Chapter 1
Optimization methods for frame conditioning and application to graph Laplacian scaling

Radu Balan, Mathew Begué, Chae Clark, Kasso Okoudjou

1.1 Introduction

The notion of scalable frame has been investigated in recent years [10, 16, 4, 15], where the focus was more on characterizing frames whose vectors can be rescaled resulting in a tight frame. For completeness, we recall that a set of vectors $F = \{f_i\}_{i=1}^M$ in some (finite dimensional) Hilbert space $\mathscr{H}$ is a frame for $\mathscr{H}$ if there exist two constants $0 < A \le B < \infty$ such that

$$A\|x\|^2 \le \sum_{i=i}^M |\langle x, f_i \rangle|^2 \le B\|x\|^2$$

for all $x \in \mathscr{H}$. When $A = B$ the frame is said to be tight and if in addition, $A = B = 1$ it is termed a Parseval frame. When $F = \{f_i\}_{i=1}^M$ is a frame, we shall abuse notations and denote by F again, the $n \times M$ matrix whose i^{th} column is f_i, and where n is the dimension of $\mathscr{H}$. Using this notation, the frame operator is the $n \times n$ matrix $S = FF^*$ where F^* is the adjoint of F. It is a folklore to note that F is a frame if and only if S is a positive definite operator and the optimal lower frame bound, A, coincides with the lowest eigenvalue of S while the optimal upper frame bound, B, equals the largest eigenvalue of S. We refer to [6, 7, 20] for more details on frame theory.

It is apparent that tight frames are optimal frames in the sense that the condition number of their frame operator is 1. We recall that, the *condition number* of a matrix A, denoted $\kappa(A)$, is defined as the ratio of the largest singular value and the smallest singular value of A, i.e., $\kappa(A) = \sigma_{\max}(A)/\sigma_{\min}(A)$. By analogy, for a frame in a Hilbert space $\{f_i\}_{i=1}^M \subseteq \mathscr{H}$ with optimal frame bounds A and B, we define the condition number of the frame to be the condition number of its associated frame operator $\kappa(\{f_i\}) := \kappa(S) = B/A$. In particular, if a frame is Parseval then its condition number equals 1. In fact, a frame is tight if and only if its condition number is 1. Scalable frames were precisely introduced to turn a non optimal (non-tight)

frame into an optimal one, by just rescaling the length of each frame vector. More precisely,

Definition 1 ([17, Definition 2.1]). A frame $\{f_i\}_{i=1}^M$ in some Hilbert space $\mathscr{H}$ is called a *scalable frame* if there exist nonnegative numbers $s_1, ..., s_M$ such that $\{s_i f_i\}_{i=1}^M$ is a Parseval frame for $\mathscr{H}$.

It follows from the definition that a frame $\{f_i\}_{i=1}^M$ is scalable if and only if there exist scalars $s_i \geq 0$ so that

$$\kappa \left(\sum_{i=1}^M s_i^2 f_i f_i^* \right) = 1.$$

To date various equivalent characterizations of scalable frames have been proved and attempts to measure how close to scalable a non-scalable frame is have been offered [16, 4, 15, 21]. In particular, if a frame is not scalable, then one can naturally measure how "not scalable" the frame is by measuring

$$\min_{s_i \geq 0} \left\| I_n - \sum_{i=1}^M s_i^2 f_i f_i^* \right\|_F, \tag{1.1}$$

as proposed in [8], where $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. Other measures of scalability were also proposed by the same authors. However, it is not clear that, when a frame is not scalable, an optimal solution to (1.1) yields a frame $\{s_i f_i\}$ that is as best conditioned as possible. Recently, the relationship between the solution to this problem and the condition number of a frame has been investigated in [5]. In particular, Casazza and Chen show that the problem of minimizing the condition number of a scaled frame

$$\min_{s_i \geq 0} \kappa \left(\sum_{i=1}^M s_i^2 f_i f_i^* \right), \tag{1.2}$$

is equivalent to solving the minimization problem

$$\min_{s_i \geq 0} \left\| I_n - \sum_{i=1}^M s_i^2 f_i f_i^* \right\|_2, \tag{1.3}$$

where $\|\cdot\|_2$ is the operator norm of a matrix. Specifically they show that any optimizer of (1.2) is also an optimizer of (1.3); vice-versa, any optimizer of (1.3) minimizes the condition number in (1.2). Furthermore, they show that the optimal solution to (1.1) does not even have to be a frame, and so would yield an undefined condition number for the corresponding system.

In this chapter, we consider numerical solutions to the scalability problem. Recall that a frame $F = \{f_i\}_{i=i}^M \subset \mathscr{H}$ is scalable if and only if the exist scalars $\{s_i\}_{i=1}^M \subset [0, \infty)$ such that

$$\sum_{i=1}^M s_i^2 f_i f_i = I.$$

Consequently, the condition number of the scaled frame $\tilde{F} = \{s_i f_i\}_{i=i}^M$ is 1. We are thus interested in investigating the solutions to the following four optimization problems:

$$\min_{s_i \geq 0, s \neq \mathbf{0}} \frac{\lambda_{\max}\left(\sum_{i=1}^M s_i^2 f_i f_i^*\right)}{\lambda_{\min}\left(\sum_{i=1}^M s_i^2 f_i f_i^*\right)}. \tag{1.4}$$

$$\min_{\substack{s_i \geq 0, s \neq \mathbf{0} \\ \sum_{i=1}^M s_i^2 \|f_i\|_2^2 = N}} \lambda_{\max}\left(\sum_{i=1}^M s_i^2 f_i f_i^*\right) - \lambda_{\min}\left(\sum_{i=1}^M s_i^2 f_i f_i^*\right). \tag{1.5}$$

$$\min_{s_i \geq 0, s \neq \mathbf{0}} \left\| I_N - \sum_{i=1}^M s_i^2 f_i f_i^* \right\|_F. \tag{1.6}$$

$$\min_{s_i \geq 0, s \neq \mathbf{0}} \left\{ \lambda_{\max}\left(\sum_{i=1}^M s_i^2 f_i f_i^*\right) \;\middle|\; \lambda_{\min}\left(\sum_{i=1}^M s_i^2 f_i f_i^*\right) \geq 1 \right\}. \tag{1.7}$$

Our motivation stems from the fact it appears from the existing literature on scalable frames that the set of all such frames is relatively small, e.g., see [16]. As a result, one is interested in scaling a frame in an optimal manner. For example, by minimizing the condition number of the scaled frame (1.4), or the gap of the spectrum of the scaled frame (1.5). Furthermore, one can try to find the relationship between the optimal solutions to these two problems with the measures of scalability introduced in [8], of which (1.1) is a typical example.

In addition, we investigate these optimization problems from a practical point of view: the existence of fast algorithms to produce optimal solutions. As such, we are naturally lead to consider these problems in the context of convex optimization. We recall that in such a setting one wants to solve for $s^* = \arg\min_s f(s)$ for a real convex function $f : X \to \mathbb{R} \cup \{\infty\}$ defined on a convex set X. Using the convexity of f and X it follows that:

1. If s^* is a local minimum of f, then it is a global minimum.
2. The set of all (global) minima is convex.
3. If f is a strictly convex function and a minimum exists, then the minimum is unique.

