Convex relaxation for the planted clique, biclique, and clustering problems

by

Brendan P. W. Ames

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Combinatorics & Optimization

Waterloo, Ontario, Canada, 2011

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

A clique of a graph G is a set of pairwise adjacent nodes of G. Similarly, a biclique (U, V) of a bipartite graph G is a pair of disjoint, independent vertex sets such that each node in U is adjacent to every node in V in G. We consider the problems of identifying the maximum clique of a graph, known as the maximum clique problem, and identifying the biclique (U, V) of a bipartite graph that maximizes the product $|U| \cdot |V|$, known as the maximum edge biclique problem. We show that finding a clique or biclique of a given size in a graph is equivalent to finding a rank one matrix satisfying a particular set of linear constraints. These problems can be formulated as rank minimization problems and relaxed to convex programming by replacing rank with its convex envelope, the nuclear norm. Both problems are NP-hard yet we show that our relaxation is exact in the case that the input graph contains a large clique or biclique plus additional nodes and edges. For each problem, we provide two analyses of when our relaxation is exact. In the first, the diversionary edges are added deterministically by an adversary. In the second, each potential edge is added to the graph independently at random with fixed probability p. In the random case, our bounds match the earlier bounds of Alon, Krivelevich, and Sudakov, as well as Feige and Krauthgamer for the maximum clique problem.

We extend these results and techniques to the k-disjoint-clique problem. The maximum node k-disjoint-clique problem is to find a set of k disjoint cliques of a given input graph containing the maximum number of nodes. Given input graph G and nonnegative edge weights $w \in \mathbf{R}_+^E$, the maximum mean weight k-disjoint-clique problem seeks to identify the set of k disjoint cliques of G that maximizes the sum of the average weights of the edges, with respect to w, of the complete subgraphs of G induced by the cliques. These problems may be considered as a way to pose the clustering problem. In clustering, one wants to partition a given data set so that the data items in each partition or cluster are similar and the items in different clusters are dissimilar. For the graph G such that the set of nodes represents a given data set and any two nodes are adjacent if and only if the corresponding items are similar, clustering the data into k disjoint clusters is equivalent to partitioning G into k-disjoint cliques. Similarly, given a complete graph with nodes corresponding to a given data set and edge weights indicating similarity between each pair of items, the data may be clustered by solving the maximum mean weight k-disjoint-clique problem.

We show that both instances of the k-disjoint-clique problem can be formulated as rank constrained optimization problems and relaxed to semidefinite programs using the nuclear norm relaxation of rank. We also show that when the input instance corresponds to a collection of k disjoint planted cliques plus additional edges and nodes, this semidefinite relaxation is exact for both problems. We provide theoretical bounds that guarantee ex-

actness of our relaxation and provide empirical examples of successful applications of our algorithm to synthetic data sets, as well as data sets from clustering applications.

Acknowledgements

I wish to thank my advisor and collaborator, Stephen Vavasis. Thank you for your guidance, encouragement, and assistance throughout the research and preparation of this thesis, and for your wise and thoughtful career advice. I am extremely fortunate to have had you as an advisor and mentor during my tenure at the University of Waterloo.

I am grateful to Henry Wolkowicz, Levent Tunçel, Shai Ben-David, and Inderjit Dhillon for serving on my examination committee and for their helpful comments and suggestions on my thesis. I would especially like to thank both Henry and Inderjit, for independently suggesting I take a look at the graph-partitioning problem, and Shai for his insight on the clustering problem.

I would also like to thank Hristo Sendov for providing me with a necessary distraction when I needed a break from thesis work, and for nudging me towards the pursuit of mathematics many years ago.

I owe the staff members, faculty members, and my fellow graduate students in the Department of Combinatorics & Optimization a tremendous amount of gratitude. I truly enjoyed my time at the University of Waterloo and have all of you to thank.

I am grateful to have received financial support in the form of an NSERC Doctoral Postgraduate Scholarship and an Ontario Graduate Scholarship.

Mom and Dad, thank you for your constant support and encouragement, and imparting on me the value of hard work and education. I also thank my sisters Kim and Megan, grandmother Pauline, parents-in-law Keith and Denise, brother-in-law Matthew, and soon to be brother-in-law Ross for all of their love and support.

Most of all, I would like to thank my wife, Lauren, for her proofreading expertise, patience, support, and love. I would not have been able to do this without you.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this dissertation, we study the behaviour of combinatorial optimization problems under certain assumptions on the structure of program inputs. In [113], Karp suggests considering the behaviour of intractable problems for random program inputs. However, a truly random model may not be a good representation of average case data for many practical applications. For many applications in information retrieval and data mining, a better representation of generic program inputs would be those where a particular hidden structure is present in the data but has been obscured by random noise. For example, natural images tend to have large regions of similar pixels separated by sharp edges rather than random arrangements of pixels. In this case, it would be more helpful to study the perfomance of an image processing algorithm for images containing this structure than it would be to consider purely random images. In matrix completion, the desired matrix to be recovered may be of low-rank. For example, Gram matrices for data embedded in a low dimensional Euclidean space [125, 170] will be of low-rank. Similarly, matrices arising from collaborative filtering applications such as the Netflix problem [16] are believed to have low rank since user preferences are believed to depend on a small number of factors relative to the number of users and items in the catalogue. For clustering applications, we may not need our clustering algorithm to obtain a partition of purely random data, which would likely not admit a meaningful clustering, but would instead need only to obtain accurate partitionings of data containing a few large distinct clusters. Due to the difficult combinatorial nature of these problems, we will not likely be able to provide an exact solution for all program inputs. We would instead like to identify algorithms or heuristics that efficiently obtain good solutions when the program inputs contain the desired structure.

We consider the maximum clique, maximum edge biclique, and the maximum node and maximum mean weight k-disjoint-clique problems. These problems arise as simple

model problems for data mining. For example, Iasemidis et al. [103] reduce a data mining problem in epilepsy prediction to an instance of the maximum clique problem. Similarly, Gillis and Glineur [83] use the biclique problem as a model problem for nonnegative matrix factorization and finding features in images. The k-disjoint-clique problem is equivalent to the clustering problem for particular program inputs. Although each of these problems is NP-hard, we show that they can be solved in polynomial time by relaxing to convex programming for input graphs structured in a specific way. In particular, our relaxation is exact in the case that the input graph consists of the desired subgraph plus a number of diversionary edges and nodes. We call such a subgraph planted. Our techniques and results rely heavily on recent results in sparse optimization. Specifically, our relaxations are based on the nuclear norm heuristic for matrix rank minimization. In matrix rank minimization, the goal is to find a matrix of minimum rank satisfying a given set of linear constraints. Several recent results [158, 36] have shown that in the case that a solution of very low-rank is known to exist *a priori*, and the set of constraints is random, then the nuclear norm relaxation recovers the unique solution of minimum rank. These results build upon recent breakthroughs in compressive sensing [38, 37, 58, 59]. In compressive sensing, the goal is to recover a sparse vector from a small number of observations of its entries. In the case that these observations are made with random matrices, the sparse solution can be solved by relaxing to linear programming using the ℓ_1-norm.

1.1 Outline of dissertation

This dissertation is organized as follows: Chapter 2 provides necessary background results on convex programming, matrix analysis, and probability theory. Chapter 3 presents an overview of the nuclear norm heuristic for matrix rank minimization. Chapters 4 and 5 present relaxations for the maximum clique and maximum edge biclique, and the k-disjoint-clique problems, respectively. For each of these problems, we provide theoretical bounds on the size of the planted subgraph and amount of noise that our algorithm can tolerate and still recover the exact solution. Finally, Chapter 6 contains the results of numerical experiments that empirically demonstrate the effectiveness of our algorithms for synthetic data sets, as well as data sets drawn from clustering and image segmentation applications.

Portions of the material in Section 2.6, and Chapters 4 and 5 appear in [7, 8].

Chapter 2

Background

We begin by reviewing terminology and classical results from convex and matrix analysis.

2.1 Vector and matrix norms

Let $\mathbf{E}$ be a Euclidean space with inner product $\langle \cdot, \cdot \rangle : \mathbf{E} \times \mathbf{E} \to \mathbf{R}$. For example, we are interested in the Euclidean spaces $\mathbf{E} = \mathbf{R}^n$ with inner product $\langle \cdot, \cdot \rangle : \mathbf{E} \times \mathbf{E} \to \mathbf{R}$ defined by $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$ and $\mathbf{E} = \mathbf{R}^{m \times n}$ with inner product $\langle X, Y \rangle = \mathrm{Tr}\,(X^T Y)$ for all $X, Y \in \mathbf{R}^{n \times n}$. A function $\| \cdot \| : \mathbf{E} \to \mathbf{R}_+$ is a *norm* if for all $\mathbf{x}, \mathbf{y} \in E$,

1. $\|\mathbf{x}\| \geq 0$ and $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = 0$,
2. $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$ for all $\alpha \in \mathbf{R}$, and
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Note that the inner product defines a norm by $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ for all $\mathbf{x} \in \mathbf{E}$. For example, for $\mathbf{E} = \mathbf{R}^n$, the inner product defines the norm $\| \cdot \|_2$, called the *Euclidean* or ℓ_2*-norm*, by $\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^T \mathbf{x}}$ for all $\mathbf{x} \in \mathbf{R}^n$. When $E = \mathbf{R}^{n \times n}$, the inner product defines the *Frobenius norm*, denoted $\| \cdot \|_F$, by $\|X\|_F = \sqrt{\mathrm{Tr}\,(X^T X)}$ for all $X \in \mathbf{R}^{m \times n}$. Other norms in $\mathbf{R}^n$ include the ℓ_1*-norm* $\| \cdot \|_1$, defined by $\|\mathbf{x}\|_1 = |\mathbf{x}_1| + \cdots + |\mathbf{x}_n|$, and the *max* or *infinity norm* $\| \cdot \|_\infty$, defined by $\|\mathbf{x}\|_\infty = \max\{|\mathbf{x}_1|, \ldots, |\mathbf{x}_n|\}$. In general, $\|\mathbf{x}\|_p = (\sum_{i=1}^n |\mathbf{x}_i|^p)^{1/p}$ is a

norm, called the ℓ_p-*norm*, for all $p \geq 1$. Note that the max norm is the limit of the ℓ_p-norm as $p \to \infty$. We have the following chain of inequalities (see [85, Equations (2.2.5)-(2.2.7)]):

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{n}\|\mathbf{x}\|_2 \leq n\|\mathbf{x}\|_\infty \tag{2.1.1}$$

for all $\mathbf{x} \in \mathbf{R}^n$. Moreover, we have

$$\|\mathbf{x}\|_2^2 \leq \|\mathbf{x}\|_1 \|\mathbf{x}\|_\infty \tag{2.1.2}$$

$\mathbf{x} \in \mathbf{R}^n$; see [101, Equation (5.4.14)].

Each norm $\|\cdot\|_v$ in $\mathbf{R}^n$ induces a norm $\|\cdot\|$ in $\mathbf{R}^{n\times n}$, called the *operator* norm induced $\|\cdot\|_v$, by

$$\|A\| = \max_{\mathbf{x}\in\mathbf{R}^n} \frac{\|A\mathbf{x}\|_v}{\|\mathbf{x}\|_v}. \tag{2.1.3}$$

In particular, the Euclidean, ℓ_1, and max norms define the following operator norms in $\mathbf{R}^{m\times n}$:

$$\begin{aligned}
\|A\|_2 &:= \max_{\mathbf{x}\in\mathbf{R}^n} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \\
\|A\|_1 &:= \max_{\mathbf{x}\in\mathbf{R}^n} \frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1} = \max_{1\leq j\leq n} \sum_{i=1}^m |A_{ij}| \\
\|A\|_\infty &:= \max_{\mathbf{x}\in\mathbf{R}^n} \frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} = \max_{1\leq j\leq m} \sum_{j=1}^n |A_{ij}|.
\end{aligned}$$

Moreover, we have the following chains of inequalities

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{n}\|A\|_F \tag{2.1.4}$$

$$\frac{1}{\sqrt{n}}\|A\|_\infty \leq \|A\|_2 \leq \sqrt{m}\|A\|_\infty \tag{2.1.5}$$

$$\frac{1}{\sqrt{m}}\|A\|_1 \leq \|A\|_2 \leq \sqrt{n}\|A\|_1 \tag{2.1.6}$$

$$\|A\|_2 \leq \sqrt{\|A\|_1\|A\|_\infty} \tag{2.1.7}$$

for all $A \in \mathbf{R}^{m\times n}$ (see [85, Equations (2.3.7), (2.3.11), (2.3.12)]). Unless otherwise noted, we will use $\|\mathbf{x}\|$ to denote the Euclidean norm $\|\mathbf{x}\|_2$ in $\mathbf{R}^n$ and $\|A\|$ to denote operator ℓ_2-norm $\|A\|_2$ in $\mathbf{R}^{n\times n}$.

2.2 Rank and the singular value decomposition

An $n \times n$ real matrix U is *orthogonal* if $U^T U = UU^T = I$. Let O^n denote the set of $n \times n$ orthogonal matrices. We may factorize each $m \times n$ matrix A as

$$A = USV^T \tag{2.2.1}$$

where $U \in O^m$, $V \in O^n$, and

$$S = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$$

where D is an $r \times r$ diagonal matrix for some integer $r \in \{1, \ldots, \min(m, n)\}$, with positive diagonal entries $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ ordered nonincreasingly. Equivalently, we have

$$A = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \tag{2.2.2}$$

where $\mathbf{u}_i, \mathbf{v}_i$ denote the ith columns of U, V respectively. We call the diagonal entries of the matrix S the *singular values* of A and the columns of U and V the *left* and *right singular vectors* of A, respectively. The factorization $A = USV^T$ is called the *singular value decomposition* of A. Note that the number of nonzero singular values of A is equal to r. The value r is called the *rank* of A, denoted $\operatorname{rank}(A)$. The rank of a matrix is equal to the number of linearly independent rows or, equivalently, columns of A. We have the following theorem.

Theorem 2.2.1 (Rank plus nullity theorem [134, (4.4.15)]) *For every $A \in \mathbf{R}^{m \times n}$, we have*

$$n = \operatorname{rank}(A) + \dim(N(A)) \tag{2.2.3}$$

where $N(A)$ denotes the nullspace of the matrix A, $N(A) = \{\mathbf{x} \in \mathbf{R}^n : A\mathbf{x} = 0\}$.

The singular value decomposition provides equivalent representations of the norms discussed in the previous chapter. For example, $\|A\|_2 = \max\{\sigma_1 A), \ldots, \sigma_r(A)\} = \sigma_1(A)$ for all $A \in \mathbf{R}^{m \times n}$. That is, the operator ℓ_2-norm is equal to the max norm of the vector of singular values of A. Similarly, $\|A\|_F = (\sigma_1(A)^2 + \cdots + \sigma(A)_r^2)^{1/2}$ is equal to the Euclidean norm of the vector of singular values for all $A \in \mathbf{R}^{m \times n}$. The nuclear norm, which will serve a major role in our heuristics for the clique, biclique, and clustering problems, is equal to the ℓ_1-norm of the vector of singular values $\|A\|_* = \sigma_1(A) + \cdots + \sigma_r(A)$.

2.3 Convex sets, functions, hulls, and cones.

Let $\mathbf{E}$ be a Euclidean space with inner product $\langle \cdot, \cdot \rangle$. We say a subset $C \subseteq \mathbf{E}$ is a *convex set* if $(t\mathbf{x} + (1-t)\mathbf{y}) \in C$ for every $\mathbf{x}, \mathbf{y}, \in \mathbf{E}$ and $t \in [0,1]$. Note that the intersection of convex sets is convex. In fact, every closed convex set in $\mathbf{E}$ can be characterized as the intersection of potentially infinitely many convex sets of the form $\{\mathbf{x} \in \mathbf{E} : \langle \alpha, \mathbf{x} \rangle \leq 0\}$ for some $\alpha \in \mathbf{E}$. We call such a set $\{\mathbf{x} \in \mathbf{E} : \langle \alpha, \mathbf{x} \rangle \leq 0\}$ a *half-space*. We have the following theorem.

Theorem 2.3.1 ([163, Theorem 11.5]) *Every closed convex set is the intersection of the closed half-spaces which contain it.*

Theorem 2.3.1 suggests a natural means of convexifying a nonconvex set. The *convex hull* of a subset $C \subseteq \mathbf{E}$, denoted $\operatorname{conv}(C)$, is the smallest convex set in E containing C. That is, $\operatorname{conv}(C)$ is equal to the intersection of all convex sets in $\mathbf{E}$ containing C:

$$\operatorname{conv}(C) = \bigcap \{S : C \subseteq S,\ S \text{ convex}\}.$$

An important class of convex sets are the convex cones. A set $C \subseteq \mathbf{E}$ is a cone if $\alpha \mathbf{x} \in C$ for all $\mathbf{x} \in C$ and $\alpha \in \mathbf{R}_+$. A set $C \subseteq E$ is a *convex cone* if it is both a cone and a convex set. This holds if and only if $\mathbf{x} + \mathbf{y} \in C$ for all $\mathbf{x}, \mathbf{y} \in C$. For example, the nonnegative cone in $\mathbf{R}^n$, $\mathbf{R}^n_+ := \{\mathbf{x} \in \mathbf{R}^n : \mathbf{x}_i \geq 0,\ i = 1, \ldots, n\}$, is a convex cone. In Section 2.5, we will see that the set of $n \times n$ symmetric, positive semidefinite matrices $\Sigma^n_+ = \{X \in \mathbf{R}^{n \times n} : X = X^T,\ \mathbf{v}^T X \mathbf{v} \geq 0\ \forall \mathbf{v} \in \mathbf{R}^n\}$ is also a convex cone, called the *semidefinite cone*.

We have a similar definition for convexity of functions. We say a function $f : \mathbf{E} \to [-\infty, \infty]$ is a *convex function* if its epigraph $\operatorname{epi} f := \{(\mathbf{x}, t) : t \geq f(\mathbf{x})\}$ is a convex set. Equivalently, f is convex if $f(t\mathbf{x} + (1-t)\mathbf{y}) \leq tf(\mathbf{x}) + (1-t)f(\mathbf{y})$ for every $\mathbf{x}, \mathbf{y}, \in E$ and $t \in [0,1]$. The *domain* of f is the set of elements of $\mathbf{E}$ where f has finite value, $\operatorname{dom} f = \{\mathbf{x} \in \mathbf{E} : f(\mathbf{x}) < +\infty\}$. A function f is *proper* if $\operatorname{dom} f$ is nonempty and f never takes the value $-\infty$. The *convex envelope* of a function $g : E \to [-\infty, \infty]$, denoted $\operatorname{conv}(g)$, is the smallest convex function f such that $f(\mathbf{x}) \geq g(\mathbf{x})$ for all $\mathbf{x} \in \mathbf{E}$. Equivalently, a function f is the convex envelope of the function g if and only if $\operatorname{epi}(f) = \operatorname{epi}(\operatorname{conv}(g)) = \operatorname{conv}(\operatorname{epi}(g))$. For a convex function $f : \mathbf{E} \to [-\infty, \infty]$, $\phi \in \mathbf{E}$ is a *subgradient* of f at $\bar{\mathbf{x}}$ if

$$\langle \phi, \mathbf{x} - \bar{\mathbf{x}} \rangle \leq f(\mathbf{x}) - f(\bar{\mathbf{x}})$$

for all $\mathbf{x} \in E$. The *subdifferential* of f at $\bar{\mathbf{x}}$ is the set of subgradients of f at $\bar{\mathbf{x}}$, denoted $\partial f(\bar{\mathbf{x}})$. Subgradients may be thought of as a generalization of the gradient for nonsmooth functions. Indeed, when f is continuously differentiable at $\mathbf{x}$ the subdifferential at $\bar{\mathbf{x}}$ is exactly the gradient $\partial f(\bar{\mathbf{x}}) = \{\nabla f(\bar{\mathbf{x}})\}$.

2.4 Convex programming

A *convex program* is an optimization problem of the form

$$\inf\{f(\mathbf{x}) : \mathbf{x} \in C\} \tag{2.4.1}$$

where $f : E \to (-\infty, \infty]$ is a proper, convex function and C is a convex subset of a Euclidean space $\mathbf{E}$. That is, convex programming seeks to minimize a given convex function over a given convex set. When the given convex set C is the intersection of a convex cone and an affine subspace, we have a *conic program*. An important example of conic programming is semidefinite programming. An overview of semidefinite programming is provided in Section 2.5.

Theorem 2.3.1 suggests that every convex program of form (2.4.1) can be rewritten in the form

$$\inf\{f(\mathbf{x}) : g_i(\mathbf{x}) \leq 0 \; \forall \; i \in I, \;\; h_j(\mathbf{x}) = 0 \; \forall \; j \in J\} \tag{2.4.2}$$

where $f, g_i : \mathbf{E} \to (-\infty, \infty]$, $i \in I$ are proper, convex functions, $h_j : E \to \mathbf{R}$ are linear for all $j \in J$ and I, J are (possibly infinite) index sets. Here, the convex set C is the set of all $\mathbf{x} \in \mathbf{E}$ such that $g_i(\mathbf{x}) \leq 0$ for all $i \in I$ and $h_j(\mathbf{x}) = 0$ for all $j \in J$. Suppose that the index sets I, J in (2.4.2) are finite; say $I = \{1, \ldots, m\}$, $J = \{1, \ldots, p\}$. The *Lagrangian* of the problem (2.4.2) is the function $L : (\mathbf{E} \times \mathbf{R}^m_+ \times \mathbf{R}^p) \to \mathbf{R}$ defined by

$$L(\mathbf{x}, \mathbf{y}, \mathbf{z}) = f(\mathbf{x}) + \sum_{i=1}^{m} \mathbf{y}_i g_i(\mathbf{x}) + \sum_{j=1}^{p} \mathbf{z}_j h_j(\mathbf{x}), \tag{2.4.3}$$

for all $\mathbf{x} \in \mathbf{E}$, $\mathbf{y} \in \mathbf{R}^m_+$, $\mathbf{z} \in \mathbf{R}^p$. We can write (2.4.2) as the *primal* problem

$$\inf_{\mathbf{x} \in \mathbf{E}} \sup_{\mathbf{y} \in \mathbf{R}^m_+, \mathbf{z} \in \mathbf{R}^p} L(\mathbf{x}, \mathbf{y}, \mathbf{z}). \tag{2.4.4}$$

Switching the order of the infimum and supremum in (2.4.4) yields the *dual* problem

$$\sup_{\mathbf{y}\in\mathbf{R}^m_+,\mathbf{z}\in\mathbf{R}^p} \inf_{\mathbf{x}\in\mathbf{E}} L(\mathbf{x},\mathbf{y},\mathbf{z}). \tag{2.4.5}$$

The dual provides a lower bound on the optimal primal value since

$$\infty \geq \inf_{\mathbf{x}\in\mathbf{E}} \sup_{\mathbf{y}\in\mathbf{R}^m_+,\mathbf{z}\in\mathbf{R}^p} L(\mathbf{x},\mathbf{y},\mathbf{z})) \geq \sup_{\mathbf{y}\in\mathbf{R}^m_+,\mathbf{z}\in\mathbf{R}^p} \inf_{\mathbf{x}\in\mathbf{E}} L(\mathbf{x},\mathbf{y},\mathbf{z}) \geq -\infty.$$

In fact, if the primal admits a strictly feasible solution, then the optimal primal and dual values coincide. We say that a convex program (2.4.2) satisfies the *Slater constraint qualification* if there exists $\hat{\mathbf{x}} \in \operatorname{dom} f$ such that $g_i(\hat{\mathbf{x}}) < 0$ for all $i = 1,\ldots,m$. We have the following theorem.

Theorem 2.4.1 ([26, Theorem 4.3.7]) *If the Slater condition holds for the primal problem* (2.4.2) *then the primal and dual values are equal, and the dual value is attained if finite.*

Duality can also be used to characterize when a particular feasible solution of (2.4.2) is optimal. We have the following theorem.

Theorem 2.4.2 ([163, Theorem 28.3]) *Suppose that* (2.4.2) *satisfies the Slater constraint qualification. Then* $\mathbf{x}^*$ *is optimal for* (2.4.2) *if and only if there exists* $\mathbf{y}^* \in \mathbf{R}^m_+$, $\mathbf{z}^* \in \mathbf{R}^p$ *such that* $\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*$ *satisfy*

1. *Primal feasibility:* $g_i(\mathbf{x}^*) \leq 0$ *for all* $i = 1,\ldots,m$, $h_j(\mathbf{x}^*) = 0$ *for all* $j = 1,\ldots,p$,
2. *Dual feasibility:* $0 \in \partial f(\mathbf{x}^*) + \sum_{i=1}^m \mathbf{y}_i^* \partial g_i(\mathbf{x}^*) + \sum_{j=1}^p \mathbf{z}_j^* \nabla h_j(x^*)$, *and*
3. *Complementary slackness:* $\mathbf{y}_i^* g_i(\mathbf{x}^*) = 0$ *for all* $i = 1,\ldots,m$.

We conclude this section by mentioning an important application of convex programming: relaxation of combinatorial problems. Convex relaxation is a popular technique in heuristics and approximation algorithms for combinatorial optimization. In *convex relaxation*, a combinatorial problem is relaxed to a convex program, such as minimizing the convex envelope of the original objective function over the convex hull of the set of feasible solutions. This relaxed problem, and convex programs of form (2.4.1) can, in general, be solved efficiently if evaluation of f and testing membership in C can be performed efficiently; see, for example, [169, 192, 142]. The optimal solution of the relaxed problem can be then used to obtain a lower bound on that of the original problem. For example, relaxing integrality constraints and solving the resulting linear program is a key component of branch and cut, and branch and bound algorithms for integer programming.

2.5 Semidefinite programming

An important class of convex programming is *semidefinite programming*. Here the convex set to be optimized over is a subset of the semidefinite cone. This section reviews several important properties of the semidefinite cone and semidefinite programs.

We say that an $n \times n$ matrix A is symmetric if $A = A^T$. For each symmetric matrix A, we have the factorization $A = UDU^T$ where $U \in O^n$ and D is an $n \times n$ diagonal matrix with diagonal entries ordered nonincreasingly. The diagonal entries of D are called *eigenvalues* of A and the columns of U are *eigenvectors*. We call such a factorization the *spectral decomposition* of A and the set of eigenvalues of A the *spectrum* of A. Note that λ is an eigenvalue, with eigenvector $\mathbf{v}$, if there exists $\mathbf{v} \in \mathbf{R}^n$, $\mathbf{v} \neq 0$, such that $A\mathbf{v} = \lambda\mathbf{v}$. Equivalently, λ is an eigenvalue of A if it is a root of the *characteristic polynomial* $\det(A - \lambda I)$. We say that a symmetric matrix A is *positive semidefinite* if each of its eigenvalues is nonnegative. Equivalently, A is positive semidefinite if $\mathbf{x}^T A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbf{R}^n$. We denote the set of $n \times n$ symmetric matrices by Σ^n, and the set of $n \times n$ symmetric positive semidefinite matrices by Σ^n_+. Note that the singular value and spectral decompositions are identical for all $A \in \Sigma^n_+$. We denote the partial order on Σ^n_+ by "$\succeq$": $A \succeq B$, if $A - B \in \Sigma^n_+$. Similarly, we say that a matrix A is *positive definite* if each of its eigenvalues is strictly positive. We denote the set of $n \times n$ symmetric positive definite matrices by Σ^n_{++} and have the partial order on Σ^n_{++} defined by $A \succ B$ if and only $A - B \succ 0$. Note that Σ^n_+ is a convex cone for all $n \in \mathbf{Z}_+$. Indeed, the set of symmetric matrices is closed under addition and $\mathbf{x}^T(A+B)\mathbf{x} = \mathbf{x}^T A\mathbf{x} + \mathbf{x}^T B\mathbf{x} \geq 0$ for all $A, B \in \Sigma^n_+$ and $\mathbf{x} \in \mathbf{R}^n$. We refer to Σ^n_+ as the *semidefinite cone*.

A *semidefinite program* (SDP) is an optimization problem of the form

$$\inf\{\langle C, X\rangle : \mathcal{A}(X) = \mathbf{b},\ X \succeq 0\} \tag{2.5.1}$$

where the symmetric matrix $C \in \Sigma^n$, linear map $\mathcal{A} : \Sigma^n_+ \to \mathbf{R}^m$, and vector $\mathbf{b} \in \mathbf{R}^m$ are given. Since the sets $\{X : \mathcal{A}(X) = \mathbf{b}\}$ and Σ^n_+ are convex, and the objective function $\langle C, \cdot\rangle$ is linear, (2.5.1) is a convex program. Hence, we may define a dual problem and optimality conditions for (2.5.1) as in the previous section. The dual of (2.5.1) is the semidefinite program

$$\sup\{\mathbf{b}^T\mathbf{y} : \mathcal{A}^*(\mathbf{y}) + S = C,\ S \succeq 0\} \tag{2.5.2}$$

where $\mathcal{A}^* : \mathbf{R}^m \to \Sigma^n_+$ is the *adjoint map* of $\mathcal{A}$ defined by

$$(\mathcal{A}(X))^T\mathbf{y} = \operatorname{Tr}(\mathcal{A}^*(\mathbf{y})X)$$

for all $X \in \Sigma_+^n$, $\mathbf{y} \in \mathbf{R}^m$. For (2.5.1), the Slater constraint qualification is satisfied if there exists $\tilde{X} \in \Sigma^n$ such that $\mathcal{A}(\tilde{X}) = \mathbf{b}$ and $\tilde{X} \succ 0$. On the other hand, the dual problem (2.5.2) satisfies the Slater condition if there exists $\tilde{S} \in \Sigma^n$ and $\tilde{\mathbf{y}} \in \mathbf{R}^m$ such that $\mathcal{A}^*(\tilde{\mathbf{y}}) + \tilde{S} = C$ and $\tilde{S} \succ 0$. As before, we have the following theorem guaranteeing equality between the optimal objective values of the primal and dual if the Slater constraint qualification holds.

Theorem 2.5.1 ([181, Corollary 2.16]) *If* (2.5.1) *and* (2.5.2) *satisfy the Slater condition, then they both attain their optimal objective values and the optimal objective values of* (2.5.1) *and* (2.5.2) *coincide.*

The necessary and sufficient conditions for optimality of a given feasible solution of a general convex program given by Theorem 2.4.2 are specialized to semidefinite programming in the following theorem.

Theorem 2.5.2 *Suppose that* (2.5.1) *and* (2.5.2) *satisfy the Slater constraint qualification. Then X^* is optimal for the primal problem* (2.5.1) *and $(S^*, \mathbf{y}^*)$ is optimal for the dual problem* (2.5.2) *if and only if $X^*, S^* \succeq 0$ and $X^*, S^*, \mathbf{y}^*$ satisfy*

1. *Primal feasibility: $\mathcal{A}(X^*) = \mathbf{b}$,*
2. *Dual feasibility: $\mathcal{A}^*(\mathbf{y}^*) + S^* = C$, and*
3. *Complementary slackness: $\langle X^*, S^* \rangle = 0$.*

Semidefinite programs can be solved efficiently using interior point algorithms (see [142, 184, 176]) or first order methods [194], and several software packages for solving semidefinite programs are available (see [173, 76, 177, 182, 194]).

2.6 Bounds on the norms of random matrices and sums of random variables

2.6.1 Bounds on the tail of the sum of random variables

We begin with a version of the well-known Chernoff bounds providing a bound on the tail distribution of a sum of independent Bernoulli trials (see [135, Theorems 4.4 and 4.5]).

Theorem 2.6.1 (Chernoff Bounds) *Let $X_1, \ldots, X_k$ be a sequence of k independent Bernoulli trials, each succeeding with probability p so that $E(X_i) = p$. Let $S = \sum_{i=1}^{k} X_i$ be the binomially distributed variable describing the total number of successes. Then for $\delta > 0$*

$$P\Big(S > (1+\delta)pk\Big) \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{pk}. \tag{2.6.1}$$

It follows that for all $a \in (0, p\sqrt{k})$,

$$P(|S - pk| > a\sqrt{k}) \leq 2\exp(-a^2/p). \tag{2.6.2}$$

We next provide a theorem of Hoeffding (see [98, Theorem 1]), which provides a bound on the tail distribution of a sum of bounded, independent random variables.

Theorem 2.6.2 (Hoeffding's Inequality) *Let $X_1, \ldots, X_m$ be independently identically distributed (i.i.d.) variables sampled from a distribution satisfying $0 \leq X_i \leq 1$ for all $i = 1, \ldots, m$. Let $S = X_1 + \cdots + X_m$. Then*

$$Pr(|S - E[S]| > t) \leq 2\exp\left(\frac{-2t^2}{m}\right) \tag{2.6.3}$$

for all $t > 0$. In particular, for $t = a\sqrt{m}$ for some scalar $a > 0$, we have

$$Pr(|S - E[S]| > a\sqrt{m}) \leq 2\exp(-2a^2). \tag{2.6.4}$$

Note that a sequence of k independent Bernoulli trials satisfies the hypothesis of Theorem 2.6.2. In this sense, Theorem 2.6.2 may be thought of as a generalization of Theorem 2.6.1.

2.6.2 Bounds on the norms of random matrices

We begin with a theorem of Füredi and Komlós [77] that bounds the operator ℓ_2-norm of a random symmetric matrix whose entries have mean zero.

Theorem 2.6.3 *Let $A \in \Sigma^n$ be a random symmetric matrix with independently identically distributed (i.i.d.) entries sampled from a distribution satisfying $0 \leq A_{ij} = A_{ji} \leq 1$ for all*

$i, j \in \{1, \ldots, n\}$. Then there exist $c_1, c_2 > 0$ depending only on the variance of the entries of A such that

$$\|A - E[A]\| \leq c_1\sqrt{n}$$

with probability at least $1 - \exp(-c_2 n^{1/6})$.

Note that this theorem is not stated in this form in [77], but can be deduced by taking $k = \sigma^{1/3} n^{1/6}$ and $v = \sigma\sqrt{n}$ in the inequality

$$P(\max|\lambda| > 2\sigma\sqrt{n} + v) < \sqrt{n}\exp(-kv/(2\sigma\sqrt{n} + v))$$

on pp. 237 of [77].

In particular, we consider the probability distribution Ω for an entry A_{ij} defined as follows:

$$A_{ij} = \begin{cases} 1 & \text{with probability } p, \\ -p/(1-p) & \text{with probability } 1-p. \end{cases}$$

It is easy to see that the variance of A_{ij} is $\sigma^2 = p/(1-p)$ and each entry of A has mean equal to 0. For random symmetric matrices drawn from the distribution Ω, Theorem 2.6.3 specializes to the following result.

Theorem 2.6.4 *For all integers i, j, $1 \leq j \leq i \leq n$, let A_{ij} be distributed according to Ω. Define symmetrically $A_{ij} = A_{ji}$ for all $i < j$.*

Then the random symmetric matrix $A = [A_{ij}]$ satisfies

$$\|A\| \leq 3\sigma\sqrt{n}$$

with probability at least $1 - \exp(-cn^{1/6})$ for some $c > 0$ that depends on σ.

A similar theorem attributed to Geman [79] is available for nonsymmetric matrices.

Theorem 2.6.5 *Let A be a $\lceil yn \rceil \times n$ matrix whose entries are chosen according to Ω for fixed $y \geq 0$. Then, with probability at least $1 - c_1\exp(-c_2 n^{c_3})$ where $c_1 > 0$, $c_2 > 0$, and $c_3 > 0$ depend on p and y,*

$$\|A\| \leq c_4\sqrt{n}$$

for some $c_4 > 0$ also depending on p, y.

As in the case of [77], this theorem is not presented by Geman in this manner explicitly, but can be deduced from the equations on pp. 255–256 of [79] by taking $k = n^q$ for a q satisfying $(2\alpha + 4)q < 1$.

Chapter 3

Nuclear Norm Relaxation for the Affine Rank Minimization Problem

3.1 The Affine Rank Minimization Problem

Recovering low-rank matrices from noisy linear measurements has many applications in machine learning. For example, low-rank matrices may represent collections of stationary objects in a series of video frames [34, Section 4], distance matrices of objects embedded in a low-dimensional Euclidean space in sensor localization or facial recognition [125, 12, 170, 54, 119], catalogues of user preferences in collaborative filtering [161, 197], and pure or nearly pure quantum states [88]. Moreover, modelling binary or $\{-1, 1\}$ variables using low-rank matrices is a classical technique in semidefinite relaxation of combinatorial optimization problems, such as the maximum cut problem [84]. We will see in Chapters 4 and 5 that rank-one matrices can be used to represent cliques in graphs and clusters of similar objects in a data set respectively. Hence, we are often interested in identifying the matrix with minimum rank satisfying a given set of constraints. When the set of feasible solutions for these constraints is convex we have the *rank minimization problem*

$$\begin{array}{ll} \min & \operatorname{rank}(X) \\ \text{s.t.} & X \in C \end{array} \tag{3.1.1}$$

where $X \in \mathbf{R}^{m \times n}$, and C is a given convex set. When the given constraint set C is affine we have the affine rank minimization problem. That is, the *affine rank minimization problem*

$$\begin{array}{ll} \min & \operatorname{rank}(X) \\ \text{s.t.} & \mathcal{A}(X) = \mathbf{b} \end{array} \tag{3.1.2}$$

seeks the matrix $X \in \mathbf{R}^{m \times n}$ with lowest rank in the affine space defined by the given linear map $\mathcal{A} : \mathbf{R}^{m \times n} \to \mathbf{R}^p$ and vector $\mathbf{b} \in \mathbf{R}^p$. In general, the affine rank minimization problem is NP-hard. Indeed, we will see in Section 4.1.2 that the maximum clique problem, well-known to be NP-hard, can be formulated as a special case of affine rank minimization. Alternately, when the decision variable X is restricted to be a diagonal matrix, affine rank minimization is equivalent to *cardinality minimization*, also known to be NP-hard [140]. We focus our discussion on the *nuclear norm minimization* (NNM) heuristic introduced by Fazel et al. in [61, 62]. This heuristic involves solving the relaxation of (3.1.2) obtained by replacing the objective function rank with the nuclear norm:

$$\begin{array}{ll} \min & \|X\|_* = \sigma_1(X) + \cdots + \sigma_{\min\{m,n\}}(X) \\ \text{s.t.} & \mathcal{A}(X) = \mathbf{b} \end{array} \tag{3.1.3}$$

We will see in Section 3.2, that the resulting optimization problem (3.1.3) is convex and may be solved in polynomial time by formulating as a semidefinite program. Recently, several software packages [33, 126, 131, 178] for solving (3.1.3) have become available. Remarkably, although the affine rank minimization problem is intractable in general, it has been established that the minimum rank solution can be recovered in polynomial time by solving the (3.1.3) for many program inputs. In particular, the nuclear norm relaxation is exact with extremely high probability when the linear map $\mathcal{A}$ is drawn from one of several families of random linear maps.

The rest of this section is organized as follows. In Section 3.2 we discuss several useful properties of the nuclear norm, which make it an attractive candidate for relaxation of rank. The nuclear norm heuristic can be thought of as an extension of the ℓ_1 heuristic for the cardinality minimization problem. In Section 3.3, we provide background on the cardinality minimization problem and highlight some parallels between the ℓ_1 heuristic and nuclear norm minimization. Conditions on the linear map $\mathcal{A}$ that ensure that the nuclear norm heuristic is exact for affine rank minimization and examples of families of random maps for which they hold can be found in Section 3.4.

3.2 Properties of the Nuclear Norm

3.2.1 Convexity

For all $X \in \mathbf{R}^{m\times n}$, $\|X\|_*/\|X\|$ provides an underestimate of $\operatorname{rank}(X)$. Indeed, suppose that X has rank $r \geq 1$ and singular value decomposition $X = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$. Then

$$\frac{\|X\|_*}{\|X\|} = \frac{1}{\sigma_1(X)}(\sigma_1(X) + \cdots + \sigma_r(X)) \leq r,$$

since $\sigma_i(X) \leq \sigma_1(X)$, for all $i = 1, \ldots, r$. In the case that $\|X\| \leq 1$, we have $\|X\|_* \leq \operatorname{rank}(X)$. Hence, the nuclear norm underestimates rank on the set $\{X \in \mathbf{R}^{m\times n} : \|X\| \leq 1\}$. In fact, this is the best convex pointwise approximation to rank on the set $\{X \in \mathbf{R}^{m\times n} : \|X\| \leq 1\}$. Indeed, we have the following theorem.

Theorem 3.2.1 ([61, Theorem 1]) *The convex envelope of* $\operatorname{rank}(X)$ *on the set* $\{X \in \mathbf{R}^{m\times n} : \|X\| \leq 1\}$ *is the nuclear norm* $\|X\|_*$.

A proof of Theorem 3.2.1 can be found in [61, Section 5.1.5]. The proof establishes the fact that the nuclear norm $\|X\|_*$ is the biconjugate of $\operatorname{rank}(X)$ on the set $\{X \in \mathbf{R}^{m\times n} : \|X\| \leq 1\}$ and then uses the fact that the biconjugate of $\operatorname{rank}(X)$ is equal to its convex envelope (see [96, Theorem 1.3.5]) to complete the argument. The following corollary extends Theorem 3.2.1 to the set $\{X \in \mathbf{R}^{m\times n} : \|X\| \leq M\}$ for all $M \geq 0$.

Corollary 3.2.1 *The convex envelope of* $\operatorname{rank}(X)$ *on the set* $\{X \in \mathbf{R}^{m\times n} : \|X\| \leq M\}$ *is* $\|X\|_*/M$ *for all* $M \geq 0$.

Corollary 3.2.1 yields the following bounds on the optimal value of an instance of the affine rank minimization problem.

Corollary 3.2.2 ([158, pg. 479]) *Let* X_0, X^* *be the minimum rank and minimum nuclear norm solutions of* $\mathcal{A}(X) = \mathbf{b}$ *for given linear map* $\mathcal{A} : \mathbf{R}^{m\times n} \to \mathbf{R}^p$ *and vector* $\mathbf{b} \in \mathbf{R}^p$. *Then*

$$\frac{\|X^*\|_*}{\|X_0\|} \leq \frac{\|X_0\|_*}{\|X_0\|} \leq \operatorname{rank}(X_0) \leq \operatorname{rank}(X^*).$$

3.2.2 Formulation as a dual norm

For a given norm $\|\cdot\|$ in a Euclidean space $\mathbf{E}$ with inner product $\langle\cdot,\cdot\rangle$, the *dual norm* $\|\cdot\|_d$ of $\|\cdot\|$ is defined by

$$\|X\|_d = \sup_{Y\in\mathbf{E}}\{\langle X, Y\rangle : \|Y\| \leq 1\}. \tag{3.2.1}$$

The dual norm is itself a norm (see [101, p. 275-276]). When $\mathbf{E}$ is the set of $m \times n$ real matrices with the trace inner product $\langle X, Y\rangle = \mathrm{Tr}\,(X^T Y)$, we have the following characterization of the nuclear norm as the dual norm of the operator norm.

Theorem 3.2.2 ([158, Proposition 2.1]) *The dual norm of the operator ℓ_2-norm $\|\cdot\|$ in $\mathbf{R}^{m\times n}$ is the nuclear norm $\|\cdot\|_*$.*

Theorem 3.2.2 implies that the nuclear norm $\|\cdot\|_*$ is indeed a norm on $\mathbf{R}^{m\times n}$. Moreover, Theorem 3.2.2 suggests that the nuclear norm $\|X\|_*$ of a given matrix $X \in \mathbf{R}^{m\times n}$ may be calculated as the optimal objective value of a semidefinite program (see [158, Equations (2.5)-(2.6)]). Indeed, $\|X\|_*$ is the optimal value of the optimization problem

$$\|X\|_* = \max\{\mathrm{Tr}\,(X^T Y) : \|Y\| \leq 1\},$$

which is equivalent to the semidefinite program

$$\begin{array}{ll} \max\limits_{Y} & \mathrm{Tr}\,(X^T Y) \\ \text{s.t.} & \begin{pmatrix} I & Y \\ Y^T & I \end{pmatrix} \succeq 0 \end{array} \tag{3.2.2}$$

since $\|Y\| \leq 1$ if and only if $I \succeq Y^T Y$ which holds if and only if

$$\begin{pmatrix} I & Y \\ Y^T & I \end{pmatrix} \succeq 0$$

by the Schur Complement Theorem (see [29, Appendix A.5.5]). The dual of (3.2.2) is given by

$$\begin{array}{ll} \min\limits_{W_1,W_2} & \frac{1}{2}(\mathrm{Tr}\,(W_1) + \mathrm{Tr}\,(W_2)) \\ \text{s.t.} & \begin{pmatrix} W_1 & X \\ X^T & W_2 \end{pmatrix} \succeq 0. \end{array} \tag{3.2.3}$$

Note that both (3.2.2) and (3.2.3) have strictly feasible solutions. Therefore, Slater's condition holds and, hence, $\|X\|_*$ may be computed using either (3.2.2) and (3.2.3) since

there is no duality gap between them. An alternate argument for the lack of duality gap between (3.2.2) and (3.2.3) can be found in the proof of Proposition 2.1 in [158].

The semidefinite formulations (3.2.2) and (3.2.3) for the nuclear norm may be used (see [158, Equations (2.7)-(2.8)]) to express each instance of nuclear norm relaxation of affine rank minimization as a semidefinite program. Minimization of the nuclear norm over the affine subspace defined by the given linear map $\mathcal{A} : \mathbf{R}^{m \times n} : \mathbf{R}^p$ and vector $\mathbf{b} \in \mathbf{R}^p$ can be formulated as the convex program

$$\min\{\|X\|_* : \mathcal{A}(X) = \mathbf{b}\} \tag{3.2.4}$$

or its dual

$$\max\{\mathbf{b}^T\mathbf{z} : \|\mathcal{A}^*(\mathbf{z})\| \leq 1\}. \tag{3.2.5}$$

Substituting the dual characterization (3.2.3) into (3.2.4) yields the following semidefinite program formulation of nuclear norm minimization

$$\begin{array}{ll} \min & \frac{1}{2}(\mathrm{Tr}\,(W_1) + \mathrm{Tr}\,(W_2)) \\ \text{s.t.} & \begin{pmatrix} W_1 & X \\ X^T & W_2 \end{pmatrix} \succeq 0 \\ & \mathcal{A}(X) = \mathbf{b}. \end{array} \tag{3.2.6}$$

Similarly, substituting the characterization of the constraint $\|X\| \leq 1$ from (3.2.2) yields the dual of (3.2.6)

$$\begin{array}{ll} \max & \mathbf{b}^T\mathbf{z} \\ \text{s.t.} & \begin{pmatrix} I & \mathcal{A}^*(\mathbf{z}) \\ \mathcal{A}^*(\mathbf{z})^T & I \end{pmatrix} \succeq 0. \end{array} \tag{3.2.7}$$

Therefore, the convex relaxation of the affine rank minimization problem obtained by replacing $\mathrm{rank}\,(X)$ with the nuclear norm $\|X\|_*$ can be approximated within arbitrary precision in polynomial-time as a semidefinite program.

3.2.3 Optimality conditions for the affine nuclear norm minimization problem

We conclude by characterizing when a feasible solution for (3.2.4) is optimal. The following lemma characterizes the subdifferential of the nuclear norm at $X \in \mathbf{R}^{m \times n}$.

Lemma 3.2.1 ([186, Example 2]) *Let $X \in \mathbf{R}^{m\times n}$ have rank equal to r and compact singular value decomposition $X = UDV^T$ where $U \in \mathbf{R}^{m\times r}$, $V \in \mathbf{R}^{n\times r}$ and $D \in \mathbf{R}^{r\times r}$ is diagonal. Then the subdifferential of the nuclear norm $\|\cdot\|_*$ at X is equal to*

$$\partial\|X\|_* = \{UV^T + W : W^TU = 0, WV = 0, \|W\| \leq 1\}. \tag{3.2.8}$$

We have the following optimality conditions for the convex problem (3.2.4).

Theorem 3.2.3 ([158, Equation (2.11)]) *A matrix $X \in \mathbf{R}^{m\times n}$ is optimal for* (3.2.4) *if there exists $\mathbf{z} \in \mathbf{R}^p$ such that*

$$\mathcal{A}(X) = \mathbf{b} \tag{3.2.9}$$
$$\mathcal{A}^*(\mathbf{z}) \in \partial\|X\|_*. \tag{3.2.10}$$

3.3 Cardinality Minimization and Compressed Sensing

Vector cardinality minimization refers to the problem of finding the sparsest vector satisfying a given set of linear constraints. That is, given matrix $A \in \mathbf{R}^{m\times n}$ and $\mathbf{b} \in \mathbf{R}^n$, we want to find the vector $\bar{\mathbf{x}}$ with the minimum number of nonzero entries satisfying $A\bar{\mathbf{x}} = \mathbf{b}$. This problem arises in applications in signal processing. Suppose that we are given a vector $\bar{\mathbf{x}} \in \mathbf{R}^n$ representing a particular sparse signal. That is, $\bar{\mathbf{x}}$ has k nonzero entries where $k << n$. We want to exploit the sparsity of $\bar{\mathbf{x}}$ and encode $\bar{\mathbf{x}}$ so that we may reduce the amount of space required to store it. One way of doing this would be to encode $\bar{\mathbf{x}}$ as a linear combination of known signals. That is, we can store $\bar{\mathbf{x}}$ as $\mathbf{b} = A\bar{\mathbf{x}}$ for some $A \in \mathbf{R}^{m\times n}$. This encoding of $\bar{\mathbf{x}}$ is only useful if we are also able to decode $\mathbf{b}$ and recover $\bar{\mathbf{x}}$. The reconstruction of a sparse signal from a small number of linear measurements is often referred to as *compressed* or *compressive sensing*. The terminology "compressed sensing" comes from the fact that, in practice, the full signal is never recorded and instead the measurement vector $\mathbf{b} = A\bar{\mathbf{x}}$ is obtained from a series of independent samplings of $\bar{\mathbf{x}}$. Hence, only a "compressed" version of the signal is "sensed". If the matrix A is chosen so that $\bar{\mathbf{x}}$ is the unique minimum cardinality solution of the system $A\mathbf{x} = \mathbf{b}$ then $\bar{\mathbf{x}}$ can be recovered by solving the cardinality minimization problem

$$\begin{array}{ll} \min & \|\mathbf{x}\|_0 \\ \text{s.t.} & A\mathbf{x} = \mathbf{b}. \end{array} \tag{3.3.1}$$

Here $\|\mathbf{x}\|_0$ denotes the number of nonzero entries of the vector $\mathbf{x}$, called the *cardinality*, or ℓ_0-*norm* of $\mathbf{x}$. Note that $\|\cdot\|_0$ is not a norm as it is neither subadditive or homogeneous. The terminology "ℓ_0-norm" arises from the fact that

$$\|\mathbf{x}\|_0 = \lim_{p \downarrow 0} \sum_{i=1}^{n} |\mathbf{x}_i|^p = \lim_{p \downarrow 0} \|\mathbf{x}\|_p^p$$

Thus, the objective in cardinality minimization is equal to the limit as p tends to 0 of the objective in p-norm minimization.

Unfortunately, the cardinality minimization problem (3.3.1) is NP-hard [140]. A popular heuristic for (3.3.1) is ℓ_1-*minimization*, also known as *basis pursuit*. Santosa and Symes [164] propose minimizing the ℓ_1-norm for detecting spikes in seismic data. Chen, Donoho, and Saunders [47] suggest extending this heuristic to signal processing and the cardinality minimization problem. Recall that the ℓ_1-norm on $\mathbf{R}^n$ is equal to

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |\mathbf{x}_i|$$

for all $\mathbf{x} \in \mathbf{R}^n$. The ℓ_1-minimization relaxation of (3.3.1) is the convex program

$$\begin{array}{ll} \min & \|\mathbf{x}\|_1 \\ \text{s.t.} & A\mathbf{x} = \mathbf{b}. \end{array} \tag{3.3.2}$$

This relaxation may be formulated as a linear program by decomposing each $\mathbf{x}$ as $\mathbf{x} = \mathbf{x}^+ - \mathbf{x}^-$ where $\mathbf{x}^+$, $\mathbf{x}^-$ are the magnitudes of the positive and negative entries of $\mathbf{x}$ respectively. That is, $\mathbf{x}_i^+ = \mathbf{x}_i$ if $\mathbf{x}_i \geq 0$ and is 0 otherwise, and $\mathbf{x}_i^- = -\mathbf{x}_i$ for all i such that $\mathbf{x}_i < 0$ and 0 otherwise. The resulting linear program is

$$\begin{array}{ll} \min & \sum_{i=1}^{n} (\mathbf{x}_i^+ + \mathbf{x}_i^-) \\ \text{s.t.} & \mathbf{x} = \mathbf{x}^+ - \mathbf{x}^- \\ & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x}^+, \mathbf{x}^- \geq 0. \end{array} \tag{3.3.3}$$

Therefore, (3.3.2) may be solved efficiently, either as a linear program using the simplex algorithm [144, Chapter 13] or interior-point methods (see [143], [144, Chapter 14]). Many algorithms [116, 50, 191, 70, 14] have recently been developed for ℓ_1-minimization. The software package `SPARSELAB` [57] contains implementations of several heuristics for compressed sensing for use in Matlab.

