is a complete graph and by Example 2 G satisfies $CD(\infty, 2)$.

Theorem 32. Suppose G is an SRG with $\mu = 2$ and $G \notin \mathcal{L}$, then G satisfies $CD(\infty, 2)$.

Proof. Suppose G is an SRG with $\mu = 2$ and $G \notin \mathcal{L}$. If G is a conference graph, then by Lemma 31 we are done. Otherwise, G has integral smallest eigenvalue $-m$. By the Claw Bound, $\mu = m(m-1)$ and G is the collinearity graph of a partial geometry with $T = 2$ and $R = 1$, in which case G is a complete graph, which we know satisfies $CD(\infty, 2)$. $\square$

Theorem 33 (partial geometry theorem)**.** Let G be the collinearity graph of a partial geometry, $pg(K, R, T)$, with strongly regular graph parameters (n, d, λ, μ) such that $\mu \geq 3$. Then G satisfies *CD*$(\infty, 2)$.

Proof. Let G be the collinearity graph of a partial geometry, $pg(K, R, T)$, with strongly regular graph parameters (n, d, λ, μ) such that $\mu \geq 3$. If $\lambda < 4\mu$, by Lemma 30 G satisfies *CD*$(\infty, 2)$. For the remainder of the proof, suppose $\lambda \geq 4\mu$. Notice if $R = 1$ or $T = 0$ then G is a union of disjoint cliques, which satisfies *CD*$(\infty, 2)$. For the remainder of the proof we assume $R \geq 2$ and $T \geq 1$. We let $\bar{\mathcal{A}}$ be the adjacency matrix of local graph of G at some fixed vertex x. We notice $\bar{\mathcal{A}} = \mathcal{A}_1 + \mathcal{A}_2$ where $\mathcal{A}_1$ is the adjacency matrix of R many $K - 1$-cliques and $\mathcal{A}_2$ is $(R-1)(T-1)$ regular. We let $\mathcal{B}_1$ be the R dimensional eigenspace of $\mathcal{A}_1$ associated with the eigenvalue $K - 2$. Notice that by Perron Frobenius we have the largest eigenvalue in absolute value of $\mathcal{A}_2$ is $(R-1)(T-1)$. Let $\gamma_1 \geq \cdots \geq \gamma_k$ be the eigenvalues of $\bar{\mathcal{A}}$. By the Courant-Fischer Theorem we have

$$\begin{aligned}\gamma_R = &\max_{\substack{\mathcal{X} \\ \dim \mathcal{X} = R}} \min_{\substack{x \in \mathcal{X} \\ ||x|| = 1}} x^T \bar{\mathcal{A}} x \\ \geq &\min_{\substack{x \in \mathcal{B}_1 \\ ||x|| = 1}} x^T \bar{\mathcal{A}} x \\ \geq &K - 2 - (R-1)(T-1).\end{aligned}$$

Similarly, we define $\mathcal{B}_2$ to be the $k - R$ dimensional eigenspace of $\mathcal{A}_1$ associated with -1. Again, by Courant-Fischer,

$$\begin{aligned}\gamma_{R+1} = &\min_{\substack{\mathcal{X} \\ \dim \mathcal{X} = k-R}} \max_{\substack{x \in \mathcal{X} \\ ||x|| = 1}} x^T \bar{\mathcal{A}} x \\ \leq &\max_{\substack{x \in \mathcal{B}_2 \\ ||x|| = 1}} x^T \bar{\mathcal{A}} x \\ \leq &-1 + (R-1)(T-1).\end{aligned}$$

From Lemma 28 we have

$$\gamma = \max_i \left[-\frac{1}{\mu}\gamma_i^2 + \frac{(\lambda-\mu)}{\mu}\gamma_i + \frac{\lambda-\mu+1}{\mu} \right].$$

Recalling $A = \dfrac{(d-\lambda-1)(\mu-1)}{\mu} + \lambda + \dfrac{\lambda}{4} - \dfrac{d}{2}$, we may rewrite $A-\gamma$ as

$$\min_i d(\frac{\mu-2}{2\mu}) + \frac{\lambda}{4} + \left[\frac{\gamma_i(\gamma_i-\lambda+\mu)}{\mu} \right]. \tag{3.24}$$

We may think of $A-\gamma$ as a function of γ_i in which case, it is an upward facing parabola. Here we wish to bound the eigenvalues of $\bar{A}$ away from from the potentially negative region – which is centered at $\gamma_i = \dfrac{\lambda-\mu}{2}$.

We start by showing that the first R eigenvalues of $\bar{A}$ are sufficiently larger than $\dfrac{\lambda-\mu}{2}$, in doing so we find suitable lower bounds for the terms in (3.24). Recall that we assume $\lambda \geq 4\mu$, and so the second term of (3.24), $\dfrac{\lambda}{4} > (R-1)(T-1)$. We see that the lower bound on γ_R, $K-2-(R-1)(T-1) = \lambda - 2(R-1)(T-1) > \dfrac{\lambda}{2} > \dfrac{\lambda-\mu}{2}$. Since $K-2-(R-1)(T-1)$ bounds γ_R from below, and $\left[\dfrac{\gamma_i(\gamma_i-\lambda+\mu)}{\mu}\right]$ is increasing when $\gamma_i > \dfrac{\lambda-\mu}{2}$, for all $j \leq R$,

$$\frac{\gamma_i(\gamma_i-\lambda+\mu)}{\mu} \geq \frac{(K-2-(R-1)(T-1))\,(K-2-(R-1)(T-1)-\lambda+\mu)}{\mu}. \tag{3.25}$$

Notice $K-2-(R-1)(T-1)-\lambda+\mu = -2(R-1)(T-1)+RT > -(R-1)(T-1)$, the right hand side of (3.25) may be bounded below by

$$\frac{(K-2-(R-1)(T-1))\,(-(R-1)(T-1))}{RT}. \tag{3.26}$$

Since $R \geq 2$ and $T \geq 1$, $(R-1)(T-1) \leq RT-2$. Using this substitution we find a lower bound for (3.26)

$$\frac{(K-2-(R-1)(T-1)\left(-(RT-2)\right)}{RT}. \tag{3.27}$$

We will use (3.27) as a lower bound for the third term of (3.24). We now wish to bound the first term of (3.24). Through substitutions we see

$$d(\frac{\mu-2}{2\mu}) = \frac{(K-1)(RT-2)}{2T}. \tag{3.28}$$

adding (3.28) to (3.27) yields

$$\frac{(K-1)(RT-2)}{2T} + \frac{(K-2-(R-1)(T-1)\left(-(RT-2)\right)}{RT} > 0. \tag{3.29}$$

(3.29) shows that the sum of the first and third terms of (3.24) is positive, and since the second term, $\frac{\lambda}{4}$ is also positive, for all $j \leq R$ we have (3.24) is positive. Therefore, if (3.24) is minimized at some $j \leq R$, G satisfies $\boldsymbol{CD}(\infty, 2)$.

We use an analogous argument for $j > R$. Since $\lambda \geq 4\mu$ we have $K-2+(R-1)(T-1) \geq 4RT$ and $K-2 > 3RT$.

With some substitution we see $\frac{\lambda-\mu}{2} = \frac{K-2+(R-1)(T-1)-RT}{2} \geq \frac{2RT}{2}$ and we have $-1+(R-1)(T-1) < \frac{\lambda-\mu}{2}$. Because $-1+(R-1)(T-1)$ bounds γ_{R+1} from above and $\left[\frac{\gamma_i(\gamma_i-\lambda+\mu)}{\mu}\right]$ is decreasing when $\gamma_i < \frac{\lambda-\mu}{2}$, for all $j \geq R+1$,

$$\frac{\gamma_i(\gamma_i-\lambda+\mu)}{\mu} \geq \frac{\left(-1+(R-1)(T-1)\right)\left(-1+(R-1)(T-1)-\lambda+\mu\right)}{\mu}.$$

Since $-1 + (R-1)(T-1) - \lambda = -1 + (R-1)(T-1) - [K - 2 - (R-1)(T-1)]$, the third term of (3.24) may be rewritten as

$$\frac{(-1 + (R-1)(T-1))\,(-(K - 1 - 2(R-1)(T-1) - RT)}{RT}. \tag{3.30}$$

Earlier we saw the first term of (3.24) may be rewritten as

$$d(\frac{\mu - 2}{2\mu}) = \frac{(K-1)(RT-2)}{2T}. \tag{3.31}$$

Since $R \geq 2$, $RT - 2 > (R-1)(T-1) - 1$ and $K - 1 > K - 1 - 2(R-1)(T-1) - RT$,

$$\frac{(-1 + (R-1)(T-1))\,(-(K - 1 - 2(R-1)(T-1) - RT)}{RT} + \frac{(K-1)(RT-2)}{2T} > 0. \tag{3.32}$$

Since (3.32) shows the sum of the first and third terms of (3.24) is positive, and $\frac{\lambda}{4}$ clearly is as well, we have for all $j \geq R+1$, (3.24) is positive. Therefore, if (3.24) is minimized at some $j > R$, G satisfies *CD*$(\infty, 2)$. Combining this with our result when $j \leq R$, we have (3.24) > 0 and by Lemma 26, G satisfies *CD*$(\infty, 2)$. □

3.4 Future Work

While we hope to extend these results to all SRG's, this will likely require additional observations regarding the structure of feasible SRGs that we have yet to uncover. We are particularly interested in the case when $\frac{\lambda}{\mu}$ is large. There are also several other families of graphs for which discrete curvature bounds are of interest. For instance, the Ollivier curvature of circulant graphs has recently received attention. It would be worthwhile to see if our combinatorial techniques could give nontrivial lower bounds on the Bakry-Émery curvature of these graphs.

Chapter 4: Drones

4.1 Introduction

In the previous two chapters we explored global implications of local structure in the setting of extremal graph theory. We now turn our attention to a the setting of computational geometry where we consider a sort of facility location problem. While the setting is different, the sentiment is the same. Using only local information, we can glean global insights. In this case, the global insights are a near optimal solution the facility location problem.

The rapid development and use of *Unmanned Aerial Vehicles* (UAVs), commonly called drones, in many activities of our daily life has created a need for the development of new algorithms to optimize their use. A factor that is common to most types of drones is their relatively short flying range due mostly to their their restricted energy capacity [41].

Patrolling a border is one typical application of air surveillance systems where the deployment of drones has become a natural choice [3, 33]. In this context, a team of UAVs can be deployed to monitor an area and send information such as images or videos to the nearest base station.

In this chapter we study the following problem which we call the *MinStation Problem*: Suppose that we want to guard the border of an island $\mathcal{I}$ whose boundary is modeled by a simple polygon $\mathcal{P}$ using a set of drones that can fly a distance d before needing to refuel. These drones can fly over the boundary of $\mathcal{I}$ or over small sectors of the sea surrounding it, but not over the interior of $\mathcal{I}$. The objective is to place a set $S = \{s_0, \ldots, s_{k-1}\}$ of k refueling base stations with *minimum cardinality* k and located on the boundary of $\mathcal{I}$, such that when a drone visits all the refueling stations it travels a closed curve that encloses

$\mathcal{P}$; the flying distance between s_i and s_{i+1} is at most d, with addition taken mod k. See Figure 4.1. We will refer to S as an *optimal solution*. A set $S' = \{s'_0, \ldots, s'_k\}$ with $k+1$ refueling stations will be called a *quasi-optimal* solution.