In addition, the convexity of f and X allows the use of convex analysis to produce fast, efficient algorithmic solvers, we refer to [2] and references therein for more details.

The condition number function κ, is not convex. As such, it is nontrivial to find the optimal solution of (1.4). However, κ is a quasiconvex function (see [1, Theorem 13.6] for a proof), meaning that its lower level sets form convex sets; that is, the set $\{X : \kappa(X) \leq a\}$ forms a convex set for any real $a \geq 0$. See [12] and references therein for a survey on some algorithms that can numerically solve certain quasiconvex problems. We refer to [19] for a survey of results on optimizing the condition number. But we note that, while minimizing the condition number κ is not a convex

problem, an equivalent convex problem was considered in [18]. For comparison and completeness we state one of the main results of [18]. First, observe that if X is a symmetric positive semidefinite matrix, then its *condition number* is defined as

$$\kappa(X) = \begin{cases} \lambda_{\max}(X)/\lambda_{\min}(X) & \text{if } \lambda_{\min}(X) > 0, \\ \infty & \text{if } \lambda_{\min}(X) = 0 \text{ and } \lambda_{\max}(X) > 0, \\ 0 & \text{if } X \equiv 0. \end{cases}$$

In this setting, it was proved in [18] that the problem of minimizing the condition number is equivalent to solving another problem with convex programming.

Theorem 1 ([18], Theorem 3.1). *Let $\Omega \subseteq \mathscr{S}^N$ be some nonempty closed convex subset of $\mathscr{S}^N$, the space of $N \times N$ symmetric matrices and let $\mathscr{S}^N_+$ be the space of symmetric positive semidefinite $N \times N$ matrices. Then the problem of solving*

$$\kappa^* = \inf\{\kappa(X) : X \in \mathscr{S}^N_+ \cap \Omega\}$$

is equivalent to the problem of solving

$$\lambda^* = \inf\{\lambda_{\max}(X) : X \in t\Omega, t \geq 0, X \succeq I\}, \tag{1.8}$$

that is, $\lambda^* = \kappa^*$.

Theorem 1 has an intuitive interpretation. Suppose $\kappa(X) = \kappa^*$. Then rescaling X by a positive scalar, t, will also scale its eigenvalues by the same factor $1/t$, thus leaving its condition number, $\kappa(X/t)$, unchanged. Therefore, without loss of generality, we can assume that X is rescaled so that $\lambda_{\min}(X/t) \geq 1$ which is imposed in the last condition of (1.8). Once we know that $\lambda_{\min}(X/t)$ is at least 1 then minimizing the condition number of X/t is equivalent to minimizing $\lambda_{\max}(X/t)$ so long as $X/t \in \Omega$ which is guaranteed by the first condition in (1.8).

   The goal of this chapter is to investigate the relationship between the solution to each of the optimization problems (1.4), (1.5), (1.6), and (1.7). In addition, we shall investigate the behavior of the optimal solution to each of these problems vis-á-vis the projection of a non-scalable frame onto the set of scalable frames. We shall also describe a number of algorithms to solve some of these problems and compare some of the performances of these algorithms. Finally, we shall apply some of the results of frame scalability to the problem of reweighing a graph in a such a way that the condition number of the resulting Laplacian is as small as possible. The chapter is organized as follow. In Section 1.2 we investigate the four problems stated above and compare their solutions, and in Section 1.3 we consider the application to finite graph reweighing.

1.2 Non-scalable frames and optimally conditioned scaled frames

We begin by showing the relationship between the four formulations of this scalability problem. We shall first show the equivalence of these problems when a frame is

exactly scalable, and present toy examples of the different solutions obtained when a frame is only approximately scalable.

Lemma 1. *Let $F = \{f_i\}_{i=1}^M$ be a frame in $\mathbb{R}^N$. Then the following statements are equivalent:*

(a) $F = \{f_i\}_{i=1}^M$ *is a scalable frame.*
(b) *Problem (1.4) has a global minimum solution, $s^* = \{s_i^*\}$, with objective function value 1.*
(c) *Problem (1.5) has a global minimum solution, $s^* = \{s_i^*\}$, with objective function value 0.*
(d) *Problem (1.6) has a global minimum solution, $s^* = \{s_i^*\}$, with objective function value 0.*
(e) *Problem (1.7) has a global minimum solution, $s^* = \{s_i^*\}$, with objective function value 1.*

Proof. Assume F is scalable with weights, $\{s_i\}_{i=1}^M$. Then $\widetilde{S} = \sum_{i=1}^M s_i^2 f_i f_i^* = I_N$, and the largest and smallest eigenvalue of the scaled frame operator is 1,

$$\frac{\lambda_{\max}\left(\sum_{i=1}^M s_i^2 f_i f_i^*\right)}{\lambda_{\min}\left(\sum_{i=1}^M s_i^2 f_i f_i^*\right)} = \frac{\lambda_{\max}\left(\widetilde{S}\right)}{\lambda_{\min}\left(\widetilde{S}\right)} = 1.$$

Assume problem (1.4) has a global minimum solution, $\{s_i\}_{i=1}^M$. As, $\lambda_{\max} \geq \lambda_{\min}$, the feasible solution must result in $\lambda_{\max} = \lambda_{\min} = A$. Applying this feasible solution as a scaling of F, we have,

$$\widetilde{S} = \sum_{i=1}^M s_i^2 f_i f_i^* = A I_N.$$

By normalizing the feasible solution by the square-root of A, we have the Parseval scaling,

$$\{\tilde{s}_i\}_{i=1}^M = \left\{\frac{1}{\sqrt{A}} s_i\right\}_{i=1}^M.$$

We have just proved that (a) and (b) are equivalent.

Assume F is scalable with weights, $\{s_i\}_{i=1}^M$. Then $\widetilde{S} = \sum_{i=1}^M s_i^2 f_i f_i^* = I_N$, and the difference between the largest and smallest eigenvalue of the scaled frame operator is 0,

$$\lambda_{\max}\left(\sum_{i=1}^M s_i^2 f_i f_i^*\right) - \lambda_{\min}\left(\sum_{i=1}^M s_i^2 f_i f_i^*\right) = \lambda_{\max}\left(\widetilde{S}\right) - \lambda_{\min}\left(\widetilde{S}\right) = 0.$$

Additionally $N = tr(I_N) = \sum_{i=1}^M s_i^2 \|f_i\|_2^2$ which shows that $\{s_i\}_{i=1}^M$ is a feasible solution for (1.5).

Assume problem (1.5) has a global minimum solution, $\{s_i\}_{i=1}^M$. As, $\lambda_{\max} \geq \lambda_{\min}$, the feasible solution must result in $\lambda_{\max} = \lambda_{\min} = A$. Applying this feasible solution

as a scaling of F, we have,

$$\widetilde{S} = \sum_{i=1}^{M} s_i^2 f_i f_i^* = A I_N.$$

But the feasibility condition $\sum_{i=1}^{M} s_i^2 \|f_i\|_2^2 = N$ implies $N = tr(A I_N)$, hence $A = 1$. We have just proved that (a) and (c) are equivalent.