Moreover, for most large underdetermined systems of linear equations, the minimal ℓ_1-norm solution is also the sparsest solution. Hence, ℓ_1-minimization may be used to efficiently recover a sparse vector from a random sampling of its entries. Indeed, suppose that the matrix $A \in \mathbf{R}^{m\times n}$ has entries independently drawn from the normal distribution with $\mu = 0$, $\sigma^2 = 1$, and $\mathbf{b} = A\bar{\mathbf{x}}$ where $\bar{\mathbf{x}} \in \mathbf{R}^n$ has $\|\bar{\mathbf{x}}\|_0 = k$ for $k << n$. Then, for $m > O(k \log n)$, the optimal solution for the ℓ_1-minimization relaxation (3.3.2) coincides with that for original cardinality minimization problem (3.3.1); see [59, 38]. Similar results hold for A with entries drawn independently from the symmetric Bernoulli distribution, and with rows equal to a random subset of the rows of the Fast Fourier Transform (see [39, 37]).

We conclude by highlighting several similarities between rank minimization and cardinality minimization, and their ℓ_1 and nuclear norm relaxations. Consider the problem of minimizing rank of a diagonal matrix subject to some linear constraints:

$$\begin{array}{ll} \min & \operatorname{rank}(\operatorname{Diag}\mathbf{x}) \\ \text{s.t.} & \mathcal{A}(\operatorname{Diag}\mathbf{x}) = \mathbf{b} \end{array} \tag{3.3.4}$$

where the decision variable is a $n \times n$ diagonal matrix, and the linear map $\mathcal{A} : \mathbf{R}^{n\times n} \to \mathbf{R}^m$ and vector $\mathbf{b} \in \mathbf{R}^m$ are given. Since the rank of a diagonal matrix is equal to the number of nonzero entries on its diagonal, we have $\operatorname{rank}(\operatorname{Diag}\mathbf{x}) = \|\mathbf{x}\|_0$ for all $\mathbf{x} \in \mathbf{R}^n$. Moreover, the operator $\mathcal{A}$ acts only on the diagonal entries of $\operatorname{Diag}\mathbf{x}$. Thus, there exists $A \in \mathbf{R}^{m\times n}$ such that $\mathcal{A}(\operatorname{Diag}\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbf{R}^n$. Cardinality minimization is a special case of rank minimization. Moreover, $\|\operatorname{Diag}\mathbf{x}\|_* = \|\mathbf{x}\|_1$ for all $\mathbf{x} \in \mathbf{R}^n$ since the nuclear norm of a matrix X is simply the ℓ_1-norm of its singular values. In this sense, the ℓ_1-norm heuristic for cardinality minimization is a vector analogue of the nuclear norm heuristic for affine rank minimization. The ℓ_1-norm and nuclear norm share several other properties as well. As can be expected, the ℓ_1-norm is the convex envelope of $\|\cdot\|_0$ on the set $\{\|\mathbf{x}\|_\infty : \|\mathbf{x}\|_\infty \leq 1\}$. When restricted to the vector of singular values $\sigma(X)$ for some matrix $X \in \mathbf{R}^{m\times n}$, this reduces to the fact that the nuclear norm is the convex envelope of rank on the set $\{X : \|X\|_2 \leq 1\}$ since the spectral norm is the max norm of the vector of singular values and rank is the cardinality of the vector of singular values. Furthermore, $\|\cdot\|_1$ is the dual norm of $\|\cdot\|_\infty$. We will see in the next two sections that a number of recovery results for compressive sensing can be extended to the matrix case to show that the minimum nuclear norm and rank solutions provided that a low rank solution exists and the linear map $\mathcal{A}$ is nicely structured.

3.4 Theoretical guarantees for the success of the nuclear norm relaxation

In this section, we provide theoretical guarantees that ensure that a particular low-rank matrix X_0 is the minimum rank matrix in the affine subspace $\{X \in \mathbf{R}^{m\times n} : \mathcal{A}(X) = \mathbf{b}\}$, where $\mathcal{A} : \mathbf{R}^{m\times n} \to \mathbf{R}^p$ is linear and $\mathbf{b} = \mathcal{A}(X_0)$, and can be found by solving convex program $\min\{\|X\|_* : X \in \mathbf{R}^{m\times n},\ \mathcal{A}(X) = \mathbf{b}\}$.

3.4.1 The restricted isometry property and nearly isometric random matrices

In [158], Recht, Fazel, and Parrilo show that if the linear operator $\mathcal{A}$ satisfies a particular restricted isometry property then the minimum rank solution on the affine space defined by $\mathcal{A}$ may be found, with probability exponentially close to 1, by minimizing the nuclear norm over this affine space. Let $\mathcal{A} : \mathbf{R}^{m\times n} \to \mathbf{R}^p$ be a linear map. For every $r \in \{1, \ldots, \min\{m, n\}\}$, the *r-restricted isometry constant* $\delta_r(\mathcal{A})$ is the smallest number such that

$$(1 - \delta_r(\mathcal{A}))\|X\|_F \leq \|\mathcal{A}(X)\| \leq (1 + \delta_r(\mathcal{A}))\|X\|_F \tag{3.4.1}$$

holds for every matrix X with rank at most r. This definition generalizes the restricted isometry property (RIP) for sparse vectors developed by Candès and Tao in [38] to one for low-rank matrices.

Let X_0 be a matrix of rank r, let $\mathcal{A} : \mathbf{R}^{m\times n} \to \mathbf{R}^p$ be a linear map, and let $\mathbf{b} := \mathcal{A}(X_0)$. Using properties of the r-restricted isometry constant $\delta_r(\mathcal{A})$, Recht, Fazel, and Parrilo provide two theorems characterizing when X_0 is equal to the minimizer X^* of $\min\{\|X\|_* : \mathcal{A}(X) = \mathbf{b}\}$. These theorems generalize analogous results for cardinality minimization to low-rank matrix recovery. The first generalizes Lemma 1.2 in [38] to provide a characterization of when X_0 is the unique minimum rank solution of $\mathcal{A}(X) = \mathbf{b}$.

Theorem 3.4.1 ([158, Theorem 3.2]) *Suppose that $\delta_{2r} < 1$ for some integer $r \geq 1$. Then X_0 is the only matrix of rank at most r satisfying $\mathcal{A}(X) = \mathbf{b}$.*

The following theorem guarantees that X_0 is the minimum nuclear norm solution of $\mathcal{A}(X) = \mathbf{b}$ provided the $(5r)$-restricted isometry constant of $\mathcal{A}$ is sufficiently small, and extends Theorems 1 and 2 in [37] to low-rank matrix recovery.

Theorem 3.4.2 ([158, Theorem 3.3]) *Suppose that $\delta_{5r} < 1/10$ for some integer $r \geq 1$. Then $X^* = X_0$.*

Recht, Fazel, and Parrilo show that if the random linear operator $\mathcal{A}$ is chosen according to a probability distribution satisfying certain properties then, with probability tending exponentially to 1 as m, n, p tend to ∞, the operator $\mathcal{A}$ will have small r-restricted isometry constant. Let $\mathcal{A}$ be a random variable that takes values in linear maps from $\mathbf{R}^{m\times n}$ to $\mathbf{R}^p$. The linear map $\mathcal{A}$ is *nearly isometrically distributed* if, for all $X \in \mathbf{R}^{m\times n}$

$$E[\|\mathcal{A}(X)\|^2] = \|X\|_F^2 \tag{3.4.2}$$

and, for all $0 < \epsilon < 1$, we have

$$Pr(|\|\mathcal{A}(X)\|^2 - \|X\|_F^2| \geq \epsilon\|X\|_F^2) \leq \exp\left(-\frac{p}{2}\left(\frac{\epsilon^2}{2} - \frac{\epsilon^3}{3}\right)\right), \tag{3.4.3}$$

and there exists constant $\gamma > 0$ such that

$$Pr\left(\|\mathcal{A}\| \geq 1 + \sqrt{\frac{mn}{p}} + t\right) \leq \exp(-\gamma p t^2) \tag{3.4.4}$$

for all $t > 0$. The following theorem states that if $\mathcal{A} : \mathbf{R}^{m\times n} \to \mathbf{R}^p$ is randomly sampled from a nearly isometric family of linear operators then $\delta_r(\mathcal{A})$ is small for sufficiently large p with probability tending exponentially to 1 as p tends to ∞.

Theorem 3.4.3 ([158, Theorem 4.2]) *Fix $0 < \delta < 1$. If $\mathcal{A} : \mathbf{R}^{m\times n} \to \mathbf{R}^p$ is a nearly isometric random variable, then, for every $1 \leq r \leq \min\{m, n\}$, there exists positive constants c_0, c_1 depending only on δ such that, with probability at least $1-\exp(-c_1 p)$, $\delta_r(\mathcal{A}) \leq \delta$ whenever $p \geq c_0 r(m + n)\log(mn)$.*

Theorem 3.4.3 implies that nearly isometric maps satisfy the hypotheses of Theorem 3.4.1 and Theorem 3.4.2 with probability exponentially close to 1. Therefore, the minimum rank solution on the affine space defined by such a map $\mathcal{A}$ can be efficiently recovered by solving the corresponding nuclear norm minimization relaxation. In particular, the family of random linear transforms whose matrix representations have independently identically distributed (i.i.d.) Gaussian entries, $A_{ij} \approx N(0, 1/p)$, is nearly isometric and, hence, the nuclear norm relaxation of rank minimization subject to affine Gaussian constraints is exact with probability exponentially close to 1 for sufficiently large p. Similarly,

the ensemble of random linear maps with matrix representations with entries sampled i.i.d. from the symmetric Bernoulli distribution

$$A_{ij} = \begin{cases} \sqrt{1/p} & \text{with probability } 1/2, \\ -\sqrt{1/p} & \text{with probability } 1/2, \end{cases}$$

and the family with matrix representations with entries sampled i.i.d. such that

$$A_{ij} = \begin{cases} \sqrt{3/p} & \text{with probability } 1/6, \\ 0 & \text{with probability } 2/3, \\ -\sqrt{3/p} & \text{with probability } 1/6 \end{cases}$$

are also nearly isometric.

The guarantees for recovery of a low-rank matrix provided by Theorems 3.4.1, 3.4.2, and 3.4.3 have been improved upon in subsequent papers. Candès and Plan [35] provide a minimal order of measurements required for recovery of a low-rank matrix, and Mohan and Fazel [136] and Oymak et al. [149] show that exact recovery occurs under weaker assumptions on the r-restricted isometry constants than those of Theorems 3.4.1 and 3.4.2. Moreover, similar results have recently been obtained for other rank minimization heuristics such as Singular Value Projection [104], Reweighted Trace Minimization [136], and ADMIRA [30, 122].

3.4.2 Nuclear norm minimization and the low-rank matrix completion problem

Low rank matrix completion is the problem of recovering a low-rank matrix from a sampling of its entries. This problem appears in many applications. In recommender systems, a vendor would like to provide recommendations to its users. These recommendations are typically based upon a small number of submitted user preferences. Users preference for unrated items must be inferred from these submitted ratings. The resulting completed matrix of all user-ratings is believed to be low-rank because it is hypothesized that only a few factors influence one's preferences. In sensor network localization, the positions of wireless sensors in a region are to be recovered from a matrix containing their relative distances. Often each sensor is only able to estimate distances using signal strength readings from its nearest neighbours. In this case, the remaining entries of the distance matrix will have to be recovered from these partial observations. In particular, the distance squared between each sensor can be obtained by recovering the missing entries of the partially observed

Gram matrix; the Gram matrix is the $n \times n$ matrix with (i, j) entry equal to $p_i^T p_j$, where p_i and p_j are the locations of the ith and jth sensors, respectively. This Gram matrix will have rank equal to two if the sensors are located in the plane or three if they are located in a three dimensional space.

In [36], Candès and Recht show that nuclear minimization can perfectly recover a low-rank matrix M from a random sampling of its entries with probability exponentially close to 1, if the sampling of M is performed in a special way. Specifically, Candès and Recht pose low-rank matrix completion as the rank minimization problem

$$\begin{array}{ll} \min & \operatorname{rank}(X) \\ \text{s.t.} & X_{ij} = M_{ij} \quad \forall\, (i,j) \in \Omega \end{array} \tag{3.4.5}$$

where $X \in \mathbf{R}^{m,n}$, $M \in \mathbf{R}^{m \times n}$ is a given rank-r matrix, and Ω is a given subset of $\{1, \ldots, m\} \times \{1, \ldots, n\}$. If M is the only matrix of rank at most r agreeing with M on Ω, then M can be recovered by solving (3.4.5). In particular, this holds, with high probability, if M is a random rank-r matrix and Ω samples sufficiently many entries of M uniformly at random. Unfortunately, (3.4.5) is NP-hard. Candès and Recht consider the nuclear norm relaxation

$$\begin{array}{ll} \min & \|X\|_* \\ \text{s.t.} & X_{ij} = M_{ij} \quad \forall\, (i,j) \in \Omega. \end{array} \tag{3.4.6}$$

When the matrix M to be recovered has rank r and satisfies some mild assumptions, Candes and Recht show that M can be recovered from a random sampling of at most a polylogarithmic, in m,n, and r, number of its entries by solving (3.4.6). This bound was subsequently improved by Candès and Tao [40], Keshavan et al. [115], Gross [87], and Recht [157].

Note that a low-rank matrix M can not be recovered if it is in the nullspace of the sampling operator. If all observed entries of M are zero then the minimum rank matrix agreeing with the observations is the $m \times n$ 0-matrix. This pathological case is characterized by the row and column spaces of M having high coherence with the standard basis. Let U be a subspace of $\mathbf{R}^n$ of dimension r and P_U be the orthogonal projection onto U. Then the *coherence* of U with respect to the standard basis is defined to be

$$\mu(U) = \frac{n}{r} \max_{1 \le i \le n} \|P_U \mathbf{e}_i\|^2.$$

For any subspace U, $\mu(U)$ is at least 1, with equality when U is spanned by vectors whose nonzero entries all have magnitude $1/\sqrt{n}$. On the other hand, $\mu(U) \le n/r$ with equality when U contains a standard basis vector. Thus, if the row or column spaces of M have high

coherence then the nonzero entries of M will be concentrated in a few rows or columns of M. When the low-rank matrix M has row and column spaces with sufficiently low coherence with the standard basis, M can be recovered from a sampling of a small number of its entries. Indeed, we have the following theorem.

Theorem 3.4.4 ([157, Theorem 1.1]) *Let M be an $m \times n$ matrix of rank r with singular value decomposition $U\Sigma V^*$. Without loss of generality, assume that $m \leq n$, $\Sigma \in \mathbf{R}^{r \times r}$, $U \in \mathbf{C}^{m \times r}$, and $V \in \mathbf{C}^{n \times r}$. Assume that*

A0 *The row and column spaces of M have coherences bounded above by some $\mu_0 > 0$.*

A1 *The matrix UV^* has a maximum entry, in absolute value, bounded by $\mu_1 \sqrt{r/(mn)}$ for some $\mu_1 > 0$.*

Suppose that p entries of M are observed with locations sampled uniformly at random. Then if

$$p \geq 32 \max\{\mu_1^2 \mu_0\} r(m+n)\beta \log^2(2n) \tag{3.4.7}$$

for some $\beta > 1$, the minimizer of (3.4.6) is unique and equal to M with probability at least $1 - 6\log(n_2)(m+n)^{2-2\beta} - n^{2-2\beta^{1/2}}$.

3.4.3 Nullspace conditions

In [159, 160], Recht et al. provide necessary and sufficient conditions for the nuclear norm heuristic for affine rank minimization based on properties of the nullspace of the linear operator $\mathcal{A}$ defining the set of constraints. That is, Recht et al. characterize properties of the nullspace of $\mathcal{A}$ that hold if and only if the minimum nuclear norm solution is equal to the minimum rank solution of $\mathcal{A}(X) = \mathbf{b}$. Moreover, Recht et al. show that the nullspace of linear operators drawn from the Gaussian ensemble satisfy these conditions with extremely high probability. Oymak and Hassibi improve upon these bounds in [147]. These results extend nullspace conditions for compressed sensing by Stojnic [172] to conditions for nuclear norm minimization. Similarly, Dvijothm and Fazel [60] show that nuclear norm minimization recovers the minimum rank solution provided that the nullspace of $\mathcal{A}$ satisfies the spherical section property. A subspace $V \subset \mathbf{R}^{m \times n}$ satisfies the Δ-*spherical section property* if

$$\frac{\|Z\|_*^2}{\|Z\|_F^2} \geq \Delta$$

for all $Z \in V, Z \neq 0$. Dvijothm and Fazel show that if the nullspace of $\mathcal{A}$ has the Δ-spherical section property for sufficiently large Δ then the minimum nuclear norm solution is unique and equal to the unique minimum rank solution of $\mathcal{A}(X) = \mathbf{b}$. This result is a generalization to nuclear norm minimization of a result of Zhang [193] for compressed sensing. Oymak et al. [149] show that if a set of conditions is sufficient for recovery of a sparse vectors then the extension of these conditions to the recovery of low rank matrices is also sufficient. This observation leads to improved nullspace and RIP conditions and provides nullspace conditions for success of the iterative reweighted least squares heuristic of Mohan and Fazel [137] for low rank matrix recovery. This algorithm involves solving the nonconvex optimization problem

$$\begin{array}{ll} \min & \sum_{i=1}^{n} (\sigma_i(X))^p \\ \text{s.t.} & \mathcal{A}(X) = \mathbf{b} \end{array} \tag{3.4.8}$$

for $0 < p < 1$ and can be thought of as a matrix analogue of the ℓ_p-minimization heuristic for cardinality minimization (see [44, 45, 46, 72]).

3.4.4 Minimizing rank while maximizing sparsity

Several recent papers have focused on the problem decomposing a given matrix into low-rank and sparse components. Chandrasekaran et al. [43] consider the relaxation obtained by minimizing a linear combination of the nuclear norm of the low rank component and sparse component of a given matrix $M \in \mathbf{R}^{m \times n}$:

$$\begin{array}{ll} \min & \|L\|_* + \lambda \|S\|_1 \\ \text{s.t.} & L + S = M. \end{array} \tag{3.4.9}$$

Here, $L, S \in \mathbf{R}^{m \times n}$ are our decision variables, the parameter $\lambda > 0$ is given, and $\|S\|_1 = \sum_{i=1}^{m} \sum_{j=1}^{n} |S_{ij}|$ is the ℓ_1-norm of the vectorization of S. Note that $\|\cdot\|_1$ in (3.4.9) is not, in general, equal to the operator norm induced by the ℓ_1-norm. The relaxation (3.4.9) is convex since each of $\|\cdot\|_*$ and $\|\cdot\|_1$ is convex, and can be rewritten as a semidefinite program (see [43, Appendix A]). Moreover, under certain assumptions on the matrix M and parameter λ, this relaxation is exact: solving (3.4.9) returns a decomposition of M into the sum of a low-rank matrix L and a sparse matrix S ([43, Theorem 2]). These results were subsequently improved by Candès et al. [34].

Oymak and Hassibi [148] consider the densest subgraph problem as a special case of finding this decomposition. The *densest subgraph problem* seeks to identify a collection of densely connected subgraphs in a given graph; this may be thought of as a general way of

posing the clustering problem (see [166]). Oymak and Hassibi note that given a graph G containing several densely connected subgraphs, the adjacency matrices of these subgraphs will have low rank while the adjacency matrix $A(G)$ will be sparse outside of the dense subgraphs. Oymak and Hassibi relax the densest subgraph problem as a modification of (3.4.9), and show that their relaxation is exact in the case that the input graph contains several densely connected subgraphs plus a moderate number of diversionary edges and nodes (see [148, Theorems 1 and 2, Corollary 1]). In Section 5.2.2, we consider the densest subgraph problem as a special case of a general graph-partitioning formulation for clustering. We will see that the optimal clustering can be obtained by solving a particular semidefinite relaxation of this general partitioning problem in the case that the input graph consists of several densely connected subgraphs plus a moderate number of diversionary edges and nodes. Similarly, Doan and Vavasis [55] formulate the problem of identifying a large rank-one submatrix of a given nonnegative matrix as simultaneously minimizing the rank and maximizing the sparsity of a nonnegative matrix X subject to linear inequality constraints. This problem arises as a model problem for nonnegative matrix factorization and identifying features in a data set. This problem is relaxed to minimizing a linear combination of the nuclear norm and ℓ_1-norm of the decision variable subject to the same inequality constraints. Under suitable assumptions on the input matrix M and the parameters in the linear combination of $\|\cdot\|_*$ and $\|\cdot\|_1$, this relaxation returns a large approximately rank-one nonnegative submatrix of M (see [55, Theorem 5]).

In [42], Chandrasekaran et al. propose a general convex optimization framework for underdetermined linear inverse problems. When structural constraints require simple solutions, the set of potential solutions may be thought of as linear combinations of a few members from an elementary atomic set. Chandrasekaran et al. propose relaxing the problem of finding a simple model agreeing with given observations to finding a solution minimizing a convex function that agrees with the observed measurements. The convex function to be minimized is the norm induced by the convex hull of the atomic set, called the *atomic norm*. When the atomic set is the set of rank-one matrices, this atomic norm is the nuclear norm. Similarly, if the atomic set is the collection of standard basis vectors, we minimize the ℓ_1-norm. Using this framework, earlier results for compressed sensing and the nuclear norm heuristic may be extended to general linear inverse problems, such as recovering binary vectors, permutation or orthogonal matrices, and low-rank tensors from a small number of generic linear measurements.

Chapter 4

Convex Relaxation for the Maximum Clique and Maximum Edge Biclique Problems

4.1 The maximum clique problem

Let $G = (V, E)$ be a simple graph. A *clique* C of G is a subset of V such that the subgraph of G induced by C, denoted $G(C)$, is complete. That is, C is a clique of G, if the vertices of C are pairwise adjacent in G, i.e. for all $i, j \in C$, $ij \in E$. The *clique number* of G is equal to the cardinality of the maximum clique of G and is denoted $\omega(G)$. The *maximum clique problem* focuses on identifying the largest cliques of the graph G. That is, the maximum clique problem seeks subsets $C \subseteq V$ such that C is a clique, and

$$\omega(G) = \max\{|S| : S \text{ is a clique of } G\} = |C|.$$

The *maximal clique problem* seeks to identify all cliques of G that are not contained as a subset of another clique of G. A subset C of V is an *independent set*, also called a *stable set*, if each pair of vertices in C are nonadjacent in G. That is, a set $C \subseteq V$ is independent if $ij \in \overline{E}$ for all $i, j \in C$, where $\overline{E} := \{ij : (i, j) \in V \times V,\ i \neq j,\ ij \notin E\}$ is the *complement* of the edge set E. The size of the maximum independent set of G, denoted $\alpha(G)$, is called the *stability number* of G. Since every clique C of G is an independent set of the *complement graph* $\overline{G} = (V, \overline{E})$ of G, we have

$$\omega(G) = \alpha(\overline{G}).$$

Therefore, the maximum clique problem and maximum independent set problem are equivalent.

The maximum clique problem arises in such applications as identifying connected components in a computer or telecommunications network [1], coding theory [32], fault diagnosis in multiprocessor systems [18, 94], and pattern recognition [20, 21, 90, 99, 145, 153, 155], among others. Unfortunately, the maximum clique problem is well-known to be NP-hard. In fact, the maximum clique problem is one of the original NP-complete problems identified by Karp [112]. Moreover, the clique number $\omega(G)$ is hard to approximate. Arora et al. [9, 10] show that $\omega(G)$ cannot be approximated for a general N-node graph $G = (V, E)$ within a factor of N^{ϵ}, for $\epsilon > 0$ unless P=NP. This result is a consequence of the observation by Feige et al. [64, 65] that if $\omega(G)$ can be approximated within a factor of $2^{\log^{1-\epsilon} N}$ in polynomial-time, then any problem in NP can be solved in time $2^{\log^{O(1)} N}$ and the fact the MAX 3-SAT problem cannot be approximated to arbitrary small constants unless P=NP. In fact, $\omega(G)$ cannot be approximated within a factor of $N^{1-\epsilon}$, for all $\epsilon > 0$ unless NP admits randomized polynomial time algorithms (see [95]). The best approximation ratio known is $N^{1-o(1)}$ and is achieved by an algorithm developed by Boppana and Halldórsson (see [25]).

Exact algorithms for the maximum clique problem typically involve either enumerating all maximal cliques in the input graph G or all maximal independent sets in its complement $\overline{G}$ (see, for example, [127, 108, 180, 31, 91]), branch and bound schemes used to enumerate only maximal cliques and/or independent sets that improve upon the current best iterate (see [175, 11, 73, 41, 179]) or directly solving integer program or 0-1 quadratic program formulations of the maximum clique problem or maximum independent set problem (see [150, 162, 27, 28, 22, 23]). The best of these algorithms are able to identify the maximum clique of G and $\omega(G)$ with time complexity of $O(2^{0.276N})$ [162], and all maximal cliques of G, and, hence, $\omega(G)$ with time complexity of $O(3^{N/3})$ [180]. This bound for identifying all maximal cliques is optimal since a graph can contain up to $O(3^{N/3})$ cliques [139]. Heuristics for the maximum clique problem include greedy algorithms based on the repeated addition or deletion of vertices from a proposed solution until a maximal clique is obtained [117, 107], randomized local search techniques to improve upon a previously identified clique (see [67, 4, 74, 80, 171, 13]), and solving continuous formulations of the maximum clique problem (see [81, 82]), among others. For a survey of techniques and applications for the maximum clique problem see [24].

4.1.1 Existing approaches for the maximum clique problem

The intractability of the maximum clique problem suggests studying this problem for randomly generated input graphs. For a random graph $G = (V, E)$ on N vertices, with each potential edge added to $E(G)$ independently with fixed probability $1/2$, we have $\omega(G) = (2 + o(1)) \log N$ with extremely high probability (see [6, Corollary 4.5.2]). In [113], Karp proposes a greedy algorithm that is guaranteed to find a clique within the random graph of size $(1 + o(1)) \log N$. It is conjectured that finding a clique of size $(1 + \epsilon) \log N$ is hard [113, 106]. This conjecture has led to finding large cliques in a random graph being used in cryptographic applications [109].

Jerrum [106] and Kučera [120] propose the problem of finding the maximum clique of size n in a random graph on N vertices that has been generated by first placing a clique of size n in the graph and then choosing each remaining pair of vertices to be adjacent independently at random with fixed probability p. We call such a clique a *planted clique*. Kučera shows that if $n = \Omega(\sqrt{N \log N})$ then the nodes of the planted clique will be those with the largest degree with extremely high probability, and, thus, the planted clique can be recovered efficiently. The bound on the size of a planted clique that can be recovered in polynomial time was subsequently improved by Alon, Krivelevich, and Sudakov [5] and Feige and Krauthgamer [66]. In this section, we present algorithms of Alon et al. [5] and Feige and Krauthgamer [66] that recover the planted clique, with extremely high probability, in polynomial time in the case that $n \geq \Omega(\sqrt{N})$. The fact that planted cliques of size $\Omega(\sqrt{N})$ can found efficiently also has cryptographic consequences; Lutomirski et al. [130] show that a particular cryptographic protocol for quantum currency is insecure because cliques of size $\Omega(\sqrt{N})$ may be found in polynomial time in a semirandom graph.

For any graph $G = (V, E)$, the symmetric matrix $A_G \in \mathbf{R}^{V \times V}$ where

$$A_G(i, j) = \begin{cases} 1, & \text{if } ij \in E \\ 0, & \text{if } ij \notin E \end{cases}$$

is called the *adjacency matrix* of G. Note that all eigenvalues of A_G are real since the adjacency matrix is symmetric. The algorithm of Alon et al. exploits properties of the ordered spectrum of the adjacency matrix A_G of a graph G containing a planted clique C. In particular, Alon et al. use the fact that if G contains a planted clique C of size $n = \Omega(N)$, then the eigenvector $\mathbf{v}_2$ corresponding to the second largest eigenvalue $\lambda_2$of A_G is close to a vector whose entries indexed by C are significantly larger in magnitude than its remaining coordinates with probability tending exponentially to 1 as $N \to \infty$. This fact implies that most (at least $5n/6$) of the n largest entries of $\mathbf{v}_2$ are indexed by the planted clique C with probability tending exponentially to 1 as $N \to \infty$. Moreover, this implies

that each vertex in $V \setminus C$ is adjacent to less than $3n/4$ of the n vertices corresponding to the n largest entries of $\mathbf{v}_2$ with probability tending exponentially to 1 as $N \to \infty$ This suggests the following algorithm for exactly solving the maximum clique problem in the case that the input graph G contains a sufficiently large planted clique.

Algorithm 4.1.1 ([5, Algorithm A]) *Given as input a graph $G = (V, E)$ containing a planted clique C of size $n = \Omega(N)$ where each potential nonclique edge is added to E with fixed probability $1/2$:*

1. *Find the eigenvector $\mathbf{v}_2$ corresponding to the second largest eigenvalue λ_2 of the adjacency matrix A_G of G.*
2. *Let $W \subseteq V$ be the set of vertices of G corresponding to the n largest (in absolute value) entries of $\mathbf{v}_2$. Let $Q \subset V$ be the set of all vertices of G that have at least $3n/4$ neighbours in W.*
3. *Output Q.*

Note that both W and Q can be constructed in polynomial time from $\mathbf{v}_2$. Moreover, λ_2 and $\mathbf{v}_2$ can be approximated to arbitrary precision in polynomial time (see [85, Chapter 9]). Moreover, the set Q is equal to the planted clique C with probability tending exponentially to 1 as $N \to \infty$. Therefore, Algorithm 4.1.1 recovers the maximum clique of G in polynomial time with probability tending exponentially to 1 as $N \to \infty$.

On the other hand, the algorithms of Feige and Krauthgamer [66] use the approximation of the stability number $\alpha(\overline{G})$ given by the Lovász theta function $\vartheta(\overline{G})$ (see [128]). For a general graph $G = (V, E)$, the ratio between $\vartheta(\overline{G})$ and the clique number $\omega(G)$ can be as large as $N^{1-o(1)}$ [63] where $N = |V|$. Moreover, for a random graph $G = (V, E)$ on N vertices, with each potential edge added to $E(G)$ independently with fixed probability $1/2$, the clique number $\omega(G)$ is close to $2 \log N$ (see [6, Corollary 4.5.2]) and $\vartheta(\overline{G}) = \Theta(\sqrt{N})$ with probability tending exponentially to 1 as $N \to \infty$ (see [110]). Therefore, the theta function may not provide a good approximation of $\omega(G)$ in general. However, in the case that the graph $G = (V, E)$ on N contains a planted clique of cardinality $\Omega(\sqrt{N})$ the approximation of $\omega(G)$ given by $\vartheta(\overline{G})$ is exact with probability tending exponentially to 1 as $N \to \infty$. Indeed, we have the following result.

Theorem 4.1.1 ([66, Lemma 2]) *Suppose that the graph $G = (V, E)$ contains a planted clique C of size $n \geq c\sqrt{N}$ for sufficiently large constant c where $N = |V|$ and each potential*

edge in E not in the complete subgraph $G(C)$ induced by C is added independently to G with fixed probability $1/2$. Then $\omega(G) = \vartheta(\overline{G}) = n$ with probability tending exponentially to 1 as $N \to \infty$.

Suppose that the graph $G = (V, E)$ satisfies the hypothesis of Theorem 4.1.1. Then Theorem 4.1.1 suggests that the planted clique C of G may be identified by checking each vertex $v \in V$ to see if v is contained in C using $\vartheta(\overline{G})$. Indeed, by Theorem 4.1.1, $\vartheta(\overline{G}) = n$ and, for each $v \in V$, we have

$$\vartheta(\overline{G \setminus v}) = \begin{cases} n-1, & \text{if } v \in C \\ n, & \text{otherwise} \end{cases}$$

with probability tending exponentially to 1 as $N \to \infty$. This fact suggests the following algorithm.

Algorithm 4.1.2 ([66, Algorithm BasicFind]) *Given as input a graph $G = (V, E)$ on N vertices containing a planted clique C of size $n = \Omega(N)$ where each potential nonclique edge is added to E with fixed probability $1/2$:*

1. *Compute $\vartheta(\overline{G})$ and $\vartheta(\overline{G \setminus v})$ for all $v \in V$.*
2. *Output $\vartheta(\overline{G})$ and $P = \{v : \vartheta(\overline{G \setminus v}) < \vartheta(\overline{G}) - 1/2\}$.*

Note that if the planted clique C of G is sufficiently large then Algorithm 4.1.2 outputs the planted clique $C = P$ and $\omega(G) = \vartheta(\overline{G})$ with probability tending exponentially to 1 as $N \to \infty$ by Theorem 4.1.1 and the discussion above. Moreover, for any graph $G = (V, E)$, $\vartheta(\overline{G})$ can be approximated within arbitrary precision in polynomial time by solving the semidefinite program

$$\vartheta_3(\overline{G}) := \max_{X \in \Sigma_+^N} \{\mathrm{Tr}\,(X\mathbf{e}\mathbf{e}^T) : \mathrm{Tr}\,(X) = 1,\ X_{ij} = 0 \text{ if } ij \in E\}; \tag{4.1.1}$$

see [129, Theorem (9.3.12)]. Therefore, the maximum clique C^* of G and the clique number $\omega(G)$ can be identified efficiently using Algorithm 4.1.2 with probability tending exponentially to 1 as $N \to \infty$, provided the input graph G contains the complete subgraph $G(C^*)$ induced by the clique C^* and a moderate amount of vertices and edges other than those of $G(C)$.

Algorithm 4.1.2 can be improved to include only one computation of the theta function. For a graph G on N vertices, a set of unit vectors $\{\mathbf{u}_i \in \mathbf{R}^N : i \in V\}$ is an *orthonormal*

representation of G if $\mathbf{u}_i^T \mathbf{u}_j = 0$ for every pair of nonadjacent vertices $i, j \in V$. For any graph G on N vertices, the theta function $\vartheta(\overline{G})$ is equal to

$$\vartheta_4(\overline{G}) = \max_{\mathbf{d}, \{\mathbf{u}_i\}_{i \in V}} \left\{ \sum_{i \in V} (\mathbf{d}^T \mathbf{u}_i)^2 : \|\mathbf{d}\| = 1, \ \{\mathbf{u}_i\}_{i \in V} \text{ is an orthonormal rep. of } G \right\} \quad (4.1.2)$$

(see [129, Theorem (9.3.12)]). Suppose that the graph $G = (V, E)$ on N vertices contains a planted clique C of size $n \geq c\sqrt{N}$ where each potential nonclique edge is added to E with fixed probability $1/2$. Feige and Krauthgamer [66] show that if $\mathbf{d}$, $\{\mathbf{u}_i\}_{i \in V}$ are optimal for (4.1.2) then the n vertices of the planted clique C correspond to the n largest terms $(\mathbf{d}^T \mathbf{u}_i)^2$ of the summand in (4.1.2) with probability tending exponentially to 1 as $N \to \infty$. Moreover, the solution of (4.1.1) can be efficiently transformed into an optimal solution $\mathbf{d}$, $\{\mathbf{u}_i\}_{i \in V}$ of (4.1.2). Since (4.1.1) can be solved to arbitrary precision in polynomial time, the following algorithm correctly identifies the planted clique C of G in polynomial time with probability tending exponentially to 1 as $N \to \infty$.

Algorithm 4.1.3 ([66, Algorithm ImprovedBasicFind]) *Given as input a graph $G = (V, E)$ on N vertices containing a planted clique C of size $n = \Omega(N)$ where each potential nonclique edge is added to E with fixed probability $1/2$:*

1. *Compute optimal solution $\mathbf{d}, \{\mathbf{u}_i\}$ and optimal value ϑ^* of (4.1.2) within additive error $\epsilon = 1/3$.*
2. *Compute $P = \{i \in V : (\mathbf{d}^T \mathbf{u}_i)^2 > 1/2\}$.*
3. *Output P and ϑ^*.*

It should be noted that the lower bound $n = \Omega(\sqrt{N})$ on the size of the planted clique C of G is the best known for polynomial time recovery of C. Recent work by Frieze and Kannan [75] shows that the maximum clique problem can be solved for an input graph G on N vertices containing a planted clique of size at least $\Omega(N^{1/3}(\log N)^4)$ by maximizing a particular three dimensional tensor, or cubic form, on the set of unit length vectors. Unfortunately, the complexity of this problem, and of maximizing cubic forms on the set of unit vectors in general, is unknown.

4.1.2 The maximum clique problem as rank minimization

Every clique C of a graph $G = (V, E)$ corresponds to a rank-one matrix. Indeed, let $\mathbf{v}$ be the characteristic vector of C:

$$\mathbf{v}_i = \begin{cases} 1, & \text{if } i \in V(C) \\ 0, & \text{otherwise} \end{cases}$$

for all $i \in V$. Clearly, the matrix $\mathbf{v}\mathbf{v}^T$ has rank equal to 1. Moreover, $\mathbf{v}\mathbf{v}^T$ is the adjacency matrix of the graph G' obtained by taking the union of the subgraph $G(C)$ of G induced by C and the set of loops for each $v \in K$. Therefore, if the clique C of G contains at least n vertices then $\mathbf{v}\mathbf{v}^T$ is optimal for the rank minimization problem

$$\begin{aligned} &\min \operatorname{rank}(X) && (4.1.3)\\ &\text{s.t.} \sum_{i\in V}\sum_{j\in V} X_{ij} \geq n^2, && (4.1.4)\\ &\quad X_{ij} = 0, \quad \text{if } ij \in (V \times V) \setminus E, \ \ i \neq j, && (4.1.5)\\ &\quad X \in \{0,1\}^{V\times V}. && (4.1.6) \end{aligned}$$

Unfortunately, the rank minimization problem (4.1.3) is NP-hard.

We consider the convex relaxation of (4.1.3) obtained by replacing the objective function with the nuclear norm. Underestimating $\operatorname{rank}(X)$ with $\|X\|_*$ and ignoring the binary constraint (4.1.6), we obtain the convex program

$$\begin{aligned} &\min \|X\|_* && (4.1.7)\\ &\text{s.t.} \sum_{i\in V}\sum_{j\in V} X_{ij} \geq n^2, && (4.1.8)\\ &\quad X_{ij} = 0 \quad \text{if } ij \in (V \times V) \setminus E, \ \ i \neq j. && (4.1.9) \end{aligned}$$

Note that the factor n^2 in (4.1.8) is the only inhomogeneity in the problem (4.1.7). This implies that if X^* is optimal for (4.1.7), we obtain the optimal solution X^*/n^2 if we replace n^2 by 1 in (4.1.7) and resolve. Therefore, we do not need to know the size of the maximum clique n prior to solving (4.1.7).

Since the maximum clique problem is hard to approximate, we cannot expect this heuristic to provide a good approximation of $\omega(G)$ for all program inputs. Moreover, note that the linear operator defining the linear constraints (4.1.4) and (4.1.5) will not satisfy the sufficient conditions for exactness of the nuclear norm heuristic given by [158] or [36] in the

planted case. Indeed, the constraint (4.1.5) is akin to a matrix completion constraint; we know that the entries of X corresponding to nonedges must be equal to 0. Since the edges of the planted clique are never sampled, in the planted case we have a large rank-one block of X^* that is left unsampled and, hence, the corresponding operator will neither satisfy the RIP condition or be sufficiently incoherent with the standard basis. In spite of this fact, we will show that the nuclear norm relaxation (4.1.7) is *exact* in the case that the input graph G contains a sufficiently large planted clique. That is, if G contains a sufficiently large planted clique C with characteristic vector $\mathbf{v}$, then $\mathbf{v}\mathbf{v}^T$ is the unique optimal solution for (4.1.3) and (4.1.7). Therefore, in this case, we can recover the maximum clique C of G by solving the convex program (4.1.7).

In the following section, we provide bounds on the size of the planted clique and number of diversionary nodes and edges in the input graph G, which guarantee that the nuclear norm relaxation (4.1.7) is exact for the formulation (4.1.3).

4.1.3 Theoretical guarantees of success for nuclear norm relaxation of the planted clique problem

We consider input graphs $G = (V, E)$ containing a planted clique constructed as follows. Let $|V| = N$. We first add the edges ij for each pair (ij) in the vertex set $V^* \subseteq V$ of size n. Then either

(Γ_A) a number r of the remaining $N(N-1)/2 - n(n-1)/2$ potential edges are added to E deterministically by an adversary, or

(Γ_R) each of the remaining potential edges is added to E independently at random with fixed probability $p \in [0, 1)$.

Notice that by our construction of E, V^* is a clique of G of size n. In this section, we provide values of r, n, N that yield G, as constructed above, such that the maximum clique V^* can be recovered by solving the convex relaxation (4.1.7) of the maximum clique problem given by (4.1.3).

In the case that the diversionary edges are added deterministically to the graph G according to (Γ_A) we have the following theorem.

Theorem 4.1.2 *Suppose that the graph $G = (V, E)$ contains a clique V^* of size n. Suppose further that G contains at most r edges not in $G(V^*)$ and each vertex in $V \backslash V^*$ is adjacent to*

at most δn vertices in V^* for some $\delta \in (0,1)$. Then there exists scalar $c_1 < 1/2$ depending only on δ such that if

$$r \leq c_1 n^2$$

then V^* is the unique maximum clique of G and can be recovered by solving the convex program (4.1.7).

These bounds are the best possible up to the constant factors. If the adversary could join a nonclique vertex v to all n clique vertices, then the adversary would have created a clique $V^* \cup v$ containing $n+1$ vertices. Thus, the restriction that a nonclique vertex is adjacent to at most $O(n)$ clique vertices is the best possible. On the other hand, if the adversary could insert $(n+1)(n+2)/2$ edges, then a new clique containing $n+1$ nodes could be created. Thus, the adversary must be a limited to adding $c_1 n^2$ edges for $c_1 < 1/2$.

We have the following theorem for the case when the diversionary edges are added independently at random.

Theorem 4.1.3 *Suppose that the graph $G = (V, E)$ contains a clique V^* of size n. Suppose further that each potential edge of G in $(V \times V) \setminus (V^* \times V^*)$ is added to E according to (Γ_R) with probability $p \in [0, 1)$. Then there exists scalar $c_2 > 0$ depending only on p such that if*

$$n \geq c_2 \sqrt{N},$$

where $N = |V|$, then V^* is the unique maximum clique of G and will correspond to the unique optimal solution of (4.1.7) *with probability tending exponentially to* 1 *as $N \to \infty$.*

Note that the bound given by Theorem 4.1.3 on the size of a planted clique recoverable by our heuristic matches, up to constants, those for Algorithms 4.1.1, 4.1.2, and 4.1.3. We believe our technique is an improvement over the algorithms of Alon et al. and Feige and Krauthgamer in several ways. First, as Theorem 4.1.2 implies, our technique can be applied to find planted cliques in graphs that are constructed deterministically. Moreover, our heuristic can be extended to the maximum edge biclique problem without modification, whereas the algorithms of Alon et al. and Feige and Krauthgamer exploit special properties of the clique problem. Indeed, a biclique in a bipartite graph G is not an independent set in the complement graph $\overline{G}$ and, hence, the theta number cannot be used to approximate the size of the maximum edge biclique. Similarly, the spectral properties of the adjacency matrix of a planted clique used by Alon et al. do not hold for the spectrum of a planted biclique. Moreover, we will show in Chapter 5 that similar techniques can be used to obtain a heuristic for the clustering problem as well.

It is important to note that we do not consider noise in the form of deletions of edges between nodes of the planted clique. If such an edge is deleted then the rank-one matrix corresponding to the planted clique is infeasible for (4.1.7); if the edge uv is deleted then $X_{uv} = 0$ for all X feasible for (4.1.7). In Section 5.2.2, we consider a heuristic for the weighted clustering problem that can recover a planted clique obscured noise in the form of random edge additions and deletions, although the minimum size of the planted clique required to guarantee success of the heuristic is significantly larger than that of Theorem 4.1.3.

4.2 The maximum edge biclique problem

A graph $G = (V, E)$ is *bipartite* if its vertex set V can be partitioned into two disjoint independent sets V_1, V_2. That is, $V_1 \cap V_2 = \emptyset$ and $E \cap (V_1 \times V_1) = E \cap (V_2 \times V_2) = \emptyset$. We use the notation $G = ((V_1, V_2), E)$ to denote a bipartite graph G with disjoint, independent vertex sets V_1, V_2 such that $V(G) = V_1 \cup V_2$ and edge set E. A pair of disjoint, independent subsets $U' \subseteq V$, $V' \subseteq V$ is a *biclique* of the graph $G = (V, E)$, denoted (U', V'), if the subgraph $G(U', V')$ of G induced by (U', V') is a bipartite complete subgraph of G; that is (U', V') is a biclique of G, if every vertex in U' is adjacent to every vertex in V' in $G(U', V')$. The *maximum edge biclique problem* takes as input a graph G and asks for the biclique (U', V') of G that maximizes the product $|U'| \cdot |V'|$ or, equivalently, the number of edges contained in the complete bipartite subgraph of G induced by (U', V'). The *maximum vertex biclique problem* is to identify a biclique (U', V') of the input graph $G = (V, E)$ such that the sum $|U'| + |V'|$ is maximized. Applications of the maximum biclique problem include the problem of identifying products that use similar components to enable the use of subassemblies to decrease assembly time in manufacturing [174], and as a model problem for nonnegative matrix factorization and finding features in images [83].

It should be noted that, unlike the maximum clique problem, the maximum edge and maximum vertex instances of the biclique problem are not equivalent. Indeed, the maximum edge biclique problem is NP-complete (see [151]) while the maximum vertex biclique problem can be solved in polynomial time for bipartite graphs (see [78]) but is NP-complete in general (see [190]). However, if we relax the definition of a biclique such that (U, V) defines a biclique of G if U, V are pairwise adjacent, although U, V may not be independent, i.e. the subgraph of G induced by (U, V) is not necessarily bipartite, then the maximum vertex biclique problem can be solved in polynomial time for general graphs (see [97]).

4.2.1 Relaxation of the maximum edge biclique problem as nuclear norm minimization

Consider a bipartite graph $G = ((U, V), E)$ where the vertex sets U and V have cardinality M and N respectively. Let (U', V') be a biclique of G such that $|U'| = m$ and $|V'| = n$. Then the $m \times n$ block of the adjacency matrix of $G(U', V')$ with rows indexed by U' and columns indexed by V' is equal to the matrix $X = \bar{\mathbf{u}}\bar{\mathbf{v}}^T \in \mathbf{R}^{U \times V}$ where $\bar{\mathbf{u}}, \bar{\mathbf{v}}$ are the characteristic vectors in $\mathbf{R}^U$, $\mathbf{R}^V$ of U' and V' respectively. Note that for every $i \in U$, $j \in V$ such that $ij \in (U \times V) \setminus E$, we have $X_{ij} = 0$. It follows that a biclique (U', V') of G such that $G(U', V')$ contains at least mn edges can be found, provided one exists, by solving the affine rank minimization problem

$$\begin{aligned} &\min \operatorname{rank}(X) && (4.2.1)\\ &\text{s.t.} \sum_{i \in U} \sum_{j \in V} X_{ij} \geq mn, && (4.2.2)\\ &\quad X_{ij} = 0 \quad \text{if } ij \in (U \times V) \setminus E && (4.2.3)\\ &\quad X \in \{0, 1\}^{U \times V}. && (4.2.4) \end{aligned}$$

On the other hand, a rank-one solution X^* satisfying (4.2.2), (4.2.3), and (4.2.4) is exactly the (U, V)-block of the adjacency matrix of the subgraph induced by some biclique of G containing at least mn edges. Like the rank minimization formulation (4.1.3) of the maximum clique problem, the rank minimization problem (4.2.1) is NP-hard. As before, we relax (4.2.1) to a convex program by underestimating $\operatorname{rank}(X)$ with the nuclear norm $\|X\|_*$. We obtain the convex optimization problem

$$\begin{array}{ll} \min & \|X\|_* \\ \text{s.t.} & \sum_{i \in U} \sum_{j \in V} X_{ij} \geq mn, \\ & X_{ij} = 0 \quad \text{if } ij \in (U \times V) \setminus E \end{array} \qquad (4.2.5)$$

Like the nuclear norm relaxation of the maximum clique problem (4.1.7), this relaxation is *exact* in the case that the bipartite graph G contains a sufficiently large planted biclique obscured by diversionary nodes and edges. We consider the cases when these diversionary edges are added deterministically by an adversary and when the diversionary edges are added independently at random.

Suppose that the edge set of the bipartite graph $G = ((U, V), E)$ is generated as follows. We begin by adding the edges of a biclique (U^*, V^*) with $|U^*| = m$, $|V^*| = n$. Then, an adversary is allowed to add at most r of the remaining $MN - mn$ potential edges to the

graph. The following theorem provides an upper bound on r that, under certain conditions, ensures that (U^*, V^*) will be the unique maximum edge biclique of G.

Theorem 4.2.1 *Suppose that the bipartite graph G contains the biclique (U^*, V^*) with $|U^*| = m$, $|V^*| = n$. Suppose further that each vertex in $V \setminus V^*$ is adjacent to at most αm vertices of U^* and each vertex in $U \setminus U^*$ is adjacent to at most βn vertices of V^* for some $\alpha, \beta \in (0, 1)$. Then there exists scalar c_1 depending only on α, β such that if the bipartite graph G contains at most $r \leq c_1 mn$ edges not in $U^* \times V^*$ then (U^*, V^*) is the unique maximum edge biclique of G and can be identified by solving* (4.2.5).

Next, let U, V be disjoint vertex sets with $|V| = N$ and $|U| = \lceil yN \rceil$ for some fixed positive scalar y. Consider $U^* \subseteq U$ and $V^* \subseteq V$ such that $|V^*| = n$ and $|U^*| = m = \lceil zn \rceil$ for fixed positive scalar z. We construct the edge set of the bipartite graph $G = ((U, V), E)$ as follows:

(Ω_1) For all $(i, j) \in U^* \times V^*$, we add $ij \in E$.

(Ω_2) For each of the remaining $(i, j) \in U \times V$, we add edge ij to E independently with fixed probability $p \in (0, 1)$.

Note that (Ω_1) implies that G contains the biclique (U^*, V^*) with mn edges. The following theorem provides a lower bound on n that ensures that (U^*, V^*) is recovered from the unique, optimal solution to (4.2.5).

Theorem 4.2.2 ([8, Theorem 5.2]) *There exists $c_2 > 0$ depending on p, y, z such that for each bipartite graph G constructed via (Ω_1), (Ω_2) with $n \geq c_2\sqrt{N}$ the biclique defined by (U^*, V^*) is the maximum edge biclique of G with probability tending exponentially to* 1 *as $N \to \infty$ and $\bar{\mathbf{u}}\bar{\mathbf{v}}^T$ is the unique solution to the convex relaxation* (4.2.5) *where $\bar{\mathbf{u}}$ is the characteristic vector of U^* in $\mathbf{R}^U$ and $\bar{\mathbf{v}}$ is the characteristic vector of V^* in $\mathbf{R}^V$.*

4.3 A general instance of nuclear norm minimization

We consider the convex optimization problem

$$\min \|X\|_* \tag{4.3.1}$$

$$\text{s.t.} \sum_{i=1}^{M}\sum_{j=1}^{N} X_{ij} \geq mn, \tag{4.3.2}$$

$$X_{ij} = 0 \text{ for } ij \in \overline{E}. \tag{4.3.3}$$

Here, $X \in \mathbf{R}^{M\times N}$, E is a subset of $\{1,\ldots,M\}\times\{1,\ldots,N\}$, and $\overline{E}$ denotes the complement of E. Note that the convex relaxation (4.1.7) of the formulation (4.1.3) for the maximum clique problem is exactly the convex program (4.3.1) when $m = n$, $M = N$, the vertices of the input graph G are labelled $V = \{1,\ldots,N\}$ and E is equal to the union of the set of edges of G and the set of loops $\{ii : i \in V\}$. Similarly, the convex relaxation (4.2.5) corresponds to (4.3.1) in the case that the vertex sets of the bipartite graph G are $U = \{1,\ldots,M\}$, $V = \{1,\ldots,N\}$ and the subset E of $\{1,\ldots,M\}\times\{1,\ldots,N\}$ is equal to the set of edges from U to V. In this section, we provide sufficent conditions for optimality and uniqueness of a solution to (4.3.1) which, in turn, will be used to prove Theorems 4.1.2, 4.1.3, 4.2.1, and 4.2.2 in Section 4.4.