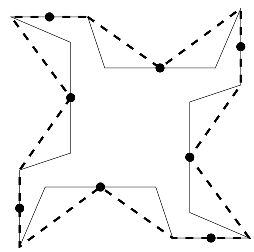

Figure 4.1: Example. A polygon $\mathcal{P}$ and d-hull $\mathcal{C}$ (black dashed line) that uses an optimal set of base stations (solid circles) for the MinStation problem.

Our main contributions are as follows: We give an algorithm, OPTSOL, with complexity $O(n^2 + Ln)$, such that if s_0 is a fixed point on the convex hull of $\mathcal{P}$, *CH*(P), finds an optimal solution S to the MinStation Problem under the restriction that $s_0 \in S$; here, L is the length of $\mathcal{P}$. This yields either an optimal or quasi-optimal solution to the unconstrained MinStation Problem (without requiring $s_0 \in S$). The problem of finding an optimal unconstrained solution is equivalent to that of finding the location of a single station in an optimal solution. We leave as an open problem the problem of designing a polynomial time algorithm for the general unconstrained case.

Then we address the following dual problem, which we call the *MinDistance Problem*: Suppose that we have a budget that allows us to build k refueling stations. Find the smallest d such that we can build k refueling stations that allow a drone with flight capacity d to guard the border of the given island. We present an algorithm which approximates an optimal solution to the MinDistance Problem up to an additive constant. The main tool in this solution is a discretization of the original MinStation Problem, via an algorithm we call APPSOL. This discretized approach also yields an easier to implement algorithm to approximate the MinStation problem.

Our results can be applied to problems such as border patrolling, where the use of fleets of small cheaper drones with limited capacity results in cheaper systems that use less resources, and increase the frequency with which the drones patrol the border.

A reviewer asked the natural question: can we find an optimal placement of base stations if instead of enclosing the polygon, we wish to travel between two arbitrary points on the polygon? We find that with some additional machinery, this problem can be answered by extending our results. We supply this proof in Section 4.6.

4.1.1 Related work. Drones have become the natural choice for the deployment of air surveillance systems [3, 33, 44]. There are three areas in which problems close to ours can be found: operations research, wireless networks and computational geometry.

In facility location problems, we are interested in finding the best places to locate a set of facilities (e.g. airports, pharmacies, gas stations, markets, etc.) to better serve a community, as well as creating optimal routes to visit them. Recently, applications in the aerial robotics community, such as finding the best places to locate drone base stations and creating flying routes for the drones, have arisen in areas such as border patrolling [55]. Cities on the borders of countries are modeled as demand points, and airports are considered as base stations or hubs. In [61] the authors studied a base location and path planning problem in maritime target reconnaissance problems. Their problem is formulated as an integer linear program where the total score obtained from visiting points of interest by flight routes of drones is maximized; a novel ant colony optimization metaheuristic approach is proposed. In a more recent paper [40], both capacity constraints on base stations and endurance limitations on drones are taken into account and two heuristic algorithms are designed to solve the problem.

Another field of research close to our work can be found in wireless sensor networks.

In [30], the authors study the k-barrier problem: how to deploy a set of sensors in a belt region surrounding a castle in such a way that any intruder is detected by at least k

sensors. In [7] the following problem is studied: Given that an intruder has been detected by a set of sensors, how can they be moved in an optimal way to prevent further intrusions? An interesting survey of problems similar to ours can be found in [1], where they study the problem of protecting several types of holes that can occur in a wireless sensor network, where a hole is a region not covered by the sensing disks of a set of sensors. In the same paper, other problems related to ours are considered, including routing in static and mobile sensor networks. See also [8, 19, 60].

Finally, in computational geometry, two classic areas of research that study problems related to ours are Art Gallery and Watchman Route problems [4, 52, 53, 59]. In the first one, we deal with the problem of finding a set of points S within an art gallery, usually modeled by a polygon $\mathcal{P}$, such that every point in the art gallery is visible from at least one point in S. The watchman problem is that of finding a path that a guard can follow in order to guarantee that any point within $\mathcal{P}$ is visible from a point in the path.

4.2 Terminology and Problem Formulation

In what follows a polygon $\mathcal{P}$ is represented by a sequence $\langle p_0, \ldots, p_{n-1} \rangle$ of its vertices given in clockwise order around its boundary. Thus, the edges of $\mathcal{P}$ are the line segments $\overline{p_i p_{i+1}}$, with addition taken mod n. We assume that our polygons are always *simple*, i.e. that no two non-consecutive edges intersect. We use *Int*$(\mathcal{P})$ and *Ext*$(\mathcal{P})$ to denote, respectively, the interior and exterior of the region enclosed by $\mathcal{P}$, and use $\mathcal{P}$ itself to refer to the boundary of this region (often referred to by $\partial\mathcal{P}$ in the literature). Accordingly, the length of $\mathcal{P}$ is the sum of the lengths of its edges. A (polygonal) path is a sequence of points $\langle q_0, \ldots, q_k \rangle$ together with the set of edges $\overline{q_i q_{i+1}}$, $i = 0, \ldots, k-1$; the length of a path is the sum of the lengths of its edges.

Given two distinct points $a, b \in \mathcal{P}$ the interval $[a, b]$ is the set of points of $\mathcal{P}$ traversed while moving from a to b in the clockwise direction along the boundary of $\mathcal{P}$. The distance

$\delta_{\mathcal{P}}(a, b)$ between a and b in $\mathcal{P}$ is the length of the interval $[a, b]$. Observe that since $a \neq b$, $[a, b] \neq [b, a]$, $[a, b] \cup [b, a] = \mathcal{P}$, and that $\delta_{\mathcal{P}}(a, b) + \delta_{\mathcal{P}}(b, a)$ is the length of $\mathcal{P}$.

The following definitions of what we will call d-paths and d-hulls arise from the restriction that the flight range of a drone is a fixed number d.

An open line segment contained in *Ext*$(\mathcal{P})$ joining two points $a, b \in \mathcal{P}$ will be called *a bridge*; if its length is at most d it is called a d-bridge of $\mathcal{P}$. Note that a drone cannot fly along a bridge of $\mathcal{P}$ with length greater than d; the base stations are restricted to be on $\mathcal{P}$, thus if a drone chooses to fly over a bridge with length greater than d it would run out of fuel and fall to the sea.

A polygonal path joining two points $a, b \in \mathcal{P}$ is called a d-path if all of its edges are d-bridges of $\mathcal{P}$, or segments of edges of $\mathcal{P}$. We say that a polygon $\mathcal{C}$ is a d-hull of $\mathcal{P}$ if it encloses $\mathcal{P}$, and all of its edges are contained in edges of $\mathcal{P}$ or are d-bridges of $\mathcal{P}$. Observe that a polygon has many (in fact an infinite number) of d-hulls, indeed $\mathcal{P}$ is a d-hull of itself.

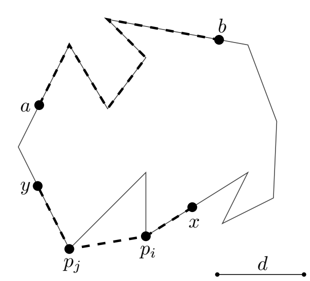

Figure 4.2: The interval $[a, b]$ and a d-path $\pi_{x,y}$ joining x and y are shown in black dashed lines. The segment $\overline{p_i p_j}$ is a d-bridge contained in $\pi_{x,y}$.

The *drone distance* $\delta(a, b)$ from a to b is the length of a shortest clockwise d-path joining a to b. As an example, Figure 4.2 shows the shortest d-path from x to y. For simplicity, we will refer to the drone distance as the distance from a to b. Observe that $\delta(a, b)$ is in general different from $\delta(b, a)$. Further observe that the drone distance and the geodesic distance from a to b (understood as the length of the shortest clockwise path from

a to b disjoint from *Int*$(\mathcal{P})$) coincide whenever this distance is at most d. Finally, note that if a drone with flight range d can fly between two points $a, b \in \mathcal{P}$ without recharging, then there is a d-path joining them of length at most d.

Our island guarding problem can now be restated as follows:

Problem 2 (MinStation)**.** Given a polygon $\mathcal{P}$ find a set of base stations $S = \{s_0, \ldots, s_{k-1}\}$ with minimum cardinality such that for every $0 \leq i \leq k-1$ there is a d-path π_i of length at most d joining s_i to s_{i+1}, and such that $\mathcal{C} = \pi_0 \cup \ldots \cup \pi_{k-1}$ is a d-hull of $\mathcal{P}$; $s_i \in \mathcal{P}$, $i = 0, \ldots, k-1$, addition taken mod k.

By a solution to the MinStation problem we refer simply to a set S of base stations together with the collection of d-paths π_i whose union is a polygon that encloses $\mathcal{P}$. Recall that a solution is optimal if it contains the least possible number of base stations, and quasi-optimal if it contains one more base station than an optimal solution.

We also study a kind of dual problem to the MinStation problem. Suppose that we have a budget that allows us to build k base stations, and want to find the locations along $\mathcal{P}$ where to build them such that the flight range of the drones used to patrol $\mathcal{P}$ is minimized, formally:

Problem 3 (MinDistance)**.** Given a simple polygon $\mathcal{P}$ and an integer $k \geq 2$, find the smallest d and a set $S = \{s_0, s_2, \ldots, s_{k-1}\}$ of k stations on $\mathcal{P}$ such that, for $i = 0, \ldots, k-1$, there is a d-path π_i of length at most d joining s_i to s_{i+1}, addition taken *mod* k, and $\mathcal{C} = \pi_0 \cup \ldots \cup \pi_{k-1}$ is a d-hull of $\mathcal{P}$.

Computing a d-hull that minimizes the number of base stations needed to solve the MinStation problem is more subtle than it may at first look. There are polygons $\mathcal{P}$ for which, given d, the smallest number of base stations needed to solve the MinStation problem, lie on the shortest d-hull enclosing $\mathcal{P}$. An example is shown in Figure 4.1. However, there are examples for which the stations of an optimal solution do not lie on the shortest

d-hull enclosing $\mathcal{P}$. An example is given in Figure 4.3. It is easy to see that placing a base station at any point other than the black points shown there, increases the number of base stations needed to solve the MinStation problem. In fact, it is not hard to construct polygons such that the number of stations required for the shortest d-hull is almost twice the number of stations given by the MinStation problem. This is the case, for example, for a star shaped polygon such that the distance between adjacent vertices on the boundary of the convex hull is $\frac{d}{2} + \epsilon$, for some arbitrarily small ϵ, as shown in Figure 4.4.

We remark that in the optimal solutions of the MinStation and the MinDistance problems, the base stations lie on $\mathcal{P}$ but not necessarily on vertices of $\mathcal{P}$ or $\mathcal{C}$.

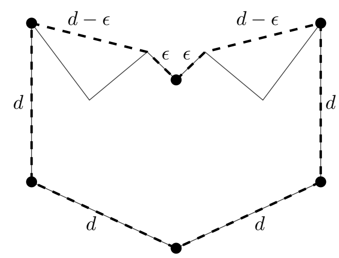

Figure 4.3: Example. The optimal d-hull requires 6 base stations. Replacing it with one with smaller perimeter increases the number of base stations to 7.