Assume F is scalable with weights, $\{s_i\}_{i=1}^{M}$. Then $\widetilde{S} = \sum_{i=1}^{M} s_i^2 f_i f_i^* = I_N$, and the objective function for (1.6) attains the global minimum ,

$$\left\| I_N - \sum_{i=1}^{M} s_i^2 f_i f_i^* \right\|_F = \|I_N - I_N\|_F = 0.$$

Assume problem (1.6) has a global minimum solution, $\{s_i\}_{i=1}^{M}$, which occurs when $\left\| I_N - \sum_{i=1}^{M} s_i^2 f_i f_i^* \right\|_F = 0$. This implies that $\widetilde{S} = \sum_{i=1}^{M} s_i^2 f_i f_i^* = I_N$, and we have a Parseval scaling. We have just proved that (a) and (d) are equivalent.

Assume F is scalable with weights, $\{s_i\}_{i=1}^{M}$. Then $\widetilde{S} = \sum_{i=1}^{M} s_i^2 f_i f_i^* = I_N$, and the largest and smallest eigenvalue of the scaled frame operator is 1,

$$\lambda_{\max}\left(\sum_{i=1}^{M} s_i^2 f_i f_i^* \right) = 1.$$

The objective function cannot take on a lower value, as the constraint, $\lambda_{\min}(\sum_{i=1}^{M} s_i^2 f_i f_i^*) \geq 1$,

Assume problem (1.7) has a global minimum solution, $\{s_i\}_{i=1}^{M}$, which occurs when $\lambda_{\max}\left(\sum_{i=1}^{M} s_i^2 f_i f_i^* \right) = \lambda_{\min}\left(\sum_{i=1}^{M} s_i^2 f_i f_i^* \right) = 1$. This implies that $\widetilde{S} = \sum_{i=1}^{M} s_i^2 f_i f_i^* = I_N$, and the solution is also a Parseval scaling. We have just proved that (a) and (e) are equivalent.

Remark 1. Lemma 1 asserts that the problem of finding optimal scalings, $\{s_i\}_{i=1}^{M}$, for a given scalable frame $F = \{f_i\}_{i=1}^{M}$ is equivalent to finding the absolute minimums of the following optimization problems:

- $\min_{s_i \geq 0, s \neq \mathbf{0}} \dfrac{\lambda_{\max}\left(\sum_{i=1}^{M} s_i^2 f_i f_i^* \right)}{\lambda_{\min}\left(\sum_{i=1}^{M} s_i^2 f_i f_i^* \right)}$

- $\min_{\substack{s_i \geq 0, s \neq \mathbf{0} \\ \sum_{i=1}^{M} s_i^2 \|f_1\|_2^2 = N}} \lambda_{\max}\left(\sum_{i=1}^{M} s_i^2 f_i f_i^* \right) - \lambda_{\min}\left(\sum_{i=1}^{M} s_i^2 f_i f_i^* \right)$

- $\min_{s_i \geq 0, s \neq \mathbf{0}} \left\| I_N - \sum_{i=1}^{M} s_i^2 f_i f_i^* \right\|_F$

- $\min_{s_i \geq 0, s \neq \mathbf{0}} \left\{ \lambda_{\max}\left(\sum_{i=1}^{M} s_i^2 f_i f_i^* \right) \mid \lambda_{\min}\left(\sum_{i=1}^{M} s_i^2 f_i f_i^* \right) \geq 1 \right\}$

Remark 1 is restrictive in that it requires the frame $F = \{f_i\}_{i=1}^{M}$ be scalable to state equivalence among problems, but there can be a wide variance in the solutions obtained when the frame is not scalable. Even nearly-tight frames vary in initial feasible solutions. We briefly consider ε-tight frames and analyze the distance from the minimum possible objective function value.

Let $F_\varepsilon = \{g_i\}_{i=1}^M$ with $\|g_i\|_2 = 1$ for all i be an ε-tight frame such that,

$$(1-\varepsilon)I_N \preceq \sum_{i=1}^M g_i g_i^* \preceq (1+\varepsilon)I_N.$$

First considering the case in which the frame cannot be conditioned any further, so the optimal scaling weights are $s_i = 1$. Analyzing the solution produced by the four optimization methods, we see the difference in solutions produced.

$$\frac{\lambda_{\max}\left(\sum_{i=1}^M s_i^2 g_i g_i^*\right)}{\lambda_{\min}\left(\sum_{i=1}^M s_i^2 g_i g_i^*\right)} = \frac{\lambda_{\max}\left(\sum_{i=1}^M g_i g_i^*\right)}{\lambda_{\min}\left(\sum_{i=1}^M g_i g_i^*\right)} = \frac{1+\varepsilon}{1-\varepsilon} = 1 + \frac{2\varepsilon}{1-\varepsilon}.$$

$$\lambda_{\max}\left(\sum_{i=1}^M s_i^2 f_i f_i^*\right) - \lambda_{\min}\left(\sum_{i=1}^M s_i^2 g_i g_i^*\right) = (1+\varepsilon) - (1-\varepsilon) = 2\varepsilon$$

$$\lambda_{\max}\left(\sum_{i=1}^M s_i^2 g_i g_i^*\right) = \lambda_{\max}\left(\sum_{i=1}^M g_i g_i^*\right) = 1 + \varepsilon.$$

We lack the information necessary to give exact results for formulation (1.6), so we instead give an upper bound when $s_i = 1$.

$$\left\| I_N - \sum_{i=1}^M s_i^2 g_i g_i^* \right\|_F = \left\| I_N - \sum_{i=1}^M g_i g_i^* \right\|_F$$

$$\leq \sqrt{N} \left\| I_N - \sum_{i=1}^M g_i g_i^* \right\|_2$$

$$\leq \varepsilon\sqrt{N}.$$

It makes sense that we could enforce this constraint, as we could re-normalize the frame elements by the reciprocal of the smallest eigenvalue of the frame operator. It is not true, though, that the scalings produced must be the same. Moreover, when not using the constraint on the smallest eigenvalue, the scalings can vary wildly.

Remark 2. For general frames, the optimization problems (1.4)-(1.7) do not produce tight frames. However they can be solved using special classes of convex optimization algorithms: problems (1.4),(1.5),(1.7) are solved by *Semi-Definite Programs* (SDP), whereas problem (1.6) is solved by a Quadratic Program (QP) – see [2] for details on SDPs and QPs. In the following we state these SDPs explicitly.

SDP 1 – Operator Norm Optimization:

$$(t^1, s^{(1)}) = argmin_{\substack{t, s_1, \ldots, s_M \geq 0 \\ \sum_{i=1}^M s_i^2 f_i f_i^* - tI_N - I_N \leq 0 \\ \sum_{i=1}^M s_i^2 f_i f_i^* + tI_N - I_N \geq 0}} t \qquad (1.9)$$

This SDP implements the optimization problem (1.3). In turn, as showed by Cassaza and Chen in [5], the solution to this problem is also an optimizer of the condition

number optimization problem (1.4). Conversely, assume $s^{(*)}$ is a solution of (1.4). Let $A = \lambda_{min}(\sum_{i=1}^{M} s_i^2 f_i f_i^*)$ and $B = \lambda_{max}(\sum_{i=1}^{M} s_i^2 f_i f_i^*)$. Let $r = \frac{2}{A+B}$. Then $s^{(*)} = (rs_i^2)_{i=1}^{M}$ is a solution of (1.9) and the optimum value of the optimization criterion is $t^1 = rB - 1 = 1 - rA$.