Let $\bar{U}$ be a subset of $\{1,\ldots,M\}$ and $\bar{V}$ be a subset of $\{1,\ldots,N\}$ such that $|\bar{U}| = m$ and $|\bar{V}| = n$. Let $\bar{\mathbf{u}}$, $\bar{\mathbf{v}}$ be the characteristic vectors of $\bar{U}$ and $\bar{V}$ respectively and let $\bar{X} := \bar{\mathbf{u}}\bar{\mathbf{v}}^T$. Note that $\operatorname{rank}(\bar{X}) = 1$. Moreover, the subdifferential of $\|\cdot\|_*$ at $\bar{X}$ is equal to

$$\partial\|\cdot\|_*(\bar{X}) = \left\{ \frac{\bar{\mathbf{u}}\bar{\mathbf{v}}^T}{\sqrt{mn}} + W : W\bar{\mathbf{v}} = 0,\ W^T\bar{\mathbf{u}} = 0,\ \|W\| \leq 1 \right\} \tag{4.3.4}$$

by Lemma 3.2.1 and the fact that $\bar{X}$ has compact singular value decomposition

$$\bar{X} = \sqrt{mn}\cdot\left(\frac{\bar{\mathbf{u}}}{\sqrt{m}}\right)\left(\frac{\bar{\mathbf{v}}}{\sqrt{n}}\right)^T = \bar{\mathbf{u}}\bar{\mathbf{v}}^T.$$

Combining (4.3.4) and the Karush-Kuhn-Tucker conditions for convex programming given by Theorem 2.4.2 yields the main theorem for this section.

Theorem 4.3.1 ([8, Theorem 3.1]) *Let U^* be a subset of $\{1,\ldots,M\}$ of cardinality m, and let V^* be a subset of $\{1,\ldots,N\}$ of cardinality n. Let $\bar{\mathbf{u}}$ and $\bar{\mathbf{v}}$ be the characteristic vectors of U^*, V^* respectively and let $X^* = \bar{\mathbf{u}}\bar{\mathbf{v}}^T$. Suppose X^* is feasible for* (4.3.1) *and that there exist $W \in \mathbf{R}^{M\times N}$, $\lambda \in \mathbf{R}^{M\times N}$ and $\mu \in \mathbf{R}_+$ such that $W\bar{\mathbf{v}} = 0$, $\bar{\mathbf{u}}^T W = 0$, $\|W\| \leq 1$ and*

$$\frac{\bar{\mathbf{u}}\bar{\mathbf{v}}^T}{\sqrt{mn}} + W = \mu\mathbf{e}\mathbf{e}^T + \sum_{ij\in\tilde{E}} \lambda_{ij}\mathbf{e}_i\mathbf{e}_j^T. \tag{4.3.5}$$

Then X^* is an optimal solution to (4.3.1). *Moreover, for any $I \subset \{1,\ldots,M\}$ and $J \subset \{1,\ldots,N\}$ such that $I \times J \subset E$ we have $|I| \cdot |J| \leq mn$.*

Furthermore, if $\|W\| < 1$ and $\mu > 0$, then X^* is the unique optimizer of (4.3.1).

Proof: Note that (4.3.1) satisfies the Slater constraint qualification. For example, $(1+\epsilon)X^*$ is strictly feasible for (4.3.1) for all $\epsilon > 0$. Applying the sufficient conditions for optimality given by Theorem 2.4.2 shows that X^* is optimal for (4.3.1) if there exist $W \in \mathbf{R}^{M\times N}$, $\lambda \in \mathbf{R}^{M\times N}$ and $\mu \in \mathbf{R}_+$ satisfying (4.3.5) such that $W\bar{\mathbf{v}} = 0$, $\bar{\mathbf{u}}^T W = 0$, and $\|W\| \leq 1$.

Now consider (I, J) such that $I \times J \subset E$. Let $X' = \bar{\mathbf{u}}'(\bar{\mathbf{v}}')^T \cdot mn/(|I|\cdot|J|)$, where $\bar{\mathbf{u}}'$ is the characteristic vector of I and $\bar{\mathbf{v}}'$ is the characteristic vector of J. Then X' is also a feasible solution to (4.3.1). Recall that for a matrix of the form $\mathbf{u}\mathbf{v}^T$, the unique nonzero singular value (and hence the nuclear norm) equals $\|\mathbf{u}\| \cdot \|\mathbf{v}\|$. Thus, $\|X'\|_* = mn/(|I| \cdot |J|)^{1/2}$ and $\|X^*\|_* = \sqrt{mn}$. Since X^* is optimal, $\|X'\|_* \geq \|X^*\|$, i.e., $\sqrt{mn} \leq mn/(|I| \cdot |J|)^{1/2}$. Simplifying yields $|I| \cdot |J| \leq mn$.

Finally, we turn to the uniqueness of X^*. The optimization problem (4.3.1) can be formulated as the semidefinite program

$$\begin{array}{ll} \min & \frac{1}{2}(\mathrm{Tr}\,(Z_1) + \mathrm{Tr}\,(Z_2)) \\ \text{s.t.} & Z = \begin{pmatrix} Z_1 & X \\ X^T & Z_2 \end{pmatrix} \succeq 0 \\ & \sum_{i=1}^M \sum_{j=1}^N X_{ij} \geq mn \\ & X_{ij} = 0 \quad \forall (i,j) \in \overline{E} \end{array} \tag{4.3.6}$$

by (3.2.6). This problem is strictly feasible, and, hence, strong duality holds. The dual of (4.3.6) is

$$\begin{array}{ll} \max & mn\mu \\ \text{s.t.} & Q = \begin{pmatrix} I & -\sum_{(i,j)\in\tilde{E}} \lambda_{ij}\mathbf{e}_i\mathbf{e}_j^T - \mu\mathbf{e}\mathbf{e}^T \\ -\sum_{(i,j)\in\tilde{E}} \lambda_{ij}\mathbf{e}_j\mathbf{e}_i^T - \mu\mathbf{e}\mathbf{e}^T & I \end{pmatrix} \succeq 0 \\ & \mu \geq 0. \end{array} \tag{4.3.7}$$

Now suppose that there exist $W, \lambda \in \mathbf{R}^{M\times N}$ and $\mu > 0$ satisfying (4.3.5) such that $W\bar{\mathbf{v}} = 0$, $\bar{\mathbf{u}}^T W = 0$ and $\|W\| < 1$. Notice that

$$X^* = \bar{\mathbf{u}}\bar{\mathbf{v}}^T, \quad Z_1^* = \frac{n}{\sqrt{mn}}\bar{\mathbf{u}}\bar{\mathbf{u}}^T, \quad Z_2^* = \frac{m}{\sqrt{mn}}\bar{\mathbf{v}}\bar{\mathbf{v}}^T$$

forms a primal feasible triple for (4.3.6) and the matrix Q^* as defined by μ and λ is dual feasible. Moreover, $\langle Z^*, Q^* \rangle = 0$, and, thus, by complementary slackness, Z^* is optimal for (4.3.6), Q^* is the corresponding dual optimal solution for (4.3.7), and X^* is optimal for (4.3.1). Also, note that $\bar{\mathbf{u}}\bar{\mathbf{v}}^T/\sqrt{mn} + W$ has maximum singular value equal to 1 with multiplicity 1 since $\|W\| < 1$. Therefore, since υ is an eigenvalue of $Q - I$ if and only if υ is an eigenvalue of $\bar{\mathbf{u}}\bar{\mathbf{v}}^T/\sqrt{mn} + W$ or $-\upsilon$ is an eigenvalue of $\bar{\mathbf{u}}\bar{\mathbf{v}}^T/\sqrt{mn} + W$, Q has exactly one singular value equal to 0. It follows immediately that Q^* has rank equal to $M + N - 1$.

Now suppose that

$$\hat{Z} = \begin{pmatrix} \hat{Z}_1 & \hat{X} \\ \hat{X}^T & \hat{Z}_2 \end{pmatrix}$$

is optimal for (4.3.6) and, hence, $\hat{X}$ is optimal for (4.3.1). Moreover, since Q^* is dual optimal, $\langle \hat{Z}, Q^* \rangle = 0$ by complementary slackness. Therefore, $\hat{Z}$ has rank 1 and must be equal to tZ^* for some nonnegative scalar t. It follows immediately that $\hat{X} = (1/t)\bar{\mathbf{u}}\bar{\mathbf{v}}^T$. However, since $\hat{X}$ is optimal for (4.3.1),

$$\sqrt{mn} = \|\hat{X}\| = \|\bar{\mathbf{u}}\bar{\mathbf{v}}^T\|/t = \sqrt{mn}/t,$$

and, hence, $t = 1$. Therefore, $X^* = \bar{\mathbf{u}}\bar{\mathbf{v}}^T$ is the unique minimizer of (4.3.1). ■

4.4 Proofs of the theoretical guarantees

4.4.1 The adversarial noise case for the maximum clique problem

Suppose that the graph G satisfies the hypothesis of Theorem 4.1.2. That is, the graph $G = (V, E)$ contains a complete subgraph induced by the clique V^* with vertex set of size n and at most r of the remaining potential edges are added deterministically such that each vertex in $V \setminus V^*$ is adjacent to at most δn vertices in V^* for fixed δ in $(0, 1)$. Let $X^* = \bar{\mathbf{v}}\bar{\mathbf{v}}^T$ be the unique optimal solution of (4.1.7) where $\mathbf{v}$ is the characteristic vector of V^*. Clearly X^* is feasible for (4.1.7). We will show in the case that $r \leq O(n^2)$, X^* is the unique optimal solution for (4.1.7) and that V^* can be recovered by solving (4.1.7). To do so, we show that the optimality and uniqueness conditions given by Theorem 4.3.1 are satisfied at X^*. The following theorem provides the necessary specialization of Theorem 4.3.1 for the maximum clique problem (4.1.3) and its nuclear norm minimization relaxation (4.1.7).

Theorem 4.4.1 ([8, Theorem 4.1]) *Let V^* be an n-node clique contained in an N-node undirected graph $G = (V, E)$. Let $\bar{\mathbf{v}} \in \mathbf{R}^V$ be the characteristic vector of V^*. Let $X^* = \bar{\mathbf{v}}\bar{\mathbf{v}}^T$.*

Suppose also that there exist $W \in \mathbf{R}^{V\times V}$, $\lambda \in \mathbf{R}^{V\times V}$ *and* $\mu \in \mathbf{R}_+$ *such that* $W\bar{\mathbf{v}} = 0$, $\bar{\mathbf{v}}^T W = 0$, $\|W\| \leq 1$ *and*

$$\frac{\bar{\mathbf{v}}\bar{\mathbf{v}}^T}{n} + W = \mu \mathbf{e}\mathbf{e}^T + \sum_{(i,j)\in\tilde{E}} \lambda_{ij}\mathbf{e}_i\mathbf{e}_j^T. \tag{4.4.1}$$

Then X^* *is an optimal solution to* (4.1.7). *Moreover,* V^* *is a maximum clique of* G. *Furthermore, if* $\|W\| < 1$ *and* $\mu > 0$, *then* X^* *is the unique optimizer of* (4.1.7), *and* V^* *is the unique maximum clique of* G.

Theorem 4.4.1 follows immediately from the optimality and uniqueness conditions given by Theorem 4.3.1 for the instance of (4.3.1) with $N = |V|$ and edge-set $E = E(G) \cup \{vv : v \in V\}$.

To see that X^* is the unique optimal solution for (4.1.7), we construct multipliers μ, λ, and W satisfying the conditions of the previous theorem. We take $\mu = 1/n$ and define W and λ by considering the following cases.

(ω_1) If $(i,j) \in V^* \times V^*$, we choose $W_{ij} = 0$ and $\lambda_{ij} = 0$. In this case, the entries on each side of (4.4.1) corresponding to this case become $1/n + 0 = 1/n + 0$.

(ω_2) If $(i,j) \in E \setminus (V^* \times V^*)$ such that $i \neq j$, then we choose $W_{ij} = 1/n$ and $\lambda_{ij} = 0$. Then the two sides of (4.4.1) become $0 + 1/n = 1/n + 0$.

(ω_3) If $i \notin V^*$, we set $W_{ii} = 1/n$. Again the two sides of (4.4.1) become $0+1/n = 1/n+0$.

(ω_4) If $(i,j) \notin E$, $i \notin V^*$, $j \notin V^*$, then we choose $W_{ij} = -\gamma/n$ and $\lambda_{ij} = -(1+\gamma)/n$ for some constant $\gamma \in \mathbf{R}$. The two sides of (4.4.1) become $0 - \gamma/n = 1/n - (1+\gamma)/n$. The value of γ is specified below.

(ω_5) If $(i,j) \notin E$, $i \in V^*$, $j \notin V^*$, then we choose

$$W_{ij} = -\frac{p_j}{n(n-p_j)}, \quad \lambda_{ij} = -\frac{1}{n} - \frac{p_j}{n(n-p_j)}$$

where p_j is equal to the number of edges in E from j to V^*.

(ω_6) If $(i,j) \notin E$, $i \notin V^*$, $j \in V^*$ then choose W_{ij}, λ_{ij} symmetrically with the previous case.

Note that for each $i \in V^*$, we have $W(i, V^*) = 0$ and, hence, $W(i,:)\bar{\mathbf{v}} = 0$. For entries $i \in V - V^*$,

$$W(i,:)\bar{v} = p_i \frac{1}{n} - (n - p_i)\frac{p_i}{n(n-p_i)} = 0,$$

by our choice of W_{ij} in cases (ω_5) and (ω_6). Therefore, $W\bar{\mathbf{v}} = 0$ as required.

It remains to show that if $r \leq cn^2$ for sufficiently small c, depending only on δ, then $\|W\| < 1$. Consider W as defined by (ω_1)–(ω_6) with $\gamma = 0$. We decompose W as $W = W^D + W^{ND}$ where $W^D \in \mathbf{R}^{V \times V}$ denotes the matrix with diagonal entries equal to the diagonal entries of W and all other entries equal to 0 and $W^{ND} \in \mathbf{R}^{V \times V}$ is the matrix with nondiagonal entries equal to those of W and diagonal entries equal to 0. By the triangle inequality,

$$\|W\|^2 \leq (\|W^D\| + \|W^{ND}\|)^2 \leq 2(\|W^D\|^2 + \|W^{ND}\|^2) = 2(1/n^2 + \|W^{ND}\|^2)$$

since $\|W^D\| = 1/n$. Since $\|W\| \leq \|W\|_F$ by (2.1.4), it suffices to determine which values of r yield

$$\|W^{ND}\|_F^2 = 2\|W(V^*, V \setminus V^*)\|_F^2 + \|W^{ND}(V \setminus V^*, V \setminus V^*)\|_F^2 < (n^2 - 2)/(2n^2)$$

since, by the symmetry of W,

$$W^{ND}(V^*, V \setminus V^*) = W(V^*, V \setminus V^*) = W(V \setminus V^*, V^*).$$

The diagonal entries of $W^{ND}(V \setminus V^*, V \setminus V^*)$ are equal to 0 and at most $2r$ of the remaining entries are equal to $1/n$. Therefore,

$$\|W^{ND}(V \setminus V^*, V \setminus V^*)\|_F^2 \leq 2r/n^2.$$

Moreover, since $n - p_j \geq (1 - \delta)n$,

$$\begin{aligned}
\|W(V^*, V \setminus V^*)\|_F^2 &= \sum_{j \in V \setminus V^*} \left(p_j \cdot \frac{1}{n^2} + (n - p_j) \cdot \frac{p_j^2}{(n-p_j)^2 n^2} \right) \\
&= \sum_{j \in V \setminus V^*} \left(\frac{p_j}{n^2} + \frac{p_j^2}{(n-p_j)n^2} \right) \\
&\leq \sum_{j \in V \setminus V^*} \left(\frac{p_j}{n^2} + \frac{\delta n p_j}{(1-\delta)n^3} \right)
\end{aligned}$$

$$= \left(\frac{1}{1-\delta}\right) \sum_{j \in V - V^*} \frac{p_j}{n^2}$$
$$\leq \left(\frac{1}{1-\delta}\right) \frac{r}{n^2}.$$

Thus, the optimality and uniqueness conditions given by Theorem 4.4.1 are satisfied by X^* if

$$\left(1 + \frac{1}{1-\delta}\right) r < (n^2 - 2)/4,$$

or, equivalently, if

$$r < \frac{1-\delta}{4(2-\delta)}(n^2 - 2).$$

Therefore, G can contain up to $O(n^2)$ edges other than those in $G(V^*)$, and yet V^* will remain the unique maximum clique of G.

4.4.2 The random noise case for the maximum clique problem

Suppose that the edge set of the graph $G = (V, E)$ is constructed as follows:

(Γ_1) First, the edge set $\{ij : i, j \in V^*, i \neq j\}$ of a clique V^* of size n is added to E.

(Γ_2) Next, each of the remaining potential edges in $(V \times V) \setminus (V^* \times V^*)$ is added to E independently at random with fixed probability $p \in [0, 1)$.

In this section, we show that if $n \geq \Omega(N)$, where $N = |V|$, then V^* is the unique maximum clique of G and can be found by solving (4.1.7) with probability tending exponentially to 1 as N tends to ∞. As in the deterministic case considered in the previous section, we construct multipliers μ, λ, and W that satisfy the conditions of Theorem 4.4.1 to show that the matrix $X^* = \bar{\mathbf{v}}\bar{\mathbf{v}}^T$, where $\bar{\mathbf{v}}$ is the characteristic vector of V^*, is the unique optimal solution for (4.1.7). As before, we take $\mu = 1/n$. We define W and λ according to (ω_1)-(ω_6) with $\gamma = -p/(1-p)$. By construction, $W\bar{\mathbf{v}} = 0$. Therefore, it suffices to show that $\|W\| < 1$ and $p_j < n$ for all $j \in V \setminus V^*$. We first show that $\|W\| < 1$ with probability tending exponentially to 1 as $N \to \infty$ in the case that $n = \Omega(\sqrt{N})$. We write $W = W_1 + W_2 + W_3 + W_4 + W_5$, where each of the five terms is defined as follows.

We first define W_1. For cases (ω_2) and (ω_4), choose $W_1(i, j) = W(i, j)$. For cases (ω_5) and (ω_6), take $W_1(i, j) = -p/((1-p)n)$. For case (ω_1), choose $W_1(i, j)$ randomly such that

$W_1(i,j)$ is equal to $1/n$ with probability p and equal to $-p/((1-p)n)$ otherwise. Similarly, in case (ω_3), take $W_1(i,i)$ to be equal to $1/n$ with probability p and equal to $-p/((1-p)n)$ otherwise. By construction, each entry of W_1 is an independent random variable with the distribution

$$W_1(i,j) = \begin{cases} 1/n & \text{with probability } p, \\ -p/((1-p)n) & \text{with probability } 1-p. \end{cases}$$

Therefore, applying Theorem 2.6.4 shows that there exists constant $c_1 > 0$ such that

$$\|W_1\| \leq 3\left(\frac{p}{1-p}\right)^{1/2} \frac{\sqrt{N}}{n} \tag{4.4.2}$$

with probability at least $1 - \exp(c_1 N^{1/6})$.

Next, W_2 is the correction matrix to W_1 in case (ω_1). That is, $W_2(i,j)$ is chosen such that

$$W_2(i,j) + W_1(i,j) = W(i,j) = 0$$

for all $(i,j) \in V^* \times V^*$ and is zero everywhere else. As before, applying Theorem 2.6.4 shows that

$$\|W_2\| \leq 3\left(\frac{p}{1-p}\right)^{1/2} \frac{1}{\sqrt{n}} \tag{4.4.3}$$

with probability at least $1 - \exp(c_1 n^{1/6})$. Similarly, W_3 is the correction to W_1 in case (ω_3), that is

$$W_3(i,i) = W(i,i) - W_1(i,i)$$

for all $i \in V \setminus V^*$ and all other entries are equal 0. Therefore, W_3 is a diagonal matrix with diagonal entries bounded by $2/n$. It follows that

$$\|W_3\| \leq \frac{2}{n}. \tag{4.4.4}$$

Finally, W_4 and W_5 are the corrections for cases (ω_5) and (ω_6) respectively. We will use the following theorem to obtain upper bounds on $\|W_4\|$ and $\|W_5\|$.

Theorem 4.4.2 ([8, Theorem 2.4]) *Let A be an $n \times N$ matrix whose entries are chosen independently at random such that*

$$A_{ij} = \begin{cases} 1 & \textit{with probability } p, \\ -p/(1-p) & \textit{with probability } 1-p, \end{cases}$$

for all $i = 1, \dots n$, $j = 1, \dots N$. *Suppose also that* $e \log N \leq \sqrt{n}$. *Let* $\tilde{A}$ *be defined as follows. For* (i, j) *such that* $A_{ij} = 1$, *we define* $\tilde{A}_{ij} = 1$. *For entries* (i, j) *such that* $A_{ij} = -p/(1-p)$, *we take* $\tilde{A}_{ij} = -n_j/(n - n_j)$, *where* n_j *is the number of* 1*'s in column* j *of* A. *Then there exist* $c_1 > 0$ *and* $c_2 \in (0, 1)$ *depending on* p *such that*

$$P(\|A - \tilde{A}\|_F^2 \leq c_1 N) \geq 1 - (2/3)^N - N c_2^n. \tag{4.4.5}$$

The proof of Theorem 4.4.2 is provided in Appendix A.1.

Note that W_4 is exactly of the form $(A - \tilde{A})/n$ as in Theorem 4.4.2, in which N in the theorem stands for $N - n$ in the present context. Examining each term of (4.4.5) shows that in the case $n = \Omega(N^{1/2})$, the probability that $\|A - \tilde{A}\|_F^2 = O(N)$ is at least $1 - k_1 \exp(-k_2 N^{k_3})$ for some $k_1, k_2, k_3 > 0$ depending on p. It follows that there exists constant $\alpha_4 > 0$ such that

$$\|W_4\|^2 \leq \|W_4\|_F^2 < \alpha_4^2 N n^{-2}$$

with probability tending exponentially to 1 as $N \to \infty$. Moreover, to obtain this bound on $\|W_4\|$ we implicitly use the fact that Condition Ψ in the proof of Theorem 4.4.2 is not violated; here n_j in the proof stands for p_j, for each $j \in V \setminus V^*$. This implies that $p_j < n$ for all $j \in V \setminus V^*$ with probability tending exponentially to 1 as $N \to \infty$. Finally, notice that, by symmetry, $W_4 = W_5^T$. Thus, since each of $W_1, W_2, \dots, W_5$ is bounded by an arbitrarily small constant if $n = \Omega(\sqrt{N})$, there exists constant $\alpha > 0$ such that $\|W\| < 1$ with probability tending exponentially to 1 as $N \to \infty$ as required.

4.4.3 The adversarial noise case for the maximum edge biclique problem

Suppose that the bipartite graph $G = ((U, V), E)$ contains the biclique (U^*, V^*) with vertex sets U^*, V^* of size m, n respectively. Suppose that each vertex in $V \setminus V^*$ is adjacent to at most αm vertices in U^* and each vertex in $U \setminus U^*$ is adjacent to at most βn vertices in V^* for some fixed $\alpha, \beta \in (0, 1)$. In this section, we show that if G contains at most $r \leq O(mn)$ of the $|U||V| - mn$ remaining possible edges then (U^*, V^*) is the maximum edge biclique of G and can be recovered by solving (4.2.5). We will use the following specialization of Theorem 4.3.1, which provides sufficient conditions for which the adjacency matrix of a subgraph induced by a biclique of G is optimal for the nuclear norm relaxation (4.2.5) of the maximum edge biclique problem.

Theorem 4.4.3 ([8, Theorem 5.1]) *Let (U^*, V^*) be a biclique of G such that $|U^*| = m$ and $|V^*| = n$. Let $\bar{\mathbf{u}} \in \mathbf{R}^M$ be the characteristic vector of U^*, and let $\bar{\mathbf{v}} \in \mathbf{R}^N$ be the characteristic vector of V^*. Let $X^* = \bar{\mathbf{u}}\bar{\mathbf{v}}^T$. Note that X^* is feasible for (4.2.5). Suppose also that there exist $W \in \mathbf{R}^{M\times N}$, $\lambda \in \mathbf{R}^{M\times N}$ and $\mu \in \mathbf{R}_+$ such that $W\bar{\mathbf{v}} = 0$, $\bar{\mathbf{u}}^T W = 0$, $\|W\| \leq 1$ and*

$$\frac{\bar{\mathbf{u}}\bar{\mathbf{v}}^T}{\sqrt{mn}} + W = \mu \mathbf{e}\mathbf{e}^T + \sum_{(i,j)\in\tilde{E}} \lambda_{ij}\mathbf{e}_i\mathbf{e}_j^T. \tag{4.4.6}$$

Then X^* is an optimal solution to (4.2.5). Moreover, every biclique (U', V') of G satisfies $|U'| \cdot |V'| \leq mn$. Furthermore, if $\|W\| < 1$ and $\mu > 0$, then X^* is the unique optimizer of (4.2.5) and (U^*, V^*) is the unique optimal biclique.

As before, Theorem 4.4.3 is obtained from the optimality and uniqueness conditions given by Theorem 4.3.1 for the instance of (4.3.1) with $M = |U|$, $N = |V|$ and edge-set $E = E(G)$.

Let $\bar{\mathbf{u}}$, $\bar{\mathbf{v}}$ be the characteristic vectors of U^* and V^* respectively. Let $X^* = \bar{\mathbf{u}}\bar{\mathbf{v}}^T$. To see that X^* is the unique optimal solution of (4.2.5) in the case that $r \leq cmn$ for some scalar c depending only on α, β, we construct multipliers μ, λ, W that satisfy the optimality and uniqueness conditions for (4.2.5) given by Theorem 4.4.3. We take $\mu = 1/\sqrt{mn}$ and consider W and λ defined according to the following cases.

(ψ_1) For $ij \in U^* \times V^*$, taking $W_{ij} = 0$ and $\lambda_{ij} = 0$ ensures the ij-entries of both sides of (4.4.6) are equal to $1/\sqrt{mn}$.

(ψ_2) For $ij \in E \setminus (U^* \times V^*)$, we take $W_{ij} = 1/\sqrt{mn}$ and $\lambda_{ij} = 0$. Again, the ij-entries of both sides of (4.4.6) are equal to $1/\sqrt{mn}$.

(ψ_3) For $ij \notin E$ such that $i \notin U^*$ and $j \notin V^*$, we select $W_{ij} = -\gamma/\sqrt{mn}$ and $\lambda_{ij} = -(1+\gamma)/\sqrt{mn}$ where γ will be defined below. In this case, the ij-entries of each side of (4.4.6) are 0.

(ψ_4) For $ij \notin E$ such that $i \notin U^*$ and $j \in V^*$, we choose

$$W_{ij} = -\frac{p_i}{(n-p_i)\sqrt{mn}} \text{ and } \lambda_{ij} = \frac{1}{\sqrt{mn}}\left(\frac{-p_i}{n-p_i} - 1\right)$$

where p_i is equal to the number of edges with left endpoint equal to i and right endpoint in V^*. Note that if $n = p_i$ then i is connected to every vertex of V^* and

thus the KKT condition cannot possibly be satisfied. If $p_i < n$, both sides of (4.4.6) are equal to $-p_i/((n-p_i)\sqrt{mn})$.

(ψ_5) For $ij \notin E$ such that $i \in U^*$ and $j \notin V^*$, we choose

$$W_{ij} = -\frac{q_j}{(m-q_j)\sqrt{mn}} \text{ and } \lambda_{ij} = \frac{1}{\sqrt{mn}}\left(\frac{-q_j}{m-q_j} - 1\right)$$

where q_j is equal to the number of edges with right endpoint equal to j and left endpoint in U^*. As before, this is appropriate only if $q_j < m$.

Consider $\gamma = 0$. As in Section 4.4.1, we use the bound $\|W\| \leq \|W\|_F$. Notice that at most r entries of $W(U \setminus U^*, V \setminus V^*)$ are equal to $1/\sqrt{mn}$ and the remainder are equal to 0. Therefore,

$$\|W(U \setminus U^*, V \setminus V^*)\|_F^2 \leq \frac{r}{mn}.$$

Moreover, for each $j \in V - V^*$, $q_j \leq \alpha m$. It follows that

$$\begin{aligned}
\|W(U^*, V \setminus V^*)\|_F^2 &= \sum_{v \in V \setminus V*} \left(\frac{q_v}{mn} + (m-q_v)\frac{q_v^2}{mn(m-q_v)^2}\right) \\
&= \sum_{v \in V^*} \frac{q_v}{mn}\left(1 + \frac{q_v}{m-q_v}\right) \\
&\leq \sum_{v \in V^*} \frac{q_v}{mn}\left(1 + \frac{\alpha}{1-\alpha}\right) \\
&= \sum_{v \in V^*} \frac{q_v}{mn(1-\alpha)} \leq \frac{r}{mn(1-\alpha)}.
\end{aligned}$$

Similarly,

$$\|W(U \setminus U^*, V^*)\|_F^2 \leq \frac{r}{(1-\beta)mn}.$$

Therefore, $\|W\| < 1$ if

$$r\left(1 + \frac{1}{1-\alpha} + \frac{1}{1-\beta}\right) < mn.$$

4.4.4 The random noise case for the maximum edge biclique problem

Let y, z be fixed positive scalars. Let U, V be two disjoint vertex sets with $|V| = N$ and $|U| = \lceil yN \rceil$. Consider $U^* \subseteq U$ and $V^* \subseteq V$ such that $|V^*| = n$ and $|U^*| = m = \lceil zn \rceil$. Suppose the edges of the bipartite graph $G = ((U, V), E)$ are determined as follows:

(β_1) For all $(i, j) \in U^* \times V^*$, $ij \in E$.

(β_2) For each of the remaining potential edges $(i, j) \in U \times V$, we add each edge ij independently to E with probability p.

Notice that G contains the biclique (U^*, V^*). Let $\bar{\mathbf{u}}$,$\bar{\mathbf{v}}$ be the characteristic vectors of U^* and V^* respectively. We show that if $n = \Omega(\sqrt{N})$ and G is constructed as in (β_1), (β_2) then $\bar{\mathbf{u}}\bar{\mathbf{v}}^T$ is optimal for the convex problem (4.2.5).

Let W be constructed as in (ψ_1)–(ψ_5) with $\gamma = -p/(1-p)$. Then $X^* = \bar{\mathbf{u}}\bar{\mathbf{v}}^T$ is the unique optimal solution of (4.2.5) if

$$\|W\| < 1, \quad q_j < \lceil zn \rceil \; \forall \, j \in V \setminus V^*, \;\; \text{and } p_j < n \; \forall \, j \in U \setminus U^*$$

by Theorem 4.4.3. To prove that $\|W\| < 1$ with high probability as $N \to \infty$ in the case that $n = \Omega(\sqrt{N})$, we write

$$W = W_1 + W_2 + W_3 + W_4$$

where each of the summands is defined as follows. We first define W_1. If $(i, j) \in U^* \times V^*$, then we set $W_1(i, j) = 1/\sqrt{mn}$ with probability p and equal to $\gamma/\sqrt{mn}$ with probability $1 - p$. For $(i, j) \in (U \times V) \setminus (U^* \times V^*)$, we set $W_1(i, j) = 1/\sqrt{mn}$ if $ij \in E$ and set $W_1(i, j) = \gamma/\sqrt{mn}$ otherwise. Note that Theorem 2.6.5 implies that $\|W_1\| \leq \alpha\sqrt{N}/\sqrt{mn}$ with probability at least $1 - c_1 \exp(-c_2 n^{c_3})$ for some $c_1, c_2, c_3 > 0$ Since $\sqrt{mn}$ equals $\sqrt{\lceil zn \rceil n}$ and hence is proportional to n, we see that $\|W_1\| \leq$ const with probability exponentially close to 1 provided $n = \Omega(N^{1/2})$.

Next, set W_2 to be the correction matrix for W_1 for $U^* \times V^*$, that is,

$$W_2(i, j) = \begin{cases} -W_1(i, j) & \text{if } (i, j) \in U^* \times V^* \\ 0 & \text{otherwise,} \end{cases}$$

Again, by Theorem 2.6.5 we conclude that

$$\|W_2\| \leq \alpha \frac{1}{\sqrt{n}}$$

with probability at least $1 - c_1' \exp(-c_2' n^{c_3'})$ for some $c_1', c_2', c_3' > 0$.

It remains to derive bounds for $\|W_3\|$ and $\|W_4\|$. Notice that the construction of $W(U^*, V \setminus V^*)$ and $W(U \setminus U^*, V^*)$ is identical to that in Case (ω_5) for the maximum clique problem. Thus, we can again apply Theorem 4.4.2, first to W_3 (in which case (n, N) in the theorem stand for $(\lceil zn \rceil, N - n)$) and second to W_4^T (in which case (n, N) in the theorem stand for $(n, \lceil yN \rceil - \lceil zn \rceil)$ to conclude that $\|W_3\|$ and $\|W_4\|$ are both strictly bounded above by constants provided $n = \Omega(N^{1/2})$ with probability tending to 1 exponentially fast as required. Moreover, as in the proof of Theorem 4.1.3, the application of Theorem 4.4.2 to bound W_3 and W_4 implies that $q_j < \lceil zn \rceil \ \forall \, j \in V \setminus V^*$ and $p_j < n \ \forall \ j \in U \setminus U^*$ with probability tending exponentially to 1 as $N \to \infty$ as required.

Chapter 5

Semidefinite Relaxations for the Clustering Problem

5.1 The clustering and k-disjoint clique problems

In this section, we show how our approach for identifying the maximum clique in a graph can be extended to identifying collections of disjoint cliques of maximum size in a graph. In particular, we consider the k-disjoint-clique problem. Let $G = (V, E)$ be a simple graph on $N = |V|$ nodes. We call a subgraph of G composed of k disjoint cliques a *k-disjoint-clique subgraph* of G. That is a subgraph $K = (V', E')$, is a k-disjoint-clique subgraph of G, if $V' = C_1 \cup C_2 \cup \cdots \cup C_k$ where $\{C_1, C_2, \ldots, C_k\}$ is a collection of k disjoint cliques of G and the edge set E' is exactly the set of edges of the complete subgraphs of G induced by $\{C_1, C_2, \ldots, C_k\}$. Given graph $G = (V, E)$ and integer k, the *maximum node k-disjoint-clique problem* (KDC) is to identify a k-disjoint-clique subgraph of G that contains the maximum number of nodes. Given graph $G = (V, E)$, integer k, and nonnegative edge weights $W \in \mathbf{R}^E_+$, the *maximum mean weight k-disjoint-clique problem* (WKDC) seeks the k-disjoint-clique subgraph K of G that maximizes the sum of the average edge-weights of the complete subgraphs of G which compose K. That is, the maximum mean weight k-disjoint-clique problem seeks the set of disjoint cliques $C_1, \ldots, C_k$ of G that maximizes

$$\sum_{i=1,\ldots,k} \frac{1}{|C_i|} \sum_{u,v \in C_i} W_{uv}. \tag{5.1.1}$$

Note that both KDC and WKDC are NP-complete. Indeed, in the special case that $k = 1$ and the weight matrix W is equal to the adjacency matrix A_G of the input graph G, both KDC and WKDC are equivalent to the maximum clique problem. However, as in our earlier discussion of the maximum clique and biclique problems, we will show that we can obtain the optimal solution in polynomial time by solving a particular convex relaxation, in the special case that the given instance of the k-disjoint-clique problem corresponds to a collection of planted cliques, obscured by a moderate number of noise edges and nodes.

Our interest in the k-disjoint-clique problem is motivated by its application in clustering. Given a set of data, *clustering* seeks to divide the data into groups of similar objects, called *clusters*. Clustering is a fundamental problem in statistics and machine learning and plays a significant role in a wide range of applications, including but not limited to information retrieval, pattern recognition, computational biology, and image processing. The complexity of the clustering problem depends significantly on the measure of similarity of items in the data set, but, in general, clustering is an intractable combinatorial problem. For most practical applications, heuristics are used to find clusters in the data sets. In particular, several optimization based heuristics have recently been proposed [168, 152, 148, 105]. We present an overview of spectral and graph based heuristics in Section 5.1.1. For a recent survey of clustering techniques and heuristics see [17].

Consider the following graph-theoretic representation of data. Suppose we have a set of N items, where each pair of items is known to be similar or dissimilar. We consider the N-node graph $G = (V, E)$ with vertex set V equal to the set of items and with each pair of nodes adjacent if and only if the corresponding items are similar. In this case, clustering of the data into k disjoint clusters such that the items in each cluster are pairwise similar is equivalent to partitioning the graph into k disjoint cliques. Such a partitioning, if one exists, can be identified by solving this instance of the maximum node k-disjoint-clique problem. On the other hand, suppose that we have a symmetric matrix $W \in \Sigma^N$ such that W_{ij} is given by some measure of similarity between objects i and j in the data set. Clustering this data set yields a partitioning such that the items in the same cluster are highly similar and items in different clusters are highly dissimilar. That is, the entries of W corresponding to edges between objects in the same cluster should be large while the entries of W corresponding to edges between objects in different clusters should be small. Therefore, we may partition the data set into k clusters by solving the corresponding instance of WKDC. Recent results by Ostrovsky et al. [146] and Ackerman and Ben-David [2, 3] show that if a data set is "clusterable" in some sense, then certain clustering heuristics will be able to find the correct partitioning of the data into clusters. Our results are of a similar form. If the underlying data set consists of k "good" clusters, then the k-disjoint-clique subgraph given by these clusters is optimal for the k-disjoint-clique problem.

The rest of this chapter is organized as follows. In Section 5.1.1 we provide an overview of graph-partitioning approaches to clustering and relate these approaches to ours. In Sections 5.2.1 and 5.2.2, we relax the maximum node and maximum mean weight k-disjoint-clique problems, respectively, to semidefinite programming. We provide general conditions for input instance of the k-disjoint-clique problem, that ensure that the optimal k-disjoint-clique subgraph can be recovered by solving our semidefinite relaxation. Specifically, the set of instances for which our algorithm successfully identifies the optimal k-disjoint-clique subgraph are exactly those corresponding to a collection of k disjoint planted cliques obscured by additional diversionary nodes and edges. Section 5.3 contains a derivation of the bounds on the amount of noise and sizes of cliques that are algorithm for KDC can tolerate. The proof that our convex relaxation exactly recovers the maximum node k-disjoint-clique subgraph constructs multipliers satisfying the Karush-Kuhn-Tucker optimality conditions. The necessary optimality and uniqueness conditions, as well as our choice of multipliers, can be found in Section 5.3.1. Similarly, the proof that our convex relaxation is exact in the planted case for WKDC is contained in Section 5.4. The required optimality conditions are developed in Section 5.4.1.

5.1.1 Graph-partitioning approaches to clustering

The maximum node and maximum mean weight k-disjoint-clique problems are examples of graph-partitioning approaches to clustering. Given a collection of N items, the *similarity graph* $G_S = (V, E, W)$ is the complete graph on N vertices with vertex set equal to the set of items, and with weight W_{uv} assigned to each edge uv indicating the similarity between items u, v. Graph-partitioning attempts to cluster the data by finding an optimal partitioning into cliques of the nodes of the similarity graph G_S, with respect to some objective function given by the edge-weights W. The objective function typically seeks to either maximize the weight of edges of the complete subgraphs of G_S induced by the cliques or to minimize the weights of the edges cut by the partitioning of the nodes of G_S. Hence, the optimal partitioning of the graph will yield a partitioning of the data in clusters such that the items in each cluster will be very similar or the items in different clusters will be very dissimilar.

Given a similarity graph $G_S = (V, E, W)$, and subsets A, B of V such that $A \cap B = \emptyset$, the value of the *cut* defined by (A, B) is equal to the weight of the edges from A to B:

$$c(A, B) = \sum_{u \in A, v \in B} W_{uv}. \tag{5.1.2}$$

Wu and Leahy [189] propose partitioning the graph using the minimum cut. Specifically,

they seek to partition the graph into k cliques such that the maximum of the minimum cuts defined by each clique is minimized. Since the minimum cut of a graph can be computed efficiently (see [71]), such a partition can be computed efficiently by recursively finding the minimum cuts bisecting the existing segments. Unfortunately, since the value of the cut increases as the size of A increases, the partitioning of G_S given by Wu and Leahy's algorithm favours cutting small sets of isolated nodes. To avoid this bias, a number of alternate cut functions normalizing the value of the cut have been proposed. Wei and Chang [187, 188] and Leighton and Rao [123] independently proposed normalizing the cut by the number of edges cut by (A, B), by replacing the objective with the *ratio cut* or *minimum mean cut*

$$Rcut(A, B) = \frac{c(A, B)}{|A| \cdot |B|}. \tag{5.1.3}$$

Shi and Malik [167] propose a clustering method based on finding the partitioning of G_S that minimizes the *normalized cut* defined by

$$Ncut(A, B) = \frac{c(A, B)}{assoc(A, V)} + \frac{c(B, A)}{assoc(B, V)} \tag{5.1.4}$$

for all $A, B \subseteq V$ where

$$assoc(A, B) = \sum_{u \in A, v \in B} W_{uv} \tag{5.1.5}$$

for all $A, B \subseteq V$. Ding et al [53] attempt to simultaneously maximize the similarity between each items within the same grouping and minimize similarity between different clusters by minimizing the *min-max cut* given by

$$Mcut(A, B) = \frac{c(A, B)}{assoc(A, A)} + \frac{c(A, B)}{assoc(B, B)} \tag{5.1.6}$$

for all $A, B \subseteq V$. On the other hand, Sarkar and Boyer [165] attempt to cluster by finding a partitioning the similarity graph into cliques $\{C_1, \ldots, C_k\}$ such that the average associations of the cliques, given by

$$\frac{assoc(C_i)}{|C_i|} \tag{5.1.7}$$

for all $i = 1, \ldots, k$, are maximized. Note that the objective value of a particular k-disjoint-clique subgraph for the maximum mean weight k-disjoint-clique problem is equal to the sum of the average associations of the cliques in its vertex set. For a comparison of graph-

partitioning and other heuristics for clustering please see [167, 111, 195, 196].

Unfortunately, finding the optimal partitionings with respect to the modified cut functions given by (5.1.3), (5.1.4), (5.1.6) is NP-hard (see [132], [167], [53] respectively). However, approximate solutions may be obtained using properties of the spectrum of the Laplacian matrix $D-W$, where $D = \operatorname{Diag} \mathbf{d}$, $\mathbf{d}_i = \sum_{j \in V} W_{ij}$ for all $i \in V$. These properties were first developed by Donath and Hoffman [56] and Fiedler [68, 69], and their application to the clustering problem is collectively referred to as *spectral clustering*. For example, Hagen and Kahng [89] show that for given vertex sets $A, B \subseteq V$, the value of the cut defined by A, B is equal to

$$c(A, B) = \frac{\mathbf{x}^T (D - W)\mathbf{x}}{2}. \tag{5.1.8}$$

where $\mathbf{x} \in \{1, -1\}^V$ is chosen such that $\mathbf{x}_u = 1$ if $u \in A$, and $\mathbf{x}_u = -1$ if $u \in B$ otherwise Relaxing the restriction that $\mathbf{x} \in \{1, -1\}^V$ to $\mathbf{x} \in S^{|V|-1}$ yields a continuous relaxation whose solution is a solution of the eigensystem

$$(D - W)\mathbf{x} = \lambda \mathbf{x}. \tag{5.1.9}$$

The matrix $D-W$ is positive semidefinite [154]. Moreover, $D-W$ has smallest eigenvalue λ_1 equal to 0 with corresponding eigenvector $\mathbf{x} = \mathbf{e}$. Hagen and Kahng prove that the value of the optimal ratio cut partition is approximated by

$$\lambda = \min_{\mathbf{x}^T \mathbf{e} = 0, \mathbf{x} \neq 0} \frac{\mathbf{x}^T (D - W)\mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \lambda_2 \tag{5.1.10}$$

with optimal value equal to the second smallest eigenvalue of $D - W$ and optimal solution equal to the corresponding unit eigenvector (see [101, Equation (4.2.7)]). Therefore, an approximation of the optimal ratio cut partition can be obtained by taking the spectral decomposition of the Laplacian $D - W$. Similarly, Shi and Malik [167] show that an approximate minimum normalized cut can be obtained by solving the generalized eigensystem $(D - W)\mathbf{x} = \lambda D\mathbf{x}$ by finding the second smallest eigenvector of the normalized Laplacian $D^{-1/2}(D-W)D^{-1/2}$ and that an approximate k-way normalized cut can be obtained from the top k eigenvectors of $D^{-1/2}(D-W)D^{-1/2}$. Dhillon et al. [51, 52] show that this approximate solution may be computed iteratively using a variant of the k-means algorithm for clustering [92, 93] that exploits information from the weight matrix W. Similarly, Rahimi and Recht [156] show that clustering using the normalized cut is equivalent to clustering using hyperplane methods and propose a variant that is less sensitive to outliers in the data. On the other hand, Ding et al. [53] obtain an approximate min-max cut using the second eigenvector of $(D-W)$ and Sarkar and Boyer [165] use the eigenvector of the largest

eigenvalue of W to approximate the subgraph with largest average association.

Our approach to the partitioning problems differs significantly from those of spectral clustering. Instead of finding an approximately optimal partitioning of the nodes of the similarity graph, we will show that our approach efficiently finds the *exact* optimal partition of the similarity graph (with respect to the k-disjoint-clique objective function), provided the data set is sufficiently clusterable. That is, if the similarity graph contains k large disjoint cliques, either in cardinality in the maximum node case or in weight in the maximum mean weight case, obscured by noise then we can find the exact solution for this instance of the k-disjoint-clique problem by solving a certain semidefinite program.

5.2 Theoretical guarantees for the success of convex relaxation of the k-disjoint clique problem

Although the k-disjoint-clique problem is known to be NP-hard, we will show it may solved in polynomial time by solving a particular semidefinite program for certain program inputs. In particular, these problems can be solved efficiently in the case that the optimal solution consists of k disjoint planted cliques plus diversionary edges and nodes. In this section, we provide theoretical bounds on the size of the planted cliques and the amount of noise that ensure that our convex relaxation is exact. We begin with the maximum node k-disjoint clique problem.

5.2.1 A convex relaxation for the maximum node k-disjoint clique problem

Given a graph $G = (V, E)$, the maximum node k-disjoint-clique problem can be formulated as the quadratically constrained binary optimization problem

$$\max_{S=\{\mathbf{v}_1,\ldots,\mathbf{v}_k\}} \sum_{i=1}^{k} \mathbf{v}_i^T \mathbf{e} \tag{5.2.1}$$

$$\text{s.t. } \mathbf{v}_i^T \mathbf{v}_j = 0, \quad \forall\, i, j = 1, \ldots, k,\ i \neq j \tag{5.2.2}$$

$$[\mathbf{v}_i \mathbf{v}_i^T]_{uv} = 0, \quad \text{if } uv \notin E,\ u \neq v, \quad \forall\, i = 1, \ldots, k \tag{5.2.3}$$

$$\mathbf{v}_i \in \{0, 1\}^V, \quad \forall\, i = 1, \ldots, k. \tag{5.2.4}$$

A feasible set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ for (5.2.1) is exactly the collection of characteristic vectors of a set of disjoint cliques of G. Indeed, the constraint (5.2.2) ensures that the sets of nodes indexed by S are disjoint and the constraint (5.2.3) ensures that the sets of nodes indexed by S are subgraphs and, hence, cliques of G. Moreover, note that a feasible solution $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ need not define a partition of V. That is, a feasible solution need not correspond to a k-disjoint-clique subgraph of K_N that contains every node in V. Unfortunately, KDC, and finding the solution to a nonlinear program with binary constraints, is NP-hard in general. The formulation (5.2.1) may be relaxed to the rank constrained semidefinite program

$$\begin{aligned} \max \ & \langle X, \mathbf{e}\mathbf{e}^T \rangle && (5.2.5)\\ \text{s.t. } & X\mathbf{e} \leq \mathbf{e}, && (5.2.6)\\ & X_{ij} = 0, \quad \text{if } (i,j) \notin E \text{ s.t. } i \neq j && (5.2.7)\\ & \operatorname{rank}(X) = k, && (5.2.8)\\ & X \succeq 0. && (5.2.9) \end{aligned}$$

Here $\langle \cdot, \cdot \rangle : \mathbf{R}^{N \times N} \to \mathbf{R}$ is the inner product on $\mathbf{R}^{N \times N}$ defined by $\langle Y, Z \rangle = \operatorname{Tr}(YZ^T)$ for all $Y, Z \in \mathbf{R}^{N \times N}$. To see that (5.2.5) is a relaxation of (5.2.1), suppose that $\{C_1, \ldots, C_k\} \subseteq V$ defines a k-disjoint-clique subgraph of G. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be the set of characteristic vectors of $\{C_1, \ldots, C_k\}$. The matrix

$$X = \sum_{i=1}^{k} \frac{\mathbf{v}_i \mathbf{v}_i^T}{|C_i|} \tag{5.2.10}$$

is positive semidefinite with rank equal to k. Note that $\sum_{i=1}^{k} \mathbf{v}_i \leq \mathbf{e}$ since $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are orthogonal binary vectors. It follows that

$$X\mathbf{e} = \sum_{i=1}^{k} \left(\frac{\mathbf{v}_i}{|C_i|} \right) \mathbf{v}_i^T \mathbf{e} = \sum_{i=1}^{k} \mathbf{v}_i \leq \mathbf{e} \tag{5.2.11}$$

since $\mathbf{v}_i^T \mathbf{e} = |C_i|$. Moreover,

$$\langle X, \mathbf{e}\mathbf{e}^T \rangle = \mathbf{e}^T X \mathbf{e} = \sum_{i=1}^{k} \mathbf{v}_i^T \mathbf{e}. \tag{5.2.12}$$

Therefore, every feasible solution S for (5.2.1) defines a feasible solution of (5.2.5) with equal objective value. The nonconvex program (5.2.5) may be relaxed further to a semidefinite program by replacing the nonconvex constraint $\operatorname{rank}(X) = k$ with the linear constraint $\operatorname{Tr}(X) = k$:

$$\begin{aligned} \max \ & \langle X, \mathbf{e}\mathbf{e}^T \rangle && (5.2.13)\\ \text{s.t. } & X\mathbf{e} \leq \mathbf{e}, && (5.2.14)\\ & X_{ij} = 0, \quad \text{if } (i,j) \notin E \text{ s.t. } i \neq j && (5.2.15)\\ & \operatorname{Tr}(X) = k, && (5.2.16)\\ & X \succeq 0. && (5.2.17) \end{aligned}$$

Recall that every positive semidefinite matrix X satisfies $\lambda_i(X) \geq 0$ for all $i = 1, \ldots, N$. Thus, for every feasible solution X for (5.2.13) we have

$$\operatorname{Tr}(X) = \sum_{i=1}^{N} \lambda_i(X) = \sum_{i=1}^{N} \sigma_i(X) = \|X\|_* \tag{5.2.18}$$

Moreover, every feasible matrix X satisfies $X\mathbf{e} \leq \mathbf{e}$. This implies that $\|X\| \leq 1$ and, thus,

$$\|X\|_* \leq \operatorname{rank}(X) \tag{5.2.19}$$

for all $X \geq 0$ feasible for (5.2.13) by (3.2.1) since (5.2.14) implies that $\|X\|_1 \leq 1$ and $\|X\|_* \leq \|X\|_1$ for every symmetric X by (2.1.7). Therefore, (5.2.13) may be thought of as the relaxation of (5.2.5) obtained by replacing the rank constraint with a nuclear norm constraint.

Note that the relaxation (5.2.13) is exactly the semidefinite programming formulation (4.1.1) for the Lovász theta function $\vartheta(\overline{G})$, with the right-hand side of the trace constraint changed from 1 to k and the additional row sum constraint (5.2.14). In the case that $k = 1$, Theorem 4.1.1 implies that the optimal solutions for (5.2.5) and (5.2.13) coincide and their nonzero eigenvector yields the maximum clique when the input graph contains a sufficiently large planted clique. This may be extended to general k. The convex relaxation of KDC given by (5.2.13) is exact in the case that the input graph G consists of planted k-disjoint-clique subgraph obscured by a moderate number of diversionary edges and nodes. That is, the solution X^* given by (5.2.10) corresponding to the planted k-disjoint-clique subgraph is the unique optimal solution of (5.2.13) and the characteristic vectors of the maximum node k-disjoint-clique subgraph are exactly the eigenvectors corresponding to the nonzero eigenvalues of the matrix X^*. As in our analysis of the maximum

clique and maximum edge biclique problems, we consider two constructions of an input graph G containing a planted k-disjoint-clique subgraph K. In the first, a number of the potential edges of G not in $E(K)$ are added deterministically by an adversary. In the second, each potential edge in $(V \times V) \setminus E(K)$ is added to the graph independently at random with fixed probability $p \in [0, 1)$.

We begin with the adversarial case. Suppose that the graph $G = (V, E)$ is generated as follows. We first add the complete subgraphs corresponding to k disjoint cliques $C_1, \ldots, C_k$ of size $r_1, \ldots, r_k$ respectively. Then an adversary is allowed to add a set C_{k+1} of additional vertices and a number of the remaining potential edges to graph. The following theorem states that our adversary can add up to $O(\hat{r}^2)$ noise edges where $\hat{r} := \min\{r_1, \ldots, r_k\}$ and the k-disjoint-clique subgraph defined by $\{C_1, \ldots, C_k\}$ will be the unique maximum k-disjoint-clique subgraph of G and will be found by solving the semidefinite relaxation (5.2.13).