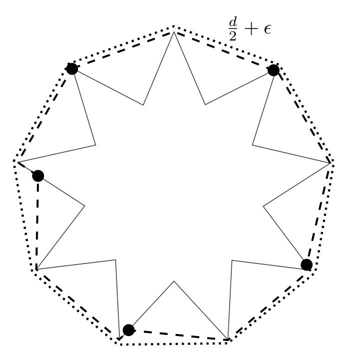

Figure 4.4: Example. The shortest d-hull (dotted) requires almost twice as many stations as the optimal d-hull that solves the MinStation problem (dashed). This example can be extended to polygons with arbitrarily many vertices.

4.3 Preliminary results

Given a fixed point $s \in \mathcal{P}$ we define a total order $O_s(\mathcal{P}, \preceq)$ on the points in $\mathcal{P}$ as follows:

1. for any point $a \in \mathcal{P}$, $s \preceq a$,
2. for any $a, b \in \mathcal{P}$, both different from s, $a \preceq b$ if $[s, a] \subseteq [s, b]$. (Note that possibly $a = b$.)

For convenience we will add an extra element s' to our order such that for any $a \in \mathcal{P}$, $a \preceq s'$; that is, s and s' are, respectively, the minimum and the maximum elements of $O_s(\mathcal{P}, \preceq)$. We can think of s' as a copy of s, and refer to $O_s(\mathcal{P}, \preceq)$ simply as $\preceq$.

Consider a point $s \in \mathcal{P}$ and the order $\preceq$ it defines on the points on $\mathcal{P}$. We define a distance δ_d on the points on $\mathcal{P}$ as follows:

1. $\delta_d(a, a) = 0$,
2. if $a \preceq b \in \mathcal{P}$, $\delta_d(a, b) = 1$ if there is a d-path of length at most d from a to b,
3. $\delta_d(a, b) = k$ if k is the smallest integer such that there is a sequence of points $p_0 = a, \cdots, p_k = b$ such that $\delta_d(p_i, p_{i+1}) = 1$, $i = 0, \ldots, k - 1$.

The following technical Lemma will be crucial in the proposed approach to solve the MinStation problem.

Lemma 34 (The Sandwich Lemma)**.** Let $w, x, y, z \in \mathcal{P}$ such that $w \preceq x \preceq y \preceq z$ on $\mathcal{P}$, such that $\delta_d(w, y) \leq 1$, and $\delta_d(x, z) \leq 1$. Then $\delta_d(w, z) \leq 2$, $\delta_d(w, x) \leq 2$ and $\delta_d(y, z) \leq 2$.

Proof. Suppose that $\delta_d(w, z) > 1$, for otherwise we are finished. Since $w \preceq x \preceq y \preceq z$ the shortest d-paths $\pi_{w,y}$ and $\pi_{x,z}$ joining w to y, and x to z intersect. Let p be a point in the intersection of $\pi_{w,y}$ and $\pi_{x,z}$. If the distance $\delta(x,p)$ along $\pi_{x,z}$ between p and x is smaller than the distance $\delta(p,y)$ between p and y along $\pi_{w,y}$, then $\delta(w,p) + \delta(p,x) \leq 1$ and therefore $\delta_d(w,z) \leq 2$. The case when $\delta(p,x) \geq \delta(p,y)$ follows the same way. The inequalities $\delta_d(w,x) \leq 2$ and $\delta_d(y,z) \leq 2$ are proved in a similar way. □

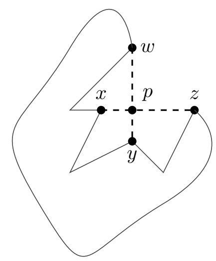

Figure 4.5: Illustration of Lemma 34.

Lemma 34 suggests that in an optimal solution to the MinStation problem, a drone flies around in a non-crossing curve $\mathcal{C}$ that encloses $\mathcal{P}$. We formalize this observation in the lemma that follows:

Lemma 35. Suppose that $s_0 \in \mathcal{P} \cap CH(\mathcal{P})$ and let $S = \{s_0, s_1, ..., s_{k-1}\}$ be a solution to the MinStation problem that goes around P in the clockwise direction and which has the least number of stations among all solutions starting from s_0. Then $s_0 \preceq s_1 \preceq s_2 \preceq \cdots \preceq s_{k-1}$.

Proof. Assume that all of the π_i paths joining s_i to s_{i+1} are of minimum length. Since $\mathcal{C} = \pi_0 \cup \cdots \cup \pi_{k-1}$ encloses $\mathcal{P}$, any point p in $CH(\mathcal{P})$ lies on $\mathcal{C}$. It is now easy to see that $\mathcal{C}$ *covers* p exactly once. It follows now that s_0 is not in the interior of π_0, and that π_0 is a simple curve that always advances in the clockwise direction along $\mathcal{C}$.

Now, suppose that $s_i \preceq s_{i-1}$ for some $i > 1$, and let j be the maximum value such that $s_j \preceq s_i$; that is, $s_j \preceq s_i \preceq s_{j+1}$. Using Lemma 34 it follows that $s_j \preceq s_i \preceq s_{j+1} \preceq s_{i-1}$. Thus, by Lemma 34, $\delta_d(s_j, s_i) \leq 2$, and since S is an optimal solution, it follows that $s_{j+1} = s_{i-1}$.

Let r be the minimum value such that $s_{i-1} \preceq s_r$. Then, we have that $s_i \preceq s_{r-1} \preceq s_{i-1} \preceq s_r$. It follows that $s_i = s_{r-1}$.

Now, since $s_{j+1} = s_{i-1}$ and $s_i = s_{r-1}$ we have that $s_j \preceq s_i \preceq s_{i-1} \preceq s_r$, where $\delta_d(s_j, s_{i-1}) = \delta_d(s_i, s_r) = 1$. Therefore, by Lemma 34, $\delta_d(s_j, s_r) \leq 2$. This is a contradiction, and thus $s_i \preceq s_{i+1}$ for all i. Hence, $\{s_0, s_1, ..., s_r\}$ continues to make forward progress and the result follows. □

A similar argument shows that for any optimal solution $S = \{s_0, s_1, ..., s_{k-1}\}$ of the MinStation Problem (no longer subject to the condition that s_0 is fixed) $\mathcal{C}$ is a simple closed curve.

4.4 OptSol

We consider the following algorithm, which constructs a solution to the MinStation problem starting at a point $v \in \mathcal{P} \cap CH(\mathcal{P})$.

Algorithm 2 OptSol

Input: Polygon $\mathcal{P}$, $s_0 \in \mathcal{P} \cap CH(\mathcal{P})$, $d > 0$.
Output: The stations in an optimal or quasi-optimal d-hull for $\mathcal{P}$.

procedure OptSol($\mathcal{P}, s_0, d$)
  Let $s_0 = s'_0 = y_{-1} = v$ and $y_0 = \max\{y : \delta_d(s_0, y) = 1\}$
  Set $S_0 = \{s_0\}$ and $i = 0$
  while $y_i \neq s'_0$ **do**
    $i = i + 1$
    $y_i = \max\{y : \delta_d(s_{i-1}, y) = 2\}$
    $s_i =$ any $s \in \{w : \delta_d(s_{i-1}, w) = 1$ and $\delta_d(w, y_i) = 1\}$
    Set $S_i = S_{i-1} \cup \{s_i\}$
  end while
  return $S = S_i$
end procedure

We claim that if we further require that $v \in S$, then the set S returned is, indeed, an optimal solution to MinStation. On the other hand, we observe that this algorithm always gives a solution that is globally optimal or quasi-optimal (no longer subject to the restriction that $v \in S$).

Theorem 36. Given a starting point $s_0 \in \mathcal{P} \cap CH(\mathcal{P})$, if k is the least value such that $s_k = s'_0$, then the set of points $S = \{s_0, \cdots, s_{k-1}\}$ returned by the OPTSOLalgorithm is an optimal solution to the MinStation problem with the additional requirement that a base station be located at s_0.

Proof. Suppose that S has more than one element, for otherwise our result is obvious. Suppose that $Z = \{z_0, \ldots, z_{n-1}\}$ is an optimal solution for the MinStation problem such that there is a base station located at $s_0 = z_0$. We prove now that $n = k$

By Lemma 35, we may assume that $v = z_0 \preceq z_1 \preceq \cdots \preceq z_{n-1}$, that for all i, $\delta_d(z_i, z_{i+1}) = 1$, and that $\delta_d(z_{n-1}, s'_0) = 1$. Consider now the set $S = \{s_0 = v, \cdots, s_{k-1}\}$ returned by the OPTSOL algorithm. Recall that $y_{k-1} = s'_0$. While the relationship between S and Z is unclear, the relationship between Z and $Y_{k-1} = \{y_{-1}, y_0, \ldots, y_{k-1}\}$ is more straightforward; indeed we prove by induction that, for all i, $z_{i+1} \preceq y_i$.

This clearly holds for $i = 0$, as $z_1 \preceq y_0$ by definition. Now, suppose $z_i \preceq y_{i-1}$. Let j be minimal so that $z_i \preceq y_{j-1}$. Then $j \leq i$, and as $v = y_{-1} \prec z_i$ we have that $j \geq 1$. Now, by the minimality of j, we know that $y_{j-2} \preceq z_i \preceq y_{j-1}$ and by definition of y_{j-2}, we have that $s_{j-1} \preceq y_{j-2}$. Combining, $s_{j-1} \preceq z_i \preceq y_{j-1}$. Then as $\delta_d(s_{j-1}, y_{j-1}) = 1$ by construction, Lemma 34 implies that $z_{i+1} \preceq y_j \preceq y_i$. This completes the inductive step, and hence the proof.

Therefore, the smallest value of k such that $\delta_d(s_k, s'_0) = 1$ is also the smallest value of n such that $\delta_d(y_n, s'_0) = 1$. Thus $S = \{s_0 = v, \cdots s_{k-1}\}$ is an optimal solution to the MinStation problem, with the additional requirement that there is a base station at $s_0 = v$.

□

Theorem 37. Let $s_0 = v \in \mathcal{P} \cap CH(\mathcal{P})$. The set $S = \{s_0, \ldots, s_{k-1}\}$ returned by the OPT-SOL algorithm is an optimal solution or a quasi-optimal solution to MinStation problem.

Proof. Suppose that $Z = \{z_0, \ldots, z_{n-1}\}$ is an optimal solution to the MinStation problem, and that $S = \{s_0, \ldots, s_{k-1}\}$ is the solution returned by the MinStation algorithm. We prove now that $k = n$ or $k = n + 1$.

Since s_0 is on the convex hull of $\mathcal{P}$ there is a shortest d-path between some z_i and z_{i+1} that contains s_0. Hence, adding s_0 to $Z = \{z_0, \ldots, z_{n-1}\}$ yields an optimal or quasi-optimal solution to the MinStation problem including s_0. □

Remark 1. Note that this algorithm can be used to solve the MinStation problem for a polygonal line instead of a simple polygon. This may be useful to patrol a section of the coastal boundary of a region that is not fully landlocked. We also note that our choice to travel clockwise is a convention and while traveling counterclockwise may result in different station locations, the same number of stations will be used.