SDP 2 – Minimum Upper Frame Bound Optimization:

$$(t^2, s^{(2)}) = argmin_{\substack{t, s_1, \ldots, s_M \geq 0 \\ \sum_{i=1}^{M} s_i^2 f_i f_i^* - I_N \geq 0 \\ \sum_{i=1}^{M} s_i^2 f_i f_i^* - tI_N \leq 0}} t \qquad (1.10)$$

This SDP implements the optimization problem (1.7). In turn, as we showed before, the solution $s^{(2)}$ is also a solution of (1.4). Conversely, assume $s^{(*)}$ is a solution of (1.4). Let $A = \lambda_{min}(\sum_{i=1}^{M} s_i^2 f_i f_i^*)$ and $B = \lambda_{max}(\sum_{i=1}^{M} s_i^2 f_i f_i^*)$. Let $r = \frac{1}{A}$. Then $s^{(*)} = (rs_i^2)_{i=1}^{M}$ is a solution of (1.10), and the optimum value of the optimization criterion is $t^2 = \frac{B}{A}$.

SDP 3 – Spectral Gap Optimization:

$$(t^3, v^3, s^{(3)}) = argmin_{\substack{t, v, s_1, \ldots, s_M \geq 0 \\ \sum_{i=1}^{M} s_i^2 f_i f_i^* - tI_N \leq 0 \\ \sum_{i=1}^{M} s_i^2 f_i f_i^* - vI_N \geq 0 \\ \sum_{i=1}^{M} s_i \|f_i\|_2^2 = N}} t - v \qquad (1.11)$$

This SDP implements the optimization problem (1.5). As remarked earlier (1.5) is not equivalent to any of (1.3),(1.4) or (1.7). A spectral interpretation of these optimization problems is as follows. The SDP 1 (and implicitly (1.4) and (1.7)) scales the frame so that the largest and smallest eigenvalues of the scaled frame operator are equidistant and closest to value 1. The SDP 3 scales the frame so that the largest and smallest eigenvalues of the scaled frame operator are closest to one another while the average eigenvalue is set to 1. Equivalently, the solution to SDP 3 also minimizes the following criterion:

$$\frac{\lambda_{max}(\tilde{S}) - \lambda_{min}(\tilde{S})}{\frac{1}{N}tr(\tilde{S})}$$

where $\tilde{S} = \sum_{i=1}^{M} s_i^2 f_i f_i^*$ is the scaled frame operator.

QP 4 – Frobenius Norm Optimization:

$$s^{(4)} = argmin_{s_1, \ldots, s_M \geq 0} \sum_{i,j=1}^{M} s_i s_j |\langle f_i, f_j \rangle|^2 - 2\sum_{i=1}^{M} s_i^2 \|f_i\|_2^2 + N \qquad (1.12)$$

This QP implements the optimization problem (1.6).

Example 1. Consider the 5-element frame, $X \subseteq \mathbb{R}^3$, generated such that each coordinate is a random integer from 0 to 5.

$$X = \begin{bmatrix} 2 & 4 & 1 & 4 & 4 \\ 3 & 1 & 2 & 0 & 2 \\ 1 & 4 & 3 & 5 & 2 \end{bmatrix}$$

We then numerically compute X_κ, X_g, X_F, which are the rescaled frames that minimize problems SDP 1, SDP 3 and QP 4, respectively. That is, X_κ is the rescaled frame, $X_\kappa = \{s_i f_i\}$, such that $s^* = \{s_i\}$ is the minimizer to Problem (1.3), which also minimizes the frame condition number, κ. Similarly, X_g is rescaled to minimize the eigenvalue gap $\lambda_{\max} - \lambda_{\min}$ while the average eigenvalue is 1, and X_F is rescaled to minimize Frobenius distance to the identity matrix.

In our numerical implementation minimizing condition number, we used the CVX toolbox in MATLAB [11] which is a solver for convex optimization problems.

Let s_κ, s_g, and s_F denote the scaling vectors that determine the frames X_κ, X_g, and X_F, respectively. That is, $X_\kappa = S_\kappa^{1/2} X$ where S_κ is the diagonal matrix with values given by s_κ, and so on. We obtained scalings

$$s_\kappa = [0.0187, 0, 0.0591, 0.0122, 0.0242],$$
$$s_g = [0.0875, 0, 0.0398, 0.0297, 0],$$
$$s_F = [0.0520, 0, 0.0066, 0.0177, 0].$$

The results comparing each of the four frames are summarized in Table 1.1.

	$\lambda_{\min}$	$\lambda_{\max}$	κ	$(\lambda_{\max} - \lambda_{\min})/\frac{1}{N}\sum_{i=1}^N \lambda_i$	$\|I_3 - \cdot\|_F$	$\|I_3 - \cdot\|_2$
X	4.1658	110.41	26.504	2.5296	109.95	109.41
X_κ	0.1716	1.8284	**10.655**	2.2888	1.4348	**0.8284**
X_g	0.0856	2.3558	27.501	**2.2701**	1.6938	1.3558
X_F	0.01672	1.1989	71.667	2.2903	**1.2048**	0.9832

Table 1.1: Comparisons of extreme eigenvalues, condition number, relative spectral gap, Frobenius distance to identity, and the operator norm distance to identity for the non scalable frame X and its rescaled versions that minimize Problems (1.4)–(1.6).

Observe that each of the three methods can produce widely-varying spectra.

We now demonstrate special conditions in which a frame's condition number can be decreased using matrix perturbation theory.

Lemma 2 (Weyl's Inequality, [23, Corollary 4.9]). *Let A be a Hermitian matrix with real eigenvalues $\{\lambda_i(A)\}_{i=1}^d$ and let B be a Hermitian matrix of the same size as A with eigenvalues $\{\lambda_i(B)\}_{i=1}^d$. Then for any $i = 1,...,d$ we have*

$$\lambda_i(A+B) \in [\lambda_i(A) + \lambda_1(B), \lambda_i(A) + \lambda_d(B)].$$

An immediate corollary of Weyl's inequality tells us that perturbing a matrix by a positive semidefinite matrix will cause the eigenvalues to not decrease.