Theorem 5.2.1 ([7, Theorem 3.1]) *Suppose that G contains a k-disjoint-clique subgraph K^* composed of cliques $C_1, \ldots, C_k$, where $|C_i| = r_i$ for all $i = 1, \ldots, k$. Let $C_{k+1} := V \setminus (\cup_{i=1}^{k} C_i)$ and let $r_{k+1} = |C_{k+1}|$. Suppose that for all $i = 1, \ldots, k$, $v \in C_i$, and $j \in \{1, \ldots, k+1\} - i$,*

$$n_v^j \leq \delta \min\{r_i, r_j\} \tag{5.2.20}$$

where $\delta \in (0, 0.382)$ and n_v^j is equal to the number of nodes adjacent to v in C_j. Then there exists scalar $\rho > 0$ depending only on δ such that if

$$|E(G) \setminus E(K^*)| \leq \rho \hat{r}^2 \tag{5.2.21}$$

then the feasible solution X^* corresponding to K^* given by (5.2.10) *is the unique optimal solution to* (5.2.13)*, and K^* is the unique maximum node k-disjoint-clique subgraph of G.*

We remark that two of the conditions imposed in this theorem are, up to the constant factors, the best possible according to the following arguments. If $n_v^j = r_j$, then node v could be inserted into clique j, so the partitioning between C_j and C_i would no longer be uniquely determined. This shows the necessity of the condition $n_v^j = O(r_j)$. The condition that $|E(G) \setminus E(G^*)| \leq \rho \hat{r}^2$ is necessary, up to a constant factor, because if $|E(G) \setminus E(G^*)| \geq \hat{r}^2(\hat{r}^2 - 1)/2$, then we could interconnect an arbitrary set of $\hat{r}$ nodes chosen from among the existing cliques with edges to make a new clique out of them, again spoiling the uniqueness of the decomposition.

We next consider the case when the nonclique edges of G are added randomly. Let $C_1, C_2, \ldots, C_{k+1}$ be disjoint vertex sets of sizes $r_1, \ldots, r_{k+1}$ respectively, and let $V = \cup_{i=1}^{k+1} C_i$. We construct the edge set of the graph $G = (V, E)$ as follows:

(Ω_1) We add (i, j) to E for each $q = 1, \ldots, k$, and each $i \in C_q$, $j \in C_q$ such that $i \neq j$.

(Ω_2) Each of the remaining possible edges is added to E independently at random with probability $p \in (0, 1)$.

Notice that the graph $G = (V, E)$ has a k-disjoint-clique subgraph K^* with cliques indexed by the vertex sets $C_1, \ldots, C_k$ by our construction of E. The following theorem provides conditions on the vertex sets $C_1, \ldots, C_{k+1}$ that ensure that K^* is the maximum k-disjoint-clique subgraph of G and is recoverable by solving (5.2.13).

Theorem 5.2.2 *Suppose that the random graph $G = (V, E)$ containing a k-disjoint-clique subgraph K^* composed of cliques $C_1, \ldots, C_k$ is constructed according to (Ω_1) and (Ω_2) with probability p. Let $N := |V|$, $C_{k+1} = V \setminus (\cup_{i=1}^{k} C_i)$ and $r_i = |C_i|$ for all $i = 1, \ldots, k+1$. Let X^* be the feasible solution for (5.2.13) corresponding to $\{C_1, \ldots, C_k\}$ given by (5.2.10). Further, suppose that $r_i \leq \hat{r}^{3/2}$ for all $i = 1, \ldots, k$ where $\hat{r} = \min_{i=1,\ldots,k}\{r_i\}$. Then there exists scalar $c_1, c_2 > 0$, depending only on p, such that if*

$$c_1 \left(\sum_{i=1}^{k} r_i^2 \right)^{1/2} \left(\sum_{j=1}^{k} \frac{1}{r_j} \right)^{1/2} + c_2 \sqrt{N} \leq \hat{r}, \tag{5.2.22}$$

then K^* is the unique maximum node k-disjoint-clique subgraph of G and X^* is the unique optimal solution of (5.2.13) with probability tending exponentially to 1 as $\hat{r} \to \infty$.

It is clear from (5.2.22) that the sufficient conditions for uniqueness and optimality given by Theorem 5.2.2 cannot be satisfied unless $N = O(\hat{r}^2)$. We now give a few examples of values for $r_1, \ldots, r_{k+1}$ that fulfill (5.2.22).

- Consider the case $k = 1$, i.e., a single large clique. In this case, taking $r_1 = O(N^{1/2})$ satisfies (5.2.22) since the first term on the left is $O(N^{1/4})$. Note that this is, up to the constants, the bound given by (4.1.3), as well as by Alon et al. [5] and Feige and Krauthgamer [66].

- Suppose $k > 1$ and $r_1 = \ldots = r_k = cN^\alpha$ for some scalar $c \geq 0$. In this case, the first parenthesized factor on the left in (5.2.22) is $O(k^{1/2}N^\alpha)$ while the second is $O(k^{1/2}N^{-\alpha/2})$. Therefore, the first term is $O(kN^{\alpha/2})$. For (5.2.22) to hold, we need this term to be $O(\hat{r}) = O(N^\alpha)$, which is valid as long as $k \leq N^{\alpha/2}$. We also need $\alpha \geq 1/2$ as noted above to handle the second term on the right. For example, for

$\alpha = 1/2$ the algorithm can find as many as $N^{1/4}$ cliques of this size. For $\alpha = 2/3$, the algorithm can find as many as $N^{1/3}$ cliques of this size, which is the maximum possible since the cliques are disjoint and N is the number of nodes.

- The cliques may also be of different sizes. For example, if there is one large clique of size $O(N^{2/3})$ and $N^{1/6}$ smaller cliques of size $O(N^{1/2})$, then $\hat{r} = O(N^{1/2})$, the first parenthesized factor in (5.2.22) is $N^{2/3}$ while the second is $N^{-1/6}$, so the entire first factor is $O(N^{1/2}) = O(\hat{r})$.

We note that the results for random noise in the k-disjoint-clique problem are much better than the results for adversary-chosen noise. In the case of adversary-chosen noise, the number of allowable noise edges is bounded above by a constant times the number of edges in the smallest clique. In the case of random noise, the number of allowable noise edges is the square of that quantity. For example, if there are $N^{1/4}$ cliques each of size $N^{1/2}$, then the smallest clique has N edges versus N^2 noise edges.

5.2.2 A convex relaxation for the maximum mean weight k-disjoint clique problem

Let $K_N = (V, E)$ be a complete graph with vertex set $V = \{1, 2, \ldots, N\}$. Given a nonnegative matrix $W \in \mathbf{R}_+^{N \times N}$ of edge weights and integer $k \in \{1, \ldots, N\}$, the maximum mean weight k-disjoint clique problem seeks to identify the k-disjoint-clique subgraph K^* of K_N that maximizes the sum of the average weight (with respect to W) covered by the edges of each clique of K^*. More precisely, this problem can be formulated as the nonlinear optimization problem

$$\max_{S=\{\mathbf{v}_1,\ldots,\mathbf{v}_k\}} \sum_{i=1}^{k} \frac{\langle W, \mathbf{v}_i \mathbf{v}_i^T \rangle}{\mathbf{v}_i^T \mathbf{e}} \tag{5.2.23}$$

$$\text{s.t. } \mathbf{v}_i^T \mathbf{v}_j = 0 \quad \text{if } i \neq j \tag{5.2.24}$$

$$\mathbf{v}_i \in \{0, 1\}^N \quad \forall\, i = 1, \ldots, k. \tag{5.2.25}$$

As in the maximum node case, each feasible solution $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ of (5.2.23) is the collection of characteristic vectors of a k-disjoint-clique subgraph K of K_N. Recall that every feasible solution $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ of (5.2.23) defines a rank-k positive semidefinite matrix

$$X^* = \sum_{i=1}^{k} \frac{\mathbf{v}_i \mathbf{v}_i^T}{\mathbf{v}_i^T \mathbf{e}}$$

by (5.2.10). Hence, (5.2.23) can be relaxed to the rank-constrained semidefinite program

$$\begin{aligned} \max \langle W, X \rangle & \quad (5.2.26) \\ \text{s.t. } X\mathbf{e} \leq \mathbf{e} & \quad (5.2.27) \\ X \geq 0 & \quad (5.2.28) \\ \operatorname{rank} X = k & \quad (5.2.29) \\ X \succeq 0. & \quad (5.2.30) \end{aligned}$$

As before, we relax further by replacing rank with trace to obtain the semidefinite program

$$\begin{aligned} \max \langle W, X \rangle & \quad (5.2.31) \\ \text{s.t. } X\mathbf{e} \leq \mathbf{e} & \quad (5.2.32) \\ X \geq 0 & \quad (5.2.33) \\ \operatorname{Tr} X = k & \quad (5.2.34) \\ X \succeq 0. & \quad (5.2.35) \end{aligned}$$

Note that the constraints (5.2.32), (5.2.33), (5.2.34), and (5.2.35) are identical to those of the semidefinite relaxation of the k-means clustering problem considered by Peng and Wei in [152].

We are interested in identifying a class of input instances of (5.2.23) that may be solved directly by solving the semidefinite programming relaxation given by (5.2.31). In particular, we will see that the SDP relaxation (5.2.31) is exact for (5.2.23) (and (5.2.26)), when the input instance corresponds to a k-disjoint-clique subgraph of the input graph K_N. Unlike in our analysis of the maximum clique, maximum edge biclique, and maximum node k-disjoint-clique problems, the planted case for WKDC is induced by the matrix of edge-weights W. In the planted case, the entries of W corresponding to the edges of the planted k-disjoint-clique subgraph of K_N are significantly larger than the remaining entries. Let K^* be a k-disjoint-clique subgraph of K_N composed of the disjoint cliques $C_1, C_2, \ldots, C_k$. We consider symmetric matrices $W \in \Sigma^N$ with random entries sampled from one of two distributions Ω_1, Ω_2 as follows:

(ω_1) The entries of each block of W corresponding to C_q, C_s, $q, s = 1, \ldots, k+1$, are independently identically distributed (i.i.d.) such that

$$E[W_{ij}] = E[W_{ji}] = \begin{cases} \alpha, & \text{if } q = s,\ 1 \leq q, s \leq k \\ \beta, & \text{otherwise} \end{cases} \quad (5.2.36)$$

for fixed scalars $\alpha > \beta > 0$ for all $i \in C_q$, $j \in C_s$, where $C_{k+1} := V \setminus (\cup_{i=1}^{k} C_i)$.

(ω_2) For all $i, j \in V$,

$$0 \leq W_{i,j} \leq 1. \tag{5.2.37}$$

That is, if the nodes i, j are in the same planted clique we sample the random variable W_{ij} from the probability distribution Ω_1 with mean α; otherwise, we sample W_{ij} from the distribution Ω_2 with mean β. This is a natural model for the similarity matrix W for a set of clustered data. Indeed, consider the following examples.

- Suppose that we have an oracle that exactly detects if items in the data are similar or not. If we require that every pair of items in each cluster is similar, then identifying the k largest disjoint clusters in the data set is equivalent to the maximum node k-disjoint-clique problem. Suppose that the data is clusterable in the sense that the similarity graph contains the planted k-disjoint-clique subgraph K composed of disjoint cliques $\{C_1, \ldots, C_k\}$ corresponding to the clusters in the data and the graph is sparse outside of K. Then the entries of the diagonal blocks of W indexed by $C_1, \ldots, C_k$ are equal to 1 and the off-diagonal blocks are sparse. In the random model, where diversionary edges are added to G with probability p, this is equivalent to having W satisfying (ω_1) and (ω_2) with $\alpha = 1$, $\beta = p$.

- A less restrictive model for clustering is to require the subgraphs of G_S corresponding to the clusters to be densely connected but not necessarily complete. Several recent papers [48, 148] have shown that if the input graph consists of a collection of k sufficiently dense (in expectation) disjoint clusters plus a moderate number of randomly inserted diversionary edges, then the clusters can be recovered by solving a convex program. The problem of identifying k dense clusters in a graph is a special case of the planted case for WKDC. In the random model, each dense (in expectation) subgraph will correspond to a planted clique in G whose block of W will have entries with expected value close to 1. On the other hand, the remaining entries of W will have small expected value since the input graph is sparse outside of the planted dense subgraphs.

- Suppose that we are given a random clustered data set D in $\mathbf{R}^n$. For example, suppose that each data point in the ith cluster C_i is placed uniformly at random in a ball of radius $\epsilon > 0$ centered at $c_i \in \mathbf{R}^n$. Suppose further that distance between each pair of cluster centers is at least $\delta > 0$; that is, $\|c_i - c_j\| > \delta$ for all $i \neq j$. Consider W with Gaussian entries $W_{\mathbf{u},\mathbf{v}} = \exp(-\|\mathbf{u} - \mathbf{v}\|^2/\sigma^2)$ for all $\mathbf{u}, \mathbf{v} \in D$ for some scalar $\sigma > 0$. For all $\mathbf{u}, \mathbf{v}$ in the same cluster, we have $W_{\mathbf{u},\mathbf{v}} \geq \exp(-2\epsilon^2/\sigma^2)$. If $\mathbf{u}, \mathbf{v}$ are

in different clusters, then $W_{\mathbf{u},\mathbf{v}} \leq \exp(-(\delta - 2\epsilon)^2/\sigma^2)$. For sufficiently large choice of δ, the entries of W indexed by edges within clusters will be significantly larger in expectation than the entries of W indexed by edges between different clusters.

The following theorem describes which partitions $\{C_1, C_2, \ldots, C_{k+1}\}$ of V yield random symmetric matrices W satisfying (ω_1) and (ω_2) such that the feasible solution corresponding to K is optimal for (5.2.23) and can be found with high probability by solving (5.2.31).

Theorem 5.2.3 *Suppose that vertex sets $C_1, \ldots, C_k$ define a k-disjoint-clique subgraph K^* of the complete graph $K_N = (V, E)$ on N vertices and let $C_{k+1} := V \setminus (\cup_{i=1}^k C_i)$. Let $r_i := |C_i|$ for all $i = 1, \ldots, k+1$, and let $\hat{r} = \min_{i=1,\ldots,k} r_i$. Let $W \in \Sigma^N$ be a random symmetric matrix sampled from distributions Ω_1, Ω_2 according to (ω_1) and (ω_2). Let X^* be the feasible solution for (5.2.31) corresponding to $C_1, \ldots, C_k$ defined by (5.2.10). Suppose that the entries of λ and η are nonnegative. Then there exist scalars $c_1, c_2, \rho_1, \rho_2 > 0$ such that if*

$$r_i \leq c_1(\alpha - \beta)^2 \hat{r}^2 \tag{5.2.38}$$

for all $i = 1, \ldots, k$, and

$$\rho_1 \left(k \sum_{s=1}^{k+1} r_s \right)^{1/2} + \rho_2 \sqrt{N} + \beta r_{k+1} \leq c_2(\alpha - \beta)\hat{r} \tag{5.2.39}$$

then X^* is the unique optimal solution for (5.2.31), and K^* is the unique maximum mean weight k-disjoint-clique subgraph of K_N corresponding to W with probability tending exponentially to 1 as $\hat{r} \to \infty$.

Note that (5.2.39) cannot be satisfied if $N = \Omega(\hat{r}^2)$ or $r_{k+1} = \Omega(\hat{r})$. We now provide a few examples of $r_1, \ldots, r_k$ satisfying the hypothesis of Thorem 5.2.3.

- Suppose that we have k cliques $C_1, \ldots, C_k$ of size $r_1 = r_2 = \cdots = r_k = N^{\epsilon/2}$. Then (5.4.61) implies that we may recover the k-disjoint-clique subgraph corresponding to $C_1, \ldots, C_k$ if $k = O(N^{\epsilon/2})$ for $\epsilon \in [1/2, 2/3]$; the lower bound $\epsilon \geq 1/2$ is a consequence of the requirement that $\sqrt{N} = O(N^{\epsilon})$, and the upper bound $\epsilon \leq 2/3$ follows from the fact that $C_1, \ldots, C_k$ are disjoint and must contain at most N nodes.

- On the other hand, we may have cliques of different sizes. For example, suppose that we wish to recover k_1 cliques of size $O(N^{3/4})$ and k_2 smaller cliques of size $N^{1/2}$.

Then the right-hand side of (5.2.39) is at most

$$O(k_1 N^{3/8}) + O(k_2 N^{1/4}).$$

Therefore, we may recover the planted cliques provided that $k_1 = O(N^{1/4})$ and $k_2 = O(N^{1/2})$.

- Consider the special case where the entries of W are distributed such that $W_{C_q,C_q} = \mathbf{e}\mathbf{e}^T$ for all $q = 1, \ldots, k$, and

$$[W_{C_q,C_s}]_{ij} = \begin{cases} 1, & \text{with probability } p \\ 0, & \text{with probability } 1-p \end{cases}$$

for some fixed probability p for all $q, s \in \{1, \ldots, k+1\}$ such that $q \neq s$ or $q = s = k+1$ corresponds to an instance of KDC. That is, $W_{ij} = 1$ if i and j are adjacent in the input graph G, and is equal to 0 otherwise. We note that the bounds on the number of planted cliques tolerated by our relaxation (5.2.31) improve upon those given by Theorem 5.2.2. However, our relaxation (5.2.31) only tolerates at most $O(\hat{r})$ noise nodes, far fewer than the bound, $O(\hat{r}^2)$, given by Theorem 5.2.2.

5.3 Proof of the theoretical bound for the maximum node k-disjoint-clique problem

5.3.1 Optimality conditions for the maximum node k-disjoint-clique problem

In this section, we provide conditions for optimality of the solution X^* corresponding to a k-disjoint-clique subgraph of G, as defined by (5.2.10), for the convex relaxation of KDC given by (5.2.13). We begin with the following sufficient condition for the optimality of X^*.

Theorem 5.3.1 *Let X^* be feasible for* (5.2.13)*. Suppose also that there exist $\lambda \in \mathbf{R}^N_+$, $\mu \in \mathbf{R}$, $\eta \in \mathbf{R}^{N \times N}$ and $S \in \Sigma^N_+$ such that*

$$-\mathbf{e}\mathbf{e}^T + \lambda \mathbf{e}^T + \mathbf{e}\lambda^T + \mu I + \sum_{\substack{(i,j) \notin E \\ i \neq j}} \eta_{ij} \mathbf{e}_i \mathbf{e}_j^T = S, \tag{5.3.1}$$

$$\lambda^T(X^*\mathbf{e} - \mathbf{e}) = 0, \tag{5.3.2}$$
$$\langle S, X^*\rangle = 0. \tag{5.3.3}$$

Then X^* is an optimal solution of (5.2.13).

Note that $X = (k/n)I$ is strictly feasible for (5.2.13). Hence, (5.2.13) satisfies Slater's constraint qualification. Similarly, choosing $\lambda = 0$, $\eta = 0$ and μ large enough in (5.3.1) so that $S = \mu I - \mathbf{ee}$ is positive definite shows that the dual of (5.2.13) also satisfies the Slater condition. Therefore, X^* is optimal if it satisfies the Karush-Kuhn-Tucker conditions. Theorem 5.3.1, provides the necessary specialization to (5.2.13) of the Karush-Kuhn-Tucker conditions for semidefinite programming given by Theorem 2.5.2 .

Our proof technique to show that X^* is optimal for (5.2.13) is to construct multipliers that satisfy Theorem 5.3.1. The difficult multiplier to construct is S, the dual semidefinite matrix. The reason is that S must simultaneously satisfy homogeneous linear equations given by $\langle S, X^*\rangle = 0$, requirements on its entries given by the gradient equation (5.3.1), and positive semidefiniteness.

Our strategy for satisfying the requirements on S is as follows. The matrix S will be constructed in blocks with sizes inherited from the blocks of X^*. In particular, let the nodes contained in the k planted cliques be denoted $C_1, \ldots, C_k$, and let the remaining nodes be C_{k+1}. Then according to (5.2.10), X^* has diagonal blocks $X^*_{C_q,C_q}$ for $q = 1, \ldots, k$ consisting of multiples of the all 1's matrix. The remaining blocks of X^* are 0's. The diagonal blocks of S will be perturbations of the identity matrix, with the rank-one perturbation chosen so that each diagonal block of S, say S_{C_q,C_q} is orthogonal to X_{C_q,C_q}.

The entries of an off-diagonal block, say S_{C_q,C_s} must satisfy, first of all, (5.3.1). This constraint, however, is binding only on the entries corresponding to edges in G, since entries corresponding to absent edges are not constrained by (5.3.1) thanks to the presence of the free multiplier η_{ij} on the left hand side. These entries that are free in (5.3.1) are chosen so that (5.3.3) is satisfied. It is a well known result in semidefinite programming that the requirements $\langle S, X\rangle = 0$, $X, S \in \Sigma^N_+$ together imply $SX = XS = 0$ (see [181, Proposition 1.19]). Thus, the remaining entries of S must be chosen so that $X^*S = SX^* = 0$. Because of the special form of X^*, this is equivalent to requiring all row and column sums of S_{C_q,C_s} to equal zero.

We parametrize the entries of S_{C_q,C_s} that are not predetermined by (5.3.1) using the entries of two vectors $\mathbf{y}^{q,s}$ and $\mathbf{z}^{q,s}$. These vectors are chosen to be the solutions to systems of linear equations, namely, those imposed by the requirement that $X^*S = SX^* = 0$. We show that the system of linear equations may be written as a perturbation of a linear

system with a known solution, and we can thus get bounds on $\mathbf{y}^{q,s}$ and $\mathbf{z}^{q,s}$. The bounds on $\mathbf{y}^{q,s}$ and $\mathbf{z}^{q,s}$ in turn translate to bounds on $\|S_{C_q,C_s}\|$, which are necessary to establish the positive semidefiniteness of S. This semidefiniteness is established by proving that the diagonal blocks, which are identity matrices plus rank-one perturbations, dominate the off-diagonal blocks.

Recalling our notation introduced earlier, $G = (V, E)$ has a k-disjoint-clique subgraph K composed of cliques $C_1, C_2, \ldots, C_k$ of sizes $r_1, r_2, \ldots, r_k$ respectively. Let $C_{k+1} := V \setminus (\cup_{i=1}^{k} C_i)$ be the set of nodes of G not in K and let $r_{k+1} := |C_{k+1}|$. Let $N := |V|$. Let $\hat{r} := \min\{r_1, r_2, \ldots, r_k\}$. For each $v \in V$, let n_v^s denote the number of nodes adjacent to v in C_s for all $s \in \{1, \ldots, k+1\}$, and let $cl(v)$ denote index $i \in \{1, \ldots, k+1\}$ such that $v \in C_i$.

Let $A_{\overline{G}} \in \mathbf{R}^{N \times N}$ be the adjacency matrix of the complement $\overline{G}$ of G; that is $[A_{\overline{G}}]_{i,j} = 1$ if $(i, j) \notin E$ and 0 otherwise. Next, fix $q, s \in \{1, \ldots, k+1\}$ such that $q \neq s$. Let $H = H_{q,s} \in \mathbf{R}^{C_q \times C_s}$ be the block of $A_{\overline{G}}$ with entries indexed by the vertex sets C_q and C_s, and let $D = D_{q,s} \in \mathbf{R}^{C_q \times C_q}$ be the diagonal matrix such that, for each $i \in C_q$, the (i, i)th entry of D is equal to the number of nodes in C_s not adjacent to i. That is, $D = r_s I - \mathrm{Diag}\,(\mathbf{n}_{C_q}^s)$ where $\mathbf{n}_{C_q}^s \in \mathbf{R}^{C_q}$ is the vector with ith entry equal to n_i^s for each $i \in C_q$. Similarly, let $F = F_{q,s} = r_q I - \mathrm{Diag}\,(\mathbf{n}_{C_s}^q)$. Next, define the scalar

$$c = c_{q,s} := \begin{cases} \dfrac{\hat{r}}{2}\left(\dfrac{1}{r_q} + \dfrac{1}{r_s}\right), & \text{if } s \leq k \\ \dfrac{\hat{r}}{2}\left(\dfrac{1}{\hat{r}} + \dfrac{1}{r_q}\right), & \text{otherwise.} \end{cases}$$

Next, for each $q, s = 1, \ldots, k+1$ such that $q \neq s$ let $\mathbf{b} = \mathbf{b}^{q,s} \in \mathbf{R}^{C_q \cup C_s}$ be defined by

$$\mathbf{b}_i = c \cdot \begin{cases} n_i^s, & \text{if } i \in C_q \\ n_i^q, & \text{if } i \in C_s. \end{cases}$$

Note that the matrix

$$\begin{pmatrix} D & H \\ H^T & F \end{pmatrix}$$

is weakly diagonally dominant since the ith row of H contains exactly $r_s - n_i^s$ 1's, and, hence, positive semidefinite. Further, let $\mathbf{y} = \mathbf{y}^{q,s}$ and $\mathbf{z} = \mathbf{z}^{q,s}$ be a solution of the perturbed system

$$\begin{pmatrix} D + \theta \mathbf{e}\mathbf{e}^T & H - \theta \mathbf{e}\mathbf{e}^T \\ H^T - \theta \mathbf{e}\mathbf{e}^T & F + \theta \mathbf{e}\mathbf{e}^T \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} = \mathbf{b} \tag{5.3.4}$$

for some scalar $\theta > 0$ to be defined later.

As mentioned earlier, it is required that all row and column sums of S_{C_q,C_s} equal zero. Consider, for example, the sum of the entries in a particular row $i \in C_q$. This sum consists of r_s terms; of these terms, n_i^s of them are equal to $-c_{q,s}$ (corresponding to edges from i to C_s) while the other $r_s - n_i^s$ have the form $[\mathbf{y}^{q,s}]_i + [\mathbf{z}^{q,s}]_j$. Thus, the requirement that the row sums to zero is written

$$-n_i^s c_{q,s} + \sum_{j \in C_s; (i,j) \notin E} ([\mathbf{y}^{q,s}]_i + [\mathbf{z}^{q,s}]_j) = 0,$$

which may be rewritten

$$(r_s - n_i^s)[\mathbf{y}^{q,s}]_i + \sum_{j \in C_s; (i,j) \notin E} [\mathbf{z}^{q,s}]_j = n_i^s c_{q,s}. \tag{5.3.5}$$

Equation (5.3.5) is exactly a row of (5.3.4) in the case $\theta = 0$ because of the formulas used to define $D, F, H, \mathbf{b}$.

In the case that θ is not zero, the equation for the ith row in (5.3.4) has an additional term of the form $\theta(\mathbf{e}^T\mathbf{y}^{q,s} - \mathbf{e}^T\mathbf{z}^{q,s})$. This additional term does not affect the result, as the following argument shows. The version of (5.3.4) with $\theta = 0$ is singular because the vector $(\mathbf{e}; -\mathbf{e})$ is in its null space. This corresponds to adding a scalar to each entry of $\mathbf{y}^{q,s}$ and subtracting the same scalar from each entry of $\mathbf{z}^{q,s}$. One particular way to fix that scalar is to require that the sum of entries of $\mathbf{y}^{q,s}$ equals the sum of entries of $\mathbf{z}^{q,s}$, i.e.,

$$\mathbf{e}^T\mathbf{y}^{q,s} - \mathbf{e}^T\mathbf{z}^{q,s} = 0. \tag{5.3.6}$$

If we are able to show that (5.3.4) is nonsingular for some $\theta > 0$ (which we will establish in Section 5.3.2 and again in Section 5.3.3) then this particular $(\mathbf{y}^{q,s}, \mathbf{z}^{q,s})$ satisfying (5.3.5) and (5.3.6) will also be a solution to (5.3.4) for nonzero θ since the additional term $\theta(\mathbf{e}^T\mathbf{y}^{q,s} - \mathbf{e}^T\mathbf{z}^{p,q})$ is zero.

For the remainder of this section, in order to formulate definitions for the remaining multipliers, assume that $\theta > 0$ and that (5.3.4) is nonsingular. Furthermore, assume that $D_{ii} > 0$ for all $i \in C_q$ and $F_{ii} > 0$ for all $i \in C_s$. Let

$$A = A(\theta) := \begin{pmatrix} D + \theta\mathbf{e}\mathbf{e}^T & 0 \\ 0 & F + \theta\mathbf{e}\mathbf{e}^T \end{pmatrix}, \quad P = P(\theta) := \begin{pmatrix} 0 & H - \theta\mathbf{e}\mathbf{e}^T \\ H^T - \theta\mathbf{e}\mathbf{e}^T & 0 \end{pmatrix}.$$

By assumption, $A+P$ is nonsingular, and

$$\begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} = (A+P)^{-1}\mathbf{b}.$$

The proof technique in Sections 5.3.2 and 5.3.3 is to show that $Q := (A+P)^{-1} - A^{-1}$ is small so that $(\mathbf{y}, \mathbf{z})$ is close to $A^{-1}\mathbf{b}$. Let $Q = (Q_1^T, Q_2^T)^T$ where $Q_1 \in \mathbf{R}^{C_q \times (C_q \cup C_s)}$ and $Q_2 \in \mathbf{R}^{C_s \times (C_q \cup C_s)}$. Then, under this notation,

$$\begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} = A^{-1}\mathbf{b} + \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}\mathbf{b} = \begin{pmatrix} (D+\theta\mathbf{e}\mathbf{e}^T)^{-1} & 0 \\ 0 & (F+\theta\mathbf{e}\mathbf{e}^T)^{-1} \end{pmatrix}\begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} + \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}\mathbf{b}$$

where $\mathbf{b}_1 \in \mathbf{R}^{C_q}$, $\mathbf{b}_2 \in \mathbf{R}^{C_s}$ are the vectors of entries of $\mathbf{b}$ corresponding to C_q and C_s respectively. Therefore, if D, F and $A+P$ are nonsingular,

$$\mathbf{y} = (D+\theta\mathbf{e}\mathbf{e}^T)^{-1}\mathbf{b}_1 + Q_1\mathbf{b} = (I+\theta D^{-1}\mathbf{e}\mathbf{e}^T)^{-1}D^{-1}\mathbf{b}_1 + Q_1\mathbf{b}$$

and

$$\mathbf{z} = (I+\theta F^{-1}\mathbf{e}\mathbf{e}^T)^{-1}F^{-1}\mathbf{b}_2 + Q_2\mathbf{b}.$$

Let $\bar{\mathbf{y}} := \mathbf{y} - Q_1\mathbf{b}$ and $\bar{\mathbf{z}} := \mathbf{z} - Q_2\mathbf{b}$. To give explicit formulas for $\bar{\mathbf{y}}$ and $\bar{\mathbf{z}}$, we use the well-known Sherman-Morrison-Woodbury formula (see, for example, [85, Equation 2.1.4]), stated in the following lemma, to calculate $(I+\theta D^{-1}\mathbf{e}\mathbf{e}^T)^{-1}$ and $(I+\theta F^{-1}\mathbf{e}\mathbf{e}^T)^{-1}$.

Lemma 5.3.1 *If A is a nonsingular matrix in $\mathbf{R}^{n\times n}$ and $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$ satisfy $\mathbf{v}^T A^{-1}\mathbf{u} \neq -1$ then*

$$(A+\mathbf{u}\mathbf{v}^T)^{-1} = A^{-1} - \frac{A^{-1}\mathbf{u}\mathbf{v}^T A^{-1}}{1+\mathbf{v}^T A^{-1}\mathbf{u}}. \tag{5.3.7}$$

As an immediate corollary of Lemma 5.3.1, notice that

$$\bar{\mathbf{y}} = \left(D^{-1} - \frac{\theta D^{-1}\mathbf{e}\mathbf{e}^T D^{-1}}{1+\theta\mathbf{e}^T D^{-1}\mathbf{e}}\right)\mathbf{b}_1 = D^{-1}\left(I - \frac{\theta\mathbf{e}\mathbf{e}^T D^{-1}}{1+\theta\mathbf{e}^T D^{-1}\mathbf{e}}\right)\mathbf{b}_1. \tag{5.3.8}$$

and

$$\bar{\mathbf{z}} = \left(F^{-1} - \frac{\theta F^{-1}\mathbf{e}\mathbf{e}^T F^{-1}}{1+\theta\mathbf{e}^T F^{-1}\mathbf{e}}\right)\mathbf{b}_2 = F^{-1}\left(I - \frac{\theta\mathbf{e}\mathbf{e}^T F^{-1}}{1+\theta\mathbf{e}^T F^{-1}\mathbf{e}}\right)\mathbf{b}_2. \tag{5.3.9}$$

Finally, we define the $(k+1)\times(k+1)$ block matrix $\tilde{S} \in \mathbf{R}^{N\times N}$ as follows:

($\tilde{\sigma}_1$) For all $q \in \{1, \ldots, k\}$, let $\tilde{S}_{C_q, C_q} = 0$.

($\tilde{\sigma}_2$) For all $q, s \in \{1, \ldots, k\}$ such that $q \neq s$, let

$$\tilde{S}_{C_q,C_s} = H_{q,s} \circ (\mathbf{y}^{q,s}\mathbf{e}^T + \mathbf{e}(\mathbf{z}^{q,s})^T) + c_{q,s}(H_{q,s} - \mathbf{e}\mathbf{e}^T). \tag{5.3.10}$$

($\tilde{\sigma}_3$) For all $q \in \{1, \ldots, k\}$ and $i \in C_q$, $j \in C_{k+1}$, let

$$[\tilde{S}_{C_q,C_{k+1}}]_{ij} = [\tilde{S}_{C_{k+1},C_q}]_{ji} = \begin{cases} -c_{q,k+1}, & \text{if } (i,j) \in E \\ c_{q,k+1}\, n_j^q/(r_q - n_j^q), & \text{otherwise.} \end{cases}$$

($\tilde{\sigma}_4$) Finally, for all $i, j \in C_{k+1}$, choose

$$[\tilde{S}_{C_{k+1},C_{k+1}}]_{ij} = \begin{cases} -1, & \text{if } (i,j) \in E \text{ or } i = j \\ \gamma, & \text{if } (i,j) \notin E \end{cases} \tag{5.3.11}$$

for some scalar γ to be defined later.

We make a couple of remarks about ($\tilde{\sigma}_2$). For $q, s \in \{1, \ldots, k\}$ such that $q \neq s$, this formula defines entries of $\tilde{S}_{C_q,C_s}$ to be $-c_{q,s}$ in positions corresponding to edges, and $[\mathbf{y}^{q,s}]_i + [\mathbf{z}^{q,s}]_j$ in other positions. The vectors $\mathbf{y}^{q,s}$ and $\mathbf{z}^{q,s}$ are defined by (5.3.4) precisely so that, when used in this manner to define $\tilde{S}_{C_q,C_s}$, its row and column sums are all 0 (so that $X^*S = SX^* = 0$; the relationship $S_{C_q,C_s} \equiv \tilde{S}_{C_q,C_s}$ is given by (5.3.17) below). The system is square because the number of constraints on $S_{q,s}$ imposed by $X^*S = SX^* = 0$ after the predetermined entries are filled in is $|C_q| + |C_s|$ (one constraint for each row and column), which is the total number of entries in $\mathbf{y}^{q,s}$ and $\mathbf{z}^{q,s}$. As mentioned earlier, there is the slight additional complexity that these $|C_q| + |C_s|$ equations have a dependence of dimension 1, which explains why we needed to regularize (5.3.4) with the addition of the $\theta\mathbf{e}\mathbf{e}^T$ terms.

As a second remark about $\tilde{\sigma}_2$, we note that $\tilde{S}_{C_q,C_s} = \tilde{S}^T_{C_s,C_q}$. This is a consequence of our construction detailed above. In particular, $\mathbf{y}^{q,s} = \mathbf{z}^{s,q}$, $H_{q,s} = H^T_{s,q}$, and $D_{q,s} = F_{s,q}$ for all $q, s = 1, \ldots, k$ such that $q \neq s$.

We finally come to the main theorem of this section, which provides a sufficient condition for when the k-disjoint-clique subgraph of G composed of the cliques $C_1, \ldots, C_k$ is the maximum node k-disjoint-clique subgraph of G.

Theorem 5.3.2 *Suppose that $G = (V, E)$ has a k-disjoint-clique subgraph K^* composed of the disjoint cliques $C_1, \ldots, C_k$ and let $C_{k+1} := V \setminus (\cup_{i=1}^k C_i)$. Let $r_i = |C_i|$ for all*

$i = 1, \ldots, k+1$, *and let* $\hat{r} = \min_{i=1,\ldots,k}\{r_i\}$. *Let* X^* *be the matrix of the form* (5.2.10) *corresponding to the* k*-disjoint-clique subgraph generated by* $C_1, \ldots, C_k$. *Moreover, suppose that the matrix* $\tilde{S}$ *as defined by* $(\tilde{\sigma}_1), \ldots, (\tilde{\sigma}_4)$ *satisfies*

$$\|\tilde{S}\| \leq \hat{r} - 1. \tag{5.3.12}$$

Then X^* *is optimal for*(5.2.13), *and* K^* *is the maximum node* k*-disjoint-clique subgraph of* G. *Moreover, if* $\|\tilde{S}\| < \hat{r} - 1$ *and*

$$n_v^q < r_q \tag{5.3.13}$$

for all $v \in V$ *and* $q \in \{1, \ldots, k\} \setminus cl(v)$ *then* X^* *is the unique optimal solution of* (5.2.13) *and* K^* *is the unique maximum node* k*-disjoint-clique subgraph of* G.

Proof: We will prove that (5.3.12) is a sufficient condition for optimality of X^* by defining multipliers μ, λ, η, and S and proving that if (5.3.12) holds then these multipliers satisfy the optimality conditions given by Theorem 5.3.1. Let us define the multipliers μ and λ by

$$\mu = \hat{r} = \min\{r_1, r_2, \ldots, r_k\}, \tag{5.3.14}$$

$$\lambda_i = \frac{(1 - \hat{r}/r_q)}{2} \quad \text{for all } i \in C_q, \tag{5.3.15}$$

for all $q = 1, \ldots, k$ and

$$\lambda_i = 0. \tag{5.3.16}$$

for all $i \in C_{k+1}$. Notice that by our choice of μ and λ we have

$$S_{C_q, C_q} = \hat{r}I - (\hat{r}/r_q)\mathbf{e}\mathbf{e}^T$$

for all $q = 1, \ldots, k$ by (5.3.1). Moreover, we choose η such that

$$\eta_{ij} = \begin{cases} \tilde{S}_{ij} - \lambda_i - \lambda_j + 1, & \text{if } (i,j) \notin E, i \neq j \\ 0, & \text{otherwise} \end{cases}$$

for all $i, j \in V$. Note that, by our choice of η, we have

$$S_{C_q, C_s} = \begin{cases} \tilde{S}_{C_q, C_s}, & \text{if } q, s \in \{1, \ldots, k+1\},\ q \neq s \\ \tilde{S}_{C_{k+1}, C_{k+1}} + \hat{r}I, & \text{if } q = s = k+1. \end{cases} \tag{5.3.17}$$

by (5.3.1).

By construction, μ, λ, η, and S satisfy (5.3.1). Since the ith row sum of X^* is equal to 1 for all $i \in C_q$ for all $q = 1, \ldots k$ and is equal to 0 for all $i \in C_{k+1}$, X^* and λ satisfy the complementary slackness condition (5.3.3). Moreover,

$$\langle X^*, S \rangle = \sum_{q=1}^{k} \left(\sum_{i \in C_q} \sum_{j \in C_q} \frac{\hat{r}}{r_q} [S_{C_q, C_q}]_{i,j} \right) = \sum_{q=1}^{k} \frac{1}{r_q} \left(1 - r_q \left(\frac{1}{r_q} \right) \right) = 0,$$

and thus X^* and S satisfy (5.3.3). It remains to prove that (5.3.12) implies that S is positive semidefinite.

To prove that S is positive semidefinite we show that $\mathbf{x}^T S \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbf{R}^N$ if $\tilde{S}$ satisfies (5.3.12). Fix $\mathbf{x} \in \mathbf{R}^N$ and decompose $\mathbf{x}$ as $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ where

$$\mathbf{x}_1(C_i) = \begin{cases} \boldsymbol{\phi}_i \, \mathbf{e}, & i \in \{1, \ldots, k\} \\ 0, & i = k+1 \end{cases}$$

for $\boldsymbol{\phi} \in \mathbf{R}^k$ chosen so that $\mathbf{x}_2(C_i)^T \mathbf{e} = 0$ for $i = 1 \ldots, k$, $\mathbf{x}_2(C_{k+1}) = \mathbf{x}(C_{k+1})$. Then, by our choice of $\mathbf{x}_1$ and $\mathbf{x}_2$,

$$\begin{aligned} \mathbf{x}^T S \mathbf{x} &= \mathbf{x}_2^T S \mathbf{x}_2 \\ &= \hat{r} \|\mathbf{x}_2(C_1 \cup \cdots \cup C_k)\|^2 + (\hat{r} - 1) \|\mathbf{x}_2(C_{k+1})\|^2 + \mathbf{x}_2^T \tilde{S} \mathbf{x}_2 \\ &\geq (\hat{r} - 1 - \|\tilde{S}\|) \|\mathbf{x}_2\|^2. \end{aligned}$$

Therefore, S is positive semidefinite, and, hence, X^* is optimal for (5.2.13) if $\|\tilde{S}\| \leq \hat{r} - 1$.

Now suppose that $\|\tilde{S}\| < \hat{r} - 1$ and, for all $i = 1, \ldots, k$, no node in C_i is adjacent to every node in some other clique. Then X^* is optimal for (5.2.13). For all $i = 1, \ldots, k$, let $\mathbf{v}_i \in \mathbf{R}^N$ be the characteristic vector of C_i. That is,

$$[\mathbf{v}_i]_j = \begin{cases} 1, & \text{if } j \in C_i \\ 0, & \text{otherwise.} \end{cases}$$

Notice that $X^* = \sum_{i=1}^{k} (1/r_i) \mathbf{v}_i (\mathbf{v}_i)^T$. By complementary slackness, $\langle X^*, S \rangle = 0$ and, thus, $\mathbf{v}_i$ is in the nullspace of S for all $i = 1, \ldots, k$. On the other hand, consider nonzero $\mathbf{x} \in \mathbf{R}^N$ such that $\mathbf{x}^T \mathbf{v}_i = 0$ for all $i = 1, \ldots, k$. That is, $\mathbf{x}$ is orthogonal to the span of $\{\mathbf{v}_i : i = 1, \ldots, k\}$. Then

$$\mathbf{x}^T S \mathbf{x} = \hat{r} \|\mathbf{x}(C_1 \cup \cdots \cup C_k)\|^2 + (\hat{r} - 1) \|\mathbf{x}(C_{k+1})\|^2 + \mathbf{x}^T \tilde{S} \mathbf{x}$$

$$\geq (\hat{r} - 1 - \|\tilde{S}\|)\|\mathbf{x}\|^2 > 0.$$

Therefore, $\text{Null}(S) = \text{span}\{\mathbf{v}_i : i = 1, \ldots, k\}$ and $\text{rank}\,(S) = N - k$.

Now suppose that $\hat{X}$ is also optimal for (5.2.13). Then, by complementary slackness, $\langle \hat{X}, S \rangle = 0$ which holds if and only if $\hat{X}S = 0$. Therefore, the row and column spaces of $\hat{X}$ lie in the null space of S. It follows immediately, since $\hat{X} \succeq 0$, that $\hat{X}$ can be written in the form

$$\hat{X} = \sum_{i=1}^{k} \sigma_i \mathbf{v}_i \mathbf{v}_i^T + \sum_{i=1}^{k} \sum_{\substack{j=1 \\ j \neq i}}^{k} \omega_{i,j} \mathbf{v}_i \mathbf{v}_j^T$$

for some $\sigma \in \mathbf{R}_+^k$ and $\omega \in \Sigma^k$, where Σ^k denotes the set of $k \times k$ symmetric matrices. Now, if $\omega_{i,j} \neq 0$ for some $i \neq j$ then every entry in the block $\hat{X}(C_i, C_j) = \hat{X}(C_j, C_i)^T$ must be equal to $\omega_{i,j}$. Since each of these entries is nonzero, this implies that each node in C_i is adjacent to every node in C_j, contradicting Assumption (5.3.13). Therefore, $\hat{X}$ has singular value decomposition $\hat{X} = \sigma_1 \mathbf{v}_1 \mathbf{v}_1^T + \cdots + \sigma_k \mathbf{v}_k \mathbf{v}_k^T$. Moreover, since $\hat{X}$ is optimal for (5.2.13) it must have objective value equal to that of X^* and thus

$$\sum_{i=1}^{k} r_i = \sum_{i=1}^{N} \sum_{j=1}^{N} X_{i,j}^* = \sum_{i=1}^{N} \sum_{j=1}^{N} \hat{X}_{i,j} = \sum_{i=1}^{k} \sigma_i r_i^2. \tag{5.3.18}$$

Further, since $\hat{X}$ is feasible for (5.2.13),

$$\sigma_i r_i \leq 1 \tag{5.3.19}$$

for all $i = 1, \ldots, k$. Combining (5.3.18) and (5.3.19) shows that $\sigma_i = 1/r_i$ for all $i = 1, \ldots, k$ and, hence, $\hat{X} = X^*$ as required. ∎

5.3.2 An upper bound on $\|\tilde{S}\|$ in the adversarial noise case

Suppose that the graph $G = (V, E)$ contains a k-disjoint-clique subgraph K^* indexed by the disjoint cliques $C_1, \ldots, C_k$. Let $C_{k+1} := V \setminus (\cup_{i=1}^{k} C_i)$, let $r_i := |C_i|$ for all $i = 1, \ldots, k+1$ and let $\hat{r} := \min\{C_1, \ldots, C_k\}$. To show that Theorem 5.2.1 holds, we show that the feasible solution X^* for (5.2.13) corresponding to K^* as defined by 5.2.10 satisfies the uniqueness and optimality conditions given by Theorem 5.3.2 in the case that K^* satisfies the conditions (5.2.20) and (5.2.21). In particular, the proof relies on establishing that $\tilde{S}$

as constructed in the previous section satisfies $\|\tilde{S}\| \leq O(\hat{r})$ in the case that K^* satisfies (5.2.20) and (5.2.21).

For the remainder of the proof, to simplify the notation, we assume that $r_{k+1} \leq 2\rho\hat{r}^2$. If $r_{k+1} > 2\rho\hat{r}^2$ then C_{k+1} would include one or more isolated nodes (i.e, nodes of degree 0) since $|E(G) \setminus E(K^*)| \leq \rho\hat{r}^2$ by assumption, and these nodes can simply be deleted in a preliminary phase of the algorithm. The algorithm still works with an arbitrary number of isolated nodes in $G \setminus K^*$, but the notation in the proof requires some needless additional complexity.

Recall that the construction of the multipliers presented in Section 5.3.1 depended on two scalars θ in (5.3.4) and γ in (5.3.11): choose $\theta = 1$ and $\gamma = 0$. Choose $q, s \in \{1, \ldots, k\}$ such that $q \neq s$ and let D, F, H, $\mathbf{b}$, and c be defined as in Section 5.3.1. Without loss of generality, we may assume that $r_q \leq r_s$. Moreover, let $\mathbf{y}$ and $\mathbf{z}$ be the solution of the system (5.3.4) and define A, Q, P as in Section 5.3.1. We impose the assumption that $\delta \in (0, 0.382)$. The constant 0.382 is chosen so that

$$0 < \delta < (1 - \delta)^2. \tag{5.3.20}$$

We will show that, under this assumption, there exists some $\beta > 0$ depending only on δ such that

$$\|\tilde{S}_{C_q, C_s}\|^2 \leq \beta \|\mathbf{b}^{q,s}\|_1$$

for all $q, s \in \{1, \ldots, k\}$ such that $q \neq s$.

We begin by showing that, under this assumption, $\mathbf{y}$ and $\mathbf{z}$ are uniquely determined. Note that, since $n_i^s = r_s - D_{ii} \leq \delta r_s$ for all $i \in C_q$ and $n_i^q = r_q - F_{ii} \leq \delta r_q$ for all $i \in C_s$ by Assumption (5.2.20), D and F are nonsingular and, hence, A is nonsingular. Moreover, $A + P = A(I + A^{-1}P)$. Therefore, $A + P$ is nonsingular if $\|A^{-1}P\| < 1$. Note that, for all $t > 0$, we have

$$\lambda_{\min}(D + t\mathbf{e}\mathbf{e}^T) \geq \lambda_{\min}(D) = \min_{i \in C_q} D_{ii} \tag{5.3.21}$$

since $\mathbf{e}\mathbf{e}^T \succeq 0$, where $\lambda_{\min}(D + t\mathbf{e}\mathbf{e}^T)$ is the smallest eigenvalue of the symmetric matrix $D + t\mathbf{e}\mathbf{e}^T$. Taking $t = 1$ in (5.3.21) shows that

$$\|(D + \mathbf{e}\mathbf{e}^T)^{-1}\| \leq \|D^{-1}\| = \frac{1}{\min_{i \in C_q} D_{ii}} \leq \frac{1}{(1 - \delta) r_s} \tag{5.3.22}$$

since, for each $i \in C_q$, we have

$$(1 - \delta) r_s \leq D_{ii} \leq r_s$$

by the asssumption that (5.2.20) holds. Similarly,

$$\|(F+\mathbf{e}\mathbf{e}^T)^{-1}\| \leq \|F^{-1}\| = \frac{1}{\min_{j\in C_s} F_{ii}} \leq \frac{1}{(1-\delta)r_q}. \tag{5.3.23}$$

Combining (5.3.22) and (5.3.23) we have

$$\|A^{-1}\| = \frac{1}{\min\{\|(D+\mathbf{e}\mathbf{e}^T)^{-1}\|, \|(F+\mathbf{e}\mathbf{e}^T)^{-1}\|\}} \leq \frac{1}{(1-\delta)r_q}. \tag{5.3.24}$$

On the other hand,

$$\|P\| = \|H-\mathbf{e}\mathbf{e}^T\| \leq \|H-\mathbf{e}\mathbf{e}^T\|_F = \left(\sum_{i\in C_q}\sum_{j\in C_s}(H_{ij}-1)^2\right)^{1/2} \leq \sqrt{\delta}r_q \tag{5.3.25}$$

since $H_{ij}-1$ is equal to -1 in the case that $(i,j)\in E$ and 0 otherwise and there at most δr_q^2 edges between C_q and C_s by Assumption (5.2.20). Therefore, since $\delta < (1-\delta)^2$ by Assumption (5.3.20), we have

$$\|A^{-1}P\| \leq \|A^{-1}\|\|P\| \leq \frac{\sqrt{\delta}}{1-\delta} < 1$$

and, thus, $A+P$ is nonsingular and $\mathbf{y}$ and $\mathbf{z}$ are uniquely determined.

Now, recall that $\tilde{S}_{C_q,C_s} = H\circ(\mathbf{y}\mathbf{e}^T+\mathbf{e}\mathbf{z}^T) - c(\mathbf{e}\mathbf{e}^T-H)$. In order to calculate an upper bound on $\|\tilde{S}_{C_q,C_s}\|$ we write $\tilde{S}_{C_q,C_s}$ as

$$\tilde{S}_{C_q,C_s} = m_1+m_2+m_3+m_4+m_5 \tag{5.3.26}$$

where

$$\begin{gathered} m_1 := H\circ(\bar{\mathbf{y}}\mathbf{e}^T), \quad m_2 := H\circ(\mathbf{e}\bar{\mathbf{z}}^T), \quad m_3 := H\circ(Q_1\mathbf{b}\mathbf{e}^T), \\ m_4 := H\circ(\mathbf{e}(Q_2\mathbf{b})^T), \quad m_5 := -c(\mathbf{e}\mathbf{e}^T-H) \end{gathered} \tag{5.3.27}$$

and apply the triangle inequality to obtain

$$\|\tilde{S}_{C_q,C_s}\| \leq \sum_{i=1}^{5}\|m_i\|. \tag{5.3.28}$$

Throughout our analysis of $\|\tilde{S}_{C_q,C_s}\|$ we will use the following series of inequalities. For

any $W \in \mathbf{R}^{m\times n}$, $\mathbf{u} \in \mathbf{R}^m$ and $\mathbf{v} \in \mathbf{R}^n$, we have

$$\begin{aligned}\|W \circ \mathbf{u}\mathbf{v}^T\| = \|\mathrm{Diag}\,(\mathbf{u}) \cdot W \cdot \mathrm{Diag}\,(\mathbf{v})\| &\leq \|\mathrm{Diag}\,(\mathbf{u})\|\|\mathrm{Diag}\,(\mathbf{v})\|\|W\| \\ &= \|\mathbf{u}\|_\infty\|\mathbf{v}\|_\infty\|W\|. \end{aligned} \tag{5.3.29}$$

On the other hand,

$$\begin{aligned}\|W \circ \mathbf{u}\mathbf{v}^T\| &= \|\mathrm{Diag}\,(\mathbf{u}) \cdot W \cdot \mathrm{Diag}\,(\mathbf{v})\| \\ &\leq \|\mathbf{v}\|_\infty\|\mathrm{Diag}\,(\mathbf{u}) \cdot W\| \leq \|\mathbf{v}\|_\infty\|\mathrm{Diag}\,(\mathbf{u}) \cdot W\|_F \\ &= \|\mathbf{v}\|_\infty \left(\sum_{i=1}^m \mathbf{u}_i^2\|W(i,:)\|^2\right)^{1/2} \qquad (5.3.30) \\ &\leq \|\mathbf{u}\|\|\mathbf{v}\|_\infty \max_{i=1,\ldots,m} \|W(i,:)\| \qquad (5.3.31)\end{aligned}$$

and

$$\|W \circ \mathbf{u}\mathbf{v}^T\| \leq \|\mathbf{u}\|_\infty\|\mathbf{v}\| \max_{j=1,\ldots,n} \|W(:,j)\| \tag{5.3.32}$$

where $W(i,:)$ and $W(:,j)$ denote the ith and jth row and column of W.