Remark 2. Although there are polygons such that the optimal solution contains no points in *CH*$(\mathcal{P})$ (a simple modification of Figure 4.1 yields one such example), from a practical point of view, it is convenient to assume that at least one station v lies in $P \cap$ *CH*$(\mathcal{P})$, as otherwise any solution that includes v could have an arbitrarily large number of stations (imagine that it is located in a large pocket where bridges cannot be established).

Remark 3. Here we also note that a naive greedy algorithm performs poorly. In Figure 4.6 we see one such example. We may further construct similar polygons by increasing the number of pockets and their depth. As the number of pockets the depth of the pockets increase we see that a naive greedy algorithm on these polygons which explores all pockets gives an arbitrarily bad approximation.

4.4.1 Time complexity. We prove now that we can implement the OPTSOL algorithm to run in $O(n^2)$ time.

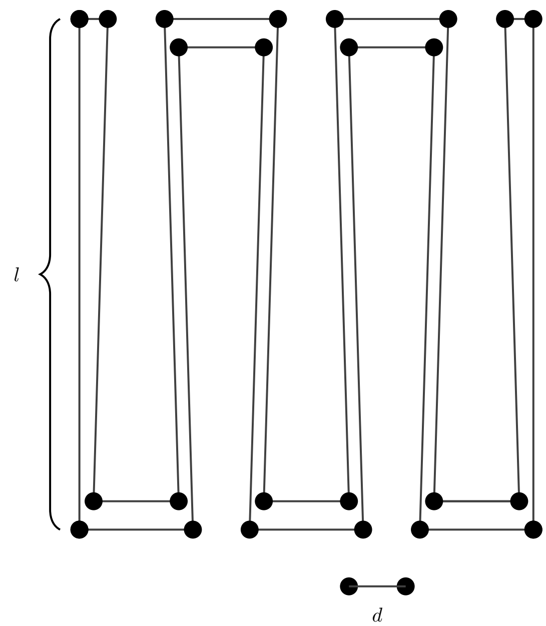

Figure 4.6: In this polygon, pockets are only avoidable if vertices of the polygon are used. The depth of this polygon is labeled as l.

Given a point s_i we want to find a point $y_{i+1} = \max\{y : \delta_d(s_i, y) = 2\}$ with respect to $\preceq$ and a point $s_{i+1} \in \{w : \delta_d(s_i, w) = 1 \wedge \delta_d(w, y_{i+1}) = 1\}$. We refer to the problem of finding s_{i+1} and y_{i+1} as the 2-*hop* problem, see Figure 4.7.

We will show that by applying a quadratic time pre-processing on $\mathcal{P}$ the 2-hop problem can be solved in linear time for each s_i.

A point x of an edge e is a *projection* of a vertex p_i on e if $x \preceq p_i$ and the line segment joining them is a d-bridge of $\mathcal{P}$ perpendicular to e. See Figure 4.7 (a).

In a similar way, we say that a point x of an edge e of $\mathcal{P}$ is called a d-*projection* of an edge f on e if there is a point $y \in f$ such that the line segment joining them is a bridge of $\mathcal{P}$ of length d perpendicular to f. See Figure 4.7 (b).

Lemma 38. Given s_i, s_{i+1} is either a vertex of $\mathcal{P}$, the projection of a vertex on an edge, the d-projection of an edge, or a point with $\delta(s_i, s_{i+1}) = d$.

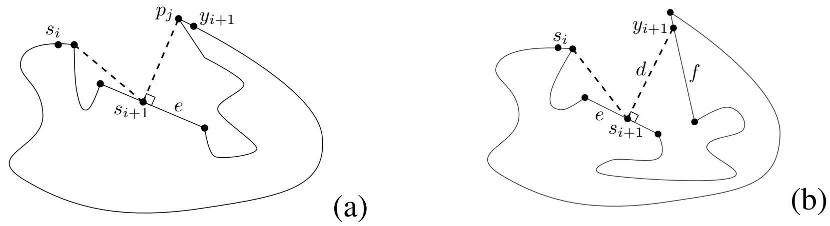

Figure 4.7: The 2-hop problem. (a) s_{i+1} is a projection of the vertex p_j on the edge e. (b) s_{i+1} is a d-projection of the edge f on the edge e.

Proof. Suppose that s_{i+1} is not a vertex of $\mathcal{P}$ and $\delta(s_i, s_{i+1}) < d$. Let e be the edge of $\mathcal{P}$ containing s_{i+1}, see Figure 4.7. Note that $\delta(s_{i+1}, y_{i+1}) = d$ by the choice of y_{i+1}. If s_{i+1} is neither the projection of a vertex on e nor the d-projection of an edge on e, then it can be moved slightly along edge e and advance y_{i+1}. This contradicts the definition of y_{i+1}. □

There might be $O(n)$ points at distance d from a previously placed station s_i. However, we only need to consider the maximum with respect to $\preceq$ among them as a candidate for placing s_{i+1}, as we prove next.

Lemma 39. Let w, x, y, z be points in $\mathcal{P}$ such that $w \preceq x \preceq y \preceq z$. Suppose that $\delta(w, x) = \delta(w, y) = d$, $\delta(w, z) > d$, and $\delta(x, z) = \ell$. Then, $\delta(y, z) \leq \ell$.

Proof. Let r be an intersection point of the shortest d-path $\pi_{w,y}$ from w to y and the shortest d-path $\pi_{x,z}$ from x to z. Note that r always exists by the choice of the four points on $\mathcal{P}$. Let $\delta(w, r)$ and $\delta(r, y)$ be the distance along $\pi_{w,y}$ between w and r, and between r and y, respectively. Let $\delta(x, r)$ and $\delta(r, z)$ be the distance along $\pi_{x,z}$ between x and r, and between r and z, respectively. Since $\delta(w, z) > d$, we have $\delta(r, y) < \delta(r, z)$. Now suppose that $\delta(r, y) > \delta(x, r)$. Then we have that $\delta(w, r) + \delta(r, x) < d$, which is as contradiction to our assumption that $\delta(w, x) = d$. Thus, $\delta(r, y) \preceq \delta(x, r)$ and $\delta(y, z) \leq \ell$. □

We claim that, although there may be $O(n^2)$ projections of vertices and d-projections of edges, $O(n)$ candidate points are sufficient to compute s_{i+1}.

Let e and f be edges of $\mathcal{P}$. We say that $e \lessdot f$ if for any point x in the interior of e and any point y in the interior of f, $x \preceq y$.

Lemma 40. For each edge e of $\mathcal{P}$ we need to store at most three points:

1. The minimum d-projection (with respect to $\preceq$) of an edge e' on e such that $e \lessdot e'$.

2. The endpoint not in e of the bridge generating the maximum d-projection (with respect to $\preceq$) on e of an edge e' such that $e' \lessdot e$. In this case the stored point lies on e'.

3. The minimum projection (with respect to $\preceq$) of a vertex on e.

Proof. Case 1. Let x and x' be d-projections of two distinct edges f and f', respectively, on e such that $e \lessdot f$ and $e \lessdot f'$. Let $\overline{xy}$ be the d-bridge perpendicular to e having x as an endpoint, i.e., $y \in f$ and $\overline{xy}$ is has length d. Define $\overline{x'y'}$ analogously. Because of the length of $\overline{xy}$ (respectively, $\overline{x'y'}$), if we place a station at x (respectively, x') then we also need to place a station at y (respectively, y'). Suppose w.l.o.g. that $x \preceq x'$, see Figure 4.8 Since all the bridges defining d-projections of edges on e are parallel, this implies that $f' \lessdot f$ and $y' \preceq y$. Moreover, as the interval $[x, y]$ contains the interval $[x', y']$, placing a station at x guarantees that both intervals of $\mathcal{P}$ are guarded. Hence, we maximize y_{i+1} with respect to $\preceq$ by choosing the minimum d-projection of an edge on e as s_{i+1}.

Case 2. This case is analogous to the first one, see Figure 4.8.

Case 3. Let x and x' be the projections of two distinct vertices p_i and p_j, respectively, on an edge e. Let $\overline{xp_i}$ and $\overline{x'p_j}$ be their corresponding d-bridges. Suppose w.l.o.g. that $x \preceq x'$. This implies that $p_j \preceq p_i$ and that placing a station at x guarantees that both intervals $[x, p_i]$ and $[x', p_j]$ are guarded, see Figure 4.9. It remains to be proven that by placing a station at

x we can advance further on $\mathcal{P}$ with respect to $\preceq$. Let $w \in \mathcal{P}$ be a point such that $p_i \preceq w$, and let $p_{x,w}$ and $p_{x',w}$ be the shortest d-paths joining x to w and x' to w. Let r be the intersection point of $\overline{xp_i}$ and $p_{x',w}$. Notice that the points x, x' and r form a right triangle that is right-angled at x. Therefore, the length of $p_{x,w}$ is smaller than the length of $p_{x',w}$, which implies that we can maximize y_{i+1} by choosing the minimum projection of a vertex on e as s_{i+1}.

□

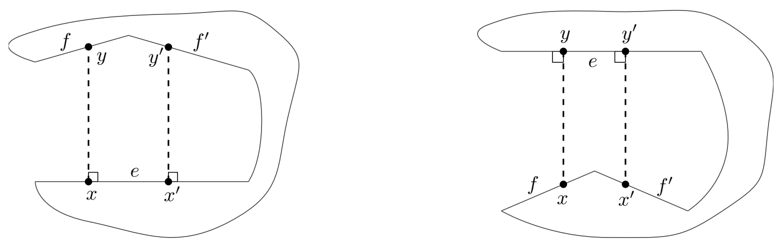

Figure 4.8: (a) Case 1: we only need to store the point x on edge e. (b) Case 2: we only need to store the point x on edge f.

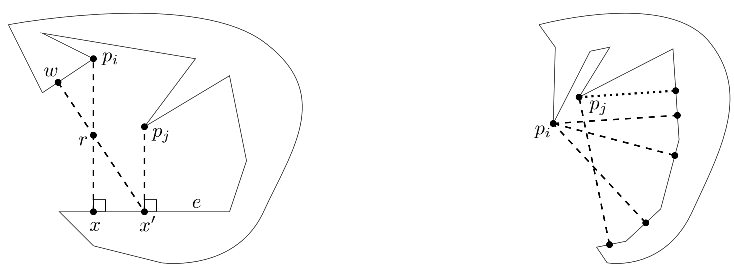

Figure 4.9: (a) Case 3: The distance from x to w is smaller than the distance from x' to w. (b) We need to store all vertex projections except the one that is the endpoint of the dotted segment.

In order to compute the candidate points on $\mathcal{P}$, we first find, for each edge $e \in P$, the subset containing each point $x \in \mathcal{P}$ for which there is a segment perpendicular to e joining x and e, and completely contained in $Ext(\mathcal{P})$. In such case we say that x is *orthogonally visible* from e.