Corollary 1. *Let A be a Hermitian matrix with real eigenvalues $\{\lambda_i(A)\}_{i=1}^d$ and let $B \succeq 0$ be Hermitian and of the same size of A. Then for any $i = 1,...,d$, we have $\lambda_i(A) \leq \lambda_i(A+B)$. The inequality is strict if $B \succ 0$ is positive definite.*

Lemma 3. *Let f be an eigenvector of A with associated eigenvalue λ. Let B be a matrix of the same size as A with the property that $Bf = 0$. Then f is an eigenvector of $A+B$ with eigenvalue λ.*

Lemma 4 ([24, Section 1.3]). *Let A and B be two $N \times N$ Hermitian matrices of same size. Then for any $i = 1,...,N$, the mapping $t \mapsto \lambda_i(A+tB)$ is Lipschitz continuous with Lipschitz constant $\|B\|_2$.*

Corollary 2. *Let A be an $N \times N$ Hermitian matrix with simple spectrum and minimum eigengap $\delta > 0$, i.e.,*

$$\delta = \min_{i \neq j} |\lambda_i - \lambda_j|.$$

Let B be a non-negative Hermitian matrix of same size as A. Then the mappings $t \mapsto \lambda_i(A+tB)$ are interlacing:

$$\lambda_1(A) \leq \lambda_1(A+tB) \leq \lambda_2(A) \leq \lambda_2(A+tB) \leq \cdots \leq \lambda_{N-1}(A+tB) \leq \lambda_N(A) \leq \lambda_N(A+tB)$$

for $t \in (0, \frac{\delta}{\|B\|_2})$.

The following theorem gives conditions in which we can guarantee that the condition number of frame can be reduced.

Theorem 2. *Let $F = \{f_i\}_{i=1}^m \subseteq \mathbb{C}^d$ be a frame that is not tight and whose frame operator has simple spectrum with minimal eigengap $\delta > 0$. Suppose that there exists some index k such that f_k is orthogonal to the eigenspace corresponding to $\lambda_{\max}(FF^*)$ and not orthogonal to the eigenspace corresponding to $\lambda_{\min}(FF^*)$. Then there exists a rescaled frame $\tilde{F} = \{s_i f_i\}_{i=1}^m$ satisfying $\kappa(\tilde{F}) < \kappa(F)$. In particular, one scaling that decreases the condition number is*

$$s_i = \begin{cases} \frac{m}{m-1+\sqrt{1+\gamma}}, & \text{for } i \neq k \\ \frac{m\sqrt{1+\gamma}}{m-1+\sqrt{1+\gamma}}, & \text{for } i = k \end{cases}$$

for $\gamma \in (0, \delta \|f_k\|^{-2})$.

Proof. Let f_k denote the frame element as described in the assumptions in the statement of the theorem. For $\gamma \in (0, \delta)$, consider the frame operator $HH^* = FF^* + \gamma f_k f_k^*$ which corresponds to the rescaled frame of F where each scale $s_i = 1$ except for $s_k = \sqrt{1+\gamma}$. The matrix $f_k f_k^*$ is Hermitian and positive semidefinite so by Corollary 1, we have $\lambda_i(FF^*) \leq \lambda_i(HH^*)$ for every $i = 1,...,N$. Then by Corollary 2, the eigenvalues of the frame operator HH^* satisfies the following interlacing property:

$$\lambda_1(FF^*) \leq \lambda_1(HH^*) \leq \lambda_2(FF^*) \leq \lambda_2(HH^*) \leq \cdots \leq \lambda_N(FF^*) = \lambda_N(HH^*),$$

where the last equality follows from Lemma 3 and the fact that f_k is orthogonal to the eigenspace corresponding to $\lambda_N(FF^*)$.

We can now compute

$$\kappa(FF^*) = \frac{\lambda_N(FF^*)}{\lambda_1(FF^*)} \geq \frac{\lambda_N(HH^*)}{\lambda_1(HH^*)} = \kappa(HH^*).$$

Finally, we renormalize the scales $\{s_i\}$ by the constant factor $m(m-1+\sqrt{1+\gamma})^{-1}$ to preserve the property that $\sum_{i=1}^m s_i = m$. This renormalization scales all eigenvalues by the same factor which leaves the condition number unchanged. The frame

$$\tilde{F} = \frac{m}{m-1+\sqrt{1+\gamma}}H$$

is the frame described in the statement of the theorem, which concludes the proof.

Remark 3. Having discussed the equivalence between the formulations above, we have seen that they do not necessarily produce similar solutions. This brings the question of which formulation we should use in general, to the forefront. One could answer this question by seeking a metric that best describes the distance of a frame to the set of tight frames. This is similar to the Paulsen problem [3], in that, after we have solved one of the formulations above, we produce a scaling and subsequent new frame and wish to determine the distance of this new frame to the canonical Parseval frame associated to our original frame. In [8], the question of distance to Parseval frames was generalized to include frames that could be made tight with a diagonal scaling, resulting in the distance between a frame and the set of scalable frames:

$$d_F = \min_{\Psi \in \mathscr{SC}(M,N)} \|F - \Psi\|_F. \tag{1.13}$$

However, due to the fact that the topology of the set of scalable frames $\mathscr{SC}(M,N)$ is not yet well-understood, computing d_F is almost impossible for a non-scalable frame. A source of future work involves finding bound on d_F using the optimal solutions to the four problems we stated above to analyze and produce bounds on the minimum distance.

1.3 Minimizing condition number of graphs

In this section we outline how to apply and generalize the problems the optimization problems from Section 1.2 in the setting of (finite) graph Laplacians. This task is not a simply as directly applying the condition number minimization problem (1.4), and the others, with graph Laplacian operators.

Recall that any finite graph has a corresponding positive semidefinite Laplacian matrix with eigenvalues $\{\lambda_k\}_{k=0}^{N-1}$ and eigenvectors $\{f_k\}_{k=0}^{N-1}$. Further any graph has smallest eigenvalue $\lambda = 0$ with multiplicity equal to number of connected components in the graph with eigenvalues equal to constant functions supported on those

connected components. Because any Laplacian's smallest eigenvalue equals 0, its
condition number $\kappa(L)$ is undefined. For simplicity, let us assume that all graphs
in this section are connected and hence $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{N-1}$. Suppose
we restricted the Laplacian operator to the $(N-1)$-dimensional space spanned by
the eigenvectors $f_1, ..., f_{N-1}$. Then this new operator, call it L_0, has eigenvalues
$\lambda_1, ..., \lambda_N$ which are all strictly positive. Now, $\kappa(L_0)$, the condition number of L_0
is a well-defined number.

Recall that the complete graph on N vertices, K_N, is the most connected a graph
on N vertices can be since one can traverse from any two vertices on precisely one
edge. It is the only graph that has all nonzero eigenvalues equal, i.e., $\lambda_0 = 0$ and
$\lambda_1 = \lambda_2 = \cdots = \lambda_{N-1} = N-1$. This graph achieves the highest possible algebraic
connectivity, λ_1, of a graph on N vertices. If we create L_0 by projecting the Laplacian
of K_N onto the $N-1$-dimensional space spanned by the eigenvectors correspond-
ing with nonzero eigenvalue then L_0 equals $N I_{N-1}$, that is a the $(N-1) \times (N-1)$
identity matrix times N.