We begin with $\|m_1\|$. Applying the bound (5.3.31) with $W = H$, $\mathbf{u} = \bar{\mathbf{y}}$, and $\mathbf{v} = \mathbf{e}$ we have

$$\|m_1\|^2 \leq \max_{i\in C_q} D_{ii}\|\bar{\mathbf{y}}\|^2. \tag{5.3.33}$$

Here, we used the fact that $\max_{i\in C_q} \|H(i,:)\| = \max_{i\in C_q} \sqrt{D_{ii}}$ since the ith row of H contains exactly $r_s - n_i^s$ equal to 1. Thus, since

$$\|\bar{\mathbf{y}}\| \leq \|(D + \mathbf{e}\mathbf{e}^T)^{-1}\|\|\mathbf{b}_1\| \leq \frac{\|\mathbf{b}_1\|}{\min_{i\in C_q} D_{ii}} \leq \frac{\|\mathbf{b}_1\|}{(1-\delta)r_s},$$

it follows immediately that

$$\|m_1\|^2 \leq \frac{1}{(1-\delta)^2 r_s}\|\mathbf{b}_1\|^2 \tag{5.3.34}$$

since $D_{ii} \leq r_s$ for all $i \in C_q$. By an identical calculation, we have

$$\|m_2\|^2 \leq \frac{1}{(1-\delta)^2 r_q}\|\mathbf{b}_2\|^2. \tag{5.3.35}$$

Next, applying (5.3.31) with $W = H$, $\mathbf{u} = Q_1\mathbf{b}$, $\mathbf{v} = \mathbf{e}$ yields

$$\|m_3\|^2 \leq \max_{i \in C_q} D_{ii} \|Q_1\mathbf{b}\|^2 \leq r_s \|Q_1\mathbf{b}\|^2 \leq r_s \|Q_1\|^2 \|\mathbf{b}\|^2,$$

since $\max_i D_{ii} \leq r_s$. To derive an upper bound on $\|m_3\|^2$, we first derive an upper bound on $\|Q_1\|$.

Note that

$$Q = (A+P)^{-1} - A^{-1} = ((I + A^{-1}P)^{-1} - I)A^{-1} = \sum_{\ell=1}^{\infty} (-A^{-1}P)^{\ell} A^{-1} \tag{5.3.36}$$

since $(I+X)^{-1} = \sum_{\ell=0}^{\infty} (-X)^{\ell}$ for all X such that $\|X\| < 1$ by Taylor's Theorem. Notice that

$$A^{-1}P = \begin{pmatrix} 0 & P_1 \\ P_2 & 0 \end{pmatrix}$$

where

$$P_1 = (D + \theta \mathbf{e}\mathbf{e}^T)^{-1}(H - \theta \mathbf{e}\mathbf{e}^T), \quad P_2 = (F + \theta \mathbf{e}\mathbf{e}^T)^{-1}(H^T - \theta \mathbf{e}\mathbf{e}^T).$$

It follows immediately that

$$Q = \sum_{\ell=0}^{\infty} \left(\begin{pmatrix} (P_1P_2)^{\ell+1} & 0 \\ 0 & (P_2P_1)^{\ell+1} \end{pmatrix} + \begin{pmatrix} 0 & (P_1P_2)^{\ell}P_1 \\ (P_2P_1)^{\ell}P_2 & 0 \end{pmatrix} \right) A^{-1} \tag{5.3.37}$$

since, for any integer $\ell \geq 1$

$$\begin{pmatrix} 0 & P_1 \\ P_2 & 0 \end{pmatrix}^{\ell} = \begin{cases} \begin{pmatrix} (P_1P_2)^{\ell/2} & 0 \\ 0 & (P_2P_1)^{\ell/2} \end{pmatrix}, & \text{if } \ell \text{ even} \\ \begin{pmatrix} 0 & (P_1P_2)^{(\ell-1)/2}P_1 \\ (P_2P_1)^{(\ell-1)/2}P_2 & 0 \end{pmatrix}, & \text{if } \ell \text{ odd.} \end{cases}$$

Therefore,

$$\|Q_1\| \leq \|(D + \theta \mathbf{e}\mathbf{e}^T)^{-1}\| \sum_{\ell=1}^{\infty} \|P_1P_2\|^{\ell} + \|P_1\|\|(F + \theta \mathbf{e}\mathbf{e}^T)^{-1}\| \sum_{\ell=0}^{\infty} \|P_1P_2\|^{\ell} \tag{5.3.38}$$

and

$$\|Q_2\| \leq \|(F+\theta\mathbf{e}\mathbf{e}^T)^{-1}\| \sum_{\ell=1}^{\infty} \|P_1P_2\|^{\ell} + \|P_2\|\|(D+\theta\mathbf{e}\mathbf{e}^T)^{-1}\| \sum_{\ell=0}^{\infty} \|P_1P_2\|^{\ell}. \tag{5.3.39}$$

Substituting (5.3.22), (5.3.23) and (5.3.25) into (5.3.38) yields

$$\|Q_1\| \leq \frac{1}{(1-\delta)r_s} \sum_{\ell=1}^{\infty} \left(\frac{\delta}{(1-\delta)^2}\right)^{\ell} + \frac{\delta^{1/2}}{(1-\delta)^2 r_s} \sum_{\ell=0}^{\infty} \left(\frac{\delta}{(1-\delta)^2}\right)^{\ell} \leq \tilde{c}/\sqrt{r_s} \tag{5.3.40}$$

where

$$\tilde{c} = \frac{2\max\{\delta/(1-\delta), \sqrt{\delta}\}}{(1-\delta)^2 - \delta}$$

since

$$\|P_1P_2\| \leq \|H - \mathbf{e}\mathbf{e}^T\|^2 \|D^{-1}\| \|F^{-1}\| \leq \frac{\delta}{(1-\delta)^2}.$$

Note that Assumption (5.3.20) ensures that the infinite series in (5.3.40) converge. It follows that

$$\|m_3\|^2 \leq \frac{\tilde{c}^2}{r_q} \|\mathbf{b}\|^2. \tag{5.3.41}$$

On the other hand,

$$\|Q_2\| \leq \frac{1}{(1-\delta)r_q} \sum_{\ell=1}^{\infty} \left(\frac{\delta}{(1-\delta)^2}\right)^{\ell} + \frac{\sqrt{\delta}}{(1-\delta)^2\sqrt{r_q r_s}} \sum_{\ell=0}^{\infty} \left(\frac{\delta}{(1-\delta)^2}\right)^{\ell} \leq \tilde{c}/r_q \tag{5.3.42}$$

since $\sqrt{r_q r_s} \geq r_q$. Thus, applying (5.3.32) with $W = H$, $\mathbf{u} = \mathbf{e}$, $\mathbf{v} = Q_2\mathbf{b}$ we have

$$\|m_4\|^2 \leq r_q \|Q_2\|^2 \|\mathbf{b}\|^2 \leq \frac{\tilde{c}^2}{r_q} \|\mathbf{b}\|^2. \tag{5.3.43}$$

Finally,

$$\begin{aligned} \|m_5\|^2 = \|c(H - \mathbf{e}\mathbf{e}^T)\|^2 &\leq \|c(H - \mathbf{e}\mathbf{e}^T)\|_F^2 \\ &= c \sum_{i \in C_q} \sum_{j \in C_s} (H_{ij} - \mathbf{e}\mathbf{e}^T)^2 \\ &= c \sum_{i \in C_q} n_i^s = \|\mathbf{b}_1\|_1. \end{aligned} \tag{5.3.44}$$

Therefore, there exists $\beta \in \mathbf{R}$ such that

$$\|\tilde{S}_{C_q,C_s}\|^2 \leq \beta \frac{\|\mathbf{b}\|^2}{r_q} + \|\mathbf{b}\|_1. \tag{5.3.45}$$

Since $\|\mathbf{b}\|^2 \leq \|\mathbf{b}\|_1 \|\mathbf{b}\|_\infty$ by (2.1.2) and

$$\|\mathbf{b}\|_\infty = c \cdot \max\left\{\max_{i \in C_q} n_i^s, \max_{i \in C_s} n_i^q\right\}\right\} \leq \delta c \min\{r_q, r_s\} = \delta c r_q \tag{5.3.46}$$

by Assumption (5.2.20), there exists $\tilde{\beta}$ depending only on δ such that

$$\|\tilde{S}_{C_q,C_s}\|^2 \leq \tilde{\beta}\|\mathbf{b}\|_1 \tag{5.3.47}$$

as required.

Next, consider $\tilde{S}_{C_q,C_{k+1}}$ for some $q \in \{1,\ldots,k\}$. Recall that

$$[\tilde{S}_{C_q,C_{k+1}}]_{ij} = \begin{cases} -c, & \text{if } (i,j) \in E \\ cn_j/(r_q - n_j), & \text{otherwise} \end{cases}$$

where $n_j = n_j^q$ is the number of edges from $j \in C_{k+1}$ to C_q for each $j \in C_{k+1}$. Hence,

$$\begin{aligned} \|\tilde{S}_{C_q,C_{k+1}}\|^2 &\leq \|\tilde{S}_{C_q,C_{k+1}}\|_F^2 \\ &= \sum_{j \in C_{k+1}} \left(n_j c^2 + (r_q - n_j)\left(\frac{n_j c}{r_q - n_j}\right)^2 \right) \\ &\leq c^2 \sum_{j \in C_{k+1}} \left(n_j + \frac{\delta n_j}{(1-\delta)} \right) \\ &= \frac{c^2}{1-\delta} |E(C_q, C_{k+1})| \end{aligned} \tag{5.3.48}$$

where $E(C_q, C_{k+1})$ is the set of edges from C_q to C_{k+1}. Similarly, by our choice of $\gamma = 0$ in $(\tilde{\sigma}_4)$, we have

$$\begin{aligned} \|\tilde{S}_{C_{k+1},C_{k+1}}\| &= \|S_{C_{k+1},C_{k+1}} - \hat{r}I\|^2 \\ &\leq \|S_{C_{k+1},C_{k+1}} - \hat{r}I\|_F^2 \\ &= r_{k+1} + 2|E(C_{k+1}, C_{k+1})|. \end{aligned} \tag{5.3.49}$$

Let $\tilde{\mathbf{b}}$ be the vector obtained by concatenating $\mathbf{b}^{q,s}$ for all $q, s \in \{1, \ldots, k\}$. Then, there exist scalars $\hat{c}_1$, $\hat{c}_2 \in \mathbf{R}$ depending only on δ such that

$$\sum_{q=1}^{k+1}\sum_{s=1}^{k+1} \|\tilde{S}_{C_q,C_s}\|^2 = \sum_{\substack{q,s\in\{1,\ldots,k\}\\ q\neq s}} \|\tilde{S}_{C_q,C_s}\|^2 + 2\sum_{q=1}^{k} \|\tilde{S}_{C_q,C_{k+1}}\|^2 + \|\tilde{S}_{k+1,k+1} - \hat{r}I\|^2$$

$$\leq \hat{c}_1\|\tilde{\mathbf{b}}\|_1 + \hat{c}_2 \sum_{q=1}^{k+1} |E(C_q, C_{k+1})| + r_{k+1}$$

by (5.3.47), (5.3.48) and (5.3.49). It follows that, since $\|\mathbf{b}^{q,s}\|_1 \leq |E(C_q, C_s)|$ for all $q, s \in \{1, \ldots, k\}$ such that $q \neq s$, there exists $\hat{c}_3 \geq 0$ depending only on δ such that

$$\sum_{q=1}^{k+1}\sum_{s=1}^{k+1} \|\tilde{S}_{C_q,C_s}\|^2 \leq \hat{c}_3 R + r_{k+1}$$

where $R := |E(G)\backslash E(K^*)|$ is the number of edges of G not contained in the k-disjoint-clique subgraph K^* composed of $C_1, \ldots, C_k$. The hypothesis of the theorem is that $R \leq \rho\hat{r}^2$. We have also assumed earlier that $r_{k+1} \leq 2\rho\hat{r}^2$. Hence, the sum of the squares of the 2-norms of the blocks of $\tilde{S}$ is at most $(\hat{c}_3 + 2)\rho\hat{r}^2$. Therefore, there exists some $\rho > 0$ depending only on δ such that the preceding inequality implies $\|\tilde{S}\| \leq \hat{r} - 1$. Applying Theorem 5.3.2 completes the proof.

5.3.3 An upper bound on $\|\tilde{S}\|$ in the random noise case

Suppose that the random graph $G = (V, E)$ containing the k-disjoint-clique subgraph K^* composed of cliques $C_1, \ldots, C_k$ is constructed according to (Ω_1) and (Ω_2) with probability p. Let $N := |V|$, $C_{k+1} = V \setminus \cup_{i=1}^{k} C_i$ and $r_i = |C_i|$ for all $i = 1, \ldots, k+1$. Further, let $\theta = 1 - p$ in (5.3.4) and $\gamma = p/(1-p)$ in (5.3.11). As before, our strategy for proving Theorem 5.2.2 is to show that the matrix X^* corresponding to K^* as defined by (5.2.10) satisfies the sufficient conditions given by Theorem 5.3.1. It suffices to show that the auxiliary matrix $\tilde{S}$ as constructed in Section 5.3.1 satisfies $\|\tilde{S}\| \leq O(\hat{r})$ with probability tending exponentially to 1 as $\hat{r} \to \infty$ whenever $r_i \leq \hat{r}^{3/2}$ for all $i = 1, \ldots, k$. Indeed, in this case, applying Theorem 5.3.2 immediately shows that X^* is the unique optimal solution of (5.2.13) and K^* is the unique maximum node k-disjoint-clique subgraph of G. The following theorem provides the necessary upper bound on $\|\tilde{S}\|$.

Theorem 5.3.3 *Suppose that $G = (V, E)$ has a k-disjoint-clique subgraph K^* composed of the cliques $C_1, \ldots, C_k$ and let $C_{k+1} := V \setminus (\cup_{i=1}^k C_i)$. Let $r_i = |C_i|$ for all $i = 1, \ldots, k+1$ and suppose that $r_q \leq \hat{r}^{3/2}$ for all $q = 1, 2, \ldots, k$ where $\hat{r} = \min_{i=1,\ldots,k}\{r_i\}$. Then there exist $\beta_1, \beta_2 > 0$ depending only on p such that*

$$\|\tilde{S}\| \leq \beta_1 \left(\sum_{s=1}^{k} r_s^2 \right)^{1/2} \left(\sum_{q=1}^{k} \frac{1}{r_q} \right)^{1/2} + \beta_2 \sqrt{N} \tag{5.3.50}$$

with probability tending exponentially to 1 as $\hat{r}$ approaches ∞.

The remainder of this section is devoted to a proof of Theorem 5.3.3. We write $\tilde{S}$ as

$$\tilde{S} = \tilde{S}_1 + \tilde{S}_2 + \tilde{S}_3 + \tilde{S}_4 + \tilde{S}_4^T$$

where $\tilde{S}_i \in \mathbf{R}^{N \times N}$, $i = 1, \ldots, 4$ are $(k+1)$ by $(k+1)$ block matrices such that

$$\tilde{S}_1(C_q, C_s) = \begin{cases} \tilde{S}(C_q, C_s), & \text{if } q, s \in \{1, \ldots, k\},\ q \neq s \\ 0, & \text{otherwise} \end{cases}$$

$$\tilde{S}_2(C_q, C_s) = \begin{cases} R(C_q, C_s), & \text{if } q, s \in \{1, \ldots, k\} \\ \hat{S}(C_q, C_{k+1}), & \text{if } s = k+1 \\ \hat{S}(C_{k+1}, C_s), & \text{if } q = k+1 \\ \tilde{S}(C_{k+1}, C_{k+1}), & \text{if } q = s = k+1 \end{cases}$$

$$\tilde{S}_3(C_q, C_s) = \begin{cases} -R(C_q, C_s), & \text{if } q, s \in \{1, \ldots, k\} \\ 0, & \text{otherwise} \end{cases}$$

$$\tilde{S}_4(C_q, C_s) = \begin{cases} S(C_q, C_{k+1}) - \hat{S}(C_q, C_{k+1}), & \text{if } s = k+1,\ q \in \{1, \ldots, k\} \\ 0, & \text{otherwise} \end{cases}$$

where $R \in \Sigma^N$ is a symmetric random matrix with independently identically distributed entries such that

$$R_{ij} = \begin{cases} -1, & \text{with probability } p \\ p/(1-p), & \text{with probability } 1-p \end{cases}$$

and $\hat{S} \in \mathbf{R}^{N \times N}$ such that

$$\hat{S}_{ij} = \begin{cases} -1, & \text{if } (i,j) \in E \\ p/(1-p), & \text{otherwise.} \end{cases}$$

Notice that, by Theorem 2.6.4, there exist some $\kappa_1, \kappa_2, \kappa_3 > 0$ such that

$$P\Big(\|\tilde{S}_2\| + \|\tilde{S}_3\| \geq \kappa_1\sqrt{N}\Big) \leq \kappa_2 \exp(-\kappa_3 N^{1/6}). \tag{5.3.51}$$

Moreover, by Theorem 4.4.2, there exists $\kappa_4 > 0$ and $\kappa_5, \kappa_6, \in (0,1)$ such that

$$P\Big(\|\tilde{S}_4\| \geq \kappa_4\sqrt{N}\Big) \leq \kappa_5^N + N\kappa_6^N. \tag{5.3.52}$$

Hence, there exists some scalar β_4 depending only on p such that

$$\|\tilde{S}\| \leq \|\tilde{S}_1\| + \beta_4\sqrt{N}$$

with probability tending exponentially to 1 as $\hat{r} \to \infty$. It remains to prove that

$$\|\tilde{S}_1\| \leq O\left(\left(\sum_{s=1}^{k} r_s^2\right)^{1/2}\left(\sum_{q=1}^{k}\frac{1}{r_q}\right)^{1/2}\right)$$

with probability tending exponentially to 1 as $\hat{r}$ approaches ∞.

To do so, consider two vertex sets C_q and C_s such that $q, s \in \{1, \ldots, k\}$, $q \neq s$. Without loss of generality we may assume that $r_q \leq r_s$. Define $H = H_{q,s}$, $D = D_{q,s}$, $F = F_{q,s}$, $\mathbf{b} = \mathbf{b}_{q,s}$, $c = c_{q,s}$, $\mathbf{y}$, $\mathbf{z}$, A, and P as in Section 5.3.1. The following theorem provides an upper bound on the spectral norm of $\tilde{S}(C_q, C_s)$ for $q \neq s$, that holds with probability tending exponentially to 1 as $\hat{r}$ approaches ∞.

Theorem 5.3.4 *Let $q, s \in \{1, \ldots, k\}$ such that $q \neq s$. Suppose that r_q and r_s satisfy*

$$r_q \leq r_s \leq r_q^{3/2}. \tag{5.3.53}$$

Then there exists $\tilde{B}_1 > 0$ depending only on p such that

$$\|\tilde{S}_1(C_q, C_s)\| = \|\tilde{S}(C_q, C_s)\| \leq \tilde{B}_1\frac{r_s}{\sqrt{r_q}} \tag{5.3.54}$$

with probability tending exponentially to 1 as $\hat{r}$ approaches ∞.

Note that Theorem 5.3.4 implies Theorem 5.3.3. Indeed, suppose that there exists $\tilde{B}_1$ depending only on p such that

$$\|\tilde{S}_1(C_q, C_s)\| \leq \tilde{B}_1 \frac{\max\{r_s, r_q\}}{(\min\{r_q, r_s\})^{1/2}} \tag{5.3.55}$$

for all $q, s \in \{1, \ldots, k\}$, $q \neq s$ with probability tending exponentially to 1 as $\hat{r} \to \infty$. Then

$$\sum_{q=1}^{k} \sum_{s=1}^{k} \|\tilde{S}_1(C_q, C_s)\|^2 \leq 2\tilde{B}_1^2 \sum_{q=1}^{k} \sum_{s=1}^{k} \frac{r_s^2}{r_q} = 2\tilde{B}_1^2 \left(\sum_{s=1}^{k} r_s^2 \right) \left(\sum_{q=1}^{k} \frac{1}{r_q} \right)$$

and, hence, there exists some $\hat{\beta}_1$ depending only on p such that

$$\|\tilde{S}_1\| \leq \hat{\beta}_1 \left(\sum_{s=1}^{k} r_s^2 \right)^{1/2} \left(\sum_{q=1}^{k} \frac{1}{r_q} \right)^{1/2}$$

with probability tending exponentially to 1 as $\hat{r} \to \infty$ as required.

The remainder of this section consists of a proof of Theorem 5.3.4. Recall that $\tilde{S}(C_q, C_s) = H \circ (\mathbf{y}\mathbf{e}^T + \mathbf{e}\mathbf{z}^T) - c(\mathbf{e}\mathbf{e}^T - H)$. We begin by showing that $A+P$ is nonsingular and, hence, $\mathbf{y}$ and $\mathbf{z}$ are uniquely determined. Let $\delta := (1-p)/(2p)$. Recall that $n_i^s = r_s - D_{ii}$ corresponds to r_s independent Bernoulli trials each succeeding with probability equal to p and, hence,

$$P(n_i^s \geq (1+\delta)pr_s) = P(r_s - D_{ii} \geq (1+\delta)pr_s) \leq \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}} \right)^{pr_s} \tag{5.3.56}$$

for each $i \in C_q$ by Theorem 2.6.1. Rearranging, we have that $D_{ii} \geq (\theta - \delta p)r_s$ with probability at least

$$1 - \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}} \right)^{pr_s}$$

for each $i \in C_q$. Similarly,

$$P(n_i^q \leq (1+\delta)pr_q) = P(F_{ii} \geq (\theta - \delta p)r_q) \geq 1 - \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}} \right)^{pr_q} \tag{5.3.57}$$

for all $i \in C_s$. Therefore, by the union bound, $r_s - D_{ii} \leq (1+\delta)pr_s$ for all $i \in C_q$ and $r_q - F_{ii} \leq (1+\delta)pr_q$ for all $i \in C_s$, and, hence, D, F are nonsingular with probability at

least

$$1-r_s\left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^{pr_q}-r_q\left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^{pr_s}\geq 1-(r_q+r_s)\left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^{p\hat{r}}. \quad (5.3.58)$$

Moreover, applying (5.3.21) shows that $D+\theta\mathbf{e}\mathbf{e}^T$ and $F+\theta\mathbf{e}\mathbf{e}^T$ are nonsingular and

$$\|(D+\theta\mathbf{e}\mathbf{e}^T)^{-1}\|\leq\|D^{-1}\|\leq\frac{1}{(\theta-\delta p)r_s}, \quad (5.3.59)$$

$$\|(F+\theta\mathbf{e}\mathbf{e}^T)^{-1}\|\leq\|F^{-1}\|\leq\frac{1}{(\theta-\delta p)r_q}, \quad (5.3.60)$$

with probability at least (5.3.58). It follows immediately that A is nonsingular and

$$\begin{aligned}\|A^{-1}\|&=\max\{\|(D+\theta\mathbf{e}\mathbf{e}^T)^{-1}\|,\|(F+\theta\mathbf{e}\mathbf{e}^T)^{-1}\|\}\\&\leq\frac{1}{(\theta-\delta p)\min\{r_q,r_s\}}=\frac{1}{(\theta-\delta p)r_q}\end{aligned} \quad (5.3.61)$$

with probability at least (5.3.58).

Recall that, in the case that A is nonsingular, it suffices to prove that $\|A^{-1}\|\|P\|<1$ to show that $A+P$ is nonsingular. Moreover, recall that $\theta=1-p$ is chosen to ensure that the entries of $H-\theta\mathbf{e}\mathbf{e}^T$ have expected value equal to 0. We can extend $H-\theta\mathbf{e}\mathbf{e}^T$ to an $r_s\times r_s$ random matrix $\tilde{P}$ with entries i.i.d. with expected value equal to 0 by adding r_s-r_q rows with entries i.i.d. such that each additional entry takes value equal to $-\theta$ with probability p and value equal to p with probability θ. Therefore, by Theorem 2.6.5

$$\|P\|=\|H-\theta\mathbf{e}\mathbf{e}^T\|\leq\|\tilde{P}\|\leq\gamma_1\sqrt{r_s} \quad (5.3.62)$$

for some $\gamma_1>0$ depending only on p with probability at least $1-\bar{c}_1\exp(-\bar{c}_2r_s^{\bar{c}_3})$ where $\bar{c}_i>0$ depend only on p. Combining (5.3.61), (5.3.62), (5.3.53) and applying the union bound shows that

$$\|A^{-1}\|\|P\|=\frac{\gamma_1\sqrt{r_s}}{(\theta-\delta p)r_q}<1$$

with probability at least

$$1-(r_q+r_s)\left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^{p\hat{r}}-\bar{c}_1\exp(-\bar{c}_2r_s^{\bar{c}_3})$$

for sufficiently large r_q. Therefore, $A + P$ is nonsingular and $\mathbf{y}$ and $\mathbf{z}$ are uniquely determined with probability tending exponentially to 1 as $\hat{r} \to \infty$.

For the remainder of the section we assume that $A + P$ is nonsingular. We define Q, Q_1, Q_2, $\bar{\mathbf{y}}$ and $\bar{\mathbf{z}}$ as in Section 5.3.1. To find an upper bound on $\|\tilde{S}(C_q, C_s)\|$, we decompose $\tilde{S}(C_q, C_s)$ as

$$\tilde{S}(C_q, C_s) = M_1 + M_2$$

where $M_1 := H \circ (\bar{\mathbf{y}}\mathbf{e}^T + \mathbf{e}\bar{\mathbf{z}}^T) - c(\mathbf{e}\mathbf{e}^T - H)$ and $M_2 := H \circ (Q_1\mathbf{b}\mathbf{e}^T + \mathbf{e}\mathbf{b}^T Q_2^T)$.

We first obtain an upper bound on the norm of M_1. We define $\mathbf{d} \in \mathbf{R}^{C_q}$ to be the vector such that $\mathbf{d}_i$ is the difference between the number of edges added between the node i and C_s and the expected number of such edges for each $i \in C_q$. That is,

$$\mathbf{d} = \mathbf{n}^s_{C_q} - E[\mathbf{n}^s_{C_q}] = \mathbf{n}^s_{C_q} - pr_s\mathbf{e}.$$

Similarly, we let $\mathbf{f} := \mathbf{n}^q_{C_s} - pr_q\mathbf{e}$. Note that, by our choice of $\mathbf{d}$ and $\mathbf{f}$, we have $r_sI - D = pr_sI + \mathrm{Diag}\,(\mathbf{d})$ and $r_qI - F = pr_qI + \mathrm{Diag}\,(\mathbf{f})$. Notice that for $\theta = 1 - p$ we have $D = \theta r_s I - \mathrm{Diag}\,(\mathbf{d})$. Expanding (5.3.8) we have

$$\begin{aligned}
\bar{\mathbf{y}} &= \left(D^{-1} - \frac{\theta D^{-1}\mathbf{e}\mathbf{e}^T D^{-1}}{1 + \theta\mathbf{e}^T D^{-1}\mathbf{e}}\right)\mathbf{b}_1 \\
&= \frac{D^{-1}}{1 + \theta\mathbf{e}^T D^{-1}\mathbf{e}}(\mathbf{b}_1 + \theta\mathbf{b}_1\mathbf{e}^T D^{-1}\mathbf{e} - \theta\mathbf{e}\mathbf{e}^T D^{-1}\mathbf{b}_1) \\
&= \frac{D^{-1}}{1 + \theta\mathbf{e}^T D^{-1}\mathbf{e}}(\mathbf{b}_1 + \theta(\mathbf{b}_1\mathbf{e}^T - \mathbf{e}\mathbf{b}_1^T)D^{-1}\mathbf{e})
\end{aligned}$$

since $\mathbf{e}^T D^{-1}\mathbf{b}_1 = \mathbf{b}_1^T D^{-1}\mathbf{e}$. Substituting $\mathbf{b}_1 = c(r_s\mathbf{e} - \bar{\mathbf{d}})$, where $\bar{\mathbf{d}} := \mathrm{diag}\,(D)$, we have

$$\mathbf{b}_1\mathbf{e}^T - \mathbf{e}\mathbf{b}_1^T = c(\mathbf{e}\bar{\mathbf{d}}^T - \bar{\mathbf{d}}\mathbf{e}^T)$$

and, hence,

$$\begin{aligned}
\bar{\mathbf{y}} &= \frac{cD^{-1}}{1 + \theta\mathbf{e}^T D^{-1}\mathbf{e}}(r_s\mathbf{e} - \bar{\mathbf{d}} + \theta(\mathbf{e}\bar{\mathbf{d}}^T - \bar{\mathbf{d}}\mathbf{e}^T)D^{-1}\mathbf{e}) \\
&= \frac{cD^{-1}}{1 + \theta\mathbf{e}^T D^{-1}\mathbf{e}}(r_s\mathbf{e} - \bar{\mathbf{d}} + \theta\mathbf{e}\mathbf{e}^T\mathbf{e} - \theta\bar{\mathbf{d}}\mathbf{e}^T D^{-1}\mathbf{e}) \\
&= \frac{cD^{-1}}{1 + \theta\mathbf{e}^T D^{-1}\mathbf{e}}(r_s\mathbf{e} + \theta r_q\mathbf{e}) - c\mathbf{e}
\end{aligned}$$

$$= \frac{c(r_s + \theta r_q)D^{-1}}{1 + \theta \mathbf{e}^T D^{-1}\mathbf{e}} \left(\frac{1}{\theta r_s}(D + \mathrm{Diag}\,(\mathbf{d})) \right) \mathbf{e} - c\mathbf{e}$$
$$= \left(\frac{c(r_s + \theta r_q)}{(1 + \theta \mathbf{e}^T D^{-1}\mathbf{e})\theta r_s} - c \right) \mathbf{e} + \frac{c(r_s + \theta r_q)}{(1 + \theta \mathbf{e}^T D^{-1}\mathbf{e})\theta r_s} D^{-1}\mathbf{d}$$

since

$$I = \frac{1}{\theta r_s}(D + \mathrm{Diag}\,(\mathbf{d})).$$

Let $\bar{\mathbf{y}}_1 := \omega_1 \mathbf{e}$, $\bar{\mathbf{y}}_2 := \upsilon_1 \mathbf{e}$ where

$$\omega_1 := \frac{c(\theta r_q + r_s)}{\theta(r_s + r_q)} - c, \qquad \upsilon_1 := \frac{c(\theta r_q + r_s)}{(1 + \theta \mathbf{e}^T D^{-1}\mathbf{e})\theta r_s} - c - \omega_1$$

and let

$$\bar{\mathbf{y}}_3 := \frac{c(\theta r_q + r_s)}{(1 + \theta \mathbf{e}^T D^{-1}\mathbf{e})\theta r_s} D^{-1}\mathbf{d}.$$

Hence, $\bar{\mathbf{y}} = \bar{\mathbf{y}}_1 + \bar{\mathbf{y}}_2 + \bar{\mathbf{y}}_3$. Similarly, $\bar{\mathbf{z}} = \bar{\mathbf{z}}_1 + \bar{\mathbf{z}}_2 + \bar{\mathbf{z}}_3$ where $\bar{\mathbf{z}}_1 := \omega_2 \mathbf{e}$, $\bar{\mathbf{z}}_2 := \upsilon_2 \mathbf{e}$ where

$$\omega_2 := \frac{c(r_q + \theta r_s)}{\theta(r_s + r_q)} - c, \qquad \upsilon_2 := \frac{c(r_q + \theta r_s)}{(1 + \theta \mathbf{e}^T F^{-1}\mathbf{e})\theta r_q} - c - \omega_2$$

and

$$\bar{\mathbf{z}}_3 := \frac{c(r_q + \theta r_s)}{(1 + \theta \mathbf{e}^T F^{-1}\mathbf{e})\theta r_q} F^{-1}\mathbf{f}.$$

Therefore, we can further decompose M_1 as $M_1 = \tilde{M}_1 + \tilde{M}_2 + \tilde{M}_3$ where

$$\tilde{M}_1 := H \circ (\bar{\mathbf{y}}_1 \mathbf{e}^T + \mathbf{e}\bar{\mathbf{z}}_1^T) - c(\mathbf{e}\mathbf{e}^T - H),$$
$$\tilde{M}_2 := H \circ (\bar{\mathbf{y}}_2 \mathbf{e}^T + \mathbf{e}\bar{\mathbf{z}}_2^T), \quad \tilde{M}_3 := H \circ (\bar{\mathbf{y}}_3 \mathbf{e}^T + \mathbf{e}\bar{\mathbf{z}}_3^T).$$

Notice that the matrix $\tilde{M}_1$ has entries corresponding to edges equal to $-c$ and remaining entries equal to $cp/(1-p)$ since

$$\omega_1 + \omega_2 = \frac{c(1+\theta)(r_q + r_s)}{\theta(r_q + r_s)} - 2c = \frac{cp}{\theta}.$$

Therefore, each entry of the matrix $\tilde{M}_1$ has expected value equal to 0. Moreover, each entry of the random block matrix $\hat{M}$ of the form

$$\hat{M} = \begin{pmatrix} \tilde{M}_1 \\ \tilde{R} \end{pmatrix}$$

has expected value equal to 0 if $\tilde{R}$ has identically independently distributed entries such that

$$\tilde{R}_{i,j} = \begin{cases} -c, & \text{with probability } p \\ cp/(1-p), & \text{with probability } 1-p. \end{cases}$$

Therefore, there exist $c_1, c_2, c_3, c_4 > 0$ such that

$$\|\tilde{M}_1\| \leq \|\hat{M}\| \leq c_4\sqrt{r_s} \tag{5.3.63}$$

with probability at least $1 - c_1 \exp(-c_2 r_s^{c_3})$ by Theorem 2.6.5.

To obtain upper bounds on $\|\tilde{M}_2\|$ and $\|\tilde{M}_3\|$ we will use the following bound on the ℓ_1 and ℓ_2-norms of random vectors with binomially distributed entries.

Theorem 5.3.5 *Let I_1, I_2 be index sets of size r_1, r_2 respectively. Let $\{X_{ij} : i \in I_1, j \in I_2\}$ be a collection of independent Bernoulli trials, each succeeding with probability p. Let $n_i = \sum_{j \in I_2} X_{ij}$ be the binomially distributed random variable describing the number of successes for each $i \in I_1$. Then there exists scalar B such that*

$$\sum_{i \in I_1} \frac{|n_i - pr_2|^\alpha}{r_2 - n_i} \leq B r_1 r_2^{\alpha/2-1} \tag{5.3.64}$$

for $\alpha = 1, 2$ with probability at least

$$1 - r_1 v_p^{r_2} - (2/3)^{r_1} \tag{5.3.65}$$

where $v_p = (e^\delta/(1+\delta)^{(1+\delta)})^p$ and $\delta = \min\{p, \sqrt{p} - p\}$.

The proof of Theorem 5.3.5 uses techniques similar to those used in the proof of Theorem 4.4.2 and is provided in Appendix A.2.

In the cases when $I_1 = C_q$, $I_2 = C_2$ and $n_i = r_s - D_{ii}$ for all $i \in C_q$ and $I_1 = C_s$, $I_2 = C_q$ and $n_i = r_q - F_{ii}$ for all $i \in C_s$, Theorem 5.3.5 specializes to the following bounds.

Lemma 5.3.2 *There exists $B > 0$ depending only on p such that*

$$\sum_{i \in C_q} \frac{|\theta r_s - D_{ii}|^\alpha}{D_{ii}} \leq B \frac{r_q}{r_s^{1-\alpha/2}} \tag{5.3.66}$$

and

$$\sum_{i \in C_s} \frac{|\theta r_q - F_{ii}|^\alpha}{F_{ii}} \leq B \frac{r_s}{r_q^{1-\alpha/2}} \tag{5.3.67}$$

for $\alpha = 1, 2$ with probability at least

$$1 - (r_q + r_s)\upsilon_p^{\hat{r}} - 2(2/3)^{\hat{r}} \tag{5.3.68}$$

where $\upsilon_p = (e^\delta/(1+\delta)^{(1+\delta)})^p$ and $\delta = \min\{p, \sqrt{p} - p\}$.

As an immediate corollary of Lemma 5.3.2, we have the following bound on $|\upsilon_1|$ and $|\upsilon_2|$.

Corollary 5.3.1 *There exists $B_1 > 0$ depending only on p such that*

$$|\upsilon_1| + |\upsilon_2| \leq B_1 \frac{r_q^{3/2} + r_s^{3/2}}{(r_q + r_s)(r_q r_s)^{1/2}}$$

with probability at least $1 - (r_q + r_s)\upsilon_p^{\hat{r}} - 2(2/3)^{\hat{r}}$.

Proof: We begin with υ_1. Notice that

$$\upsilon_1 = \frac{c(\theta r_q + r_s)}{\theta} \left(\frac{1}{(1 + \theta \mathbf{e}^T D^{-1} \mathbf{e}) r_s} - \frac{1}{r_q + r_s} \right) = \frac{c(\theta r_q + r_s)(r_q - \theta r_s \mathbf{e}^T D^{-1} \mathbf{e})}{\theta r_s (r_q + r_s)(1 + \theta \mathbf{e}^T D^{-1} \mathbf{e})}.$$

Moreover,

$$|\theta r_s \mathbf{e}^T D^{-1} \mathbf{e} - r_q| = \left| \sum_{i \in C_q} \frac{\theta r_s}{D_{ii}} - \sum_{i \in C_q} 1 \right| \leq \sum_{i \in C_q} \left| \frac{\theta r_s}{D_{ii}} - 1 \right| = \sum_{i \in C_q} \frac{|\theta r_s - D_{ii}|}{D_{ii}}$$

and, since $D_{ii} \leq r_s$ for all $i \in C_q$, we have

$$r_s(1 + \theta \mathbf{e}^T D^{-1} \mathbf{e}) \geq r_s \left(1 + \frac{\theta r_q}{r_s} \right) = \theta r_q + r_s. \tag{5.3.69}$$

Therefore, setting $\alpha = 1$ in (5.3.66) shows that

$$|\upsilon_1| \leq \frac{c(\theta r_q + r_s)Br_q}{\theta r_s^{3/2}(r_q + r_s)(1 + \theta \mathbf{e}^T D^{-1}\mathbf{e})} \leq \frac{cB(\theta r_q + r_s)r_q}{\theta\sqrt{r_s}(r_q + r_s)(\theta r_q + r_s)} \leq B_1 \frac{r_q}{\sqrt{r_s}(r_q + r_s)}$$

holds with probability at least $1 - (2/3)^{\hat{r}} - r_q v_p^{\hat{r}}$, where $B_1 := B/\theta$. By an identical calculation

$$|\upsilon_2| \leq \frac{B_1 r_s}{\sqrt{r_q}(r_q + r_s)}$$

with probability at least $1 - (2/3)^{\hat{r}} - r_s v_p^{\hat{r}}$. Applying the union bound yields the desired probabilistic bound on $|\upsilon_1| + |\upsilon_2|$. ■

Observe that, as an immediate consequence of Corollary 5.3.1 and the facts that $H \circ \mathbf{e}\mathbf{e}^T = H$ and $\|H\|_F \leq \sqrt{r_q r_s}$, we have

$$\|\tilde{M}_2\| = \|H \circ (\bar{\mathbf{y}}_2 \mathbf{e}^T + \mathbf{e}\bar{\mathbf{z}}_2^T)\| \leq (|\upsilon_1| + |\upsilon_2|)\|H\|_F \leq B_1 \frac{{r_q}^{3/2} + {r_s}^{3/2}}{r_q + r_s} \leq 2B_1\sqrt{r_s} \quad (5.3.70)$$

with probability at least $1 - (r_q + r_s)v_p^{\hat{r}} - 2(2/3)^{\hat{r}}$.

The following corollary of Lemma 5.3.2 provides an an upper bound on $\|\tilde{M}_3\|$.

Corollary 5.3.2 *There exists B_2 depending only on p such that*

$$\|\tilde{M}_3\| \leq \|H \circ (\bar{\mathbf{y}}_3 \mathbf{e}^T + \mathbf{e}\bar{\mathbf{z}}_3^T)\| \leq B_2(\sqrt{r_q} + \sqrt{r_s}) \quad (5.3.71)$$

with probability at least $1 - (r_q + r_s)v_p^{\hat{r}} - 2(2/3)^{\hat{r}}$.

Proof: To obtain an upper bound on $\|\tilde{M}_3\|$, we first obtain upper bounds on $\|H \circ (\bar{\mathbf{y}}_3 \mathbf{e}^T)\|$ and $\|H \circ (\mathbf{e}\bar{\mathbf{z}}_3^T)\|$. We begin with $\|H \circ (\bar{\mathbf{y}}_3 \mathbf{e}^T)\|$. Since

$$\sum_{i \in C_q} D_{ii}(D^{-1}\mathbf{d})_i^2 = \sum_{i \in C_q} \frac{|\theta r_s - D_{ii}|^2}{D_{ii}}$$

applying (5.3.66) with $\alpha = 2$ and (5.3.30) with $W = H$, $\mathbf{u} = \bar{\mathbf{y}}_3$, and $\mathbf{v} = \mathbf{e}$ shows that

$$\|H \circ (\bar{\mathbf{y}}_3 \mathbf{e}^T)\| \leq \left(\sum_{i \in C_q} \bar{\mathbf{y}}_3(i)^2 \|H(i,:)\|^2 \right)^{1/2}$$

$$= \frac{c(\theta r_q + r_s)}{(1+\theta \mathbf{e}^T D^{-1}\mathbf{e})\theta r_s} \cdot \left(\sum_{i \in C_q} D_{ii}[D^{-1}\mathbf{d}]_i^2 \right)^{1/2}$$

$$\leq \frac{c(\theta r_q + r_s)}{(1+\theta \mathbf{e}^T D^{-1}\mathbf{e})\theta r_s} (B r_q)^{1/2} \tag{5.3.72}$$

$$\leq B_2 \sqrt{r_q} \tag{5.3.73}$$

where $B_2 := \sqrt{B}/\theta$, (5.3.73) follows from (5.3.69) and (5.3.72) holds with probability at least $1 - (2/3)^{\hat{r}} - r_q v_p^{\hat{r}}$. Similarly,

$$\|H \circ (\mathbf{e}\bar{\mathbf{z}}_3^T)\| \leq B_2 \sqrt{r_s} \tag{5.3.74}$$

with probability at least $1 - ((2/3)^{\hat{r}} + r_s v_p^{\hat{r}})$. Applying the union bound shows that

$$\|\tilde{M}_3\| \leq \|H \circ (\bar{\mathbf{y}}_3 \mathbf{e}^T)\| + \|H \circ (\mathbf{e}\bar{\mathbf{z}}_3^T)\| \leq B_2(\sqrt{r_q} + \sqrt{r_s}) \tag{5.3.75}$$

with probability at least $1 - (r_q + r_s)v_p^{\hat{r}} - 2(2/3)^{\hat{r}}$ as required. ■

We complete the proof of Theorem 5.3.4 by showing that $M_2 = H \circ (Q_1 \mathbf{b}\mathbf{e}^T + \mathbf{e}\mathbf{b}^T Q_2^T)$ has norm at most a constant multiple of $r_s/\sqrt{r_q}$ with high probability. The following lemma provides an upper bound on $\|Q_1\mathbf{b}\|$ and $\|Q_2\mathbf{b}\|$.

Lemma 5.3.3 *There exist B_3, B_4 and $\bar{c}_i > 0$, $i = 1, 2, 3$, depending only on p such that*

$$\|Q_1\mathbf{b}\| \leq B_3 \frac{r_s^{1/2}}{r_q^{1/2}} \tag{5.3.76}$$

$$\|Q_2\mathbf{b}\| \leq B_4 \frac{r_s^{1/2}(r_q + r_s^{1/2})}{r_q^{3/2}} \tag{5.3.77}$$

with probability at least

$$1 - \bar{c}_1 \exp(-\bar{c}_2 \hat{r}^{\bar{c}_3}) - (r_q + r_s)\left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^{p\hat{r}} \tag{5.3.78}$$

where $\delta = (1-p)/(2p)$.

Proof: We first derive a bound on each of $\|Q_1\|$, $\|Q_2\|$ and $\|\mathbf{b}\|$ and consequently a bound on each of $\|Q_1\mathbf{b}\|$ and $\|Q_2\mathbf{b}\|$ by applying the inequalities $\|Q_1\mathbf{b}\| \leq \|Q_1\|\|\mathbf{b}\|$ and

$\|Q_2\mathbf{b}\| \leq \|Q_2\|\|\mathbf{b}\|$. Recall that

$$\|Q_1\| \leq \|(D+\theta\mathbf{e}\mathbf{e}^T)^{-1}\| \sum_{\ell=1}^{\infty} \|P_1P_2\|^\ell + \|P_1\|\|(F+\theta\mathbf{e}\mathbf{e}^T)^{-1}\| \sum_{\ell=0}^{\infty} \|P_1P_2\|^\ell$$

and

$$\|Q_2\| \leq \|(F+\theta\mathbf{e}\mathbf{e}^T)^{-1}\| \sum_{\ell=1}^{\infty} \|P_1P_2\|^\ell + \|P_2\|\|(D+\theta\mathbf{e}\mathbf{e}^T)^{-1}\| \sum_{\ell=0}^{\infty} \|P_1P_2\|^\ell$$

where $P_1 = (D+\theta\mathbf{e}\mathbf{e}^T)^{-1}(H-\theta\mathbf{e}\mathbf{e}^T)$, $P_2 = (F+\theta\mathbf{e}\mathbf{e}^T)^{-1}(H^T-\theta\mathbf{e}\mathbf{e}^T)$. Applying the upper bounds on $\|(D+\theta\mathbf{e}\mathbf{e}^T)^{-1}\|$, $\|(F+\theta\mathbf{e}\mathbf{e}^T)^{-1}\|$, and $\|H-\theta\mathbf{e}\mathbf{e}^T\|$ given by (5.3.59), (5.3.60), and (5.3.62) shows that

$$\|P_1P_2\| \leq \frac{\|H-\theta\mathbf{e}\mathbf{e}^T\|^2}{(\min_{i\in C_q} D_{ii})(\min_{i\in C_s} F_{ii})} \leq \frac{\gamma_1^2}{(\theta-\delta p)^2 r_q} \tag{5.3.79}$$

with probability at least (5.3.78). Therefore, there exists $\gamma_2 > 0$ depending only on p such that

$$\begin{aligned}\|Q_1\| &\leq \frac{1}{(\theta-\delta p)r_s} \sum_{\ell=1}^{\infty} \left(\frac{\gamma_1^2}{(\theta-\delta p)^2 r_q}\right)^\ell + \frac{\gamma_1}{(1-\delta p)^2 r_q\sqrt{r_s}} \sum_{\ell=0}^{\infty} \left(\frac{\gamma_1^2}{(\theta-\delta p)^2 r_q}\right)^\ell \\ &\leq \frac{\gamma_2}{r_q\sqrt{r_s}} \end{aligned} \tag{5.3.80}$$

with probability at least (5.3.78) since

$$\sum_{\ell=0}^{\infty} \left(\frac{\gamma_1^2}{(\theta-\delta p)^2 r_q}\right)^\ell \leq O(1)$$

and

$$\sum_{\ell=1}^{\infty} \left(\frac{\gamma_1^2}{(\theta-\delta p)^2 r_q}\right)^\ell \leq O(r_q^{-1}).$$

with probability at least (5.3.78) in the case that $r_q > (\gamma_1/(\theta-\delta p))^2$. Similarly, there exists $\gamma_3 > 0$ depending only on p such that

$$\|Q_2\| \leq \frac{\gamma_3}{r_q}(r_q^{-1} + r_s^{-1/2}) = \frac{\gamma_3(r_q + r_s^{1/2})}{r_q^2 r_s^{1/2}} \tag{5.3.81}$$

with probability at least (5.3.78). Finally, recall that

$$\mathbf{b}_i = c \cdot \begin{cases} n_i^s, & \text{if } i \in C_q \\ n_i^q, & \text{if } i \in C_s. \end{cases}$$

Therefore,

$$\|\mathbf{b}\| = c \left(\sum_{i \in C_q} (n_i^s)^2 + \sum_{i \in C_s} (n_i^q)^2 \right)^{1/2} \leq (1+\delta) pc (r_q r_s)^{1/2} (r_q + r_s)^{1/2}$$

with probability at least $1 - (r_q + r_s) \left(e^{\delta} / (1+\delta)^{(1+\delta)} \right)^{p\hat{r}}$ by (5.3.56) and (5.3.57). Thus, applying the union bound shows that there exist $B_3, B_4 > 0$ depending only on p such that

$$\|Q_1 \mathbf{b}\| \leq \frac{\gamma_2 (1+\delta) pc (r_q r_s)^{1/2} (r_q + r_s)^{1/2}}{r_q r_s^{1/2}} \leq B_3 \frac{r_s^{1/2}}{r_q^{1/2}}$$

$$\|Q_2 \mathbf{b}\| \leq \frac{\gamma_3 (1+\delta) pc (r_q r_s)^{1/2} (r_q + r_s)^{1/2} (r_q + r_s^{1/2})}{r_q^{3/2} r_s^{1/2}} \leq B_4 \frac{r_s^{1/2} (r_q + r_s^{1/2})}{r_q^{3/2}}$$

with probability at least (5.3.78) since $(r_q + r_s)^{1/2} \leq 2(r_s^{1/2})$. ∎

To obtain an upper bound on $\|M_2\|$ we decompose M_2 as

$$M_2 = (H - \theta \mathbf{e}\mathbf{e}^T) \circ (Q_1 \mathbf{b}\mathbf{e}^T) + \theta (Q_1 \mathbf{b}\mathbf{e}^T) + (H - \theta \mathbf{e}\mathbf{e}^T) \circ (\mathbf{e}(Q_2 \mathbf{b})^T) + \theta (\mathbf{e}(Q_2 \mathbf{b})^T).$$

As an immediate corollary of Lemma 5.3.3 we have

$$\|(Q_1 \mathbf{b})\mathbf{e}^T\| = \|Q_1 \mathbf{b}\| \|\mathbf{e}\| = \sqrt{r_s} \|Q_1 \mathbf{b}\| \leq B_3 \frac{r_s}{r_q^{1/2}} \tag{5.3.82}$$

and

$$\|\mathbf{e}(Q_2 \mathbf{b})^T\| = \|\mathbf{e}\| \|Q_2 \mathbf{b}\| = \sqrt{r_q} \|Q_2 \mathbf{b}\| \leq B_4 \frac{r_s^{1/2} (r_q + r_s^{1/2})}{r_q} \tag{5.3.83}$$

with probability at least (5.3.78). Moreover, applying (5.3.29) with $W = H - \theta \mathbf{e}\mathbf{e}^T$, $\mathbf{u} = Q_1 \mathbf{b}$, and $\mathbf{v} = \mathbf{e}$ we have

$$\|(H - \theta \mathbf{e}\mathbf{e}^T) \circ (Q_1 \mathbf{b}\mathbf{e}^T)\| \leq \|H - \theta \mathbf{e}\mathbf{e}^T\| \|Q_1 \mathbf{b}\|_\infty \leq \|H - \theta \mathbf{e}\mathbf{e}^T\| \|Q_1 \mathbf{b}\|. \tag{5.3.84}$$

Thus, combining (5.3.84), (5.3.62), and (5.3.76) we have

$$\|(H - \theta \mathbf{e}\mathbf{e}^T) \circ (Q_1 \mathbf{b}\mathbf{e}^T)\| \leq B_3 \gamma_1 \frac{r_s}{r_q^{1/2}} \tag{5.3.85}$$

with probability at least (5.3.78). Similarly,

$$\|(H - \theta \mathbf{e}\mathbf{e}^T) \circ (\mathbf{e}(Q_2 \mathbf{b})^T)\| \leq B_4 \gamma_1 \frac{r_s(r_q + r_s^{1/2})}{r_q^{3/2}} \tag{5.3.86}$$

with probability at least (5.3.78). Therefore, there exists $\hat{c}$ depending only on p such that

$$\|M_2\| \leq \hat{c} \frac{r_s}{r_q^{1/2}} \tag{5.3.87}$$

with probability at least (5.3.78) since $r_s \leq O(r_q^2)$ and, hence $(r_q + r_s^{1/2})/r_q \leq O(1)$.

5.4 Proof of Theorem 5.2.3

5.4.1 Optimality conditions for the maximum mean weight k-disjoint clique problem

As in our analysis of the maximum node k-disjoint-clique problem, the proof of Theorem 5.2.3 relies on showing that a proposed optimal solution satisfies the sufficient conditions for optimality given by the Karush-Kuhn-Tucker conditions. The following theorem provides the necessary specialization of these optimality conditions to (5.2.31).