We define a *lid* as an edge of the convex hull of $\mathcal{P}$ that is not an edge of $\mathcal{P}$. Each lid $h = \overline{ab}$ defines a polygon $\mathcal{P}_h$, which is the union of h and the interval of $\mathcal{P}$ determined by a and b which has no points in the convex hull of $\mathcal{P}$ besides a and b. Note that any projection of a vertex or d-projection of an edge is defined by a segment whose endpoints are contained in the same $\mathcal{P}_h$, for otherwise the segment would intersect $Int(\mathcal{P})$. Therefore, we only need to compute the set of points orthogonally visible from each edge e contained in a $\mathcal{P}_h$; moreover, we only need to look at the polygon $\mathcal{P}_h$ containing e to find these points.

For the next lemma, we assume that we have computed the polygons defined by all the lids of $\mathcal{P}$, as well as the triangulation of each such polygon. This can be done in $O(n)$ time overall, see [47] and [12].

Lemma 41. We can find the set containing all the segments of $\mathcal{P}$ orthogonally visible from any edge of $\mathcal{P}$ in $O(n)$ time. Moreover, each such set has $O(n)$ size.

Proof. Let $h = \overline{ab}$ be a lid of $\mathcal{P}$ and let $e = \overline{uv}$ be an edge of $\mathcal{P}_h$. We proceed as follows: Compute the set $VP(\mathcal{P}_h, e)$ of points of $\mathcal{P}_h$ visible from a point in e. $VP(\mathcal{P}_h, e)$ can be computed in $O(n)$ time [21].

Suppose w.l.o.g. that $u \prec v$. Let R be the region contained between the lines perpendicular to e through u and v, and to the left of the line directed from u to v. It is easy to see that any point of $\mathcal{P}$ orthogonally visible from e must lie in $\mathcal{R}_e = VP(\mathcal{P}_h, e) \cap R$, which can be computed in $O(n)$ time by intersecting $VP(\mathcal{P}_h, e)$ with both lines. We suppose w.l.o.g. that e is horizontal and that the interior of $\mathcal{R}_e$ lies above e.

We say that a vertex $p \in \mathcal{R}_e$ is a *turn vertex* if the maximal vertical segment $\overline{xy}$ through p and completely contained in $\mathcal{R}_e$ separates $\mathcal{R}_e$ into three subpolygons, see Figure 4.10(a). If two of these subpolygons lie to the right (left) of $\overline{xy}$, we say that p is a *right (left)* turn vertex. Let x be the top endpoint of $\overline{xy}$. The segment $\overline{px}$ separates $\mathcal{R}_e$ into two subpoygons, one of them containing e. Let $R_e(p)$ denote the subpolygon generated by $\overline{px}$ not containing e. It is easy to see that any point in $\mathcal{R}_e$ not being orthogonally visible from

e lies in the subpolygon $\mathcal{R}_e(p)$ for some turn vertex p, and that any point in $\mathcal{R}_e(p) \setminus \overline{px}$ is not orthogonally visible from e.

Note that the internal angles at both vertices of $e = \overline{uv}$ are convex in $\mathcal{R}_e$. Ghosh et al. [20] proved that for any vertex p in R_e, the shortest path from u to p, denoted as $\rho_{u,p}$, makes a left turn at every vertex of the path, and $\rho_{v,p}$ makes a right turn at every vertex of the path. This also holds true for the points in the interior of any edge of $\mathcal{R}_e$.

Let p be a turn vertex of $\mathcal{R}_e$ and let x be the top endpoint of the maximal vertical segment through p completely contained in $\mathcal{R}_e$. We claim that the vertical line through p, ℓ_p, does not intersect any point of $\mathcal{R}_e(p) \setminus \overline{px}$. Suppose otherwise that there is a point x' in $\mathcal{R}_e(p) \setminus \overline{px}$ contained in ℓ_p. Then, there exists a vertex q in $\mathcal{R}_e(p) \setminus \overline{px}$ such that $\rho_{v,x'}$ makes a left turn at q or $\rho_{u,x'}$ makes a right turn at q, which is a contradiction [20], see Figure 4.10 (b).

It follows that $\mathcal{R}_e(p) \cap \ell_p = \overline{px}$. This fact yields the following algorithm for removing $\mathcal{R}_e(p)$ from $\mathcal{R}_e$ for each turn vertex p.

We deal with the right turn vertices by traversing the edges of $\mathcal{R}_e$ clockwise from v to u. We set a variable *edgeIsVisible* to *true*. Let $f = \overline{qr}$, $q \prec r$, be the current edge in the traversal.

- If *edgeIsVisible* is *true* we check if r is a right turn vertex. In the affirmative case, we set *edgeIsVisible* to *false* and store the vertical line through r, ℓ_r and the edge f.

- If *edgeIsVisible* is *false*, then we had previously stored the last visible edge $g = \overline{op}$, where p is a right turn vertex, and the vertical segment through p, ℓ_p. We check if $x = f \cap \ell_p$ is not empty. In such a case, we replace the interval $[p, x]$ of $\mathcal{R}_e$ with the vertical segment $\overline{px}$, set *edgeIsVisible* to *true*, and discard g and ℓ_p.

We can remove the sub-polygons defined by the left turn vertices analogously by traversing $\mathcal{R}_e$ counter-clockwise from u to v. As each edge of $\mathcal{R}_e$ is visited at most twice, the removal of the sub-polygons defined by all the turn vertices takes $O(n)$ time. Let $\mathcal{R}'_e$ be the polygon obtained by this traversal.

To obtain the subset of $\mathcal{P}$ orthogonally visible from e, we only need to discard e, the segment contained in the lid of $\mathcal{P}_h$, and the vertical segments added in the previous process (at most one per turn vertex) from $\mathcal{R}'_e(p)$.

Since $\mathcal{P}_h$ has no holes, each edge of $\mathcal{P}_h$ provides at most one segment to $\mathcal{R}'_e$. Therefore, the set of segments of $\mathcal{P}$ orthogonally visible from any edge e has $O(n)$ size. $\square$

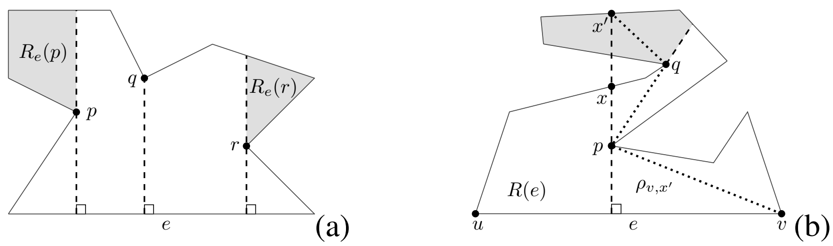

Figure 4.10: (a) p is a left turn vertex, r is a right turn vertex, and q is not a turn vertex. (b) Neither x' nor any point in the shaded region is in R_e: the shortest path from v to x' makes a left turn at q.

Lemma 42. For any edge e of $\mathcal{P}$ we can find the projections described in Lemma 40 in $O(n)$ time.

Proof. Suppose that $e = \overline{uv}$, $u \preceq v$. By Lemma 41, we can find the set W of all the segments of edges and vertices of $\mathcal{P}$ orthogonally visible from e in $O(n)$ time; moreover, W has $O(n)$ size. Let W_B be the subset of elements of W smaller than u and let W_A be the subset of the elements of W greater than v with respect to $\preceq$.

We find the d-projections corresponding to the first two cases of Lemma 40 as follows. Let ℓ be the line parallel to e, to the left of the line directed from u to v and at distance d from e. We first compute the intersection of ℓ with both W_B and W_A, which by the size of W can be obtained in $O(n)$ time. To obtain the point described in the first case of the proof of Lemma 40 we take the maximum point q with respect to $\preceq$ in $\ell \cap W_A$ and store the intersection point of e with the line through q perpendicular to e. To obtain the described in the second case of the proof of Lemma 40 we store the minimum point in $\ell \cap W_B$ with respect to $\preceq$, if any.

We find the projection of the maximum vertex on e described in the third case of Lemma 40 as follows. For each vertex of $\mathcal{P}$ in W_B we compute its distance with respect to e. We then store the maximum with respect to $\preceq$ of the vertices at distance less or equal than d from e. □

Now we need to solve the following subproblem: given a point $x \in \mathcal{P}$, find the maximum w, $x \preceq w$, such that $\delta(x, w) = d$. Guibas et al. [21] proved that, given the triangulation of a polygon $\mathcal{R}$ and a point $p \in \mathcal{R}$, the euclidean shortest paths from p to all the vertices of $\mathcal{R}$ can be found in linear time (see also [32]). The union of all the shortest paths from the source point p to the vertices of $\mathcal{R}$ is a planar tree called the *shortest-path tree* of $\mathcal{R}$ with respect to p.

Let $\mathcal{R}$ be the polygon obtained by enclosing $\mathcal{P}$ in a sufficiently large rectangle and connecting one of the sides of the rectangle to the starting point of the sequence, x_0, with a thin corridor. The polygon $\mathcal{R}$ can be obtained in $O(n)$ time, see [52]. Note that $\mathcal{R}$ has $m \leq n + 8$ vertices and $\mathcal{P}$ is contained in the exterior of $\mathcal{R}$. We assign to the points in $\mathcal{R}$ that are also points in $\mathcal{P}$ the same order as in $\mathcal{P}$.

Henceforth we assume that $\mathcal{R}$ has been computed along with its triangulation, which as proven by Chazelle [12] can be found in $O(n)$ time.

Lemma 43. Given any point $x \in \mathcal{P}$, the point $w \in \mathcal{P}$ with $\delta(x, w) = d$ such that w is maximum with respect to $\preceq$ can be found in $O(n)$ time.

Proof. Let x be a point in $\mathcal{P}$ and let x' be its corresponding point in $\mathcal{R}$. We compute the shortest path $\rho(x', y)$ from x' to every vertex $y \in R$ such that y is also a vertex of P and $x' \prec y$. Let T be the shortest-path tree obtained by the union of these shortest paths. Let

M be the set of vertices of T such that, for any $w \in M$, $\delta(x', w) \leq d$, and w shares an edge of $\mathcal{R}$ with a vertex y such that $\delta(x', y) > d$. The set M can be found in $O(n)$ time by traversing T from its root x'.

Observe that any point of $\mathcal{R}$ at distance d from x' is one of the following:

- An element of M.
- A point in an edge $e = \overline{uv}$, $u \prec v$, of $\mathcal{R}$ such that $e \in E(T)$. In this case, $u \in M$ and $\delta(x', v) > d$.
- A point in an edge $e = \overline{uv}$, $u \prec v$, of $\mathcal{R}$ such that $e \notin E(T)$. Notice that, in this case, $\delta(x', v) > d$. Moreover, there is exactly one $z \in M$ such that $(z, u) \in E(T)$.

Hence, in order to find all the points at distance exactly d from x' it is sufficient to check the edges having a neighbour of an element of M in T as an endpoint. Since each vertex is adjacent to at most one element of M, this can be done in $O(n)$ time. At the final step we need to find the maximum among all the points at distance d from x', which can also be done in $O(n)$ time. Our result follows. □

Theorem 44. Let $\mathcal{P}$ be a polygon with n vertices and let $s_0 \in \mathcal{P}$ be a point on the convex hull of $\mathcal{P}$. Then OPTSOL returns an optimal solution S to the MinStation Problem such that $s_0 \in S$ in $O(n^2 + Ln)$ time, where L is the length of $\mathcal{P}$.