Lemma 5. *Let G be a connected graph with eigenvalues $\{\lambda_k\}_{k=0}^{N-1}$ and eigenvectors
$\{f_k\}_{k=0}^{N-1}$ of the graph Laplacian L. Let $\tilde{F} = [f_1 \, f_2 \, \cdots \, f_{N-1}]$ be the $N \times (N-1)$ matrix
of eigenvectors excluding the constant vector f_0. Then the $(N-1) \times (N-1)$ matrix*

$$L_0 = \tilde{F}^* L \tilde{F} \tag{1.14}$$

has eigenvalues $\{\lambda_k\}_{k=1}^{N-1}$ and associated orthonormal eigenvectors $\{\tilde{F}^* f_k\}_{k=1}^{N-1}$.

Proof. We first show that $\{\tilde{F}^* f_k\}_{k=1}^{N-1}$ are eigenvectors to L_0 with eigenvalues λ_k.
For any $k = 1, ..., N-1$ we have

$$L_0 \tilde{F}^* f_k = \tilde{F}^* L \tilde{F} \tilde{F}^* f_k.$$

But since $\tilde{F}$ is an orthonormal basis for the eigenspace that its vectors span,
then $\tilde{F}\tilde{F}^*$ is simply the orthogonal projection onto the eigenspace spanned by
$\{f_1, ..., f_{N-1}\}$. That is, for any vector f, we have $\tilde{F}\tilde{F}^* f = f - \langle f, f_0 \rangle f_0$, which is
simply the function f minus its mean value. For each $k = 1, ..., N-1$, the eigenvec-
tors f_k have zero mean, i.e., $\langle f_k, f_0 \rangle = 0$. Hence $\tilde{F}\tilde{F}^* f_k = f_k$ and therefore

$$L_0 \tilde{F}^* f_k = \tilde{F}^* L f_k = \tilde{F}^* (\lambda_k f_k) = \lambda_k \tilde{F}^* f_k.$$

The orthonormality of the eigenvectors $\{\tilde{F}^* f_k\}_{k=1}^{N-1}$ follows directly from the or-
thonormality of $\{f_k\}_{k=0}^{N-1}$ and the computation

$$\langle \tilde{F}^* f_k, \tilde{F}^* f_j \rangle = (\tilde{F}^* f_k)^* \tilde{F}^* f_j = f_k^* \tilde{F}\tilde{F}^* f_j = f_k^* f_j = \delta(k, j).$$

Unlike the Laplacian, the operator in (1.14) is full rank and its rank equals the rank
of the Laplacian. We denote it L_0 because it behaves as the Laplacian after the
projection of the function onto the zero'th eigenspace is removed.

For a general finite graph, the Laplacian can be written as the sum of rank-one
matrices $L = \sum_{i=1}^{m} v_i v_i^*$ where v_i is the i'th column in the incidence matrix B asso-

ciated to the i'th edge in the graph and m is the total number of edges in the graph. Thus, the Laplacian can be formed by the product $L = BB^*$. The columns of the incidence matrix, B, as vectors in $\mathbb{R}^N$ do not form a frame; B has rank $N-1$. However, the restriction B to the $(N-1)$-dimensional space spanned by $f_1, ..., f_{N-1}$, call it B_0, is a frame in that space. Then and the methods of Section 1.2 do apply to the frame B_0 with corresponding frame operator $L_0 = B_0 B_0^*$. Therefore the operator L_0 can also be written as one matrix multiplication $L_0 = (\tilde{F}^* B)(\tilde{F}^* B)^*$. For other related results on graphs and frames we refer to [22]. We seek scalars $s_i \geq 1$ so that the rescaled frame $\{s_i \tilde{F}^* v_i\}_{i=1}^m$ is tight or as close to tight as possible. In terms of matrices, we seek a nonnegative diagonal matrix $X = \mathrm{diag}(s_i)$ so that $\tilde{L}_0 := \tilde{F}^* B X^2 B^* \tilde{F}$ has minimal condition number. The resulting graph Laplacian, denoted $\tilde{L}_\kappa = BX^2 B^*$, is the operator with minimal condition number, $\tilde{L}_0$, without the projection onto $(N-1)$ eigenspaces, thus acting on the entire N-dimensional space. One can interpret this problem as rescaling weights of graph edges to not only make $\tilde{L}_0$ as close as possible to the $(N-1)$-identity matrix but also make the $N \times N$ Laplacian, $\tilde{L}$, as close to possible as the Laplacian of the complete graph K_N.

We present the pseudocode for the algorithm, `GraphCondition`, that produces $\tilde{L}_\kappa$, the Laplacian of the graph that minimizes the condition number of L.

L_κ=`GraphCondition`(L, F, B)
where L is the Laplacian matrix of the graph G,
F is the $N \times N$ eigenvector matrix of L
B is the incidence matrix of L.

1. Set $\tilde{F} = F(:, 2:N)$.
2. Use `cvx` to solve for X that minimizes $\lambda_{\max}(\tilde{F}^* B X^2 B^* \tilde{F})$.
   subject to: $X \succeq 0$ is diagonal, $\mathrm{trace}(X) \geq t \geq 0$, and $\tilde{F}^* B X^2 B^* \tilde{F} \succeq I$.
3. Create $L_\kappa = BX^2 B^*$.

Example 2. We consider the barbell graph G which consists of two complete graphs on 5 vertices that are connected by exactly one edge. The Laplacian for G has eigenvalues $\lambda_1 \approx 0.2984$ and $\lambda_9 \approx 6.7016$, thus giving a condition number of $\kappa(G) \approx 22.45$. We rescale the edges via the `GraphCondition` algoritihm and obtained a rescaled weighted graph $\tilde{G}_\kappa$ which has eigenvalues $\lambda_1 \approx 0.3900$ and $\lambda_{10} \approx 6.991$, thus giving a condition number $\kappa(\tilde{G}_\kappa) \approx 17.9443$.

Both graphs, G and $\tilde{G}_\kappa$, are shown in Figure 1.1. The edge bridging the two complete clusters is assigned the highest weight of 1.8473. All other edges eminating from those two vertices are assigned the smallest weights of 0.7389. All other edges not connected to either of the two "bridge" vertices are assigned a weight of 1.1019.

We show in the following example that the scaling coefficients $\{s_i\}_{i=1}^m$ that minimize the condition number of a graph are not necessarily unique.

Example 3. Consider the graph G complete graph on four nodes with the edge $(3,4)$ removed. Then G was rescaled and conditioned via `GraphCondition`; both
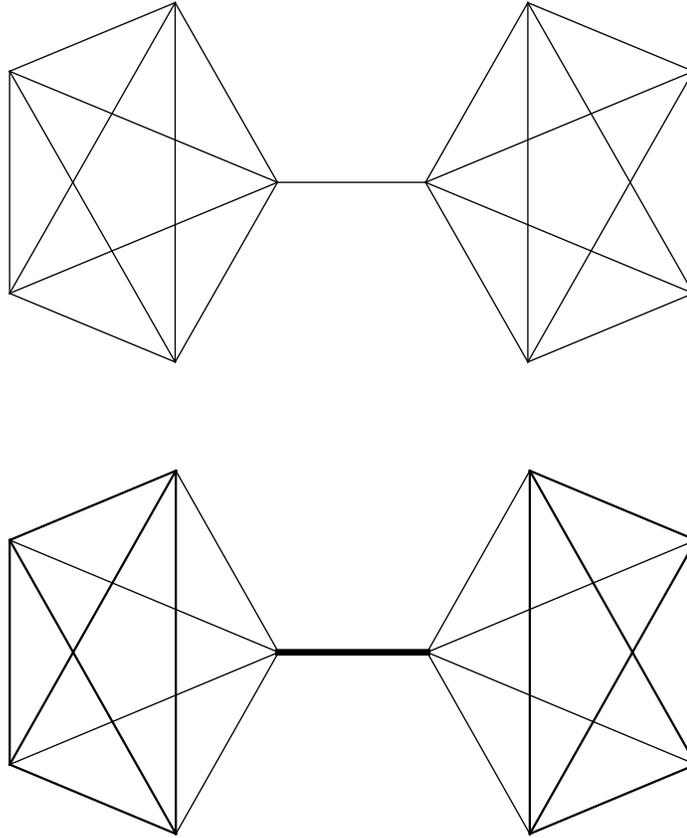
Fig. 1.1: Top: The barbell graph G. Bottom: The condiitoned graph with rescaled weights that minimizes the condition number. The width of the edges are drawn to be proportional to the weight assigned to that edge.