Theorem 5.4.1 *Let X^* be feasible for* (5.2.31) *and suppose that there exist some $\mu \geq 0$, $\lambda \in \mathbf{R}_+^N$, $\eta \in \mathbf{R}_+^{N \times N}$ and $S \in \Sigma_+^N$ such that*

$$-W + \lambda \mathbf{e}^T + \mathbf{e}\lambda^T - \eta + \mu I = S \tag{5.4.1}$$

$$\langle \lambda, X\mathbf{e} - \mathbf{e} \rangle = 0 \tag{5.4.2}$$

$$\langle X, \eta \rangle = 0 \tag{5.4.3}$$

$$\langle X, S \rangle = 0. \tag{5.4.4}$$

Then X^* is optimal for (5.2.31).

Let K^* be a k-disjoint-clique subgraph of K_N composed of the vertex sets $C_1, \ldots, C_k$ of sizes $r_1, \ldots, r_k$ and let X^* be the corresponding feasible solution of (5.2.31) defined by (5.2.10). Let $C_{k+1} := V \setminus (\cup_{i=1}^k C_i)$ and $r_{k+1} := N - \sum_{i=1}^k r_i$. Let $\hat{r} := \min_{i=1,\ldots,k} r_i$. Let $W \in \Sigma^N$ be a random symmetric matrix sampled from a distribution satisfying (ω_1) and (ω_2). Our approach to show that X^* is optimal for (5.2.31) is to construct multipliers $\mu \geq 0$, $\lambda \in \mathbf{R}^N_+$, $\eta \in \mathbf{R}^{N \times N}_+$, and $S \in \Sigma^N_+$ satisfying (5.4.1), (5.4.2), (5.4.3), and (5.4.4). Note that once the multipliers μ, λ, η are chosen, the gradient equation (5.4.1) provides an explicit formula for the multiplier S.

The matrix S and, hence, λ and η will be constructed in blocks indexed by the vertex sets $C_1, \ldots, C_{k+1}$. According to (5.2.10), the diagonal blocks $X^*_{C_q,C_q}$ of X^* for $q = 1, \ldots, k$, consist of multiples of the all ones matrix $\mathbf{e}\mathbf{e}^T$. The remaining blocks of X^* are equal to 0. It follows that we have $\eta_{C_q,C_q} = 0$ for all $q = 1, \ldots, k$ by (5.4.3). Moreover, we have $\lambda_{C_{k+1}} = 0$ by (5.4.2). For each $q = 1, \ldots, k$, we choose λ_{C_q} such that S_{C_q,C_q} is orthogonal to $X^*_{C_q,C_q}$. In particular, it suffices to choose λ such that

$$0 = S_{C_q,C_q}\mathbf{e} = \mu\mathbf{e} + r_q\lambda_{C_q} + (\lambda^T_{C_q}\mathbf{e})\mathbf{e} - W_{C_q,C_q}\mathbf{e} \tag{5.4.5}$$

for all $q = 1, \ldots, k$. Rearranging (5.4.5) shows that λ_{C_q} is the solution to the system

$$(r_q I + \mathbf{e}\mathbf{e}^T)\lambda_{C_q} = W_{C_q,C_q}\mathbf{e} - \mu\mathbf{e} \tag{5.4.6}$$

for all $q = 1, \ldots, k$. Applying the Sherman-Morrison-Woodbury formula (5.3.7) with $A = r_q I$, $\mathbf{u} = \mathbf{v} = \mathbf{e}$ shows that choosing

$$\lambda_{C_q} = \frac{1}{r_q}\left(W_{C_q,C_q}\mathbf{e} - \frac{1}{2}\left(\mu + \frac{\mathbf{e}^T W_{C_q,C_q}\mathbf{e}}{r_q}\right)\mathbf{e}\right) \tag{5.4.7}$$

ensures that $\langle S_{C_q,C_q}, X^*_{C_q,C_q}\rangle = 0$ for all $q = 1, \ldots, k$.

Recall that the requirements that $\langle S, X\rangle = 0$ and $S \in \Sigma^N_+$ imply that $XS = SX = 0$. Therefore, we must choose the entries of each off-diagonal block S_{C_q,C_s} so that $X^*S = SX^* = 0$. For our particular choice of X^*, this is equivalent to requiring all row and column sums of S_{C_q,C_s} to be equal to 0 for all $q, s \in \{1, \ldots, k+1\}$ such that $q \neq s$. To ensure that $S_{C_q,C_s}\mathbf{e} = 0$ and $S_{C_s,C_q}\mathbf{e} = 0$, we parametrize the entries of η_{C_q,C_s} using the vectors $\mathbf{y}^{q,s}$ and $\mathbf{z}^{q,s}$. In particular, we take

$$\eta_{C_q,C_s} = \left(\frac{\bar{\delta}_{q,k+1}}{2}\left(\alpha - \frac{\mu}{r_q}\right) + \frac{\bar{\delta}_{s,k+1}}{2}\left(\alpha - \frac{\mu}{r_s}\right) - \beta\right)\mathbf{e}\mathbf{e}^T + \mathbf{y}^{q,s}\mathbf{e}^T + \mathbf{e}(\mathbf{z}^{q,s})^T. \tag{5.4.8}$$

Here $\bar{\delta}_{ij} := 1 - \delta_{ij}$, where δ_{ij} is the Kronecker delta function defined by $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. That is, we take η_{C_q,C_s} to be the expected value of $\lambda_{C_q}\mathbf{e}^T + \mathbf{e}\lambda_{C_s}^T - W_{C_q,C_s}$ plus the parametrizing terms $\mathbf{y}^{q,s}\mathbf{e}^T$ and $\mathbf{e}(\mathbf{z}^{q,s})^T$. The vectors $\mathbf{y}^{q,s}$ and $\mathbf{z}^{q,s}$ are chosen to be the solutions to the systems of linear equations imposed by the requirement that $X^*S = SX^* = 0$. As before, we show that this system of linear equations is a perturbation of a linear system with known solution. Using the solution of the perturbed system we obtain bounds on $\mathbf{y}^{q,s}$ and $\mathbf{z}^{q,s}$, which are in turn used to establish that η is nonnegative and S is positive semidefinite.

Fix $q, s \in \{1, \ldots, k+1\}$ such that $q \neq s$. Let

$$\tilde{\eta}_{C_q,C_s} := \lambda_{C_q}\mathbf{e}^T + \mathbf{e}\lambda_{C_s}^T - W_{C_q,C_s}. \tag{5.4.9}$$

Note that the symmetry of W implies that $\tilde{\eta}_{C_s,C_q} = \tilde{\eta}_{C_q,C_s}^T$. Let $\mathbf{b} = \mathbf{b}^{q,s} \in \mathbf{R}^{C_q \cup C_s}$ be defined by

$$\mathbf{b}_{C_q} = \tilde{\eta}_{C_q,C_s}\mathbf{e} - E[\tilde{\eta}_{C_q,C_s}]\mathbf{e}, \tag{5.4.10}$$

$$\mathbf{b}_{C_s} = \tilde{\eta}_{C_s,C_q}\mathbf{e} - E[\tilde{\eta}_{C_s,C_q}]\mathbf{e}. \tag{5.4.11}$$

We choose $\mathbf{y}$ and $\mathbf{z}$ to be solutions of the system

$$\begin{pmatrix} r_s I + \theta\mathbf{e}\mathbf{e}^T & (1-\theta)\mathbf{e}\mathbf{e}^T \\ (1-\theta)\mathbf{e}\mathbf{e}^T & r_q I + \theta\mathbf{e}\mathbf{e}^T \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} = \mathbf{b} \tag{5.4.12}$$

for some scalar $\theta > 0$ to be defined later.

The requirement that the row sums of S_{C_q,C_s} are equal to zero is equivalent to $\mathbf{y} = \mathbf{y}^{q,s}$ and $\mathbf{z} = \mathbf{z}^{q,s}$ satisfying the system of linear equations

$$\begin{aligned} 0 = -r_s\mathbf{y}_i - \mathbf{z}^T\mathbf{e} + r_s\left(\lambda_i - \frac{\bar{\delta}_{q,k+1}}{2r_q}(\alpha r_q - \mu)\right) + \left(\lambda_{C_s}^T\mathbf{e} - \frac{\bar{\delta}_{s,k+1}}{2}(\alpha r_s - \mu)\right) \\ - ([W_{C_q,C_s}\mathbf{e}]_i - r_s\beta) \end{aligned} \tag{5.4.13}$$

for all $i \in C_q$. Similarly, the column sums of S_{C_q,C_s} are equal to zero if and only if $\mathbf{y}$ and $\mathbf{z}$ satisfy

$$\begin{aligned} 0 = -r_q\mathbf{z}_i - \mathbf{y}^T\mathbf{e} + r_q\left(\lambda_i - \frac{\bar{\delta}_{s,k+1}}{2r_s}(\alpha r_s - \mu)\right) + \left(\lambda_{C_q}^T\mathbf{e} - \frac{\bar{\delta}_{q,k+1}}{2}(\alpha r_q - \mu)\right) \\ - ([W_{C_s,C_q}\mathbf{e}]_i - r_q\beta) \end{aligned} \tag{5.4.14}$$

for all $i \in C_s$. Note that the system of equations defined by (5.4.13) and (5.4.14) is equivalent to (5.4.12) in the special case that $\theta = 0$. However, when $\theta = 0$, the system of equations in (5.4.12) is singular, with nullspace equal to the vector $(\mathbf{e}; -\mathbf{e})$. It follows that $(\mathbf{y} + c\mathbf{e}; \mathbf{z} - c\mathbf{e})$ is a solution of (5.4.12) for any scalar c if $(\mathbf{y}; \mathbf{z})$ is a solution of (5.4.12). In particular, there exists solution $(\mathbf{y}; \mathbf{z})$ of (5.4.12) such that

$$\mathbf{e}^T\mathbf{y}^{q,s} - \mathbf{e}^T\mathbf{z}^{q,s} = 0. \tag{5.4.15}$$

When θ is nonzero, each row of the system (5.4.12) has an additional term of the form $\theta(\mathbf{e}^T\mathbf{y} - \mathbf{e}^T\mathbf{z})$. Therefore, for $\theta > 0$ such that (5.4.12) is nonsingular, the solution $(\mathbf{y}; \mathbf{z})$ satisfying (5.4.13), (5.4.14), and (5.4.15) is also the unique solution to (5.4.12) since the term $\theta(\mathbf{e}^T\mathbf{y} - \mathbf{e}^T\mathbf{z})$ is zero. In particular, note that (5.4.12) is nonsingular for $\theta = 1$. For this choice of θ, $\mathbf{y}$ and $\mathbf{z}$ are the unique solutions of the systems

$$(r_s I + \mathbf{e}\mathbf{e}^T)\mathbf{y} = \mathbf{b}_1 \tag{5.4.16}$$

$$(r_q I + \mathbf{e}\mathbf{e}^T)\mathbf{z} = \mathbf{b}_2 \tag{5.4.17}$$

where $\mathbf{b}_1 := \mathbf{b}(C_q)$ and $\mathbf{b}_2 := \mathbf{b}(C_s)$. Applying the Sherman-Morrison-Woodbury formula (5.3.7) with $A = r_s I$, $\mathbf{u} = \mathbf{v} = \mathbf{e}$ and $A = r_q I$, $\mathbf{u} = \mathbf{v} = \mathbf{e}$ yields

$$\mathbf{y} = \frac{1}{r_s}\left(\mathbf{b}_1 - \frac{(\mathbf{b}_1^T\mathbf{e})}{r_q + r_s}\mathbf{e}\right) \tag{5.4.18}$$

$$\mathbf{z} = \frac{1}{r_q}\left(\mathbf{b}_2 - \frac{(\mathbf{b}_2^T\mathbf{e})}{r_q + r_s}\mathbf{e}\right) \tag{5.4.19}$$

respectively.

In summary, we choose the multipliers $\mu \in \mathbf{R}$, $\lambda \in \mathbf{R}^N$, $\eta \in \mathbf{R}^{N\times N}$ as follows:

$$\mu = (\alpha - \beta)\hat{r}/2 \tag{5.4.20}$$

$$\lambda_{C_q} = \begin{cases} \dfrac{1}{r_q}\left(W_{C_q,C_q}\mathbf{e} - \dfrac{1}{2}\left(\mu + \dfrac{\mathbf{e}^T W_{C_q,C_q}\mathbf{e}}{r_q}\right)\mathbf{e}\right), & \text{if } q \in \{1,\ldots,k\} \\ 0, & \text{if } q = k+1 \end{cases} \tag{5.4.21}$$

$$\eta_{C_q,C_s} = \begin{cases} E[\tilde{\eta}_{C_q,C_s}] + \mathbf{y}^{q,s}\mathbf{e}^T + \mathbf{e}(\mathbf{z}^{q,s})^T, & \text{if } q,s \in \{1,\ldots,k+1\}, q \neq s \\ 0, & \text{otherwise} \end{cases} \tag{5.4.22}$$

where $\tilde{\eta}_{C_q,C_s}$ is defined as in (5.4.9) and $\mathbf{y}^{q,s}, \mathbf{z}^{q,s}$ are given by (5.4.18) and (5.4.19) for all $q, s \in \{1, \ldots, k+1\}$ such that $q \neq s$. We choose S according to (5.4.1). Finally, we define the $(k+1) \times (k+1)$ block matrix $\tilde{S}$ in Σ^N by

$$\tilde{S}_{C_q,C_s} = \begin{cases} S_{C_q,C_s}, & \text{if } q, s \in \{1, \ldots, k+1\}, q \neq s \\ -W_{C_{k+1},C_{k+1}}, & \text{if } q = s = k+1 \\ 0, & \text{otherwise.} \end{cases} \tag{5.4.23}$$

We conclude by providing the following theorem, which provides a sufficient condition for when the proposed solution X^* is the unique optimal solution for (5.2.31) and when K^* is the maximum mean weight k-disjoint-clique subgraph of K_N corresponding to W.

Theorem 5.4.2 *Suppose that the vertex sets $C_1, \ldots, C_k$ define a k-disjoint-clique subgraph K^* of the complete graph $K_N = (V, E)$ on N vertices and let $C_{k+1} := V \setminus (\cup_{i=1}^k C_i)$. Let $r_i := |C_i|$ for all $i = 1, \ldots, k+1$, and let $\hat{r} = \min_{i=1,\ldots,k} r_i$. Let $W \in \Sigma^N$ be a random symmetric matrix sampled from distributions Ω_1, Ω_2 satisfying (ω_1) and (ω_2). Let X^* be the feasible solution for (5.2.31) corresponding to $C_1, \ldots, C_k$ defined by (5.2.10). Let $\mu \geq 0$, $\lambda \in \mathbf{R}^N$, $\eta \in \mathbf{R}^{N \times N}$ be chosen according to (5.4.20), (5.4.21), and (5.4.22), and let S be chosen according to (5.4.1). Suppose that the entries of λ and η are nonnegative. Then there exist scalars $c_1, c_2 > 0$ such that if*

$$r_i \leq c_1(\alpha - \beta)^2 \hat{r}^2 \tag{5.4.24}$$

for all $i = 1, \ldots, k$, and

$$\|\tilde{S}\| \leq c_2(\alpha - \beta)\hat{r} \tag{5.4.25}$$

then X^* is optimal for (5.2.31), and K^* is the maximum mean weight k-disjoint-clique subgraph of K_N corresponding to W. Moreover, if

$$r_s \mathbf{e}^T W_{C_q,C_q} \mathbf{e} > r_q \mathbf{e}^T W_{C_q,C_s} \mathbf{e} \tag{5.4.26}$$

for all $q, s \in \{1, \ldots, k\}$ such that $q \neq s$, then X^* is the unique optimal solution of (5.2.31) and K^* is the unique maximum mean weight k-disjoint-clique subgraph of K_N.

Proof: By construction, μ, λ, η, and S satisfy (5.4.1), (5.4.2), (5.4.3), and (5.4.4). Moreover, μ, λ, and η are nonnegative by assumption. Therefore, to prove that X^* is optimal for (5.2.31), it suffices to show that S is positive semidefinite. To do so, we fix

$\mathbf{x} \in \mathbf{R}^N$ and decompose $\mathbf{x}$ as $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ where

$$\mathbf{x}_1(C_i) = \begin{cases} \phi_i \mathbf{e}, & \text{if } i \in \{1, \ldots, k\} \\ 0, & \text{if } i = k+1 \end{cases}$$

for some $\phi \in \mathbf{R}^k$ chosen such that $\mathbf{x}_2(C_i)$ is orthogonal to $\mathbf{e}$ for all $i = 1, \ldots, k$, and $\mathbf{x}_2(C_{k+1}) = \mathbf{x}(C_{k+1})$. By our choice of $\mathbf{x}_1$ and $\mathbf{x}_2$, we have

$$\begin{aligned} \mathbf{x}^T S \mathbf{x} &= \mathbf{x}_2^T S \mathbf{x}_2 \\ &= \sum_{i=1} (\mathbf{x}_2(C_i)^T (\alpha \mathbf{e}\mathbf{e}^T - W_{C_i,C_i}) \mathbf{x}_2(C_i) + \mathbf{x}_2^T (\tilde{S} + \mu I) \mathbf{x}_2 \\ &\geq \left(\mu - \max_{i=1,\ldots,k} \|\alpha \mathbf{e}\mathbf{e}^T - W_{C_i,C_i}\| - \|\tilde{S}\| \right) \|\mathbf{x}_2\| \\ &\geq \left(\frac{\alpha - \beta}{2} \hat{r} - \gamma \max_{i=1,\ldots,k} \sqrt{r_i} - \|\tilde{S}\| \right) \|\mathbf{x}_2\|. \end{aligned}$$

with probability tending exponentially to 1 as $\hat{r} \to \infty$, since there exists $\gamma > 0$ such that

$$\|\alpha \mathbf{e}\mathbf{e}^T - W_{C_i,C_i}\| \leq \gamma \sqrt{r_i}.$$

with probability tending exponentially to 1 as $\hat{r} \to \infty$ by Theorem 2.6.3. Therefore, there exists scalars c_1, c_2 such that if $r_i \leq c_1(\alpha - \beta)^2 \hat{r}^2$ and $\|\tilde{S}\| \leq c_2(\alpha - \beta)\hat{r}$, then $\mathbf{x}^T S \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbf{R}^N$ with equality if and only if $\mathbf{x}_2 = 0$ with probability tending exponentially to 1 as $\hat{r} \to \infty$. Therefore X^* is optimal for (5.2.31) with probability tending exponentially to 1 as $\hat{r} \to \infty$. Moreover, $\mathbf{v}_i$ is in the nullspace of S for all $i = 1, \ldots, k$ by (5.4.4) and the fact that $X^* = \sum_{i=1}^k \mathbf{v}_i \mathbf{v}_i^T / r_i$. Since $\mathbf{x}^T S \mathbf{x} = 0$ if and only if $\mathbf{x}_2 = 0$, the nullspace of S is exactly equal to the span of $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ and S has rank equal to $N - k$.

To see that X^* is the unique optimal solution for (5.2.31) if Assumption (5.4.26) holds, suppose, on the contrary, that $\tilde{X}$ is also optimal for (5.2.31). By (5.4.4), we have $\langle \tilde{X}, S \rangle = 0$, which holds if and only if $\tilde{X} S = 0$. That is, the row and column spaces of $\tilde{X}$ lie in the nullspace of S. Since $\tilde{X} \succeq 0$ and $\tilde{X} \geq 0$, we may write $\tilde{X}$ as

$$\tilde{X} = \sum_{i=1}^k \sum_{j=1}^k \sigma_{ij} \mathbf{v}_i \mathbf{v}_j^T \tag{5.4.27}$$

for some $\sigma \in \mathbf{R}_+^{k\times k}$. The fact that $\tilde{X}$ satisfies $\tilde{X}\mathbf{e} \leq \mathbf{e}$ implies that

$$\sigma_{qq} r_q + \sum_{\substack{s=1 \\ s\neq q}}^{k} \sigma_{qs} r_s \leq 1 \tag{5.4.28}$$

for all $q =, 1 \ldots, k$. Moreover, since $\langle \tilde{X}, W\rangle = \langle X^*, W\rangle$, there exists some $q \in \{1, \ldots, k\}$ such that

$$\sigma_{qq} \mathbf{v}_q^T W \mathbf{v}_q + \sum_{\substack{s=1 \\ s\neq q}}^{k} \sigma_{qs} \mathbf{v}_q^T W \mathbf{v}_s \geq \frac{\mathbf{v}_q^T W \mathbf{v}_q}{r_q}. \tag{5.4.29}$$

Combining (5.4.28) and (5.4.29) shows that

$$\begin{aligned} 0 &\leq \mathbf{v}_q^T W \mathbf{v}_q \left(\frac{1}{r_q} - \sum_{\substack{s=1 \\ s\neq q}}^{k} \frac{\sigma_{qs} r_s}{r_q} \right) + \sum_{\substack{s=1 \\ s\neq q}}^{k} \sigma_{qs} \mathbf{v}_q^T W \mathbf{v}_s - \frac{\mathbf{v}_q^T W \mathbf{v}_q}{r_q} \\ &= \sum_{\substack{s=1 \\ s\neq q}}^{k} \frac{\sigma_{qs}}{r_q} (r_q \mathbf{v}_q^T W \mathbf{v}_s - r_s \mathbf{v}_q^T W \mathbf{v}_q), \end{aligned}$$

contradicting Assumption (5.4.26). Therefore, X^* is the unique optimal solution of (5.2.31) as required. ■

5.4.2 A lower bound on λ and η in the planted case

Let $C_1, \ldots, C_k$ denote the vertex sets of a k-disjoint-clique subgraph of the complete graph $K_N = (V, E)$ on N vertices. Let $C_{k+1} := V \setminus (\cup_{i=1}^k C_i)$ and let $r_i := |C_i|$ for all $i = 1, \ldots, k+1$. Let $\hat{r} := \min\{r_1, \ldots, r_k\}$. Let $W \in \Sigma^N$ be a random symmetric matrix sampled from distributions Ω_1, Ω_2 according to (ω_1) and (ω_2). Let μ, λ, η be chosen as in (5.4.20), (5.4.21), and (5.4.22) respectively. We now establish that the entries of λ and η are nonnegative with probability tending exponentially to 1 as $\hat{r}$ approaches ∞. To do so, we will use the following bounds on the ℓ_1 and ℓ_2-norms of random vectors whose entries are equal to the sum of independently identically distributed (i.i.d.) random variables.

Theorem 5.4.3 *Let I_1, I_2 be index sets of size r_1, r_2 respectively. Let $\{X_{ij} : i \in I_1, j \in I_2\}$ be a collection of independently identically distributed (i.i.d.) random variables with mean*

$\bar{\mathbf{x}}$ *such that* $0 \leq X_{ij} \leq 1$ *for all* $i \in I_1$, $j \in I_2$. *Then there exists scalar* B *such that*

$$\sum_{i \in I_1} \left| \sum_{j \in I_2} X_{ij} - r_2 \bar{\mathbf{x}} \right|^{\delta} \leq B r_1 r_2^{\delta/2} \tag{5.4.30}$$

with probability at least

$$1 - (2/3)^{r_1} - 2r_1 \exp(-2r_2) \tag{5.4.31}$$

for $\delta = 1, 2$.

The proof of Theorem 5.4.3 is a modification of the proof of Theorem 5.3.5 and is included in Appendix A.3.

We begin by deriving lower bounds on the entries of η. To show that $\eta_{ij} \geq 0$ for all $i, j \in V$ with high probability, we will use the following lemma, which provides an upper bound on $\|\mathbf{y}^{q,s}\|_\infty$ and $\|\mathbf{z}^{q,s}\|_\infty$ for all $q, s \in \{1, \ldots, k+1\}$ such that $q \neq s$, holding with probability tending to 1 as $\hat{r}$ tends to ∞.

Lemma 5.4.1 *There exists scalar* $\tilde{c} > 0$ *such that*

$$\|\mathbf{y}^{q,s}\|_\infty + \|\mathbf{z}^{q,s}\|_\infty \leq \tilde{c} \hat{r}^{-1/4} \tag{5.4.32}$$

for all $q, s \in \{1, \ldots, k+1\}$ *such that* $q \neq s$ *with probability at least*

$$1 - k^2(10((2/3)^{\hat{r}} - 2\hat{r}\exp(-2\hat{r})) - 8\hat{r}\exp(-2\hat{r}^{1/2})). \tag{5.4.33}$$

Proof: Fix $q, s \in \{1, \ldots, k\}$ such that $q \neq s$. The proof for the case when either q or s is equal to $k+1$ is analogous. We first obtain an upper bound on $\|\mathbf{y}\|_\infty = \|\mathbf{y}^{q,s}\|_\infty$. By the triangle inequality, we have

$$\|\mathbf{y}\|_\infty \leq \frac{1}{r_s} \left\| \mathbf{b}_1 + \frac{|\mathbf{b}_1^T \mathbf{e}|}{r_q + r_s} \mathbf{e} \right\|_\infty \leq \frac{1}{r_s} \left(\|\mathbf{b}_1\|_\infty + \frac{|\mathbf{b}_1^T \mathbf{e}|}{r_q + r_s} \right). \tag{5.4.34}$$

Hence, to obtain an upper bound on $\|\mathbf{y}\|_\infty$, it suffices to obtain bounds on $\|\mathbf{b}_1\|_\infty$ and $|\mathbf{b}_1^T \mathbf{e}|$. We begin with $\|\mathbf{b}_1\|_\infty$. Recall that we have

$$\mathbf{b}_i = r_s \left(\lambda_i - \frac{1}{2r_q}(\alpha r_q - \mu) \right) + \left(\lambda_{C_s}^T \mathbf{e} - \frac{1}{2}(\alpha r_s - \mu) \right) - \left(\sum_{j \in C_s} W_{ij} - \beta r_s \right). \tag{5.4.35}$$

for each $i \in C_q$. Note that

$$\lambda_{C_s}^T \mathbf{e} = \frac{1}{r_s}\left(\mathbf{e}^T W_{C_s,C_s}\mathbf{e} - \frac{1}{2}r_s\mu - \frac{1}{2}\mathbf{e}^T W_{C_s,C_s}\mathbf{e}\right) = \frac{1}{2r_s}(\mathbf{e}^T W_{C_s,C_s}\mathbf{e} - r_s\mu).$$

Applying (5.4.30) with $\delta = 1$, $I_1 = I_2 = C_s$, and $X_{\ell_1,\ell_2} = W_{\ell_1,\ell_2}$ for all $\ell_1, \ell_2 \in C_s$, we have

$$\left|\lambda_{C_s}^T \mathbf{e} - \frac{1}{2}(\alpha r_s - \mu)\right| = \frac{1}{2r_s}|\mathbf{e}^T W_{C_s,C_s}\mathbf{e} - \alpha r_s^2| \leq \frac{B}{2}\sqrt{r_s} \tag{5.4.36}$$

with probability at least

$$1 - (2/3)^{r_s} - 2r_s\exp(-2r_s) \geq 1 - p_1 \tag{5.4.37}$$

where

$$p_1 := (2/3)^{\hat{r}} + 2\hat{r}\exp(-2\hat{r}). \tag{5.4.38}$$

Next, applying (2.6.3) with $S = \sum_{\ell \in C_s} W_{i\ell}$ and $t = r_s^{3/4}$ shows that

$$\left|\sum_{\ell \in C_s} W_{i\ell} - \beta r_s\right| \leq r_s^{3/4} \tag{5.4.39}$$

with probability at least $1 - p_2$ where

$$p_2 := 2\exp(-2\hat{r}^{1/2}). \tag{5.4.40}$$

Finally, by applying (2.6.3) with $S = \sum_{\ell \in C_q} W_{i\ell}$, $t = r_q^{3/4}$ and (5.4.30) with $\delta = 1$, $I_1 = I_2 = C_q$ and $X_{\ell_1,\ell_2} = W_{\ell_1,\ell_2}$ for all $\ell_1, \ell_2 \in C_q$.

$$\begin{aligned}|\lambda_i - \frac{1}{2r_q}(\alpha r_q - \mu)| &\leq \frac{1}{r_q}\left|\sum_{\ell \in C_q} W_{i\ell} - r_q\alpha\right| + \frac{1}{2r_q^2}\sum_{\ell_1 \in C_q}\left|\sum_{\ell_2 \in C_q} W_{\ell_1,\ell_2} - r_q\alpha\right| \\ &\leq r_q^{-1/4} + \frac{B}{2}r_q^{-1/2} \leq \left(1 + \frac{B}{2}\right)r_q^{-1/4}\end{aligned} \tag{5.4.41}$$

with probability at least $1 - p_1 - p_2$. Combining (5.4.36), (5.4.39) and (5.4.41) and applying the union bound shows that there exists scalar $c_1 > 0$ such that

$$\|\mathbf{b}_1\|_\infty \leq c_1 r_q^{-1/4} r_s \tag{5.4.42}$$

with probability at least $1 - 2p_1 - 2p_2$. By a similar argument, there exists scalar $c_2 > 0$ such that

$$\|\mathbf{b}_2\|_\infty \leq c_2 r_q^{3/4} \tag{5.4.43}$$

with probability at least $1 - 2p_1 - 2p_2$.

We next obtain an upper bound on $|\mathbf{b}_1^T \mathbf{e}|$ and $|\mathbf{b}_2^T \mathbf{e}|$. We have

$$\mathbf{b}_1^T \mathbf{e} = r_s \left(\lambda_{C_q}^T \mathbf{e} - \frac{1}{2}(\alpha r_q - \mu) \right) + r_q \left(\lambda_{C_s}^T \mathbf{e} - \frac{1}{2}(\alpha r_s - \mu) \right) + (\beta r_s r_q - \mathbf{e}^T W_{C_q, C_s} \mathbf{e}). \tag{5.4.44}$$

By (5.4.36) and the union bound, we have

$$\left| \lambda_{C_s}^T \mathbf{e} - \frac{1}{2}(\alpha r_s - \mu) \right| \leq \frac{B}{2} \sqrt{r_s} \tag{5.4.45}$$

$$\left| \lambda_{C_q}^T \mathbf{e} - \frac{1}{2}(\alpha r_q - \mu) \right| \leq \frac{B}{2} \sqrt{r_q} \tag{5.4.46}$$

with probability at least $1 - 2p_1$. Moreover, applying (5.4.30) with $\delta = 1$, $I_1 = C_q$, $I_2 = C_s$, $X_{\ell_1, \ell_2} = W_{\ell_1, \ell_2}$ for all $\ell_1 \in C_q, \ell_2 \in C_s$, we have

$$|\mathbf{e}^T W_{C_q, C_s} \mathbf{e} - \beta r_s r_q| \leq \sum_{\ell_1 \in C_s} \left| \sum_{\ell_2 \in C_q} W_{\ell_1, \ell_2} - \beta r_q \right| \leq B r_q \sqrt{r_s} \tag{5.4.47}$$

with probability at least $1 - p_1$. Substituting (5.4.45), (5.4.46), and (5.4.47) into (5.4.44), we have

$$|\mathbf{b}_1^T \mathbf{e}| \leq c_3 r_s \sqrt{r_q} \tag{5.4.48}$$

for some scalar $c_3 > 0$ with probability at least $1 - 3p_1$ by the union bound. Similarly, there exists scalar $c_4 > 0$ such that

$$|\mathbf{b}_2^T \mathbf{e}| \leq c_4 r_q \sqrt{r_s} \tag{5.4.49}$$

with probability at least $1 - 3p_1$. Substituting (5.4.42) and (5.4.48) in (5.4.34) yields

$$\|\mathbf{y}\|_\infty \leq \tilde{c}_1 r_q^{-1/4}. \tag{5.4.50}$$

for some scalar $\tilde{c}_1 > 0$ with probability at least

$$1 - 5p_1 - 2p_2. \tag{5.4.51}$$

Similarly, there exists scalar $\tilde{c}_2 > 0$ such that

$$\|\mathbf{z}^{q,s}\|_\infty \leq \tilde{c}_2 r_q^{-1/4} \tag{5.4.52}$$

with probability at least (5.4.51). Combining (5.4.50) and (5.4.52) and applying the union bound once more completes the proof. ■

As an immediate consequence of Lemma 5.4.1, we have the following corollary that states that η is nonnegative with probability tending exponentially to 1 for sufficiently large values of $\hat{r}$.

Corollary 5.4.1 *For all $\hat{r}$ such that $\hat{r} \geq (2\tilde{c}/(\alpha - \beta))^4$, the entries of the matrix η are nonnegative with probability at least* (5.4.33).

Proof: Fix $q, s \in \{1, \ldots, k+1\}$ such that $q \neq s$. Lemma (5.4.1) implies that

$$\eta_{ij} \geq \frac{1}{2}(\alpha - \beta) - \|\mathbf{y}\|_\infty - \|\mathbf{z}\|_\infty \geq \frac{1}{2}(\alpha - \beta) - \tilde{c}\hat{r}^{-1/4} \tag{5.4.53}$$

for all $i \in C_q$, $j \in C_s$, with probability at least (5.4.33). Therefore, $\eta_{ij} \geq 0$ for any $\hat{r}$ greater than $(2\tilde{c}/(\alpha - \beta))^4$ for all $i, j \in V$ with probability at least (5.4.33). ■

The following theorem provides a lower bound on the entries of λ_{C_q} for all $q = 1, \ldots, k$.

Theorem 5.4.4 *There exist scalars $\bar{c}_1, \bar{c}_2 > 0$ such that*

$$\lambda_i \geq \hat{r}(\bar{c}_1 - \bar{c}_2 \hat{r}^{-1/4}) \tag{5.4.54}$$

for all $i \in V \setminus C_{k+1}$ with probability at least

$$1 - N((2/3)^{\hat{r}} - 2\hat{r}\exp(-2\hat{r}) - 2\hat{r}\exp(-2\hat{r}^{1/2})). \tag{5.4.55}$$

Proof: Fix $q \in \{1, \ldots, k\}$ and $i \in C_q$. Recall that

$$\lambda_i = \sum_{j \in C_q} W_{ij} - \frac{1}{2r_q}\mathbf{e}^T W_{C_q, C_q}\mathbf{e} - \frac{\mu}{2}.$$

Applying (2.6.3) with $S = \sum_{j \in C_q} W_{ij}$ and $t = r_q^{3/4}$ yields

$$\sum_{j \in C_q} W_{ij} \geq \alpha r_q - r_q^{3/4} \tag{5.4.56}$$

with probability at least $1 - p_2$. Moreover, (5.4.46) implies that

$$\frac{1}{2r_q} \mathbf{e}^T W_{C_q, C_q} \mathbf{e} \leq \frac{1}{2}(\alpha r_q + B\sqrt{r_q}) \tag{5.4.57}$$

with probability at least $1 - p_1$. Combining (5.4.56) and (5.4.57) and applying the union bound shows that there exist scalars $\bar{c}_1, \bar{c}_2$ depending only on α, β such that

$$\lambda_i \geq \alpha r_q - r_q^{3/4} - \frac{1}{2}(\alpha r_q + B\sqrt{r_q}) - \frac{\mu}{2} \geq r_q(\bar{c}_1 - \bar{c}_2 r_q^{-1/4})$$

with probability at least $1 - p_1 - p_2$. Applying the union bound over all $i \in V \setminus C_{k+1}$ completes the proof. ■

Note that Theorem 5.4.4 implies that $\lambda \geq 0$ with probability tending exponentially to 1 as $\hat{r}$ tends to ∞. Therefore, μ, λ, η constructed according to (5.4.20), (5.4.21), and (5.4.22) are dual feasible for (5.2.31) with probability tending exponentially to 1 as $\hat{r} \to \infty$. The following theorem states the uniqueness condition given by (5.4.26) is satisfied with high probability for sufficiently large $\hat{r}$.

Theorem 5.4.5 *There exists scalar $c > 0$ such that if $\hat{r} > 4c/(\alpha - \beta)^2$ then*

$$r_s \mathbf{e}^T W_{C_q, C_q} \mathbf{e} > r_q \mathbf{e}^T W_{C_q, C_s} \mathbf{e}$$

for all $q, s \in \{1, \ldots, k\}$ such that $q \neq s$ with probability at least

$$1 - 2k^2\big((2/3)^{\hat{r}} + 2\hat{r}\exp(-2\hat{r})\big). \tag{5.4.58}$$

Proof: Fix $q \neq s$ such that $r_q \leq r_s$. Applying (5.4.30) with $\delta = 1$, $I_1 = I_2 = C_q$, and $X_{\ell_1, \ell_2} = W_{\ell_1, \ell_2}$ for all $\ell_1, \ell_2 \in C_q$ shows that

$$\mathbf{e}^T W_{C_q, C_q} \mathbf{e} \geq \alpha r_q^2 - B r_q^{3/2} \tag{5.4.59}$$

with probability at least $1 - p_1$. Similarly, applying (5.4.30) with $\delta = 1$, $I_1 = C_q$, $I_2 = C_s$, and $X_{\ell_1,\ell_2} = W_{\ell_1,\ell_2}$ for all $\ell_1 \in C_q, \ell_2 \in C_s$ yields

$$\mathbf{e}^T W_{C_q,C_s} \mathbf{e} \leq \beta r_q r_s + B r_q r_s^{1/2} \tag{5.4.60}$$

with probability at least $1 - p_1$. Combining (5.4.59) and (5.4.60) yields

$$\begin{aligned} r_s \mathbf{e}^T W_{C_q,C_q} \mathbf{e} - r_q \mathbf{e}^T W_{C_q,C_s} \mathbf{e} &\geq r_s r_q^2 (\alpha - \beta - B(r_q^{-1/2} + r_s^{-1/2}) \\ &\geq r_s r_q^2 (\alpha - \beta - 2B\hat{r}^{-1/2}) > 0 \end{aligned}$$

if $\hat{r} > 4B^2/(\alpha - \beta)^2$, with probability at least $1 - 2p_1$. Applying the union bound over all choices of q, s completes the proof. ■

We have shown that μ, λ, η constructed according to (5.4.20), (5.4.21), and (5.4.22) are dual feasible for (5.2.31) and the uniquness condition (5.4.26) is satisfied with probability tending exponentially to 1 as $\hat{r} \to \infty$. In the next subsection, we derive an upper bound on the norm of $\tilde{S}$ and use this bound to obtain conditions ensuring dual feasibility of S and, hence, optimality of X^* for (5.2.31).

5.4.3 A bound on $\|\tilde{S}\|$ in the planted case

Suppose that the random matrix W is sampled from distributions Ω_1, Ω_2 satisfying (ω_1) and (ω_2) corresponding to partitioning $C_1, \ldots, C_{k+1}$ of the vertices of the complete graph $K_N = (V, E)$ on $N = |V|$ vertices. Let $r_i = |C_i|$ for all $i = 1, \ldots, k+1$. Let $\hat{r} := \min_{i=1,\ldots,k} r_i$. Let $\mu \in \mathbf{R}_+$, $\lambda \in \mathbf{R}^N$, $\eta \in \mathbf{R}^{N \times N}$, $S \in \Sigma^N$, $\tilde{S} \in \Sigma^N$ be defined as in Section 5.4.1. In this section, we derive an upper bound on $\|\tilde{S}\|$, which will be used to verify that the conditions on the partitioning $C_1, \ldots, C_{k+1}$ imposed by (5.2.38) and (5.2.39) ensure that the k-disjoint-clique subgraph of K_N composed of the cliques $C_1, \ldots, C_k$ is the unique maximum mean weight k-disjoint-clique of K_N with respect to W and can be recovered by solving (5.2.31) with probability tending exponentially to 1 as $\hat{r} \to \infty$. In particular, we will prove the following theorem.

Theorem 5.4.6 *There exist scalars $\rho_1, \rho_2 > 0$ such that*

$$\|\tilde{S}\| \leq \rho_1 \left(k \sum_{s=1}^{k+1} r_s \right)^{1/2} + \rho_2 \sqrt{N} + \beta r_{k+1} \tag{5.4.61}$$

with probability tending exponentially to 1 *as* $\hat{r}$ *approaches* ∞*.*

This theorem, along with Theorems 5.4.2 and 5.4.4, and Corollary 5.4.1, establishes Theorem 5.2.3. Indeed, if the right-hand side of (5.4.61) is at most $c_2\hat{r}$ and $\mathrm{s}r_i \leq c_1\hat{r}^2$ for each $i = 1, \ldots, k$ then Theorems 5.4.2 and 5.4.4, and Corollary 5.4.1 imply that the k-disjoint-clique subgraph corresponding to the partitioning $C_1, \ldots, C_k$ is optimal for (5.2.23) and can be recovered by solving (5.2.31).

The remainder of this section consists of a proof of Theorem 5.4.6. We decompose $\tilde{S}$ as

$$\tilde{S} = \tilde{S}_1 + \tilde{S}_2 + \tilde{S}_3 + \tilde{S}_4$$

where $\tilde{S}_i \in \Sigma^N$, $i = 1, \ldots, 4$, are $(k+1)$ by $(k+1)$ block matrices such that

$$\tilde{S}_1(C_q, C_s) = \begin{cases} \tilde{S}(C_q, C_s), & \text{if } q, s \in \{1, \ldots, k+1\},\ q \neq s \\ 0, & \text{otherwise} \end{cases}$$

$$\tilde{S}_2(C_q, C_s) = \begin{cases} \beta \mathbf{e}\mathbf{e}^T - W, & \text{if } q = s = k+1 \\ \beta \mathbf{e}\mathbf{e}^T - R(C_q, C_s), & \text{otherwise} \end{cases}$$

$$\tilde{S}_3(C_q, C_s) = \begin{cases} 0, & \text{if } q = s = k+1 \\ R(C_q, C_s) - \beta \mathbf{e}\mathbf{e}^T, & \text{otherwise} \end{cases}$$

$$\tilde{S}_4(C_q, C_s) = \begin{cases} -\beta \mathbf{e}\mathbf{e}^T, & \text{if } q = s = k+1 \\ 0, & \text{otherwise} \end{cases}$$

where $R \in \Sigma^N$ is a random symmetric random matrix with independently identically distributed (i.i.d.) entries satisfying (ω_2) with mean equal to β. By Theorem 2.6.3, there exist $\rho_2, \kappa_1, \kappa_2 > 0$ such that

$$\|\tilde{S}_2\| + \|\tilde{S}_3\| \leq \rho_2 \sqrt{N} \tag{5.4.62}$$

with probability at least

$$1 - \kappa_1 \exp(-\kappa_2 N^{-1/6}). \tag{5.4.63}$$

Morever, we have

$$\|\tilde{S}_4\| = \beta \|\mathbf{e}\mathbf{e}^T\| = \beta r_{k+1}. \tag{5.4.64}$$

The fact that

$$\|\tilde{S}_1\|^2 = O\left(k \sum_{s=1}^{k+1} r_s\right)$$

with probability tending exponentially to 1 as $\hat{r} \to \infty$ is an immediate consequence of the the following theorem, which provides an upper bound on the norm of $\tilde{S}(C_q, C_S)$ holding

with probability tending exponentially to 1 as $\hat{r}$ approaches ∞.

Theorem 5.4.7 *There exists $t > 0$*

$$\|\tilde{S}_1(C_q, C_s)\| = \|\tilde{S}(C_q, C_s)\| \leq t\sqrt{\max\{r_q, r_s\}} \tag{5.4.65}$$

for all $q, s \in \{1, \ldots, k{+}1\}$, $q \neq s$, with probability tending exponentially to 1 as $\hat{r}$ approaches ∞.

Proof:

We consider $q, s \in \{1, \ldots, k\}$ such that $q \neq s$. The derivation of the bound on $\|\tilde{S}_1(C_q, C_s)\|$ for the case that $q = k+1$ or $s = k+1$ is analogous. Without loss of generality we may assume that $r_q \leq r_s$. We decompose S_{C_q,C_s} as $S_{C_q,C_s} = M_1 + M_2 + M_3$ where

$$\begin{aligned} M_1 &= \left(\lambda_{C_q} - \frac{1}{2r_q}(\alpha r_q - \mu)\mathbf{e}\right)\mathbf{e}^T \\ M_2 &= \mathbf{e}\left(\lambda_{C_s} - \frac{1}{2r_s}(\alpha r_s - \mu)\mathbf{e}\right)^T \\ M_3 &= \mathbf{y}\mathbf{e}^T + \mathbf{e}\mathbf{z}^T. \end{aligned}$$

We first obtain a bound on the norm of M_1. Recall that

$$\lambda_{C_q} = \frac{1}{r_q}\left(W_{C_q,C_q}\mathbf{e} - \frac{1}{2}\left(\mu + \frac{\mathbf{e}^T W_{C_q,C_q}\mathbf{e}}{r_q}\right)\mathbf{e}\right)$$

by (5.4.7). Rearranging, we have

$$\begin{aligned} M_1 &= \frac{1}{r_q}\left(W_{C_q,C_q} - \frac{1}{2}\left(\mu + \frac{\mathbf{e}^T W_{C_q,C_q}\mathbf{e}}{r_q}\right)\mathbf{e}\right) - \frac{1}{2r_q}(\alpha r_q - \mu) \\ &= \frac{1}{r_q}(W_{C_q,C_q}\mathbf{e} - \alpha r_q \mathbf{e})\mathbf{e}^T + \frac{1}{2r_q}\left(\alpha r_q - \frac{\mathbf{e}^T W_{C_q,C_q}\mathbf{e}}{r_q}\right)\mathbf{e}\mathbf{e}^T. \end{aligned} \tag{5.4.66}$$

Note that we have

$$\left|\alpha r_q - \frac{\mathbf{e}^T W_{C_q,C_q}\mathbf{e}}{r_q}\right| \leq B\sqrt{r_q} \tag{5.4.67}$$

with probability at least (5.4.37) by (5.4.36) On the other hand, applying (5.4.30) with $\delta = 2$, $I_1 = I_2 = C_q$, and $X_{\ell_1,\ell_2} = W_{\ell_1,\ell_2}$ for all $\ell_1, \ell_2 \in C_q$, shows that

$$\frac{1}{r_q}\|W_{C_q,C_q}\mathbf{e} - \alpha r_q \mathbf{e}\|^2 \leq B r_q \tag{5.4.68}$$

with probability at least (5.4.37). Substituting (5.4.68) and (5.4.67) into (5.4.66) and applying the union bound, we have

$$\|M_1\| \leq \frac{1}{\sqrt{r_q}}\sqrt{B r_q r_s} + \frac{1}{2r_q} B r_q \sqrt{r_s} = \frac{\sqrt{B}}{2}(\sqrt{B}+2)\sqrt{r_s} \tag{5.4.69}$$

with probability at least $1 - 2p_1$ where

$$p_1 = (2/3)^{\hat{r}} - 2\hat{r}\exp(-2\hat{r}) \tag{5.4.70}$$

Similarly, we have

$$\|M_2\| \leq \frac{\sqrt{B}}{2}(\sqrt{B}+2)\sqrt{r_s} \tag{5.4.71}$$

with probability at least $1 - 2p_1$.

It remains to obtain an upper bound on $\|M_3\|$. Applying the triangle inequality, we have

$$\|M_3\| \leq \sqrt{r_s}\|\mathbf{y}\| + \sqrt{r_q}\|\mathbf{z}\|. \tag{5.4.72}$$

We begin by obtaining a bound on $\|\mathbf{y}\|$. Note that there exists scalar $c > 0$ such that

$$\|\mathbf{y}\| \leq \frac{1}{r_s}\left(\|\mathbf{b}_1\| + \frac{|\mathbf{b}_1^T\mathbf{e}|}{r_q + r_s}\sqrt{r_q}\right) \leq \frac{1}{r_s}\left(\|\mathbf{b}_1\| + \frac{c r_s r_q}{r_q + r_s}\right) \tag{5.4.73}$$

with probability at least $1 - 3p_1$ by the triangle inequality and (5.4.48). We next obtain a bound on $\|\mathbf{b}_1\|$. Recall that

$$\|\mathbf{b}_1\| \leq r_s\left\|\lambda_{C_q} - \frac{1}{2r_q}(\alpha r_q - \mu)\mathbf{e}\right\| + \left|\lambda_{C_s}^T\mathbf{e} - \frac{1}{2}(\alpha r_s - \mu)\right|\sqrt{r_q} + \|W_{C_q,C_s}\mathbf{e} - \beta r_s \mathbf{e}\|. \tag{5.4.74}$$

Note that (5.4.69) implies that

$$r_s\left\|\lambda_{C_q} - \frac{1}{2r_q}(\alpha r_q - \mu)\mathbf{e}\right\| = \sqrt{r_s}\|M_1\| \leq \frac{\sqrt{B}}{2}(\sqrt{B}+2)r_s \tag{5.4.75}$$

with probability at least $1 - 2p_1$. Next, applying (5.4.30) with $\delta = 2$, $I_1 = C_q$, $I_2 = C_s$, and $X_{\ell_1\ell_2} = W_{\ell_1\ell_2}$ for all $\ell_1 \in C_q$, $\ell_2 \in C_s$ yields

$$\|W_{C_q,C_s}\mathbf{e} - \beta r_s \mathbf{e}\|^2 \leq B r_q r_s \tag{5.4.76}$$

with probability at least $1 - p_1$. Therefore, substituting (5.4.45), (5.4.75), (5.4.76) in (5.4.73), there exists scalar $t_1 > 0$ such that

$$\|\mathbf{b}_1\| \leq t_1 r_s \tag{5.4.77}$$

with probability at least $1 - 3p_1$ by the union bound. Substituting (5.4.77) in (5.4.73) yields

$$\|\mathbf{y}\| \leq t_1 + c \tag{5.4.78}$$

with probability at least $1 - 6p_1$ by the union bound. Similarly, there exists $t_2 > 0$ such that

$$\|\mathbf{z}\| \leq t_2 \frac{\sqrt{r_s}}{\sqrt{r_q}} \tag{5.4.79}$$

with probability at least $1-6p_1$. Substituting (5.4.78) and (5.4.79) in (5.4.72) and applying the union bound shows that

$$\|M_3\| \leq (t_1 + t_2 + c)\sqrt{r_s} \tag{5.4.80}$$

with probability at least $1-12p_1$. Finally, combining (5.4.69), (5.4.71), and (5.4.80) shows that there exists scalar $t > 0$ such that

$$\|S_{C_q,C_s}\| \leq t\sqrt{r_s} \tag{5.4.81}$$

with probability at least $1 - 16p_1$ as required. ■

Chapter 6

Numerical results

In this chapter, we provide empirical evidence that our heuristics for the maximum clique, KDC, WKDC, and clustering problems are exact for a variety of program inputs. Our tests were performed using the 64-bit version of Matlab R2011a on a PC running Mac OSX 10.6.7, with a 2.26 GHz Intel Core 2 Duo processor and 2GB of RAM. The Matlab code used to perform these tests is available from http://www.math.uwaterloo.ca/∼bpames, under a GNU General Public License.

6.1 The Maximum Clique Problem

For $N = 200$ and each $n = 2, 4, \cdots, 170$, we repeated the following procedure 15 times. First, a random $N \times N$ matrix A, corresponding to a random graph containing planted clique of size n, was generated as follows. Let $V = \{1, \ldots, N\}$ and $V^* = \{1, \ldots, n\}$. We set $A_{ij} = 1$ for all $(i, j) \in V^* \times V^*$. For each remaining pair $(i, j) \in (V \times V) \backslash (V^* \times V^*)$ such that $i < j$ we set $A_{ij} = 1$ with probability $1/2$ and $A_{ij} = 0$ with probability $1/2$, independently. We then set $A_{ji} = A_{ij}$. Note that A is the adjacency matrix of a graph containing a planted clique V^* of size n and loops for all $i \in V^*$ constructed according to (Γ_R) in Section 4.1.3. We attempt to recover the planted clique V^* using three different algorithms. We first attempt to identify V^* using Algorithm 4.1.1. We implement Algorithm 4.1.1 in Matlab as the function `Alon`. We use the software package `PROPACK` [121] to compute the second eigenvector $\mathbf{v}_2$ of A and declare V^* recovered if the vector output by `Alon` is equal to the characteristic vector of V^*. Next, we attempt to recover V^* by solving the nuclear norm

minimization problem

$$\min\{\|X\|_* : \langle X, \mathbf{e}\mathbf{e}^T\rangle \geq n, \;\; X_{ij} = 0 \text{ if } A_{ij} = 0, \; i \neq j\}. \tag{6.1.1}$$

We solve (6.1.1) using the nuclear norm minimization software package `PPAPack` [126] in Matlab. We formulate and solve this problem using the Matlab function `MaxCliquePPA`, which is a modification of the example `RunMaxClique` provided with `PPAPack`. We declare V^* to be recovered if the output matrix X^* satisfies $\|X^* - \mathbf{v}\mathbf{v}^T\|_F / \|\mathbf{v}\mathbf{v}^T\|_F < 10^{-2}$. We test theoretical recovery of V^* by checking if the multiplier W, as constructed according to (ω_1)-(ω_6) in Section 4.4.1 with $\gamma = -1$, satisfies $\|W\| < 1$. Finally, we attempt to recover V^* by solving the semidefinite program

$$\min\{\langle X, \mathbf{e}\mathbf{e}^T\rangle : X \succeq 0, \; X\mathbf{e} \leq \mathbf{e}, \; \mathrm{Tr}\,(X) = 1, \; X_{ij} = 0 \text{ if } A_{ij} = 0, \; i \neq j\}. \tag{6.1.2}$$

We solve (6.1.2) in Matlab using the semidefinite programming software package `SDPNAL` [194]. We formulate and solve (6.1.2) in `SDPNAL` using the Matlab functions `kdc_sdpnal_1` and `kdc_sdp_nal_2` with input $G = A$ and $k = 1$. The function `kdc_sdpnal_1` generates the appropriate program inputs $\mathcal{A}$, $\mathbf{b}$, C for `SDPNAL`, calls `SDPNAL` to solve (6.1.2), and outputs the obtained optimal solution X^*, as well as the coefficient matrix C and the rows of $\mathcal{A}$ and $\mathbf{b}$ corresponding to the row sums and trace constraints. The function `kdc_sdpnal_2` generates the rows of the program inputs $\mathcal{A}$, $\mathbf{b}$ not already saved in memory and solves the resulting semidefinite program using `SDPNAL`, and outputs X^*. We declare V^* to be recovered in this case if the output X^* satisfies $\|nX^* - \mathbf{v}\mathbf{v}^T\|_F / \|\mathbf{v}\mathbf{v}^T\| < 10^{-3}$. Moreover, we test theoretical recovery of V^* by checking if the dual variable S, as constructed in Section 5.3.1 with $\theta = 1/2$ and $\gamma = 1$, is semidefinite. We say that V^* is theoretically recovered if $\lambda_{\min}(S) > -10^{-8}$. Figures 6.1.1 and 6.1.2 plot the average recovery and average error, respectively, for each trial of each heuristic. The average value of $\|W\|$ and $\lambda_{\min}(S)$ for each trial is plotted in Figure 6.1.3. Table 6.1.1 summarizes the computational time required for each algorithm.