Proof. By Lemma 38, given s_i, the point s_{i+1} is either a point on $\mathcal{P}$ at distance exactly d from s_i, a vertex of $\mathcal{P}$, the projection of a vertex onto an edge, or the d-projection of an edge onto another edge.

By Lemma 39, we only need to consider the maximum point with respect to $\preceq$ at distance d from x_i, which can be found in $O(n)$ time as stated in Lemma 43.

There might be $O(n^2)$ projections of vertices and d-projections of edges. However, Lemma 40 states that in the set of candidates we need to store at most three projections for each edge of $\mathcal{P}$. Moreover, these projections can be found in $O(n)$ time for each edge.

The set of candidate points to compute all the elements of the set S has $O(n)$ size. For each candidate x, we compute the maximum point at distance d from x and associate this point to x, which by Lemma 43 takes $O(n)$ time per candidate.

It is easy to see that we only need to consider the candidates contained in the interval of $\mathcal{P}$ from s_i to the maximum point with respect to $\preceq$ at distance d from s_i. From all these candidates, we choose as s_{i+1} the candidate which maximizes y_{i+1}, which can be done in $O(n)$ time. Since we might need to place $O(\frac{L}{d})$ stations, this step takes time $O(\frac{L}{d}n)$. Therefore, the set S can be found in $O(n^2 + \frac{L}{d}n)$ time. □

4.5 Discretization

In this section, we present a discretization algorithm that is easy to implement for the MinStation problem, and then show how it can be utilized to obtain a solution to the MinDistance problem which is close to optimal. This algorithm avoids computing projections, drone distances (geodesic paths) and orthogonal visibility, which makes it very practical. The idea is to construct a graph and apply a slight modification of Dijkstra algorithm.

Fix $0 < \epsilon \leq d$ and let $X = \{s_0 = x_0 \preceq \ldots \preceq x_{r-1}\} \subseteq \mathcal{P}$ be a set of points so that s_0 lies on $\mathcal{P} \cap \mathit{CH}(\mathcal{P})$ and the distance between x_i and x_{i+1} along $\mathcal{P}$ is at most ϵ, addition taken mod r. For technical reasons that will become apparent later, we also ask that the vertices of $\mathcal{P}$ are contained in X. Consider the graph $G_d(X)$ such that $V(G_d(X)) = X$ in which two elements $x_i, x_j \in X$ are adjacent if the length of the geodesic path π_{x_i,x_j} in $\mathcal{P} \cup Ext(\mathcal{P})$ connecting them is at most d (as we will show soon, computing $G_d(X)$ does not require the shortest-path trees mentioned in Lemma 10). We then solve the problem of finding a shortest cycle in $G_d(X)$ from x_0 to itself going around $\mathcal{P}$. The set of vertices of that cycle, including x_0, is a valid solution to our problem, but not necessarily an optimal one.

Note that the problem of finding a shortest cycle from x_0 to itself can be reduced to that of finding a shortest path from x_0 to a copy $x'_0 = x_r$ of x_0. To this end, we insert x'_0 in $V(G_d(X))$ in such a way that, if the length of the interval $[x_i, x_0]$ is at most d, then x_i is adjacent to x'_0 instead of x_0.

Now we show in detail how the algorithm works, including how to compute $G_d(X)$.

It is possible to check whether a directed edge (x_i, x_j) belongs to E_2 in $O(n)$ time. This leads to a total time complexity of $O((\frac{L}{\epsilon})^3 + (\frac{L}{\epsilon})^2 n)$ for AppSol, where L denotes the total length of $\mathcal{P}$.

This algorithm, while simpler to implement than OptSol, does not directly yield an approximation to the MinStation problem (this is discussed in more detail in the next section). We now show how we can improve on this by applying this algorithm more than once: two applications of the MinStation AppSol algorithm can be used to *certify* the sharpness of a single application of this result, and a logarithmic number of applications can be used to give an additive approximation for MinDistance.

Denote by $\alpha(\mathcal{P}, s_0, d, \epsilon)$ be the number of base stations found by the AppSol algorithm for given $\mathcal{P}$, $s_0 \in \mathcal{P}$, flight range d, and $\epsilon > 0$. Let k be the minimum number of base stations among all solutions that have s_0 as one of their base stations. The key is the following result:

Theorem 45. $\alpha(\mathcal{P}, s_0, d+\epsilon, \epsilon) \leq k \leq \alpha(\mathcal{P}, s_0, d, \epsilon)$. In particular, if $\alpha(\mathcal{P}, s_0, d+\epsilon, \epsilon) = \alpha(\mathcal{P}, s_0, d, \epsilon)$, the solution is best possible among those containing s_0.

Proof. Clearly, $\alpha(\mathcal{P}, s_0, d, \epsilon) \geq k$. It suffices to show that $\alpha(\mathcal{P}, s_0, d+\epsilon, \epsilon) \leq k$. Consider an optimal set of k base stations S^*. Let S be a set of base stations obtained by selecting the nearest point in X for each point in S^*, then the geodesic distance between consecutive base stations in S is at most $d+\epsilon$. Therefore $\alpha(\mathcal{P}, s_0, d+\epsilon, \epsilon) \leq k$. □

Algorithm 3 AppSol

Input: Polygon $\mathcal{P}$, $s_0 \in \mathcal{P} \cap CH(\mathcal{P})$, $d > 0$, $\epsilon > 0$.

Output: List of stations in a d-hull of $\mathcal{P}$.

procedure AppSol$(\mathcal{P}, s_0, d, \epsilon)$
  if $s_0 = x_0$ is not a vertex of $\mathcal{P}$ **then**
    we make it so by splitting the edge that contains it into two edges
  end if
  Let V be the set of vertices of $\mathcal{P}$ and set $X = V$
  Add a copy x'_0 of x_0 to X
  $N \leftarrow n$
  for each edge of $\mathcal{P}$ of length $\ell > \epsilon$ **do**
    Add $\lceil \frac{l}{\epsilon} \rceil$ points to X dividing the edge into segments of length $\leq \epsilon$
  end for
  Let $X = \{x_0, x_1, \ldots, x_{m-1}, x_m = x'_0\}$ be the set of points in clockwise order around $\mathcal{P}$
  Construct a weighted directed graph $H_d(X) = (X, E)$ with $E = E_1 \cup E_2$ defined as follows:
    a: $(x_i, x_j) \in E_1$ if $j = i + 1$ and x_i, x_{i+1} are on the same edge of $\mathcal{P}$
    b: $(x_i, x_j) \in E_2$ if $i < j$, and the open segment from x_i to x_j has length $\leq d$ and is contained in *Ext*$(\mathcal{P})$
    c: The weight of each edge $(x_i, x_j) \in E$ is the Euclidean distance between x_i and x_j
  for each $x_i \in X$ **do**
    use Dijkstra's algorithm to compute X_i, the set of vertices of X that can be reached from x_i by a directed path of total weight $\leq d$
  end for
  Construct a graph with vertex set X where x_i is adjacent to x_j iff $x_j \in X_i$ or $x_i \in X_j$. Since the vertices of $\mathcal{P}$ belong to X, one can easily check that this graph is actually $G_d(X)$.
  Using Dijkstra's algorithm, compute a shortest path P from x_0 to $x'_0 = x_m$ of minimum length in $G_d(X)$
  return vertices of P
end procedure

In practice, Theorem 45 can be used to certify that the approximation given by a discretization corresponds to the solution computed by the OPTSOL algorithm, for fixed s_0. Indeed, Figure 4.11 shows two optimal solutions, for different values of d with $s_0 = 0$, that are obtained and certified by this method.

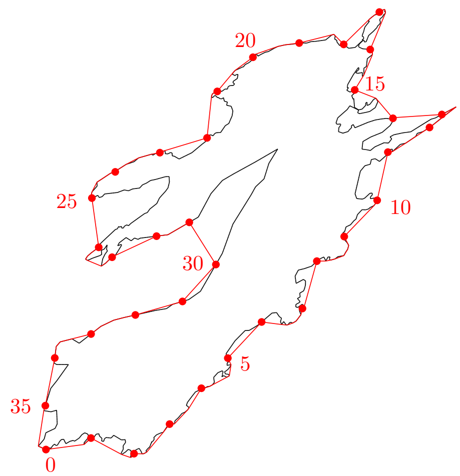

Figure 4.11: Salamis Island in the Saronic Gulf. (a) 36 base stations for $d = 2000$. (b) 26 base stations for $d = 2400$. The number of base stations in (a) and (b) is optimal among those containing s_0 by Theorem 45.

On the other hand, since s_0 lies on the boundary of the convex hull of P, every solution to MinStation must contain a station on a point $z \in \mathcal{P}$ such that $\delta(x_0, z) \leqslant d$. This can easily be seen to imply that Theorem 13 can be adapted to work for general solutions (and not only those that contain x_0) by modifying the algorithm so that it searches for the shortest path in $G_d(X)$ from x_i to itself for all x_i with $\delta(x_0, x_i) \leqslant d + \frac{\epsilon}{2}$, and then returns the shortest one among all of these. This slight variant of APPSOL will be called APPSOL2.

This has immediate implications for MinDistance; if the least number of stations in a solution in $G_{d+\epsilon}(X)$ is at most k, then the optimal solution to the MinDistance Problem (find the smallest flight range such that k stations are sufficient) lies between d and $d + \epsilon$. Thus by using binary search on d, the optimal flight range can be approximated up to an additive constant.

Theorem 46. Given a positive integer k and an $\epsilon > 0$, it is possible to find a solution to the MinDistance problem using k base stations such that the flight capacity of the drones is at most ϵ larger than the optimal one. This is achieved by running $O(\log |X|) = O(\log(\frac{L}{\epsilon} + n))$ iterations of APPSOL2 to perform a binary search on the set of all distinct drone (geodesic) distances between pairs of points of X.

Corollary 47. Given a positive integer k and an $\epsilon > 0$,

- An additive ϵ-approximation for the MinDistance problem with one fixed base station can be computed in $O((n^2 + Ln)\log(\frac{L}{\epsilon} + n))$ time.
- A quasi-optimal additive ϵ-approximation for the MinDistance problem (i.e. with k or $k+1$ base stations) can be computed in $O((n^2 + Ln)\log(\frac{L}{\epsilon} + n))$ time.

4.6 Extending Results

We have an algorithm for encircling an island with close to an optimal number of drones. A very natural question is, what if we only wish to connect two points on this polygon? We now consider the following variant.

Problem 4 (MinStationVariant)**.** Given a polygon $\mathcal{P}$ and two points on the boundary s, f, find a set of base stations $S = \{s = s_0, \ldots, s_k = f\}$ with minimum cardinality such that for every $0 \leq i \leq k-1$ there is a d-path π_i of length at most d joining s_i to s_{i+1}.

At first glance it may seem that Problem 4 may be solved directly by OPTSOL. Under many circumstances this is indeed the case. However, we find that navigating 'pockets' can become tricky. In Figure 4.12 we see where problems arise. An optimal solution to Problem 4 clearly requires base stations on both the clockwise and anticlockwise arcs from s to t. In OPTSOL we argued that under $\preceq$ an optimal S was clearly monotonic. Unfortunately, as we see in Figure 4.12, defining $\preceq$ as before would not allow optimal S to have such a property. It is for this reason we introduce two orderings, which as it turns out, ends up being enough.