graphs are shown in Figure 1.2. The orignal Laplacian, L, and the rescaled conditioned Laplacian, $\tilde{L}_\kappa$, produced by the `GraphCondition` algorithm are given as

$$L = \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 2 & 0 \\ -1 & -1 & 0 & 2 \end{bmatrix}, \quad \tilde{L}_\kappa \approx \begin{bmatrix} 2.8406 & -0.6812 & -1.0797 & -1.0797 \\ -0.6812 & 2.8406 & -1.0797 & -1.0797 \\ -1.0797 & -1.0797 & 2.1594 & 0 \\ -1.0797 & -1.0797 & 0 & 2.1594 \end{bmatrix},$$

with spectra

$$\sigma(L) = \{0, 2, 4, 4\}, \quad \sigma(\tilde{L}_\kappa) = \{0, 2.1594, 3.5218, 4.3188\}.$$

Both Laplacians have a condition number $\kappa(L) = \kappa(\tilde{L}_\kappa) = 2$ which shows that the scaling of edges that minimize condition number are not necessarily unique.
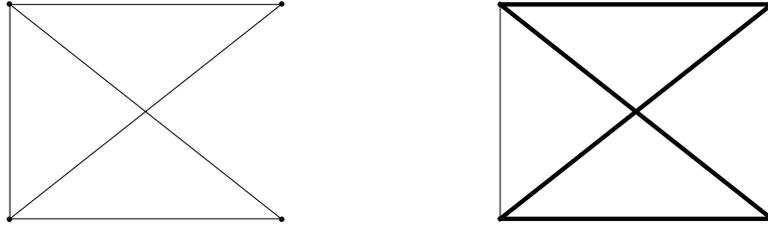

Fig. 1.2: The unweighted graph G (left) and its rescaled version $\tilde{G}_\kappa$ (right) yet both graphs have a condition number equal to 2.

We prove that the `GraphCondition` algorithm will not disconnect a connected graph.

Proposition 1. *Let $G = G(V, E, \omega)$ be a connected graph and let $\tilde{G}_\kappa = \tilde{G}_\kappa(V, \tilde{E}, \tilde{\omega})$ be the rescaled version of G that minimizes graph condition number. Then $\tilde{G}_\kappa$ is also a connected graph.*

Proof. Let $\kappa_0 := \kappa(G) \geq 1$ and suppose that $\tilde{G}_\kappa$ is disconnected. This implies that $\tilde{G}_\kappa$ has eigenvalue 0 with multiplicity at least 2 (one for each of its connected components). This violates the condition $\tilde{F}^* B X^2 B^* \tilde{F} \succeq I$ in the `GraphCondition` algorithm, which yields the unique minimizer.

We next consider the analogue of minimizing the spectral gap, $\lambda_{N-1} - \lambda_1$, for graphs. Just as before with condition number, we create the positive definite matrix L_0 and its incidence matrix, B_0, and minimize its spectral gap by the methods in Section 1.2 to minimize problem (1.5). We denote the rescaled graph that minimizes the spectral gap by $\tilde{G}_g$.

Example 4. We present numerical results of each of the graph rescaling techniques for the barbell graph shown in Figure 1.1. Each of the rescaled graphs are pictured in Figure 1.3 and numerical data is summarized in Table 1.2.

As discussed in the motivation of this section, reducing the condition number of a graph makes the graph more "complete", that is, more like the complete graph in
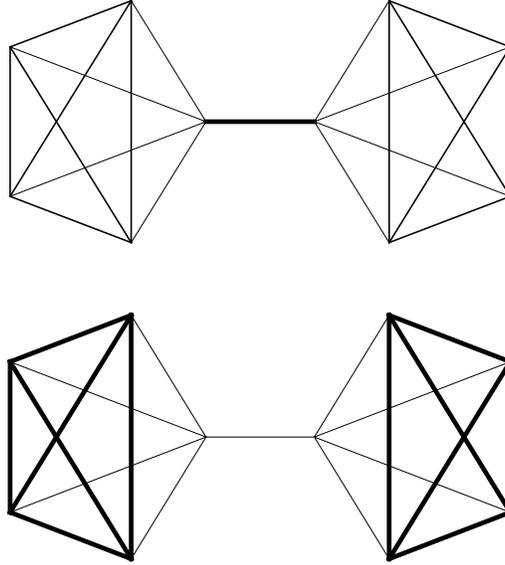
Fig. 1.3: From top to bottom: $\tilde{G}_\kappa$ and $\tilde{G}_g$, which minimize the condition number and spectral gap, respectively.

	λ_1	λ_{N-1}	κ	$\lambda_{N-1} - \lambda_1$
G	0.2984	6.7016	22.4555	6.4031
$\tilde{G}_\kappa$	1.0000	17.9443	**17.9443**	16.9443
$\tilde{G}_g$	0.0504	1.1542	22.8794	**1.1038**

Table 1.2: Comparison of condition number and spectral gap of the barbell graph, G, shown in Figure 1.1 and its rescaled versions, respectively.

terms of its spectrum. Since the algebraic connectivity λ_1 is as great as possible, it is the only graph for which $\lambda_1 = \lambda_{N-1}$, the graph is the most connected a graph can possibly be, and as such the distance between any two points is minimal. As previously discussed, the effective resistance is a natrual metric on graphs and one can compute that for any two distinct vertices, i and j, on the complete graph on N vertices we have

$$R(i,j) = \sum_{k=1}^{N-1} \frac{1}{\lambda_k}(f_k(i) - f_k(j))^2 = \frac{1}{N}\sum_{k=1}^{N-1}(f_k(i) - f_k(j))^2$$

$$= \frac{1}{N}(e_i - e_j)^* F F^*(e_i - e_j) = \frac{1}{N}(e_i - e_j)^*(e_i - e_j)$$

$$= \frac{1}{N}\|e_i - e_j\|^2 = \frac{2}{N}.$$

Conjecture 1. The process of conditioning a graph reduces the average resistance between any two vertices on the graph.