Note that we have a sharp transition to theoretical perfect recovery at $n \approx 30 \approx \sqrt{200}$ for both NNM and SDP heuristics, although we do not obtain perfect recovery in practice until n is much larger. Moreover, we have a sharp decrease in the error for the NNM heuristic at $n \approx 30$. The lack of recovery in the NNM case appears to be due to the solver `PPAPack` not providing a solution within the desired error tolerance, 10^{-2}, of the optimal solution of (6.1.2). The desired optimal solution X_0 can be recovered by rounding each entry of X^* to the nearest integer. On the other hand, the lack of recovery by the SDP heuristic appears to be due to the SDP solver `SDPNAL` failing to obtain an optimal solution for n less than 160. In this case, `SDPNAL` appears to be terminating at a suboptimal

Figure 6.1.1: Success rate (out of 15 trials) of Algorithm 4.1.1, and our heuristics based on nuclear norm minimization and semidefinite programming for recovery of a planted clique of size n in a graph on $N = 200$ nodes.

(a) Alg 4.1.1 (b) NNM (c) SDP

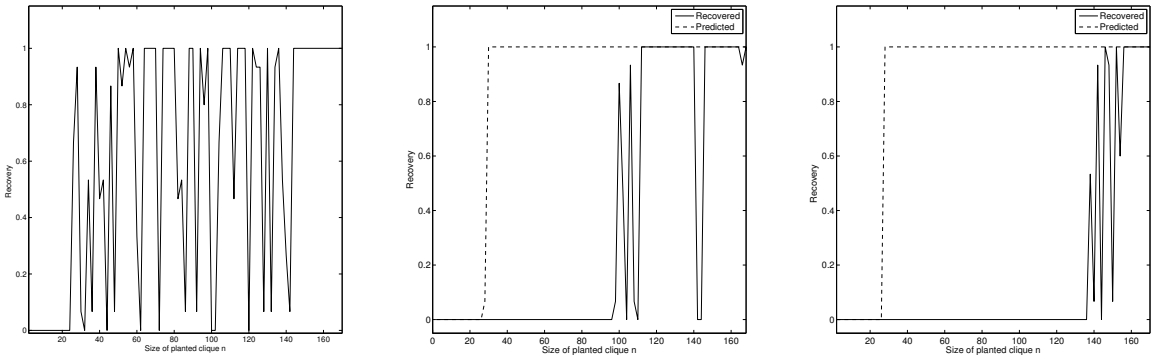


Figure 6.1.2: Average error, as measured by the Frobenius norm, for the NNM and SDP heuristics for recovery of a planted clique of size n in a graph on $N = 200$ nodes.

(a) NNM (b) SDP

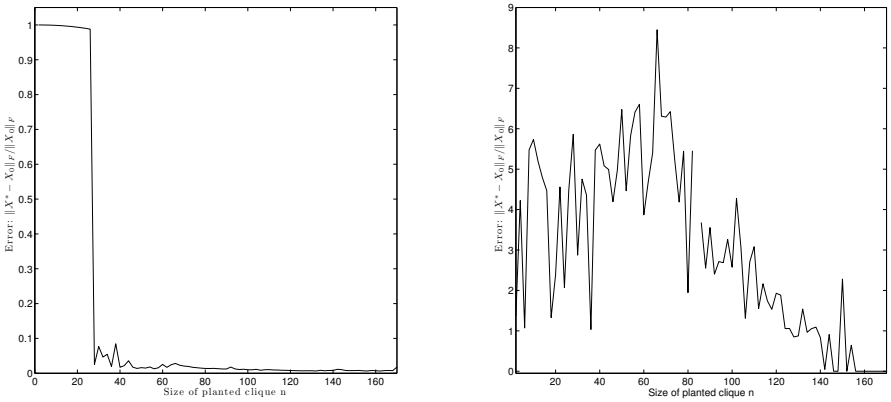
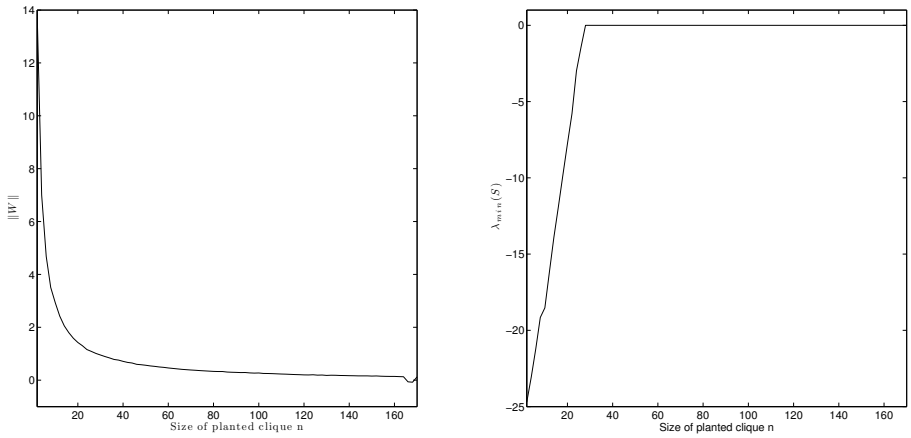


Table 6.1.1: CPU times of the spectral, NNM, and SDP heuristics for max clique

n	CPU time								
	Algorirthm 4.1.1			NNM			SDP		
	Best	Worst	Average	Best	Worst	Average	Best	Worst	Average
2	0.128	0.294	0.197	1.95	3.08	2.32	172.0	214.8	196.8
10	0.060	0.121	0.068	1.47	1.76	1.58	185.4	224.8	195.9
18	0.061	0.148	0.073	1.10	1.43	1.22	237.5	245.1	241.7
26	0.047	0.061	0.051	0.93	1.46	1.10	145.8	274.3	205.1
34	0.034	0.035	0.035	4.76	6.29	5.47	155.9	247.1	206.8
42	0.023	0.037	0.028	3.67	9.95	7.49	108.1	257.7	186.8
50	0.024	0.147	0.038	3.71	5.23	4.36	172.0	244.1	192.9
58	0.025	0.094	0.030	3.27	3.59	3.38	172.5	226.8	221.7
66	0.025	0.026	0.025	2.05	2.31	2.11	249.9	252.0	250.9
74	0.017	0.061	0.020	1.92	2.09	2.03	200.3	313.5	227.5
82	0.018	0.057	0.025	2.12	3.88	2.83	162.0	206.8	186.4
90	0.018	0.027	0.020	2.08	3.07	2.32	191.4	257.1	239.2
98	0.026	0.031	0.027	2.81	2.98	2.88	152.9	252.3	179.4
106	0.026	0.040	0.027	2.31	2.72	2.49	239.6	330.4	297.1
114	0.026	0.090	0.032	1.63	2.52	2.05	146.0	259.4	217.6
122	0.026	0.027	0.026	1.35	1.93	1.64	164.6	275.4	209.4
130	0.026	0.027	0.026	2.02	2.35	2.22	150.9	429.4	282.1
138	0.026	0.089	0.036	1.74	2.22	1.96	60.4	223.3	116.8
146	0.027	0.128	0.049	1.80	2.35	2.02	46.0	65.1	49.1
154	0.038	0.056	0.044	1.65	2.14	1.84	43.5	259.0	127.8
162	0.038	0.059	0.041	1.38	4.98	1.86	41.8	55.0	47.6
170	0.045	0.066	0.047	2.72	3.74	2.99	22.9	26.7	24.7

Figure 6.1.3: Behaviour of the multipliers W and S predicting theoretical recovery of a planted clique of size n in a graph on $N = 200$ nodes for the NNM and SDP heuristics.

(a) $\|W\|$ (b) $\lambda_{\min}(S)$



solution that is not close in any sense to X_0. From these trials, it appears that our NNM heuristic for the maximum clique problem outperforms the existing spectral and ϑ-function heuristics; we quickly (< 10 seconds) obtain a good approximation of the planted clique via NNM for much smaller values of n than we do using Algorithm 4.1.1 or by solving the SDP relaxation.

We repeat these trials for the nuclear norm heuristic for $N = 500$ and $N = 1000$. It takes, on average, approximately 15 seconds to solve each instance (6.1.2) for $N = 500$ and 45 seconds to solve each instance for $N = 1000$.s As before, we have a sharp transition to predicted recovery at $n = 50 \approx 2\sqrt{N}$ and $n = 65 \approx 2\sqrt{N}$ for $N = 500$ and $N = 1000$ respectively, but do not consistently obtain the correct solution (within the desired error bounds) until n is much larger. We also have a sharp decrease in error shortly after these transition points, which suggests that the lack of recovery of the predicted solution is caused by inaccurate solutions to (6.1.2) returned by `PPApack` and not because the predicted solution is suboptimal. The results of these trials are summarized in Figures 6.1.4, 6.1.5, and 6.1.6.

Figure 6.1.4: Success rate (out of 15 trials) of the NNM heuristic for recovery of a planted clique of size n in a graph on in a graph on N nodes.

(a) $N = 500$ (b) $N = 1000$

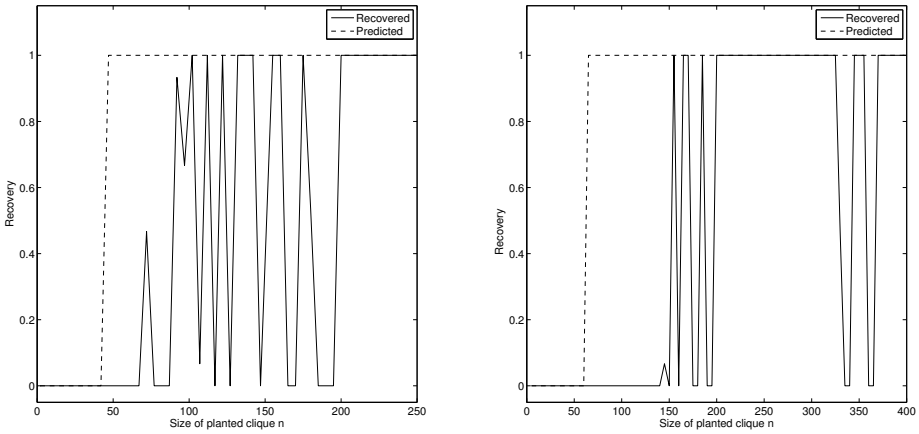


Figure 6.1.5: Average error, as measured by the Frobenius norm, for the NNM heuristic for recovery of a planted clique of size n in a graph on N nodes.

(a) $N = 500$ (b) $N = 1000$

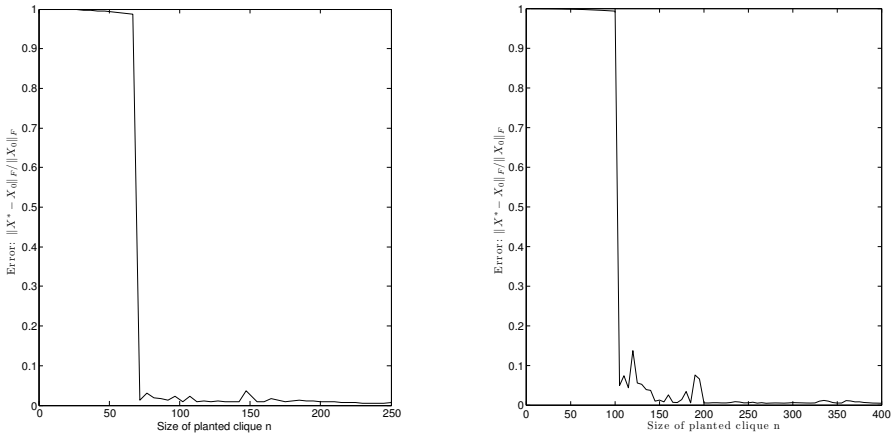
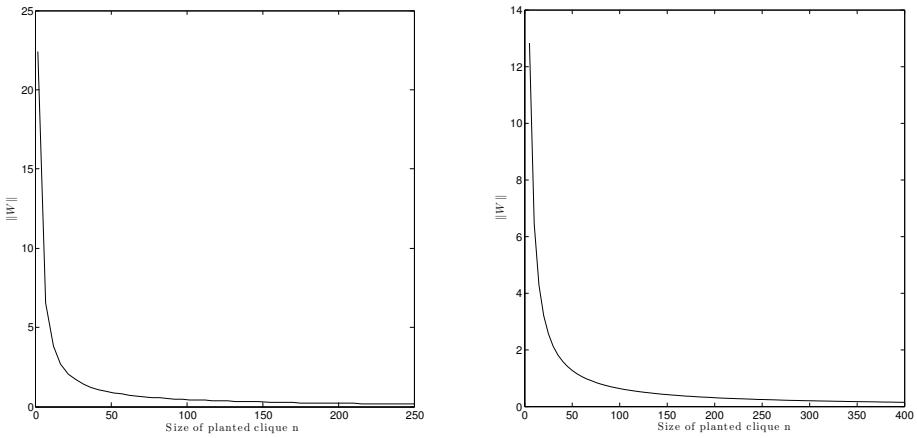


Figure 6.1.6: Behaviour of the multiplier W predicting theoretical recovery of a planted clique of size n in a graph on in a graph on N nodes.

(a) $N = 500$ (b) $N = 1000$



6.2 Random data for KDC

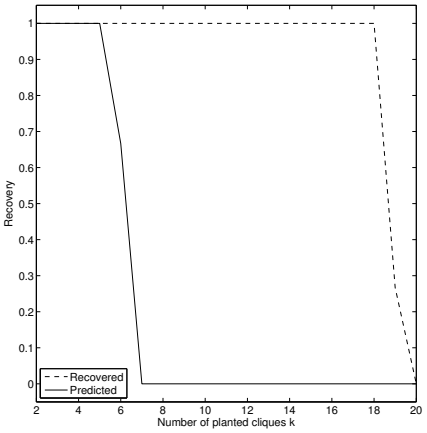
For each $k = 2, \ldots, 20$, we performed the following procedure 15 times. We generate the adjacency matrix A of a random graph containing a k-disjoint-clique subgraph composed of cliques $C_1, \ldots, C_k$ of size $r_1 = r_2 = \cdots = r_k = \hat{r} = 25$ and $r_{k+1} = 0$. We do so by setting all entries of the 25×25 diagonal blocks indexed by $C_1, \ldots, C_k$ equal to 1 and then choosing each remaining upper triangular entry of A to be equal to 1 with probability $1/2$, and the lower triangular entries of A by symmetry. and equal to 0 with probability $1/2$. We then solve the semidefinite program

$$\min\{\langle X, \mathbf{e}\mathbf{e}^T\rangle : X \succeq 0,\ X\mathbf{e} \leq \mathbf{e},\ \mathrm{Tr}\,(X) = k,\ X_{ij} = 0 \text{ if } A_{ij} = 0,\ i \neq j\} \tag{6.2.1}$$

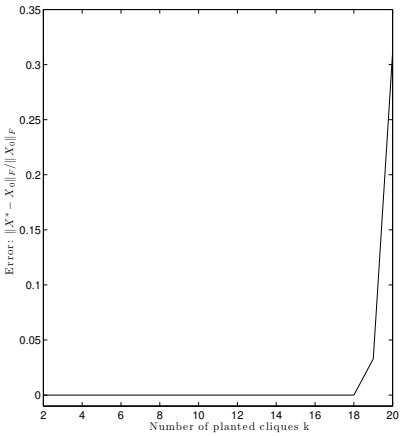
using `SDPNAL` in Matlab. As before, we construct the necessary program inputs $\mathcal{A}, \mathbf{b}$ and C using the functions `kdc_sdpnal_1` and `kdc_sdp_nal_2`. We declare the planted k-disjoint-clique subgraph to be recovered if the optimal solution X^* returned by `SDPNAL` satisfies $\|X^* - X_0\|_F / \|X_0\|_F < 10^{-3}$, where X_0 is the predicted optimal solution given by (5.2.10). We also recorded the minimum eigenvalue of the predicted semidefinite dual multiplier S, constructed as in Section 5.3.1 with $\theta = 1/2$, $\gamma = 1$, to test theoretical recovery of X_0. As before, we say that we have theoretical recovery if $\lambda_{min}(S) > 10^{-8}$ and calculate the

Figure 6.2.1: Results of trials for the SDP relaxation of KDC for $\hat{r} = 25$

(a) Recovery



(b) Error



minimum eigenvalue of S using `PROPACK`. Our results are summarized in Figure 6.2.1 and Table 6.2.1. Note that our bounds on the number of cliques that our algorithm can recover appear to be conservative compared to those encountered in practice. There is a sharp transition from predicted recovery of X_0 at $k = 5 = \sqrt{\hat{r}}$, yet we still recover the proposed optimal solution X_0 for all k less than 19.

Note that solving each instance of KDC as generated above requires solving a semidefinite program with $k\hat{r}(k\hat{r}-1)/2$ decision variables and approximately $k\hat{r}+1+k(k-1)\hat{r}^2/2$ constraints. For larger values of k or $\hat{r}$, it becomes impractical to solve this large SDP. Hence, we do not check performance of our heuristic for $\hat{r}$ larger than 25. Instead we check the *predicted* performance. That is, we test our theoretical bounds for larger values ensuring optimality of X^*, for random graphs containing k planted cliques of size $\hat{r} = 50, 100$. We repeat the following procedure 15 times for each k. We first generate the random matrix A as above with all-ones diagonal blocks of size $\hat{r}$. We then calculate the proposed choice of S (with $\theta = 1/2$ and $\gamma = 1$) and check the smallest eigenvalue of S using `PROPACK`. We declare that S is semidefinite, and, hence, we would recover X_0 if we solved (6.2.1), if $\lambda_{\min}(S) > -10^{-8}$. Figures 6.2.2 and 6.2.3 plot the number of predicted recoveries and average value of $\lambda_{\min}(S)$ for each k.

Table 6.2.1: Results of the SDP relaxation for KDC: Here k denotes the number of planted cliques. Recovery is the number of instances (out of 15) where the predicted and recovered solutions agree (within error tolerance). Theoretical recovery is the number of times that the multiplier S is positive semidefinite. The error and $\lambda_{\min}(S)$ are the average over all 15 trials of $\|X^* - X_0\|_F / \|X_0\|_F$ and the smallest eigenvalue of the predicted multiplier S. CPU times include the time required to construct the inputs for the SDP solver, as well as the time required to solve the SDP.

k	Recovery	Predicted Recovery	$\lambda_{\min}(S)$	Error	CPU time		
					Best	Average	Worst
2	15	15	-9.1057e-15	1.1819e-06	0.34583	0.4705	1.0359
3	15	15	-1.7146e-14	1.0815e-06	0.63443	0.74517	0.94566
4	15	15	-1.7458e-14	2.5751e-06	1.3334	1.4266	1.5638
5	15	15	-2.953e-14	1.5193e-06	1.7705	2.0589	2.4045
6	15	10	-0.17539	1.0473e-06	2.3517	2.6684	3.1613
7	15	0	-2.0738	1.6333e-06	3.9691	4.4433	5.6791
8	15	0	-4.3379	1.9826e-06	6.6964	7.2963	8.4287
9	15	0	-5.8891	3.8052e-06	9.0208	10.033	11.247
10	15	0	-8.2213	4.3938e-06	12.033	13.265	16.077
11	15	0	-9.9102	2.5215e-06	14.749	16.894	20.724
12	15	0	-11.724	1.8852e-06	19.408	23.265	25.447
13	15	0	-13.417	2.3138e-06	26.058	27.049	28.429
14	15	0	-14.989	3.3615e-06	33.059	35.694	43.768
15	15	0	-16.402	3.2636e-06	47.341	51.233	54.927
16	15	0	-17.749	3.639e-06	59.124	62.782	64.848
17	15	0	-19.387	7.6072e-06	76.937	82.384	89.216
18	15	0	-20.854	5.2288e-06	107.05	115.01	127.81
19	4	0	-22.056	0.032889	219.18	576.66	901.86
20	0	0	-23.174	0.31126	355.72	435.97	651.17

Figure 6.2.2: Predicted recovery rates for the SDP relaxation of KDC.

(a) $\hat{r} = 50$ (b) $\hat{r} = 100$

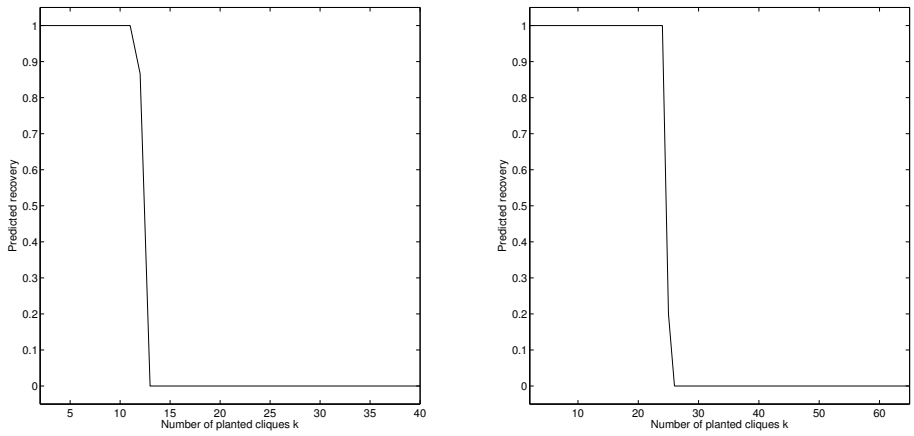
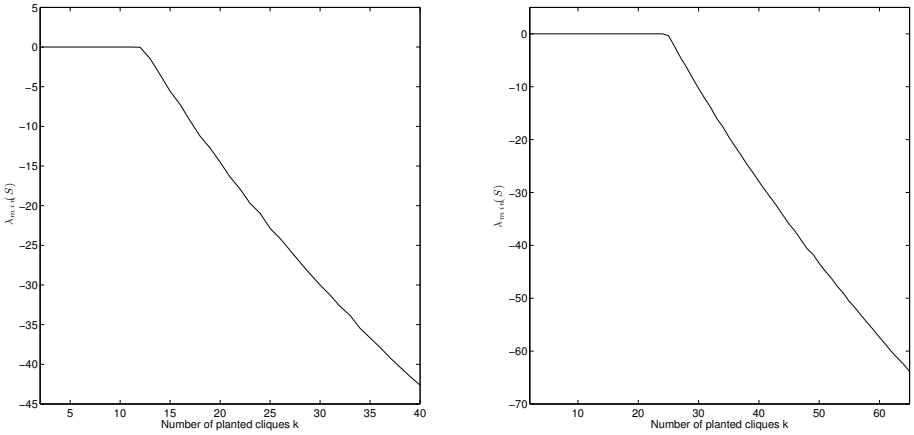


Figure 6.2.3: The minimum eigenvalue of the proposed multiplier S.

(a) $\hat{r} = 50$ (b) $\hat{r} = 100$



6.3 Random data for WKDC

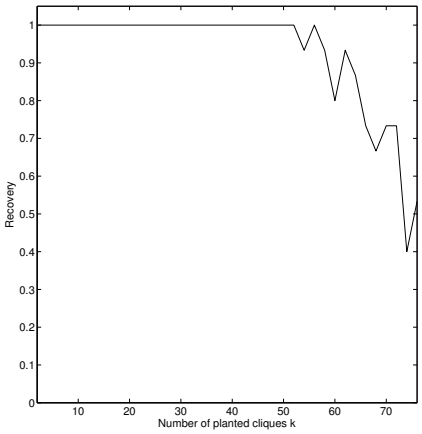
We considered random matrices $W \in \Sigma^{k\hat{r}}$ generated as follows. Let $C_1 = \{1, \ldots, \hat{r}\}$, $C_2 = \{\hat{r}+1, \ldots, 2\hat{r}\}$, $\ldots$, $C_k = \{(k-1)\hat{r}+1, \ldots, k\hat{r}\}$. Fix $\alpha > 2\beta$. For each $q = 1, \ldots, k$, we choose the entries of W_{ij}, $i, j \in C_q$, $i \leq j$ uniformly at random (independently) from the interval $[1-2(1-\alpha), 1]$ so that these entries of W have expected value α. The off-diagonal upper triangular entries of W are drawn independently at random from the uniform distribution on the interval $[0, 2\beta]$. These entries of W are i.i.d. random variables with expected value β. We choose the lower triangular entries of W by symmetry. Note that this W is drawn from a distribution satisfying Conditions (ω_1) and (ω_2) as in Section 5.2.2.

Let $\hat{r} = 15$, $\alpha = 0.65$, and $\beta = 0.3$. We repeated the following process 15 times for each k. We generate the random symmetric matrix $W \in \Sigma^{15k}$ as above and solve the semidefinite relaxation of WKDC given by (5.2.31) using `SDPNAL` in Matlab. We use the Matlab functions `wkdc_sdpnal_1` and `wkdc_sdpnal_2` to formulate and solve (5.2.31). These functions are the WKDC analogue of the functions `kdc_sdpnal_1` and `kdc_sdpnal_2` described in the previous sections. Let X^* be the solution obtained from solving (5.2.31). We consider the planted k-disjoint-clique subgraph recovered if X^* satisfies $\|X^* - X_0\|_F / \|X_0\|_F < 10^{-3}$ where X_0 is the feasible solution corresponding to $\{C_1, \ldots, C_k\}$ given by (5.2.10). We also test the predicted recovery rate, by checking if the proposed multipliers S, λ, and η as constructed in Section 5.4.1, are dual feasible. Figure 6.3.1 plots the empirical and theoretical success rates of our heuristic as well as the average error for each trial, and Figure 6.3.2 plots the average behaviour of the dual variables S, λ, and η for each k. Our results are summarized in Table 6.3.1. Note that our theoretical guarantees for optimality of X^* appear to be rather conservative; the proposed multiplier S is never semidefinite in our trials, yet we have perfect recovery for all k up to 54. Moreover, our results do not exhibit a sharp transition from perfect recovery. Although we do not recover X_0 in all trials for $k \geq 54$, we still have recovery in some trials for all k less than 76.

As before, it was impractical to test the actual performance of our heuristic for larger choices of $\hat{r}$. Instead we tested the predicted recovery, according to the proposed dual variables S, λ, and η. For each $\hat{r} = 50, 100$, we repeated the following process 15 times for each choice of k. We generated the random symmetric matrix $W \in \Sigma^{k\hat{r}}$, and S, λ, η as above. We say that our algorithm would recover the correct optimal solution if each of $\lambda_{min}(S)$, $\min(\lambda)$ and $\min(\eta)$ is at least -10^{-8}. Figure 6.3.3 plots the fraction of trials in which all three multipliers are feasible for each k. We have expected perfect recovery for all k less than 19 and 38 for $\hat{r} = 50$ and 100 respectively, at which point there is a sharp transition to zero recovery. The minimum entry of η was equal to 0 for each trial. Figures 6.3.4 and 6.3.5 plot the average values of $\lambda_{\min}(S)$ and $\min(\lambda)$ for each k.

Figure 6.3.1: Results for the SDP relaxation of WKDC for $\hat{r} = 15$

(a) Recovery



(b) Error

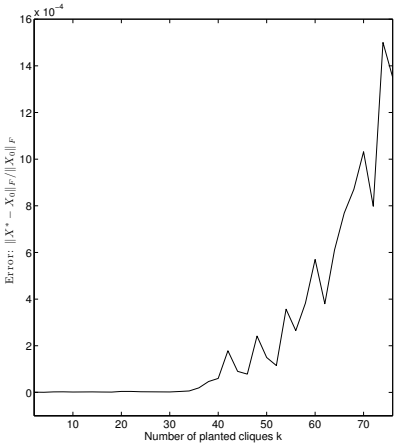
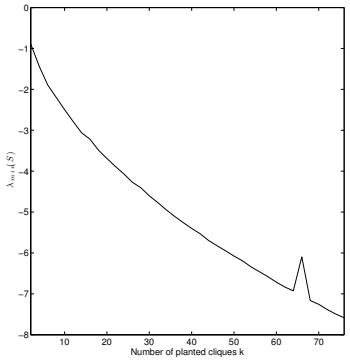
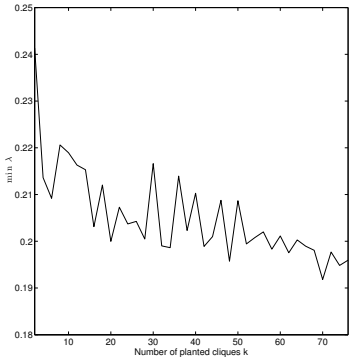


Figure 6.3.2: Average behaviour (over all 15 trials) of the dual variables S, λ, η for the SDP relaxation of WKDC

(a) $\lambda_{\min}(S)$



(b) $\min \lambda$



(c) $\min \eta$

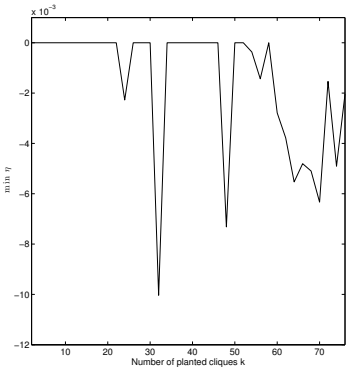


Table 6.3.1: Results of the SDP relaxation for WKDC: k denotes the number of planted cliques. Recovery is the number of instances (out of 15) where the predicted and recovered solutions agree (within error tolerance). Theoretical recovery is the number of times that the multiplier S is positive semidefinite, and the multipliers λ and η are nonnegative. The error, $\lambda_{\min}(S)$, $\min \lambda$, $\min \eta$ are the average over all 15 trials of $\|X^* - X_0\|_F / \|X_0\|_F$, the smallest eigenvalue of the predicted multiplier S, and smallest entries of λ and η respectively. CPU times include the time required to construct the inputs for the SDP solver, as well as the time required to solve the SDP.

k	Recovery	Predicted Recovery	Error	$\lambda_{\min}(S)$	$\min \lambda$	$\min \eta$	CPU time		
							Best	Average	Worst
2	15	0	1.27E-06	-0.89	0.24	0	0.37954	0.45645	0.53273
6	15	0	2.33E-06	-1.89	0.21	0	0.85935	0.99998	1.1072
10	15	0	1.73E-06	-2.50	0.22	0	1.5955	2.2866	3.425
14	15	0	2.20E-06	-3.06	0.22	0	5.4248	6.9991	9.4268
18	15	0	1.44E-06	-3.49	0.21	0	11.3	14.449	17.766
22	15	0	4.26E-06	-3.89	0.21	0	22.713	25.458	28.729
26	15	0	2.51E-06	-4.27	0.20	0	28.964	37.073	45.767
30	15	0	1.98E-06	-4.60	0.22	0	35.352	48.372	60.963
34	15	0	6.41E-06	-4.94	0.20	0	78.588	108.59	173.68
38	15	0	4.67E-05	-5.26	0.20	0	124.39	221.45	317.52
42	15	0	1.79E-04	-5.53	0.20	0	181.09	607.23	1321.1
46	15	0	7.84E-05	-5.82	0.21	0	211.84	629.26	1449.7
50	15	0	1.50E-04	-6.08	0.21	0	268.67	1014.6	1857.4
54	14	0	3.57E-04	-6.34	0.20	-3.68E-04	326.28	1045.9	2753.7
58	14	0	3.83E-04	-6.58	0.20	0	259.59	1220	3955
62	14	0	3.80E-04	-6.83	0.20	-3.78E-03	461.43	1128.5	2916.1
66	11	0	7.69E-04	-6.10	0.20	-4.81E-03	1086.6	2213.1	4410.4
70	11	0	1.03E-03	-7.26	0.19	-6.33E-03	1183.7	3129.4	5365.1
74	6	0	1.50E-03	-7.49	0.19	-4.90E-03	2375.7	3932.9	4979.7

Figure 6.3.3: Predicted recovery rates for the SDP relaxation of WKDC

(a) $\hat{r} = 50$ (b) $\hat{r} = 100$

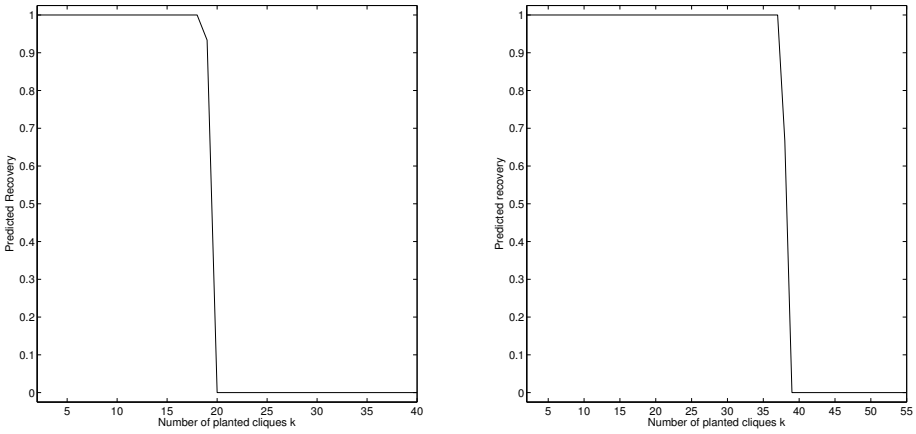


Figure 6.3.4: Average behaviour (over all 15 trials) of the dual variable S

(a) $\hat{r} = 50$ (b) $\hat{r} = 100$

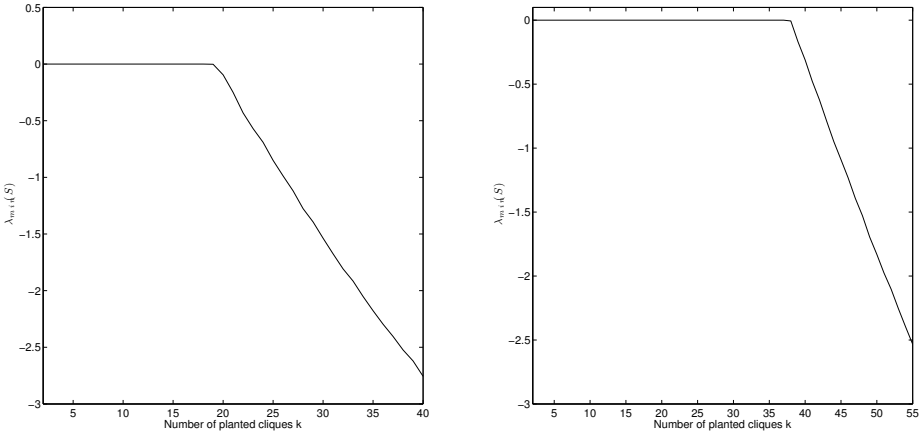
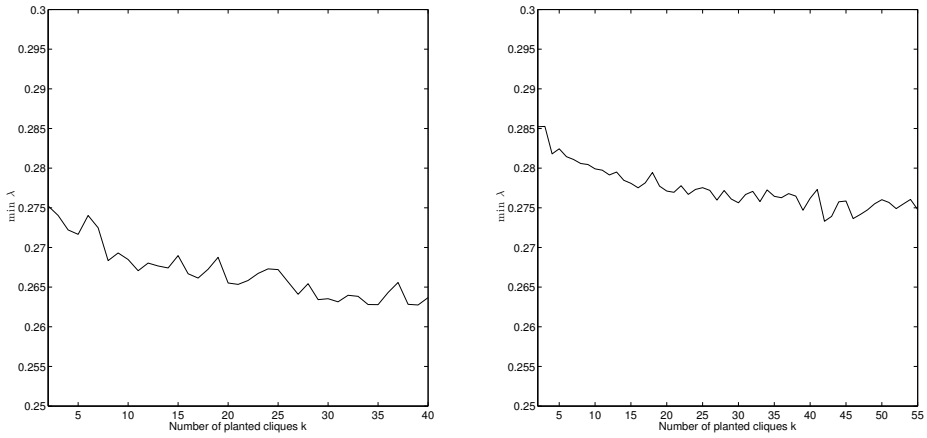


Figure 6.3.5: Average behaviour (over all 15 trials) of the dual variable λ

(a) $\hat{r} = 50$ (b) $\hat{r} = 100$



6.4 Clustering data sets

In this section, we apply our heuristic for KDC and WKDC to data sets from clustering applications.

6.4.1 The Supreme Court data set

We consider the liberal-conservative divide of the Rehnquist Supreme Court. The data used is the 9×9 similarity matrix W giving the percentages of cases in which the justices agreed, contained in Table 6.4.1. Here, clusters within the data set will correspond to subsets of the Supreme Court whose decisions on cases, and hence, legal and political ideologies, are similar. This data set was previously analyzed by Hubert and Stanley in [102], as well as Ben-Israel and Iyigun in [15]. We will show that this data set may be clustered using our semidefinite relaxation of WKDC.

We use the following algorithm for clustering data.

Algorithm 6.4.1 *Given data set of N items, similarity matrix $W \in \Sigma^N$, and proposed number of clusters k.*

Table 6.4.1: Agreement among the nine Supreme Court justices

	St	Br	Gi	So	Oc	Ke	Re	Sc	Th
1 St	1	0.62	0.66	0.63	0.33	0.36	0.25	0.14	0.15
2 Br	0.62	1	0.72	0.71	0.55	0.47	0.43	0.25	0.24
3 Gi	0.66	0.72	1	0.78	0.47	0.49	0.43	0.28	0.26
4 So	0.63	0.71	0.78	1	0.55	0.5	0.44	0.31	0.29
5 Oc	0.33	0.55	0.47	0.55	1	0.67	0.71	0.54	0.54
6 Ke	0.36	0.47	0.49	0.5	0.67	1	0.77	0.58	0.59
7 Re	0.25	0.43	0.43	0.44	0.71	0.77	1	0.66	0.68
8 Sc	0.14	0.25	0.28	0.31	0.54	0.58	0.66	1	0.79
9 Th	0.15	0.24	0.26	0.29	0.54	0.59	0.68	0.79	1

1. *Solve* (5.2.31) *with given* k *and* W*. Let* X^* *be optimal solution obtained.*

2. *Refine* X^* *to* $\tilde{X}$*:*

   (a) *For each* $i = 1, \ldots, N$*, if the* i*th row of* X^* *has row sum less than* 0.9*, set* $X^*(:, i) = 0$*,* $X^*(i, :) = 0$*.*

   (b) *For all* $i, j \in \{1, \ldots, N\}$ *let* $\tilde{X}_{ij} = \left\lceil X^*_{ij} - 1/N \right\rceil$*.*

3. *Obtain the eigenvectors* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ *corresponding to the* k *largest eigenvalues of* $\tilde{X}$*.*

4. *Refine these eigenvectors to* $\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_k$ *by taking* $\tilde{\mathbf{v}}_i(j) = \lceil \mathbf{v}_i(j) - 1/N \rceil$ *for all* $i = 1, \ldots, k$ *and* $j = 1, \ldots, N$*. These refined eigenvectors are the characteristic vectors of the clusters identified by the algorithm.*

We applied Algorithm 6.4.1 with $k = 2$ and W as given by Table 6.4.1. The Matlab function `weighted_kdc_cvx` was used to formulate and solve (5.2.31) using `CVX` [86] and `SeDuMi` [173]. We identified two disjoint clusters in the data. The names of the Supreme Court justices in each cluster are listed in Table 6.4.2. Our 2-clustering of the Rehnquist Supreme Court agrees with that identified in [102] and [15]. Note that our algorithm does not provide any information about the political ideologies of either cluster; the labels "Liberal" and "Conservative" used to identify each cluster were assigned in [102], in which a ranking of justices from most liberal to most conservative was obtained.

Note that our heuristic for clustering is sensitive to the choice of k. For example, if we instead solve (5.2.31) with $k = 3$, we obtain a clustering of the Supreme Court into three voting blocks, labelled in Table 6.4.3 as "Liberal", "Moderate Conservative", and "Most Conservative", in agreement with the ranking provided in [102]. Thus, the choice of $k = 3$ causes the "Conservative" cluster to be split into two smaller clusters. Larger choices of k result in clusters of size 1. Hence, only $k = 2$ and $k = 3$ yield meaningful clusterings of the Supreme Court data set.

Table 6.4.2: A 2-clustering of the Rehnquist Supreme Court

1: "Liberal"	2: "Conservative"
Stevens (St) Breyer (Br) Ginsberg (Gi) Souter (So)	O'Connor (Oc) Kennedy (Ke) Rehnquist (Re) Scalia (Sc) Thomas (Th)

Table 6.4.3: A 3-clustering of the Rehnquist Supreme Court

1: "Most Conservative"	2: "Moderate Conservative"	3: "Liberal"
Thomas (Th) Scalia (Sc)	O'Connor (Oc) Kennedy (Ke) Rehnquist (Re)	Stevens (St) Breyer (Br) Ginsberg (Gi) Souter (So)

6.4.2 Birth and death rates

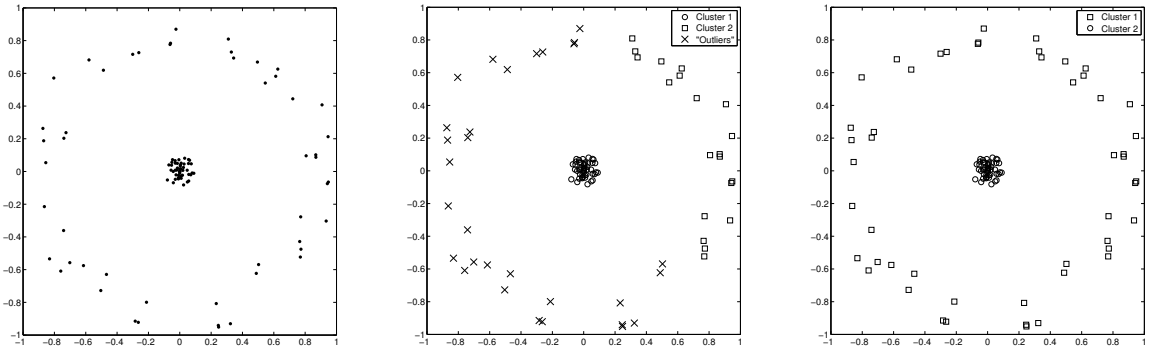
We consider the problem of clustering subsets of a Euclidean space $\mathbf{E}$. In such a space, we have a natural model of similarity; $\mathbf{x}, \mathbf{y} \in \mathbf{E}$ are similar if they are sufficiently close with respect to the norm in $\mathbf{E}$. We define the similarity matrix $W \in [0, 1]^{n \times n}$ by

$$W_{\mathbf{x},\mathbf{y}} = \exp(-\|\mathbf{x} - \mathbf{y}\|^2/\sigma^2) \tag{6.4.1}$$

for all $\mathbf{x}, \mathbf{y} \in S$, where S is the set to be clustered, $n = |S|$ is the cardinality of S, and $\|\cdot\|$ is the norm defined by the inner product associated with $\mathbf{E}$, and $\sigma > 0$ is the variance of the distances between the points in S. Note that the clustering of the data is highly dependent on the choice of similarity matrix W. For example, consider the data set in $\mathbf{R}^2$ depicted in Figure 6.4.1. This data set consists of two clusters, one centered at the origin and another arranged in a ring surrounding the cluster centered at the origin. We obtain a 2-clustering of the data using Algorithm 6.4.1 for $k = 2$ and W defined as in (6.4.1) where $\|\cdot\|$ is the Euclidean norm in $\mathbf{R}^2$. We solve the relaxation of (5.2.31) using `SDPNAL`. Our heuristic correctly identifies the cluster at the origin but fails to recover the entire second cluster; this failure to correctly cluster the data is due to the use of the incorrect measure of similarity $\|\cdot\|$ in W. We perform the change of basis obtained by converting the data set to polar coordinates, and then rescale the θ-coordinates of the data set by 1/10. Applying Algorithm 6.4.1 with Euclidean W to this transformed data set yields the correct clustering of the original data; see Figure 6.4.1. In general, choosing the correct similarity matrix is a hard problem and will restrict ourselves to *ad hoc* choices of W.

Figure 6.4.1: Clustering of a sample data set in $\mathbf{R}^2$. The data is not clusterable with respect to the Euclidean norm in $\mathbf{R}^2$. After change of basis, we obtain the correct clustering

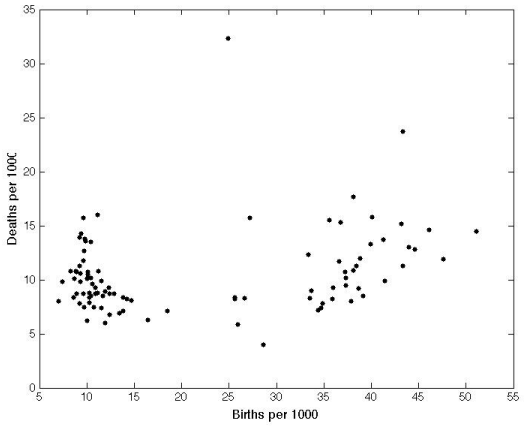
(a) Original data (b) Incorrect clustering (c) Correct clustering



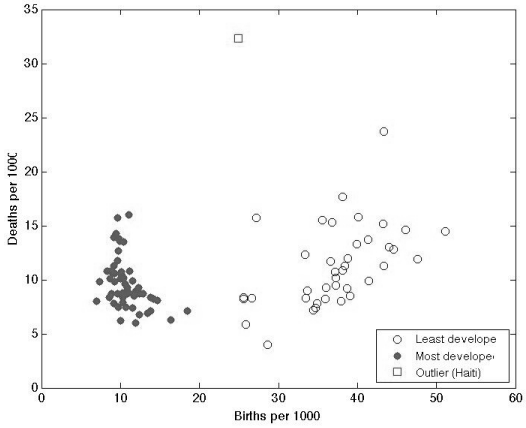
We analyze crude birth and death rate data drawn from the United States Census Bureau's International Database [183]. Figure 6.4.2 plots the number of births and number of deaths per 1000 people for 97 countries drawn from either the set of "More developed countries" or "Less developed countries" as identified by the United Nations. From Figure 6.4.2(a), it appears as though this data consists of two clusters in $\mathbf{R}^2$; one corresponding to the more developed nations and one corresponding to the less developed nations. We

Figure 6.4.2: Clustering of nations based on crude birth and death rates

(a) Original data



(b) Clustered data



attempt to cluster this data by applying Algorithm 6.4.1 with $k = 2$ and W as defined by (6.4.1) using `SDPNAL` in Matlab. Solving (5.2.31) for this data set takes less than 5 seconds. We obtain the partitioning of the data into two clusters and one outlier given by Figure 6.4.2 (b).

We next consider expanding the data to the full set of 227 countries provided by the United States Census Bureau's International Database. Solving (5.2.31) with $k = 2$ and W as above does not yield a good partitioning of the data. This is because the separation between the two predicted clusters is not distinct; there are many countries that could, based solely on their crude birth and death rates, be consider members of either the most developed or least developed. After solving (5.2.31) and refining, these countries are classified as members of both clusters. However, if we solve our relaxation (5.2.13) of KDC using the refined optimal solution X^* as the adjacency matrix of G, we obtain a clustering that identifies two tight clusters in the data set; see Figure 6.4.3.

6.4.3 Image segmentation

In this section, we apply our clustering heuristic to image segmentation. *Image segmentation* refers to the process of partitioning the pixels of a given image into segments or regions such that the pixels in each segment share certain characteristics. Hence, image segmentation seeks a clustering of the pixels of a given image.

We construct the similarity graph $G = (V, E, W)$ corresponding to the pixels of a given image I as in [167]. The set of nodes of G is the set of pixels of the input image I. We add an edge for every pair of nodes in V, and assign weight

$$W_{ij} = \begin{cases} \exp\left(-\frac{\|I(i)-I(j)\|^2}{\sigma_I^2}\right) \cdot \exp\left(-\frac{\|X(i)-X(j)\|^2}{\sigma_X^2}\right), & \text{if } \|X(i) - X(j)\| < r, \\ 0, & \text{otherwise} \end{cases} \tag{6.4.2}$$

for each edge $ij \in E$, where $I(i)$ and $X(i)$ denote the intensity and position of the ith pixel of the input image, and σ_I, σ_X, r are fixed, positive scalars to be chosen later. Note that W_{ij} is large if the i and jth pixels are close to each other and similar in intensity, and $W_{ij} = 0$ if i and j are more than r pixels apart.

Using Algorithm 6.4.1, we obtain a clustering into 3 distinct segments of a 30×30 synthetic image obscured by noise. Figure 6.4.4 contains the noisy image, as well as the identified segments. In general, the computational cost of our algorithm is prohibitively high. To segment an $m \times n$ image, we must solve a convex program with an $mn \times mn$ semidefinite decision variable and $O(m^2n^2)$ nonnegative decision variables and constraints. This is impractical, or impossible, for even moderate choices of m, n; say $m, n > 50$.

Figure 6.4.3: Clustering of nations based on crude birth and death rates

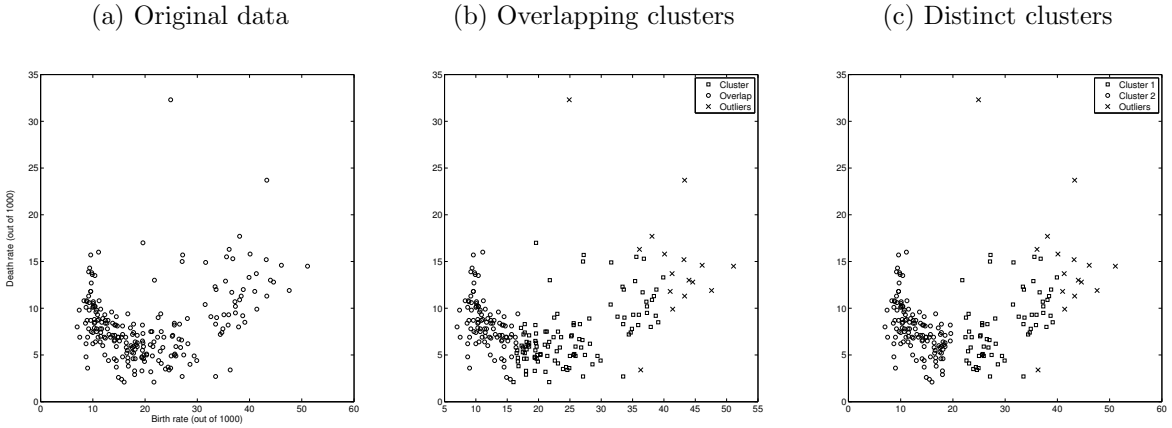
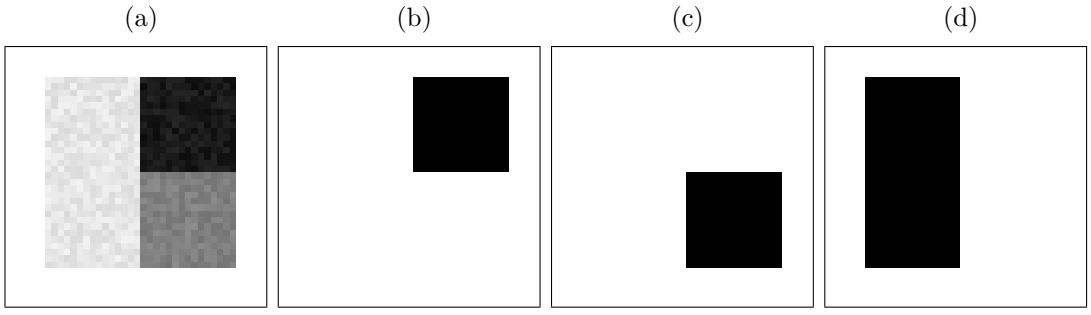


Figure 6.4.4: (a) A synthetic noisy image with three segments. (b)-(d) show the recovered segments, obtained using Alg 6.4.1 using $k = 3$, W as in (6.4.2) with $\sigma_I = 0.2$, $\sigma_X = 50$, $r = 25$.