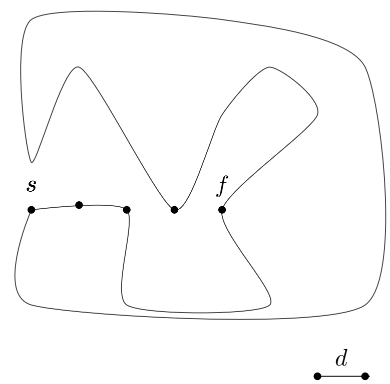

Figure 4.12: Optimally placed based stations will be located on both the clockwise and counterclockwise arcs from s to t.

4.6.1 The 2 step algorithm for pockets. We choose two distinct points on ∂P, name one s (for start) and the other f (for finish). We then define a clockwise ordering $\prec^+$ on the points of ∂P, where $a \prec b$ if b lies on the clockwise arc (a, f). We also define a counterclockwise ordering $\prec^-$, where $a \prec^- b$ if b lies on the counterclockwise arc (a, f). Notice that in both orderings we think of f as being maximal. Like before, we say that $d(x, y)$ is the geodesic distance on $\mathbb{R}^2 \setminus \text{int}(P)$. We define a metric d_h on ∂P. Given $a \in P$ we let $d_h(a, a) = 0$. We say that $d_h(a, b) = 1$ if a drone at the location a may legally fly to a base station located at b, and $d_h(a, b) = k$ if k is the minimal value such that there is a location $c \in \partial P$ such that $d_h(c, b) = 1$ and $d_h(a, c) = k - 1$.

4.6.2 The Algorithm. We define $x_0^+ = x_0^- = s \neq f$. We define y_0^+ as the point y maximum with respect to $\prec^+$ such that $d(y, s) \leq d$ and y_0^- as the point y maximum with respect to $\prec^-$ such that $d(y, s) \leq d$. For $i \geq 1$ we define y_i^+ as the point on (s, f) maximum with respect to $\prec^+$ such that either $d_h(x_{i-1}^+, y_i^+) \leq 2$ or $d_h(x_{i-1}^-, y_i^+) \leq 2$. We choose x_i^+ as a point that is one hop from y_i^+ and one hop from either x_{i-1}^- or x_{i-1}^+. Similarly for $i \geq 1$ we define y_i^- as the point on (s, f) maximum with respect to $\prec^-$ such that either $d_h(x_{i-1}^+, y_i^-) \leq 2$ or $d_h(x_{i-1}^-, y_i^-) \leq 2$. We choose x_i^- as a point that

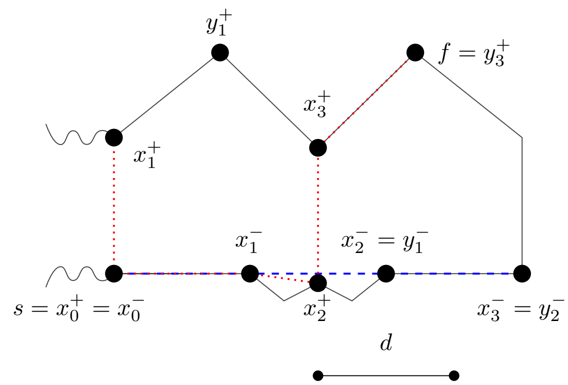

Figure 4.13: In this example x_1^+ is also y_0^+, and x_1^- is also y_0^-. y_1^+ is also a candidate for x_2^+, as both are exactly distance d from x_3^+. However, a slight modification of the polygon could make one of these two candidates part of the unique optimal solution. Here the optimal $S = \{s, x_1^-, x_2^+, x_3^+, f\}$.

is one hop from y_i^- and one hop from either x_{i-1}^- or x_{i-1}^+. Notice that for all $i \geq 1$, $\min\{d_h(x_i^+, x_{i-1}^+), d_h(x_i^+, x_{i-1}^-)\} \leq 1$ and $\min\{d_h(x_i^-, x_{i-1}^+), d_h(x_i^-, x_{i-1}^-)\} \leq 1$. It follows that $d_h(s, y_i^+) \leq i$ and $d_h(x_0, y_i^-) \leq i$. We terminate the algorithm when y_j^+ or y_j^- is f and let S be the set consisting of s, f, and the $x_i's$ which witness $d_h(s, f) \leq j$.

In Figure 4.13 we see a small example of this algorithm in practice. Notice that while there are two solutions to OPTSOLVAR for this example, small changes in the polygon can quickly make solutions unique.

We can picture the set of x_i forming a tree with root s where x_i's are vertices and edges are present when some x_{i+1} is formed through some x_i. (y_i^+) and are non-decreasing with respect to $\prec^+$ and $\prec^-$, respectively.

Like before, we will show that S from OPTSOLVAR is optimal by relating some optimal solution to y_i values (in this case y_i^+ and y_i^-).

Label some set of points (z_i) which are a solution of OPTSOLVARsuch that $z_0 = s$, $z_m = f$ and for all $i < m$, $d(z_i, z_{i+1}) \leq 1$. We will show that that our set S is of size m. In

Lemma 48 we argue that if z_k is properly bounded by y_i^+ and y_i^-, then for some $j \leq i$, it is properly sandwiched by some x_j and y_j where $j \leq i$.

Lemma 48 (Sandwich Lemma Variant)**.** Fix some k. If for some i, $z_k \preceq^- y_i^-$ and $z_k \preceq^+ y_i^+$, then there exists some $j \leq i$ such that

1. $x_j^+ \prec^- z_k \preceq^- y_j^-$ and $d_h(x_j^+, y_j^-) \leq 1$ or
2. $x_j^- \prec^- z_k \preceq^- y_j^-$ and $d_h(x_j^-, y_j^-) \leq 1$ or
3. $x_j^+ \prec^+ z_k \preceq^+ y_j^+$ and $d_h(x_j^+, y_j^+) \leq 1$ or
4. $x_j^- \prec^+ z_k \preceq^+ y_j^+$ and $d_h(x_j^-, y_j^+) \leq 1$.

Proof. Consider the minimal i with the property $z_k \preceq^- y_i^-$ and $z_k \preceq^+ y_i^+$. By the minimality of i, $y_{i-1}^+ \prec^+ z_k$ or $y_{i-1}^- \prec^- z_k$. WLOG suppose $y_{i-1}^+ \prec^+ z_k$. Notice that by definition y_{i-1}^+ is the maximal two step from x_{i-2}^+ and x_{i-2}^-, and notice that x_i^+ and x_i^- are 2 steps from x_{i-2}^+ or x_{i-2}^-, therefore $x_i^- \preceq^+ y_{i-1}^+$ and $x_i^+ \preceq^+ y_{i-1}^+$. By definition, either $d_h(x_i^+, y_i^+) \leq 1$ or $d_h(x_i^-, y_i^+) \leq 1$. Therefore one of the four cases is satisfied. □

In Lemma 49 we assume that z_k is properly sandwiched by some x_j and y_j and show that z_{k+1} is properly bounded by some y_{j+1}.

Lemma 49 (Bounding z_{k+1} when z_k is sandwiched.)**.** If for some k and some j,

1. $x_j^+ \prec^- z_k \preceq^- y_j^-$ and $d_h(x_j^+, y_j^-) \leq 1$ or
2. $x_j^- \prec^- z_k \preceq^- y_j^-$ and $d_h(x_j^-, y_j^-) \leq 1$ or
3. $x_j^+ \prec^+ z_k \preceq^+ y_j^+$ and $d_h(x_j^+, y_j^+) \leq 1$ or
4. $x_j^- \prec^+ z_k \preceq^+ y_j^+$ and $d_h(x_j^-, y_j^+) \leq 1$,

then $z_{k+1} \preceq^- y_{j+1}^-$ and $z_{k+1} \preceq^+ y_{j+1}^+$.

Proof. Fix some k and some j, with the desired property. By 48

1. $x_j^+ \prec^- z_k \preceq^- y_j^-$ and $d_h(x_j^+, y_j^-) \leq 1$ or
2. $x_j^- \prec^- z_k \preceq^- y_j^-$ and $d_h(x_j^-, y_j^-) \leq 1$ or
3. $x_j^+ \prec^+ z_k \preceq^+ y_j^+$ and $d_h(x_j^+, y_j^+) \leq 1$ or
4. $x_j^- \prec^+ z_k \preceq^+ y_j^+$ and $d_h(x_j^-, y_j^+) \leq 1$.

For cases 1 and 2, By sandwich lemma and the definition of y_{j+1}^+ and y_{j+1}^-, $z_{k+1} \preceq^- y_{j+1}^-$ and $z_{k+1} \preceq^+ y_{j+1}^+$.
For cases 3 and 4, By sandwich lemma and the definition of y_{j+1}^+ and y_{j+1}^-, $z_{k+1} \preceq^- y_{j+1}^-$ and $z_{k+1} \preceq^+ y_{j+1}^+$. □

Like in the original OPTSOLwe use induction to show that for for all $i \geq 0$, $z_{i+1} \preceq^+ y_i^+$ and $z_{i+1} \preceq^- y_i^-$, which will finish our proof.

Theorem 50 (Proof that the OPTSOLVARis optimal)**.** For all $i \geq 0$, $z_{i+1} \preceq^+ y_i^+$ and $z_{i+1} \preceq^- y_i^-$.

Proof. We proceed by induction. We see this holds when $i = 0$.
Suppose this holds for some $i \geq 0$. By Lemma 1 there exists some $j \leq i$ such that

1. $x_j^+ \prec^- z_{i+1} \prec^- y_j^-$ and $d_h(x_j^+, y_j^-) \leq 1$ or
2. $x_j^- \prec^- z_{i+1} \prec^- y_j^-$ and $d_h(x_j^-, y_j^-) \leq 1$ or
3. $x_j^+ \prec^+ z_{i+1} \prec^+ y_j^+$ and $d_h(x_j^+, y_j^+) \leq 1$ or
4. $x_j^- \prec^+ z_{i+1} \prec^+ y_j^+$ and $d_h(x_j^-, y_j^+) \leq 1$,

By Lemma 2, $z_{i+2} \prec^- y_{j+1}^-$ and $z_{i+2} \prec^+ y_{j+1}^+$ and since (y_i^+) and (y_i^-) are increasing, $z_{i+2} \prec^- y_{i+1}^-$ and $z_{i+2} \prec^+ y_{i+1}^+$ concluding our inductive step. □

4.7 Acknowledgements

The result of this chapter are also found in the paper 'Optimal Placement of Base Stations in Border Surveillance using Limited Capacity Drones' which was completed during the X Spanish Workshop on Geometric Optimization. This was written with the following coauthors: S. Bereg, J.M. Díaz-Báñez, M. Haghpanah, P. Horn, M.A. Lopez, N. Marín, A. Ramírez-Vigueras, F. Rodríguez, O. Solé-Pi, A. Stevens, and J. Urrutia.

BIBLIOGRAPHY

[1] Nadeem Ahmed, Salil S Kanhere, and Sanjay Jha. The holes problem in wireless sensor networks: a survey. *ACM SIGMOBILE Mobile Computing and Communications Review*, 9(2):4–18, 2005.