The intuition behind Conjecture 1 can be motivated by studying the quantity $\sum_{k=1}^{N-1} 1/\lambda_k$. Consider a sequence of positive numbers $\{a_k\}_{k=1}^{N}$ with average $\bar{a} = 1/N\sum_{k=1}^{N} a_k$. Then since the function $h(t) = 1/t$ is continous and convex on the set of positive numbers, it is also midpoint convex on that set, i.e.,

$$\frac{N}{\bar{a}} = Nh(\bar{a}) \leq \sum_{k=1}^{N} h(a_k) = \sum_{k=1}^{N} \frac{1}{a_k}.$$

With this fact, let $\{\lambda_k\}_{k=1}^{N-1}$ denote the eigenvalues of connected graph G and $\{\tilde{\lambda}_k\}_{k=0}^{N-1}$ denote the eigenvalues of the conditioned graph $\tilde{G}_\kappa$, both satisfying $\bar{\lambda} = 1/N\sum_{k=1}^{N-1}\lambda_k = 1/N\sum_{k=1}^{N-1}\tilde{\lambda}_k$. Since $\tilde{G}_\kappa$ is better conditioned than G, then $\left\|\sum_{k=1}^{N-1}\tilde{\lambda}_k - \bar{\lambda}\right\| \leq \left\|\sum_{k=1}^{N-1}\lambda_k - \bar{\lambda}\right\|$. In other words, the eigenvalues $\{\tilde{\lambda}_k\}_{k=1}^{N-1}$ are closer to the average $\bar{\lambda}$ than the eigenvaleus $\{\lambda_k\}_{k=1}^{N-1}$ are. Hence

$$\sum_{k=1}^{N-1} \frac{1}{\tilde{\lambda}_k} \leq \sum_{k=1}^{N-1} \frac{1}{\lambda_k}. \tag{1.15}$$

Equation (1.15) almost resembles the effective resistance $R(i,j) = \sum_{k=1}^{N-1} 1/\lambda_k(f_k(i) - f_k(j))^2$ except for the term $(f_k(i) - f_k(j))^2$. This term will be difficult to account for since little is known about the eigenvectors of $\tilde{G}_\kappa$. Analysis of eigenvectors of perturbed matrices is a widely open area of research and results are very limited, see [14, 24, 23, 9].

We remark that Conjecture 1 claims that conditoning a graph will reduce the average effective resistance between points; it is not true that the resistance between all points will be reduced. If the weight on edge (i,j) is reduced, then its effective resistance between points i and j is increased. Since we impose that the trace of the Laplacians be preserved, if any edge weights are increased, then by conservation at least one other edge's weight must be decreased. The vertex pairs for those edges will then have an increased effective resistance between them.

While we lack the theoretical justification, numerical simulations support Conjecture 1 and this is a source of future work.

The authors of [13] approach a similar way. They propose using convex optimization to minimize the total effective resistance of the graph,

$$R_{tot} = \sum_{i,j=1}^{N} R(i,j).$$

They show that the optimization problem is related to the problem of reweighting edges to maximize the algebraic connectivity λ_1.

Acknowledgment

References

1. Owe Axelsson, *Iterative Solution Methods*, Cambridge University Press, 1996.
2. Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
3. Jameson Cahill and Peter G. Casazza, *The Paulsen problem in operator theory*, Operators and Matrices **7** (2013), no. 1, 117–130.
4. Jameson Cahill and Xuemei Chen, *A note on scalable frames*, 10th International Conference on Sampling Theory and Applications (SampTA 2013) (Bremen, Germany), July 2013, pp. 93–96.
5. Peter G Casazza and Xuemei Chen, *Frame scalings: A condition number approach*, arXiv preprint arXiv:1510.01653 (2015).
6. Peter G. Casazza and Gitta Kutyniok (eds.), *Finite Frames: Theory and Applications*, Springer-Birkhäuser, New York, 2013.
7. Petter G. Casazza and Gitta Kutyniok, *Introduction to finite frames*, Finite Frames, Theory and Applications (Petter G. Casazza and Gitta Kutyniok, eds.), Springer-Birkhäuser, New York, 2013, pp. 1–53.
8. Xuemei Chen, Gitta Kutyniok, Kasso A Okoudjou, Friedrich Philipp, and Rongrong Wang, *Measures of scalability*, Information Theory, IEEE Transactions on **61** (2015), no. 8, 4410–4423.
9. Alexander Cloninger, *Exploiting Data-Dependent Structure for Improving Sensor Acquisition and Integration*, Ph.D Thesis, University of Maryland, College Park (2014).
10. M. S. Copenhaver, Y. H. Kim, C. Logan, K. Mayfield, S. K. Narayan, and J. Sheperd, *Diagram vectors and tight frame scaling in finite dimensions*, Operators and Matrices **8** (2014), no. 1.
11. Inc. CVX Research, *CVX: Matlab software for disciplined convex programming, version 2.0*, http://cvxr.com/cvx, August 2012.
12. David Eppstein, *Quasiconvex programming*, Combinatorial and Computational Geometry **52** (2005), 287–331.
13. Arpita Ghosh, Stephen Boyd, and Amin Saberi, *Minimizing effective resistance of a graph*, SIAM review **50** (2008), no. 1, 37–66.
14. Tosio Kato, *Perturbation theory for linear operators*, vol. 132, Springer Science & Business Media, 1976.
15. Gitta Kutyniok, Kasso A Okoudjou, and Friedrich Philipp, *Preconditioning of frames*, SPIE Optical Engineering+ Applications, International Society for Optics and Photonics, 2013, pp. 88580G–88580G.
16. Gitta Kutyniok, Kasso A. Okoudjou, and Friedrich Philipp, *Scalable frames and convex geometry*, Contemp. Math **626** (2014), 19–32.
17. Gitta Kutyniok, Kasso A. Okoudjou, Friedrich Philipp, and Elizabeth K Tuley, *Scalable frames*, Linear Algebra and its Applications **438** (2013), no. 5, 2225–2238.
18. Zhaosong Lu and Ting Kei Pong, *Minimizing condition number via convex programming*, SIAM Journal on Matrix Analysis and Applications **32** (2011), no. 4, 1193–1211.
19. Pierre Maréchal and Jane J Ye, *Optimizing condition numbers*, SIAM Journal on Optimization **20** (2009), no. 2, 935–947.
20. Kasso A. Okoudjou (ed.), *Finite Frame Theory: A Complete Introduction to Overcompleteness*, Proceedings of Symposia in Applied Mathematics, AMS, Providence, RI, 2016.
21. _____ , *Preconditioning techniques in frame theory and probabilistic frames*, Finite Frame Theory: A Complete Introduction to Overcompleteness (Kasso A. Okoudjou, ed.), Proceedings of Symposia in Applied Mathematics, AMS, Providence, RI, 2016.
22. Isaac Pesenson, *Sampling in paley-wiener spaces on combinatorial graphs*, Trans. Amer. Math. Soc. **360** (2008), no. 10.
23. Gilbert W Stewart, *Matrix perturbation theory*, Academic Press, Inc., 1990.
24. Terrence Tao, *Topics in random matrix theory*, vol. 132, American Mathematical Soc., 2012.