[2] N. Alon and F.R.K. Chung. Explicit construction of linear sized tolerant networks. *Discrete Mathematics*, 72(1):15–19, 1988.

[3] Bander Alzahrani, Omar Sami Oubbati, Ahmed Barnawi, Mohammed Atiquzzaman, and Daniyal Alghazzawi. Uav assistance paradigm: State-of-the-art in applications and challenges. *Journal of Network and Computer Applications*, 166:102706, 2020.

[4] Esther M Arkin, Joseph SB Mitchell, and Christine D Piatko. Minimum-link watchman tours. *Information Processing Letters*, 86(4):203–207, 2003.

[5] Bakry, Dominique, Émery, and Michel. Diffusions hypercontractives. *Séminaire de probabilités de Strasbourg*, 19:177–206, 1985.

[6] Frank Bauer, Paul Horn, Yong Lin, Gábor Lippner, Dan Mangoubi, and Shing-Tung Yau. Li-yau inequality on graphs. *arXiv: Analysis of PDEs*, 2013.

[7] Binay Bhattacharya, Mike Burmester, Yuzhuang Hu, Evangelos Kranakis, Qiaosheng Shi, and Andreas Wiese. Optimal movement of mobile sensors for barrier coverage of a planar region. *Theoretical Computer Science*, 410(52):5515–5528, 2009.

[8] Prosenjit Bose, Pat Morin, Ivan Stojmenović, and Jorge Urrutia. Routing with guaranteed delivery in ad hoc wireless networks. *Wireless networks*, 7(6):609–616, 2001.

[9] R. C. Bose. Strongly regular graphs, partial geometries and partially balanced designs. *Pacific Journal of Mathematics*, 13:389–419, 1963.

[10] Andries E. Brouwer and H. Van Maldeghem. *Strongly Regular Graphs*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2022.

[11] R. H. Bruck. Finite nets. ii. uniqueness and imbedding. *Pacific Journal of Mathematics*, 13:421–457, 1963.

[12] Bernard Chazelle. Triangulating a simple polygon in linear time. *Discrete & Computational Geometry*, 6(3):485–524, 1991.

[13] Fan Chung, Yong Lin, and Shing-Tung Yau. Harnack inequalities for graphs with non-negative ricci curvature, 2012.

[14] Fan R. K. Chung and Shing-Tung Yau. Logarithmic harnack inequalities. *Mathematical Research Letters*, 3:793–812, 1996.

[15] D. Cushing, Supanat Kamtue, Riikka Kangaslampi, Shiping Liu, and Norbert Peyerimhoff. Curvatures, graph products and ricci flatness. *J. Graph Theory*, 96:522–553, 2021.

[16] David Cushing, Riikka Kangaslampi, Valtteri Lipiäinen, Shiping Liu, and George William Stagg. The graph curvature calculator and the curvatures of cubic graphs, 2017.

[17] David Cushing, Shiping Liu, and Norbert Peyerimhoff. Bakry–Émery curvature functions on graphs. *Canadian Journal of Mathematics*, 72(1):89–143, Jan 2019.

[18] Miroslav Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1973.

[19] Saurabh Ganeriwal, Aman Kansal, and Mani B Srivastava. Self aware actuation for fault repair in sensor networks. In *IEEE International Conference on Robotics and*

Automation, 2004. Proceedings. ICRA'04. 2004, volume 5, pages 5244–5249. IEEE, 2004.

[20] Subir Kumar Ghosh, Anil Maheshwari, Sudebkumar Prasant Pal, Sanjeev Saluja, and CE Veni Madhavan. Characterizing and recognizing weak visibility polygons. *Computational Geometry*, 3(4):213–233, 1993.

[21] Leonidas Guibas, John Hershberger, Daniel Leven, Micha Sharir, and Robert E Tarjan. Linear-time algorithms for visibility and shortest path problems inside triangulated simple polygons. *Algorithmica*, 2(1):209–233, 1987.

[22] Paul Horn, Yong Lin, Shuang Liu, and Shing-Tung Yau. Volume doubling, poincaré inequality and gaussian heat kernel estimate for non-negatively curved graphs. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2019(757):89–130, 2019.

[23] Bobo Hua and Yong Lin. Stochastic completeness for graphs with curvature dimension conditions. *Advances in Mathematics*, 306:279–302, 2017.

[24] Bobo Hua and Florentin Münch. Ricci curvature on birth-death processes. *arXiv preprint arXiv:1712.01494*, 2017.

[25] Matthias Keller and Florentin Münch. Gradient estimates, bakry-emery ricci curvature and ellipticity for unbounded graph laplacians. *arXiv preprint arXiv:1807.10181*, 2018.

[26] Gustav R. Kirchhoff. Ueber die auflösung der gleichungen, auf welche man bei der untersuchung der linearen vertheilung galvanischer ströme geführt wird. *Annalen der Physik*, 148:497–508, 1847.

[27] Bo'az Klartag, Gady Kozma, Peter Ralli, and Prasad Tetali. Discrete curvature and abelian groups, 2015.

[28] Michael Krivelevich and Benny Sudakov. Sparse pseudo-random graphs are hamiltonian. *J. Graph Theory*, 42:17–33, 2003.

[29] Michael Krivelevich and Benny Sudakov. Pseudo-random graphs. In *More sets, graphs and numbers*, pages 199–262. Springer, 2006.

[30] Santosh Kumar, Ten H Lai, and Anish Arora. Barrier coverage with wireless sensors. *Wireless Networks*, 6(13):817–834, 2007.

[31] Sajjad Lakzian and Zachary Mcguirk. A global poincaré inequality on graphs via a conical curvature-dimension condition. *Analysis and Geometry in Metric Spaces*, 6(1):32–47, 2018.

[32] Der-Tsai Lee and Franco P Preparata. Euclidean shortest paths in the presence of rectilinear barriers. *Networks*, 14(3):393–410, 1984.

[33] Miao Li, Lu Zhen, Shuaian Wang, Wenya Lv, and Xiaobo Qu. Unmanned aerial vehicle scheduling problem for traffic monitoring. *Computers & Industrial Engineering*, 122:15–23, 2018.

[34] Yong Lin and Shuang Liu. Equivalent properties of cd inequality on graph. *arXiv preprint arXiv:1512.02677*, 2015.

[35] Shiping Liu, Florentin Münch, and Norbert Peyerimhoff. Rigidity properties of the hypercube via bakry-emery curvature. *arXiv preprint arXiv:1705.06789*, 2017.

[36] Shiping Liu, Florentin Münch, and Norbert Peyerimhoff. Bakry–émery curvature and diameter bounds on graphs. *Calculus of Variations and Partial Differential Equations*, 57(2):1–9, 2018.

[37] Shiping Liu, Florentin Münch, Norbert Peyerimhoff, and Christian Rose. Distance bounds for graphs with some negative bakry-emery curvature. *Analysis and Geometry in Metric Spaces*, 7(1):1–14, 2019.

[38] Shiping Liu, Florentin Münch, and Norbert Peyerimhoff. Bakry-emery curvature and diameter bounds on graphs. *Calculus of Variations and Partial Differential Equations*, 57, 03 2018.

[39] Shiping Liu and Norbert Peyerimhoff. Eigenvalue ratios of non-negatively curved graphs. *Combinatorics, Probability and Computing*, 27(5):829–850, 2018.

[40] Yao Liu, Zhong Liu, Jianmai Shi, Guohua Wu, and Chao Chen. Optimization of base location and patrol routes for unmanned aerial vehicles in border intelligence, surveillance, and reconnaissance. *Journal of Advanced Transportation*, 2019, 2019.

[41] Teng Long, Mustafa Ozger, Oktay Cetinkaya, and Ozgur B Akan. Energy neutral internet of drones. *IEEE Communications Magazine*, 56(1):22–28, 2018.

[42] L. Lovasz and M. Simonovits. The mixing rate of markov chains, an isoperimetric inequality, and computing the volume. In *Proceedings [1990] 31st Annual Symposium on Foundations of Computer Science*, pages 346–354 vol. 1, 1990.

[43] László Miklós Lovász and Miklós Simonovits. Random walks in a convex body and an improved volume algorithm. *Random Struct. Algorithms*, 4:359–412, 1993.

[44] Satyanarayana G Manyam, Steven Rasmussen, David W Casbeer, Krishnamoorthy Kalyanam, and Suresh Manickam. Multi-uav routing for persistent intelligence surveillance & reconnaissance missions. In *2017 international conference on unmanned aircraft systems (ICUAS)*, pages 573–580. IEEE, 2017.

[45] Manton M. Matthews and David P. Sumner. Longest paths and cycles in k1,3-free graphs. *Journal of Graph Theory*, 9(2):269–277, 1985.

[46] Jose M Mazon, Marcos Solera, and Julián Toledo. The heat flow on metric random walk spaces. *Journal of Mathematical Analysis and Applications*, 483(2):123645, 2020.

[47] Avraham A Melkman. On-line construction of the convex hull of a simple polyline. *Information Processing Letters*, 25(1):11–12, 1987.

[48] Florentin Münch. Remarks on curvature dimension conditions on graphs. *Calculus of Variations and Partial Differential Equations*, 56(1):1–8, 2017.

[49] Florentin Münch. Li-yau inequality under $cd(0, n)$ on graphs. *arXiv preprint arXiv:1909.10242*, 2019.

[50] Florentin Münch and Christian Rose. Spectrally positive bakry-emery ricci curvature on graphs. *Journal de Mathématiques Pures et Appliquées*, 143:334–344, 2020.

[51] Arnold Neumaier. Strongly regular graphs with smallest eigenvalue-m. *Archiv der Mathematik*, 33(1):392–400, 1979.

[52] Simeon Ntafos and Laxmi Gewali. External watchman routes. *The visual computer*, 10(8):474–483, 1994.

[53] Joseph O'Rourke. *Art gallery theorems and algorithms*. Oxford University Press Oxford, 1987.

[54] Justin Salez. Sparse expanders have negative ollivier-ricci curvature. *arXiv:2101.08242*, 01 2021.

[55] İnci Sarıçiçek and Yasemin Akkuş. Unmanned aerial vehicle hub-location and routing for monitoring geographic borders. *Applied Mathematical Modelling*, 39(14):3939–3953, 2015.

[56] Adrian Spener, Frederic Weber, and Rico Zacher. Curvature-dimension inequalities for non-local operators in the discrete setting. *Calculus of Variations and Partial Differential Equations*, 58, 09 2019.

[57] Daniel A. Spielman and Shang-Hua Teng. Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems, 2008.

[58] Paul Turán. Eine extremalaufgabe aus der graphentheorie. *Mat. Fiz. Lapok*, 48(436-452):61, 1941.

[59] Jorge Urrutia. Art gallery and illumination problems. In *Handbook of computational geometry*, pages 973–1027. Elsevier, 2000.

[60] Guiling Wang, Guohong Cao, and Thomas F La Porta. Movement-assisted sensor deployment. *IEEE Transactions on Mobile Computing*, 5(6):640–652, 2006.

[61] Ertan Yakıcı. Solving location and routing problem for uavs. *Computers & Industrial Engineering*, 102:294–301, 2016.