Chapter 7

Conclusions

We have shown that the maximum clique, maximum edge biclique, and maximum node and maximum mean weight k-disjoint-clique problems can be relaxed to convex programming by replacing matrix rank with its convex envelope, the nuclear norm. In the special case that the input graph consists of a single large instance of the desired subgraph and a moderate number of additional nodes and edges then these relaxations are exact. Therefore, in this planted case, these hard combinatorial problems can be solved efficiently by solving these tractable relaxations. For each problem, we provide theoretical bounds on the number of diversionary edges and nodes our algorithm can tolerate and still recover the desired hidden subgraph. We provide separate analyses for deterministically and randomly inserted diversionary edges. Our bounds for the amount of random noise tolerated by our nuclear norm minimization (4.1.7) and semidefinite relaxation (5.2.13) of the maximum clique problem match those in the literature by Alon et al. [5] and Feige and Krauthgamer [66]. We believe our technique provides an improvement over these existing heuristics. We have shown that it may be extended to the problems of identifying the maximum edge biclique in a bipartite graph and of identifying the maximum k-disjoint-clique in an undirected graph without modification. Moreover, our relaxation technique may be extended to any rank minimization or rank-constrained optimization problem, although possibly without the accompanying guarantee of exact recovery. Furthermore, we have provided empirical evidence that our algorithm for recovering a planted clique is more robust in practice than those of Alon et al. and Feige and Krauthgamer, and that our heuristic for KDC and WKDC provides meaningful partitions of data sets from clustering applications.

- *Extensions to other combinatorial problems:* A common relaxation technique for combinatorial optimization is to lift vectors of integer variables to the positive semidef-

inite cone. The resulting positive semidefinite variable in the lifted problem will necessarily be low-rank. For example, Goemans and Williamson [84] obtain semidefinite relaxations of the maximum cut and maximum 2-satisfiability problems by replacing integer variables $\mathbf{u} \in \{-1, 1\}^N$ with semidefinite variables $X \in \Sigma_+^N$ satisfying $\operatorname{diag}(X) = \mathbf{e}$, $\operatorname{rank}(X) = 1$. In general, this lifting procedure yields a rank-constrained semidefinite programming formulation for the original combinatorial problem. This formulation is then relaxed to semidefinite programming by dropping the rank constraints on the decision variable. We have seen that relaxing this rank constraint using the nuclear norm instead of removing it entirely yields an exact relaxation for the maximum clique problem in the special case that the input graph contains a sufficiently large planted clique. It would be interesting to see if our relaxation and analytical techniques can be applied to other combinatorial optimization problems to improve upon existing approximation bounds or identify classes of program inputs for which the resulting relaxation is exact.

- *Relaxations for other graph-partitioning objectives:* In particular, we are interested in obtaining exact relaxations for the graph-partitioning problems discussed in Section 5.1.1. The maximum mean weight k-disjoint-clique problem seeks to identify a partition the nodes of the complete graph K_N on N nodes into k disjoint cliques and a set of outliers such that the weight of the edges between nodes in each clique is maximized. However, this partition does not exploit information about the weights of edges between nodes in different vertex sets. When the input graph is the similarity graph for some data set to be clustered, this partition yields a clustering of the data set where similarity between items in each cluster is high but not necessarily one where items in different clusters are dissimilar. Therefore, clustering data by identifying the maximum mean weight k-disjoint-clique subgraph of its similarity graph is prone to incorrectly labelling items as outliers. We would like to cluster data using graph-partitioning that simultaneously maximizes the similarity between items in the clusters and minimizes the similarity between items in different clusters such as those given by the normalized cut (5.1.4) or min-max cut (5.1.6) objectives. Ideally, we would like to extend our results to semidefinite relaxations of these problems, although it is unclear what these relaxations would look like.

- *Theoretical bounds for spectral clustering:* Computing the partitioning of the input graph given by our relaxations of KDC and WKDC is significantly more expensive than obtaining the partitioning yielded by spectral clustering techniques. Indeed, after obtaining an optimal solution X^* for (5.2.13) or (5.2.31) we must extract the eigenvectors corresponding to the k largest eigenvalues of X^* to obtain characteristic

vectors for the k disjoint cliques composing the optimal k-disjoint-clique subgraph of the input graph. In this sense, solving (5.2.13) and (5.2.31) can be thought of as a preprocessing step for a spectral method where the similarity matrix W is "corrected" to X^*. That is, we learn the correct choice of W by solving for X^* and then obtain a partitioning of the input graph using the spectrum of X^*. In the planted case, we are guaranteed that the partition of the input graph given by X^* is exactly the characteristic vectors of the optimal k disjoint cliques with respect to W (and X^*). For especially large input graphs, it may take hours to obtain X^* by solving (5.2.13) or (5.2.31) directly while taking a spectral or singular value decomposition of X^* may take only a few minutes. Hence, we would like to obtain the optimal k-disjoint-clique subgraph directly from the adjacency matrix A_G or the similarity matrix W. Alon et al. [5] show that the maximum clique of may be recovered in the planted case using only spectral information; recall Algorithm 4.1.1. Similarly, McSherry [133] shows that the graph bisection and k-colouring problems may be solved exactly using spectral methods for certain input graphs. It would be interesting to determine if these results can be extended to other spectral approaches to clustering and graph-partitioning.

References

[1] J. Abello, P.M. Pardalos, and M.G.C. Resende. On maximum clique problems in very large graphs. *External memory algorithms*, 50:119–130, 1999. 30

[2] M. Ackerman and S. Ben-David. Which data sets are 'clusterable'? - a theoretical study of clusterability. 2008. 53

[3] M. Ackerman and S. Ben-David. Clusterability: A theoretical study. *Proceedings of AISTATS-09, JMLR: W&CP*, 5:1–8, 2009. 53

[4] N. Alon, L. Babai, and A. Itai. A fast and simple randomized parallel algorithm for the maximal independent set problem. *Journal of Algorithms*, 7(4):567–583, 1986. 30

[5] N. Alon, M. Krivelevich, and B. Sudakov. Finding a large hidden clique in a random graph. *Random Structures and Algorithms*, 13:457–466, 1998. 31, 32, 61, 132, 134

[6] N. Alon and J.H. Spencer. *The probabilistic method*. Wiley-Interscience, 2008. 31, 32

[7] B. Ames and S. Vavasis. Convex optimization for the planted k-disjoint-clique problem, 2010. Available from http://arxiv.org/abs/1008.2814v1. 2, 60

[8] B. Ames and S. Vavasis. Nuclear norm minimization for the planted clique and biclique problems. *Mathematical Programming*, pages 1–21, 2011. 2, 40, 41, 43, 47, 48

[9] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and hardness of approximation problems. In *Proc. 33rd Ann. IEEE Symp. on Found. of Comp. Sci*, pages 14–23. IEEE Comput. Soc. Press, 1992. 30

[10] S. Arora and S. Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM (JACM)*, 45(1):70–122, 1998. 30

[11] E. Balas and CS Yu. Finding a maximum clique in an arbitrary graph. *SIAM J. COMPUT.*, 15(4):1054–1068, 1986. 30

[12] R. Basri and D.W. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 218–233, 2003. 14

[13] R. Battiti and M. Protasi. Reactive Local Search for the Maximum Clique Problem 1. *Algorithmica*, 29(4):610–637, 2001. 30

[14] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009. 20

[15] A. Ben-Israel and C. Iyigun. Probabilistic d-clustering. *Journal of Classification*, 25(1):5–26, 2008. 125, 126

[16] J. Bennett and S. Lanning. The netflix prize. In *Proceedings of KDD Cup and Workshop*, volume 2007, page 35. Citeseer, 2007. 1

[17] P. Berkhin. Survey of clustering data mining techniques. *Grouping Multidimensional Data: Recent Advances in Clustering*, pages 25–71, 2006. 53

[18] P. Berman and A. Pelc. Distributed probabilistic fault diagnosis for multiprocessor systems. In *Fault-Tolerant Computing, 1990. FTCS-20. Digest of Papers., 20th International Symposium*, pages 340–346. IEEE, 1990. 30

[19] R. Bhatia. *Matrix analysis*. Springer Verlag, 1997.

[20] R.C. Bolles and R.A. Cain. Recognizing and locating partially visible objects: The local-feature-focus method. *The International Journal of Robotics Research*, 1(3):57, 1982. 30

[21] R.C. Bolles and P. Horaud. 3DPO: A three-dimensional part orientation system. *The International Journal of Robotics Research*, 5(3):3, 1986. 30

[22] I.M. Bomze. Global escape strategies for maximizing quadratic forms over a simplex. *Journal of Global Optimization*, 11(3):325–338, 1997. 30

[23] I.M. Bomze. On standard quadratic optimization problems. *Journal of Global Optimization*, 13(4):369–387, 1998. 30

[24] I.M. Bomze, M. Budinich, P.M. Pardalos, and M. Pelillo. The maximum clique problem. *Handbook of combinatorial optimization*, 4(1):1–74, 1999. 30

[25] R. Boppana and M.M. Halldórsson. Approximating maximum independent sets by excluding subgraphs. *BIT Numerical Mathematics*, 32(2):180–196, 1992. 30

[26] J.M. Borwein and A.S. Lewis. *Convex analysis and nonlinear optimization: theory and examples*, volume 3. Springer Verlag, 2006. 8

[27] J.M. Bourjolly, P. Gill, G. Laporte, and H. Mercure. An exact quadratic 0-1 algorithm for the stable set problem. *Cliques, Coloring, and Satisfiability, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, eds. DS Johnson and MA Trick. American Mathematical Society, Providence, Rhode Island*, pages 53–74, 1996. 30

[28] J.M. Bourjolly, G. Laporte, and H. Mercure. A combinatorial column generation algorithm for the maximum stable set problem. *Operations research letters*, 20(1):21–29, 1997. 30

[29] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, UK, 2004. 17

[30] Y. Bresler and K. Lee. Efficient and guaranteed rank minimization by atomic decomposition. In *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*, pages 314–318. IEEE, 2009. 24

[31] C. Bron and J. Kerbosch. Algorithm 457: finding all cliques of an undirected graph. *Communications of the ACM*, 16(9):575–577, 1973. 30

[32] A.E. Brouwer, J.B. Shearer, N.J.A. Sloane, and W.D. Smith. A new table of constant weight codes. *Information Theory, IEEE Transactions on*, 36(6):1334 –1380, November 1990. 30

[33] J.F. Cai, E.J. Candès, and Z. Shen. A singular value thresholding algorithm for matrix completion. *Arxiv preprint arXiv:0810.3286*, 2008. 15

[34] E.J. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *Arxiv preprint arXiv:0912.3599*, 2009. 14, 27

[35] E.J. Candès and Y. Plan. Tight oracle bounds for low-rank matrix recovery from a minimal number of random measurements. *Arxiv preprint arXiv:1001.0339*, 2010. 24

[36] E.J. Candès and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9(6):717–772, 2009. 2, 25, 35

[37] E.J. Candès, J.K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006. 2, 21, 22

[38] E.J. Candès and T. Tao. Decoding by linear programming. *Information Theory, IEEE Transactions on*, 51(12):4203–4215, 2005. 2, 21, 22

[39] E.J. Candès and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *Information Theory, IEEE Transactions on*, 52(12):5406–5425, 2006. 21

[40] E.J. Candès and T. Tao. The power of convex relaxation: Near-optimal matrix completion. *Information Theory, IEEE Transactions on*, 56(5):2053–2080, 2010. 25

[41] R. Carraghan and P.M. Pardalos. An exact algorithm for the maximum clique problem. *Operations Research Letters*, 9(6):375–382, 1990. 30

[42] V. Chandrasekaran, B. Recht, P.A. Parrilo, and A.S. Willsky. The convex geometry of linear inverse problems. *Arxiv preprint arXiv:1012.0621*, 2010. 28

[43] V. Chandrasekaran, S. Sanghavi, P.A. Parrilo, and A.S. Willsky. Rank-sparsity incoherence for matrix decomposition. *Arxiv preprint arXiv:0906.2220*, 2009. 27

[44] R. Chartrand. Exact reconstruction of sparse signals via nonconvex minimization. *Signal Processing Letters, IEEE*, 14(10):707–710, 2007. 27

[45] R. Chartrand and V. Staneva. Restricted isometry properties and nonconvex compressive sensing. *Inverse Problems*, 24:035020, 2008. 27

[46] R. Chartrand and W. Yin. Iteratively reweighted algorithms for compressive sensing. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pages 3869–3872. IEEE, 2008. 27

[47] S.S. Chen, D.L. Donoho, and M.A. Saunders. Atomic decomposition by basis pursuit. *SIAM journal on scientific computing*, 20(1):33–61, 1999. 20

[48] X. Chen, J. Peng, and S. Zhang. Existence of sparse solutions to the standard quadratic programming with random matrices, 2010. Available from http://netfiles.uiuc.edu/pengj/www/papers/Co-positive_MPversion1.pdf. 64

[49] C.K. Cheng. *Maximum concurrent flow and minimum ratio cut.* Dept. of Computer Science and Engineering, University of California, San Diego, 1988.

[50] I. Daubechies, M. Fornasier, and I. Loris. Accelerated projected gradient method for linear inverse problems with sparsity constraints. *Journal of Fourier Analysis and Applications*, 14(5):764–792, 2008. 20

[51] I.S. Dhillon, Y. Guan, and B. Kulis. Kernel k-means: spectral clustering and normalized cuts. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 551–556. ACM, 2004. 56

[52] I.S. Dhillon, Y. Guan, and B. Kulis. Weighted graph cuts without eigenvectors: A multilevel approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1944–1957, 2007. 56

[53] C.H.Q. Ding, X. He, H. Zha, M. Gu, and H.D. Simon. A min-max cut algorithm for graph partitioning and data clustering. In *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, pages 107–114. IEEE, 2001. 55, 56

[54] Y. Ding, N. Krislock, J. Qian, and H. Wolkowicz. Sensor network localization, euclidean distance matrix completions, and graph realization. *Optimization and Engineering*, 11(1):45–66, 2010. 14

[55] X.V. Doan and S.A. Vavasis. Finding approximately rank-one submatrices with the nuclear norm and l1 norm. *Arxiv preprint arXiv:1011.1839*, 2010. 28

[56] W.E. Donath and A.J. Hoffman. Lower bounds for the partitioning of graphs. *IBM Journal of Research and Development*, 17(5):420–425, 1973. 56

[57] D. Donoho, V. Stodden, and Y. Tsaig. Sparselab. *Software: http://sparselab. stanford. edu*, 2005. 20

[58] D.L. Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, 52(4):1289–1306, 2006. 2

[59] D.L. Donoho. For most large underdetermined systems of linear equations the minimal l1-norm solution is also the sparsest solution. *Communications on pure and applied mathematics*, 59(6):797–829, 2006. 2, 21

[60] K. Dvijotham and M. Fazel. A nullspace analysis of the nuclear norm heuristic for rank minimization. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 3586–3589. IEEE, 2010. 26

[61] M. Fazel. *Matrix rank minimization with applications*. PhD thesis, Stanford University, 2002. 15, 16

[62] M. Fazel, H. Hindi, and S.P. Boyd. A rank minimization heuristic with application to minimum order system approximation. In *American Control Conference, 2001. Proceedings of the 2001*, volume 6, pages 4734–4739. IEEE, 2001. 15

[63] U. Feige. Randomized graph products, chromatic numbers, and the Lovász ϑ-function. *Combinatorica*, 17(1):79–90, 1997. 32

[64] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy. Approximating clique is almost np-complete. In *Foundations of Computer Science, 1991. Proceedings., 32nd Annual Symposium on*, pages 2 –12, October 1991. 30

[65] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy. Interactive proofs and the hardness of approximating cliques. *Journal of the ACM (JACM)*, 43(2):268–292, 1996. 30

[66] U. Feige and R. Krauthgamer. Finding and certifying a large hidden clique in a semirandom raph. *Random Structures Algorithms*, 16(2):195–208, 2000. 31, 32, 33, 34, 61, 132

[67] T.A. Feo, M.G.C. Resende, and S.H. Smith. A greedy randomized adaptive search procedure for maximum independent set. *Operations Research*, pages 860–878, 1994. 30

[68] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1973. 56

[69] M. Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Mathematical Journal*, 25(4):619–633, 1975. 56

[70] M.A.T. Figueiredo, R.D. Nowak, and S.J. Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *Selected Topics in Signal Processing, IEEE Journal of*, 1(4):586–597, 2007. 20

[71] L.R. Ford and D.R. Fulkerson. *Flows in networks*. Princeton University Press, Princeton, NJ, 1962. 55

[72] S. Foucart and M.J. Lai. Sparsest solutions of underdetermined linear systems via l_q-minimization for $0 < q < 1$. *Applied and Computational Harmonic Analysis*, 26(3):395–407, 2009. 27

[73] C. Friden, A. Hertz, and D. De Werra. TABARIS: an exact algorithm based on Tabu Search for finding a maximum independent set in a graph. *Computers & Operations Research*, 17(5):437–445, 1990. 30

[74] C. Friden, A. Hertz, and D.D.E. Werra. Stabulus: a technique for finding stable sets in large graphs with tabu search. *Computing*, 42(1):35–44, 1989. 30

[75] A. Frieze and R. Kannan. A new approach to the planted clique problem. In *Proc. of FST & TCS*. Citeseer, 2008. 34

[76] K. Fujisawa, Y. Futakata, M. Kojima, S. Matsuyama, S. Nakamura, K. Nakata, and M. Yamashita. Sdpa-m (semidefinite programming algorithm in matlab) users manual: version 6.2. 0. *Research Reports on Mathematical and Computing Sciences, Series B: Operation Res., Dep. Math. and Computing Sci., Tokyo Institute of Technol., Japan*. 10

[77] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981. 12, 13

[78] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979. 38

[79] S. Geman. A limit theorem for the norm of random matrices. *Ann. Probab.*, 8(2):252–261, 1980. 12, 13

[80] M. Gendreau, P. Soriano, and L. Salvail. Solving the maximum clique problem using a tabu search approach. *Annals of Operations Research*, 41(4):385–403, 1993. 30

[81] L.E. Gibbons, D.W. Hearn, and P.M. Pardalos. A continuous based heuristic for the maximum clique problem. *Clique, Graph Coloring and Satisfiability: Second DIMACS Implementation Challenge*, 1996. 30

[82] L.E. Gibbons, D.W. Hearn, P.M. Pardalos, and M.V. Ramana. Continuous characterizations of the maximum clique problem. *Mathematics of Operations Research*, pages 754–768, 1997. 30

[83] N. Gillis and F. Glineur. Nonnegative factorization and the maximum edge biclique problem. Available from http://arxiv.org/abs/0810.4225, 2008. 2, 38

[84] M.X. Goemans and D.P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995. 14, 133

[85] G.H. Golub and C.F. Van Loan. *Matrix computations*. Johns Hopkins University Press, 1996. 4, 32, 70

[86] M. Grant and S. Boyd. Cvx: Matlab software for disciplined convex programming. *Available httpstanford edu boydcvx*, 2008. 126

[87] D. Gross. Recovering low-rank matrices from few coefficients in any basis. *Arxiv preprint arXiv:0910.1879*, 2009. 25

[88] D. Gross, Y.K. Liu, S.T. Flammia, S. Becker, and J. Eisert. Quantum state tomography via compressed sensing. *Physical review letters*, 105(15):150401, 2010. 14

[89] L. Hagen and A.B. Kahng. New spectral methods for ratio cut partitioning and clustering. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 11(9):1074–1085, 1992. 56

[90] M.H. Han and D. Jang. The use of maximum curvature points for the recognition of partially occluded objects. *Pattern Recognition*, 23(1-2):21–33, 1990. 30

[91] F. Harary and I.C. Ross. A procedure for clique detection using the group matrix. *Sociometry*, 20(3):205–215, 1957. 30

[92] J.A. Hartigan. *Clustering algorithms*. John Wiley & Sons, Inc. New York, NY, USA, 1975. 56

[93] J.A. Hartigan and M.A. Wong. A k-means clustering algorithm. *Journal of the Royal Statistical Society C*, 28(1):100–108, 1979. 56

[94] J. Hasselberg, P.M. Pardalos, and G. Vairaktarakis. Test case generators and computational results for the maximum clique problem. *Journal of Global Optimization*, 3(4):463–482, 1993. 30

[95] J. Hastad. Clique is hard to approximate within n. In *Proc. 37th Annu. Symp. Found. Comput. Sci*, pages 627–636. Citeseer, 1996. 30

[96] J.B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms: Advanced theory and bundle methods*. Springer, 1993. 16

[97] D.S. Hochbaum. Approximating clique and biclique problems. *J. Algorithms*, 29(1):174–200, 1998. 38

[98] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. American Statistical Assoc.*, 58:13–30, 1962. 11, 155

[99] R. Horaud and T. Skordas. Stereo correspondence through feature grouping and maximal cliques. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(11):1168–1180, 1989. 30

[100] R.A. Horn and C.R. Johnson. *Topics in matrix analysis*. Cambridge University Press, 1994.

[101] R.A. Horn and C.R. Johnson. *Matrix analysis*. Cambridge University Press, 2005. 4, 17, 56

[102] L. Hubert and D. Steinley. Agreement among supreme court justices: Categorical vs. continuous representation. *SIAM News*, 38(7), 2005. 125, 126, 127

[103] L.D. Iasemidis, P. Pardalos, J. C. Sackellares, and D.-S. Shiau. Quadratic binary programming and dynamical system approach to determine the predictability of epileptic seizures. *Journal of Combinatorial Optimization*, 5(1):9–26, 2001. 2

[104] P. Jain, R. Meka, and I. Dhillon. Guaranteed rank minimization via singular value projection. *Advances in Neural Information Processing Systems*, 23:937–945, 2010. 24

[105] A. Jalali, Y. Chen, S. Sanghavi, and H. Xu. Clustering partially observed graphs via convex optimization. *Arxiv preprint arXiv:1104.4803*, 2011. 53

[106] M. Jerrum. Large cliques elude the metropolis process. *Random Structures & Algorithms*, 3(4):347–359, 1992. 31

[107] D.S. Johnson. Approximation algorithms for combinatorial problems. *Journal of Computer and System Sciences*, 9(3):256–278, 1974. 30

[108] D.S. Johnson, M. Yannakakis, and C.H. Papadimitriou. On generating all maximal independent sets. *Information Processing Letters*, 27(3):119–123, 1988. 30

[109] A. Juels and M. Peinado. Hiding cliques for cryptographic security. In *Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (San Francisco, CA, 1998)*, pages 678–684. Citeseer. 31

[110] F. Juhász. The asymptotic behaviour of Lovász ϑ function for random graphs. *Combinatorica*, 2(2):153–155, 1982. 32

[111] R. Kannan, S. Vempala, and A. Vetta. On clusterings: Good, bad and spectral. *Journal of the ACM (JACM)*, 51(3):497–515, 2004. 56

[112] RM Karp. Reducibility among Combinatorial Problems. *Complexity of Computer Computations*, pages 85–103, 1972. 30

[113] R.M. Karp. The probabilistic analysis of some combinatorial search algorithms, 1976. 1, 31

[114] R.H. Keshavan, A. Montanari, and S. Oh. Matrix completion from a few entries. *Information Theory, IEEE Transactions on*, 56(6):2980–2998, 2010.

[115] R.H. Keshavan, A. Montanari, and S. Oh. Matrix completion from noisy entries. *The Journal of Machine Learning Research*, 99:2057–2078, 2010. 25

[116] S.J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky. An interior-point method for large-scale l1-regularized least squares. *Selected Topics in Signal Processing, IEEE Journal of*, 1(4):606–617, 2007. 20

[117] R. Kopf and G. Ruhe. A computational study of the weighted independent set problem for general graphs. *FOUND. CONTROL ENG.*, 12(4):167–180, 1987. 30

[118] N. Krislock. *Semidefinite facial reduction for low-rank euclidean distance matrix completion*. PhD thesis, University of Waterloo, 2010.

[119] N. Krislock and H. Wolkowicz. Explicit sensor network localization using semidefinite representations and facial reductions. *Arxiv preprint arXiv:1002.0013*, 2010. 14

[120] L. Kucera. Expected complexity of graph partitioning problems. *Discrete Applied Mathematics*, 57(2-3):193–212, 1995. 31

[121] RM Larsen. Propack: A software package for the symmetric eigenvalue problem and singular value problems on lanczos and lanczos bidiagonalization with partial reorthogonalization, sccm. *URL http://soi. stanford. edu/rmunk/PROPACK*. 111

[122] K. Lee and Y. Bresler. Admira: Atomic decomposition for minimum rank approximation. *Information Theory, IEEE Transactions on*, 56(9):4402–4416, 2010. 24

[123] T. Leighton and S. Rao. An approximate max-flow min-cut theorem for uniform multicommodity flow problems with applications to approximation algorithms. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science*, pages 422–431. IEEE Computer Society, 1988. 55

[124] A.S. Lewis. The mathematics of eigenvalue optimization. *Mathematical Programming*, 97(1):155–176, 2003.

[125] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995. 1, 14

[126] Y.J. Liu, D. Sun, and K.C. Toh. An implementable proximal point algorithmic framework for nuclear norm minimization. *Mathematical Programming*, pages 1–38, 2009. 15, 112

[127] E. Loukakis and C. Tsouros. A depth first search algorithm to generate the family of maximal independent sets of a graph lexicographically. *Computing*, 27(4):349–366, 1981. 30

[128] L. Lovász. On the Shannon capacity of a graph. *Information Theory, IEEE Transactions on*, 25(1):1–7, 1979. 32

[129] L. Lovász, M. Grötschel, and A. Schrijver. *Geometric algorithms and combinatorial optimization*. Berlin: Springer-Verlag, 1988. 33, 34

[130] A. Lutomirski, S. Aaronson, E. Farhi, D. Gosset, A. Hassidim, J. Kelner, and P. Shor. Breaking and making quantum money: toward a new quantum cryptographic protocol. *Arxiv preprint arXiv:0912.3825*, 2009. 31

[131] S. Ma, D. Goldfarb, and L. Chen. Fixed point and bregman iterative methods for matrix rank minimization. *Mathematical Programming*, pages 1–33, 2009. 15

[132] D.W. Matula and F. Shahrokhi. Sparsest cuts and bottlenecks in graphs. *Discrete Applied Mathematics*, 27(1-2):113–123, 1990. 56

[133] F. McSherry. Spectral partitioning of random graphs. In *Foundations of Computer Science, 2001. Proceedings. 42nd IEEE Symposium on*, pages 529–537. IEEE. 134

[134] C.D. Meyer. *Matrix analysis and applied linear algebra*, volume 2. Society for Industrial and Applied Mathematics, 2000. 5

[135] M. Mitzenmacher and E. Upfal. *Probability and computing*. Cambridge University Press, Cambridge, 2005. Randomized algorithms and probabilistic analysis. 11

[136] K. Mohan and M. Fazel. New restricted isometry results for noisy low-rank recovery. In *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, pages 1573–1577. IEEE. 24

[137] K. Mohan and M. Fazel. Iterative reweighted least squares for matrix rank minimization. In *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, pages 653–661. IEEE, 2010. 27

[138] K. Mohan and M. Fazel. Reweighted nuclear norm minimization with application to system identification. In *American Control Conference (ACC), 2010*, pages 2953–2959. IEEE, 2010.

[139] J.W. Moon and L. Moser. On cliques in graphs. *Israel journal of Mathematics*, 3(1):23–28, 1965. 30

[140] B.K. Natarajan. Sparse approximate solutions to linear systems. *SIAM J. Comput.*, 24(2):227–234, 1995. 15, 20

[141] Y. Nesterov. Smooth minimization of non-smooth functions. *Mathematical Programming*, 103(1):127–152, 2005.

[142] Y. Nesterov and A. Nemirovskii. *Interior-point polynomial algorithms in convex programming*. Society for Industrial and Applied Mathematics, 1994. 8, 10

[143] Y.E. Nesterov and M.J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*, pages 1–42, 1997. 20

[144] J. Nocedal and S.J. Wright. *Numerical optimization*. Springer verlag, 1999. 20

[145] H. Ogawa. Labeled point pattern matching by Delaunay triangulation and maximal cliques. *Pattern Recognition*, 19(1):35–40, 1986. 30

[146] R. Ostrovsky, Y. Rabani, L. Schulman, and C. Swamy. The effectiveness of Lloyd-type methods for the k-means problem. In *Proceedings of 47st Annual IEEE Symposium on the Foundations of Computer Science*, 2006. 53

[147] S. Oymak and B. Hassibi. New null space results and recovery thresholds for matrix rank minimization. *Arxiv preprint arXiv:1011.6326*, 2010. 26

[148] S. Oymak and B. Hassibi. Finding dense clusters via "low rank + sparse" decomposition. *Arxiv preprint arXiv:1104.5186*, 2011. 27, 28, 53, 64

[149] S. Oymak, K. Mohan, M. Fazel, and B. Hassibi. A simplified approach to recovery conditions for low rank matrices. *Arxiv preprint arXiv:1103.1178*, 2011. 24, 27

[150] P.M. Pardalos and G.P. Rodgers. A branch and bound algorithm for the maximum clique problem. *Computers & operations research*, 19(5):363–375, 1992. 30

[151] R. Peeters. The maximum edge biclique problem is NP-complete. *Discrete Applied Mathematics*, 131:651–654, 2003. 38

[152] J. Peng and Y. Wei. Approximating k-means-type clustering via semidefinite programming. *SIAM Journal on Optimization*, 18(1):186–205, 2007. 53, 63

[153] F. Pla and J.A. Marchant. Matching feature points in image sequences through a region-based method. *Computer vision and image understanding*, 66(3):271–285, 1997. 30

[154] A. Pothen, H.D. Simon, K.P. Liou, et al. Partitioning sparse matrices with eigenvectors of graphs. *SIAM J. MATRIX ANAL. APPLIC.*, 11(3):430–452, 1990. 56

[155] B. Radig. Image sequence analysis using relational structures. *Pattern Recognition*, 17(1):161–167, 1984. 30

[156] A. Rahimi and B. Recht. Clustering with normalized cuts is clustering with a hyperplane. *Statistical Learning in Computer Vision*, 2004. 56

[157] B. Recht. A simpler approach to matrix completion. *Arxiv preprint arXiv:0910.0651*, 2009. 25, 26

[158] B. Recht, M. Fazel, and P.A. Parrilo. Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization. *SIAM Review*, 52(471), 2010. 2, 16, 17, 18, 19, 22, 23, 35

[159] B. Recht, W. Xu, and B. Hassibi. Necessary and sufficient conditions for success of the nuclear norm heuristic for rank minimization. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 3065–3070. IEEE, 2008. 26

[160] B. Recht, W. Xu, and B. Hassibi. Null space conditions and thresholds for rank minimization. *Mathematical Programming*, pages 1–28, 2010. 26

[161] J.D.M. Rennie and N. Srebro. Fast maximum margin matrix factorization for collaborative prediction. In *Proceedings of the 22nd international conference on Machine learning*, pages 713–719. ACM, 2005. 14

[162] J.M. Robson. Algorithms for maximum independent sets. *Journal of Algorithms*, 7(3):425–440, 1986. 30

[163] R.T. Rockafellar. *Convex analysis*, volume 28. Princeton Univ Pr, 1997. 6, 8

[164] F. Santosa and W.W. Symes. Linear inversion of band-limited reflection seismograms. *SIAM Journal on Scientific and Statistical Computing*, 7:1307, 1986. 20

[165] S. Sarkar and K.L. Boyer. Quantitative measures of change based on feature organization: Eigenvalues and eigenvectors. *Computer vision and image understanding*, 71(1):110–136, 1998. 55, 56

[166] S.E. Schaeffer. Graph clustering. *Computer Science Review*, 1(1):27–64, 2007. 27

[167] J. Shi and J. Malik. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8):888–905, 2000. 55, 56, 130

[168] R. Shioda and L. Tunçel. Clustering via minimum volume ellipsoids. *Computational Optimization and Applications*, 37(3):247–295, 2007. 53

[169] N.Z. Shor. Cut-off method with space extension in convex programming problems. *Cybernetics and systems analysis*, 13(1):94–96, 1977. 8

[170] A.M.C. So and Y. Ye. Theory of semidefinite programming for sensor network localization. *Mathematical Programming*, 109(2):367–384, 2007. 1, 14

[171] P. Soriano and M. Gendreau. Diversification strategies in tabu search algorithms for the maximum clique problem. *Annals of Operations Research*, 63(2):189–207, 1996. 30

[172] M. Stojnic, W. Xu, and B. Hassibi. Compressed sensing-probabilistic analysis of a null-space characterization. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pages 3377–3380. IEEE, 2008. 26

[173] J. Sturm, P. Imre, and T. Terlaky. Sedumi 1.3, 2006. *URL http://sedumi. ie. lehigh. edu*. 10, 126

[174] J.M. Swaminathan and S.R. Tayur. Managing broader product lines through delayed differentiation using vanilla boxes. *Management Science*, pages 161–172, 1998. 38

[175] R.E. Tarjan and A.E. Trojanowski. Finding a maximum independent set. *SIAM J. Comput.*, 6(3):537–546, 1977. 30

[176] M.J. Todd. Semidefinite optimization. *Acta Numerica 2001*, 10(-1):515–560, 2001. 10

[177] K.C. Toh, M.J. Todd, and R.H. Tütüncü. Sdpt3 - a matlab software package for semidefinite programming. *Optimization Methods and Software*, 11(12):545–581, 1999. 10

[178] K.C. Toh and S. Yun. An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems. *Pacific Journal of Optimization*, 6:615–640, 2010. 15

[179] E. Tomita and T. Seki. An efficient branch-and-bound algorithm for finding a maximum clique. *Discrete Mathematics and Theoretical Computer Science*, pages 220–220, 2003. 30

[180] E. Tomita, A. Tanaka, and H. Takahashi. The worst-case time complexity for generating all maximal cliques and computational experiments. *Theoretical Computer Science*, 363(1):28–42, 2006. 30

[181] L. Tunçel. *Polyhedral and semidefinite programming methods in combinatorial optimization*. American Mathematical Society, 2010. 10, 67

[182] R.H. Tütüncü, K.C. Toh, and M.J. Todd. Solving semidefinite-quadratic-linear programs using sdpt3. *Mathematical Programming*, 95(2):189–217, 2003. 10

[183] U.S. Census Bureau, International Data Base. 2010 crude birth & death, net migration, & growth rates, births, deaths, & migrants, December 2010. Available from http://www.census.gov/ipc/www/idb/informationGateway.php. 128

[184] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM review*, 38(1):49–95, 1996. 10

[185] S. Wang and J.M. Siskind. Image segmentation with ratio cut. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 675–690, 2003.

[186] G. A. Watson. Characterization of the subdifferential of some matrix norms. *Linear Algebra Appl.*, 170:33–45, 1992. 19

[187] Y.C. Wei and C.K. Cheng. Towards efficient hierarchical designs by ratio cut partitioning. In *Computer-Aided Design, 1989. ICCAD-89. Digest of Technical Papers., 1989 IEEE International Conference on*, pages 298–301. IEEE, 1989. 55

[188] Y.C. Wei and C.K. Cheng. Ratio cut partitioning for hierarchical designs. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 10(7):911–921, 1991. 55

[189] Z. Wu and R. Leahy. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, pages 1101–1113, 1993. 54

[190] M. Yannakakis. Node-and edge-deletion NP-complete problems. In *Proceedings of the tenth annual ACM symposium on Theory of computing*, pages 253–264. ACM, 1978. 38

[191] W. Yin, S. Osher, D. Goldfarb, and J. Darbon. Bregman iterative algorithms for l1-minimization with applications to compressed sensing. *SIAM Journal on Imaging Sciences*, 1(1):143–168, 2008. 20

[192] D.B. Yudin and A.S. Nemirovski. Informational complexity and efficient methods for solving complex extremal problems. *Matekon*, 13:25–45, 1977. 8

[193] Y. Zhang. On theory of compressive sensing via l1-minimization: Simple derivations and extensions. *TR08-11, CAAM, Rice University*. 27

[194] X.Y. Zhao, D. Sun, and K.C. Toh. A newton-cg augmented lagrangian method for semidefinite programming. *SIAM J. Optim*, 20(4):1737–1765, 2010. 10, 112

[195] Y. Zhao and G. Karypis. Criterion functions for document clustering: Experiments and analysis. 2002. 56

[196] Y. Zhao and G. Karypis. Empirical and theoretical comparisons of selected criterion functions for document clustering. *Machine Learning*, 55(3):311–331, 2004. 56

[197] Y. Zhou, D. Wilkinson, R. Schreiber, and R. Pan. Large-scale parallel collaborative filtering for the netflix prize. *Algorithmic Aspects in Information and Management*, pages 337–348, 2008. 14

Appendix A

Proofs of probabilistic bounds

A.1 Proof of Theorem 4.4.2

From the definition of $\tilde{A}$, for column j, there are exactly $n-n_j$ entries of $\tilde{A}$ that differ from those of A. Furthermore, the difference of these entries is exactly $(n_j-pn)/((1-p)(n-n_j))$. Therefore, for each $j = 1, \ldots, N$, the contribution of column j to the square norm difference $\|A - \tilde{A}\|_F^2$ is given by

$$\|A(:,j) - \tilde{A}(:,j)\|_F^2 = \frac{(n_j - pn)^2}{(1-p)^2(n-n_j)}.$$

Recall that the random variables $n_1, \ldots, n_N$ are independent, and each is the result of n Bernoulli trials done with probability p.

We now define Ψ to be the event that at least one n_j is very far from the mean. In particular, Ψ is the event that there exists a $j \in \{1, \ldots, N\}$ such that $n_j > qn$, where $q = \min(\sqrt{p}, 2p)$. Let $\tilde{\Psi}$ be its complement, and let $\tilde{\psi}(n_j)$ be the indicator of this complement (i.e., $\tilde{\psi}(n_j) = 1$ if $n_j \leq qn$ else $\tilde{\psi}(n_j) = 0$). Let c be a positive scalar depending on p to be determined later. Observe that

$$\begin{aligned} P(\|A - \tilde{A}\|_F^2 \geq cN) &= P(\|A - \tilde{A}\|_F^2 \geq cN \ \wedge \ \tilde{\Psi}) + P(\|A - \tilde{A}\|_F^2 \geq cN \ \wedge \ \Psi) \\ &\leq P(\|A - \tilde{A}\|_F^2 \geq cN \ \wedge \ \tilde{\Psi}) + P(\Psi). \end{aligned} \tag{A.1.1}$$

We now analyze the two terms separately. For the first term, we use the same technique as in the proofs of Theorems 5.3.5 and 5.4.3. Let ϕ be the indicator function of nonnegative

reals. Then

$$P(\|A-\tilde{A}\|_F^2 \geq cN \,\&\wedge\, \tilde{\Psi}) = P(\|A-\tilde{A}\|_F^2 - cN \geq 0 \,\wedge\, \tilde{\psi}(n_1)=1 \,\wedge\, \cdots \,\wedge\, \tilde{\psi}(n_N)=1)$$
$$= E(\phi(\|A-\tilde{A}\|_F^2 - cN)\cdot\tilde{\psi}(n_1)\cdots\tilde{\psi}(n_N)).$$

Let h be a positive scalar depending on p to be determined later. Observe that for any such h and for all $x \in \mathbf{R}$, $\phi(x) \leq \exp(hx)$. Thus,

$$\begin{aligned}
P(\|A-\tilde{A}\|_F^2 \geq cN \,\wedge\, \tilde{\Psi}) &\leq E(\exp(h\|A-\tilde{A}\|_F^2 - hcN)\cdot\tilde{\psi}(n_1)\cdots\tilde{\psi}(n_N)) \\
&= E\left(\exp\left(h\sum_{j=1}^{N}\left(\|A(:,j)-\tilde{A}(:,j)\|_F^2 - c\right)\right)\cdot\tilde{\psi}(n_1)\cdots\tilde{\psi}(n_N)\right) \\
&= E\left(\exp\left(h\sum_{j=1}^{N}\left(\frac{(n_j-pn)^2}{(1-p)^2(n-n_j)} - c\right)\right)\cdot\tilde{\psi}(n_1)\cdots\tilde{\psi}(n_N)\right) \\
&= E\left(\prod_{j=1}^{N}\exp\left(h\left(\frac{(n_j-pn)^2}{(1-p)^2(n-n_j)} - c\right)\right)\tilde{\psi}(n_j)\right) \\
&= \prod_{j=1}^{N}E\left(\exp\left(h\left(\frac{(n_j-pn)^2}{(1-p)^2(n-n_j)} - c\right)\right)\tilde{\psi}(n_j)\right) \qquad \text{(A.1.2)} \\
&= f_1\cdots f_N, \qquad \text{(A.1.3)}
\end{aligned}$$

where

$$f_j = E\left(\exp\left(h\left(\frac{(n_j-pn)^2}{(1-p)^2(n-n_j)} - c\right)\right)\tilde{\psi}(n_j)\right).$$

To obtain (A.1.2), we used the independence of the n_j's. Let us now analyze f_j in isolation.

$$\begin{aligned}
f_j &= \sum_{i=0}^{n}\exp\left(h\left(\frac{(i-pn)^2}{(1-p)^2(n-i)} - c\right)\right)\tilde{\psi}(n_j)P(n_j=i) \\
&= \sum_{i=0}^{\lfloor qn\rfloor}\exp\left(h\left(\frac{(i-pn)^2}{(1-p)^2(n-i)} - c\right)\right)P(n_j=i) \\
&\leq \sum_{i=0}^{\lfloor qn\rfloor}\exp\left(h\left(\frac{(i-pn)^2}{(1-p)^2(n-\sqrt{p}n)} - c\right)\right)P(n_j=i).
\end{aligned}$$

To derive the last line, we used the fact that $i \leq \sqrt{p}n$ since $i \leq qn$. Now let us reorganize

this summation by considering first i such that $|i-pn| < \sqrt{n}$, and next i such that $|i-pn| \in [\sqrt{n}, 2\sqrt{n})$, etc. Notice that, since $i \le qn \le 2pn$, we need consider intervals only until $|i-pn|$ reaches pn. Then

$$\begin{aligned}
f_j &\le \sum_{k=0}^{\lfloor p\sqrt{n}\rfloor} \sum_{i:|i-pn|\in[k\sqrt{n},(k+1)\sqrt{n})} \exp\left(h\left(\frac{(i-pn)^2}{(1-p)^2(n-\sqrt{p}n)} - c\right)\right) P(n_j = i) \\
&\le \sum_{k=0}^{\lfloor p\sqrt{n}\rfloor} \sum_{i:|i-pn|\in[k\sqrt{n},(k+1)\sqrt{n})} \exp\left(h\left(\frac{(k+1)^2 n}{(1-p)^2(n-\sqrt{p}n)} - c\right)\right) P(n_j = i) \\
&= \sum_{k=0}^{\lfloor p\sqrt{n}\rfloor} \sum_{i:|i-pn|\in[k\sqrt{n},(k+1)\sqrt{n})} \exp\left(h\left(\frac{(k+1)^2}{(1-p)^2(1-\sqrt{p})} - c\right)\right) P(n_j = i) \\
&= \sum_{k=0}^{\lfloor p\sqrt{n}\rfloor} \exp\left(h\left(\frac{(k+1)^2}{(1-p)^2(1-\sqrt{p})} - c\right)\right) \sum_{i:|i-pn|\in[k\sqrt{n},(k+1)\sqrt{n})} P(n_j = i) \\
&\le 2\sum_{k=0}^{\lfloor p\sqrt{n}\rfloor} \exp\left(h\left(\frac{(k+1)^2}{(1-p)^2(1-\sqrt{p})} - c\right)\right) \exp(-k^2/p),
\end{aligned}$$

where, for the last line, we have applied (2.6.2). The theorem is valid since $k \le p\sqrt{n}$.

Continuing this derivation and overestimating the finite sum with an infinite sum,

$$\begin{aligned}
f_j &\le 2\exp(-hc) \cdot \sum_{k=0}^{\infty} \exp\left(\frac{h(k+1)^2}{(1-p)^2(1-\sqrt{p})} - k^2/p\right) \\
&= 2\exp\left(\frac{h}{(1-p)^2(1-\sqrt{p})} - hc\right) \\
&\qquad + 2\exp(-hc) \cdot \sum_{k=1}^{\infty} \exp\left[\frac{h(k+1)^2}{(1-p)^2(1-\sqrt{p})} - k^2/p\right].
\end{aligned}$$

Choose h so that $h/((1-p)^2)(1-\sqrt{p}) < 1/(8p)$, i.e., $h < (1-p)^2(1-\sqrt{p})/(8p)$. Then the second term in the square-bracket expression at least twice the first term for all $k \ge 1$, hence

$$f_j \le 2\exp\left(\frac{h}{(1-p)^2(1-\sqrt{p})} - hc\right) + 2\exp(-hc) \cdot \sum_{k=1}^{\infty} \exp\left(-k^2/(2p)\right). \qquad \text{(A.1.4)}$$

Observe that $\sum_{k=1}^{\infty} \exp(-k^2/(2p))$ is dominated by a geometric series and hence is a finite number depending on p. Thus, once h is selected, it is possible to choose c sufficiently large so that each of the two terms in (A.1.4) is at most $1/3$. Thus, with appropriate choices of h and c, we conclude that $f_j \leq 2/3$. Thus, substituting this into (A.1.3) shows that

$$P(\|A - \tilde{A}\|_F^2 \geq cN \ \wedge \ \tilde{\Psi}) \leq (2/3)^N. \tag{A.1.5}$$

We now turn to the second term in (A.1.1). For a particular j, the probability that $n_j > qn$ is bounded using (2.6.1) by v_p^n where $v_p = (e^\delta/(1+\delta)^{(1+\delta)})^p$, where $\delta = q/p - 1$, i.e., $\delta = \min(p, \sqrt{p} - p)$. Then the union bound asserts that the probability that any j satisfies $n_j > qn$ is at most Nv_p^n. Thus,

$$P(\|A - \tilde{A}\|_F^2 \geq cN) \leq (2/3)^N + Nv_p^n.$$

This concludes the proof.

A.2 Proof of Theorem 5.3.5

The random numbers $\{n_i : i \in I_1\}$ are independent, and each is the result of r_2 Bernoulli trials, each with probability of success equal to p. We define Ψ to be the event that at least one n_i is very far from its expected value. That is, Ψ is the event that there exists $i \in I_1$ such that $n_i > tr_2$, where $t = \min\{\sqrt{p}, 2p\}$. Moreover, we define $\tilde{\Psi}$ to be its complement, and let $\tilde{\psi}(n_i)$ be the indicator function such that

$$\tilde{\psi}(n_i) = \begin{cases} 1, & \text{if } n_i \leq tr_2 \\ 0, & \text{otherwise.} \end{cases}$$

Let B be a positive scalar depending on p to be determined later. Then

$$P\left(\sum_{i \in I_1} \frac{|n_i - pr_2|^\alpha}{r_2 - n_i} \geq B \frac{r_1}{r_2^{1-\alpha/2}}\right) \leq P\left(\sum_{i \in I_1} \frac{|n_i - pr_2|^\alpha}{r_2 - n_i} \geq B \frac{r_1}{r_2^{1-\alpha/2}} \wedge \tilde{\Psi}\right) + P(\Psi). \tag{A.2.1}$$

We will analyze the two terms separately. For the first term we use a technique of Bernstein (see [98]). Let ϕ be the indicator function of the nonnegative reals. Then,

$$\begin{aligned}
&P\left(\sum_{j\in I_1}\frac{|n_j-pr_2|^\alpha}{r_2-n_j}\geq B\frac{r_1}{r_2^{1-\alpha/2}}\wedge\tilde{\Psi}\right)\\
&\quad=P\left(\sum_{j\in I_1}\frac{|n_j-pr_2|^\alpha}{r_2-n_j}-B\frac{r_1}{r_2^{1-\alpha/2}}\geq 0\wedge\tilde{\psi}(n_j)=1\ \ \forall\, j\in I_1\right)\\
&\quad=P\left(\sum_{j\in I_1}\frac{r_2^{1-\alpha/2}|n_j-pr_2|^\alpha}{r_2-n_j}-Br_1\geq 0\wedge\tilde{\psi}(n_j)=1\ \ \forall\, j\in I_1\right)\\
&\quad=E\left(\phi\left(\sum_{j\in I_1}\frac{r_2^{1-\alpha/2}|n_j-pr_2|^\alpha}{r_2-n_j}-Br_1\right)\cdot\prod_{j\in I_1}\tilde{\psi}(n_j)\right).
\end{aligned}$$

Let h be a positive scalar depending on p to be determined later. Notice that for any $h>0$ and all $\mathbf{x}\in\mathbf{R}$, $\phi(\mathbf{x})\leq\exp(h\mathbf{x})$. Thus, by the independence of the n_j's,

$$\begin{aligned}
&P\left(\sum_{i\in I_1}\frac{|n_i-pr_2|^\alpha}{r_2-n_i}\geq B\frac{r_1}{r_2^{1-\alpha/2}}\wedge\tilde{\Psi}\right)\\
&\quad\leq E\left(\exp\left(h\sum_{j\in I_1}\left(\frac{|n_j-pr_2|^\alpha}{r_2^{\alpha/2-1}(r_2-n_j)}-B\right)\right)\cdot\prod_{j\in I_1}\tilde{\psi}(n_j)\right)\\
&\quad=\prod_{j\in I_1}E\left(\exp\left(h\left(\frac{|n_j-pr_2|^\alpha}{r_2^{\alpha/2-1}(r_2-n_j)}-B\right)\right)\tilde{\psi}(n_j)\right)\\
&\quad=\prod_{j\in I_1}f_j,
\end{aligned}$$

where

$$f_j=E\left[\exp\left(h\left(\frac{|n_j-pr_2|^\alpha}{r_2^{\alpha/2-1}(r_2-n_j)}-B\right)\right)\tilde{\psi}(n_j)\right].$$

We now analyze each f_j individually. Fix $j\in I_1$. Then

$$f_j=\sum_{i=0}^{\lfloor tr_2\rfloor}\exp\left(h\left(\frac{|n_j-pr_2|^\alpha}{r_2^{\alpha/2-1}(r_2-n_j)}-B\right)\right)P(n_j=i)$$

$$\leq \sum_{i=0}^{\lfloor tr_2 \rfloor} \exp\left(h\left(\frac{|i-pr_2|^\alpha}{(1-\sqrt{p}){r_2}^{\alpha/2}} - B\right)\right) P(n_j = i)$$

since $i \leq tr_2$ and, hence, $i \leq \sqrt{p}pr_2$. We now reorganize this summation by considering i such that $|i - pr_2| < \sqrt{r_2}$, then i such that $\sqrt{r_2} \leq |i - pr_2| < 2\sqrt{r_2}$, and so on. Notice, since $i \leq tr_2 \leq 2pr_2$, we need only to consider intervals until $|i - pr_2|$ reaches pr_2. Hence,

$$\begin{aligned} f_j &\leq \sum_{k=0}^{\lfloor p\sqrt{r_2} \rfloor} \sum_{i:|i-pr_2| \in [k\sqrt{r_2},(k+1)\sqrt{r_2})} \exp\left(h\left(\frac{|i-pr_2|^\alpha}{(1-\sqrt{p})r_2^{\alpha/2}} - B\right)\right) P(n_j = i) \\ &\leq \sum_{k=0}^{\lfloor p\sqrt{r_2} \rfloor} \sum_{i:|i-pr_2| \in [k\sqrt{r_2},(k+1)\sqrt{r_2})} \exp\left(h\left(\frac{(k+1)^\alpha}{1-\sqrt{p}} - B\right)\right) P(n_j = i) \\ &\leq 2 \sum_{k=0}^{\lfloor p\sqrt{r_2} \rfloor} \exp\left(h\left(\frac{(k+1)^\alpha}{1-\sqrt{p}} - B\right)\right) \exp(-k^2/p) \end{aligned}$$

by (2.6.2). Overestimating the finite sum with an infinite sum, we have

$$f_j \leq 2\exp(-hB) \cdot \sum_{k=0}^{\infty} \exp\left(\frac{h(k+1)^\alpha}{1-\sqrt{p}} - k^2/p\right).$$

Choosing h such that $h \leq (1-\sqrt{p})/(8p)$ ensures that

$$\frac{h(k+1)^\alpha}{1-\sqrt{p}} - k^2/p \leq -k^2/(2p)$$

for all r_1, r_2 and $k \geq 1$. Hence, splitting off the $k = 0$ term, we have

$$f_j \leq 2\exp\left(\frac{h}{1-\sqrt{p}} - hB\right) + 2\exp(-hB) \cdot \sum_{k=1}^{\infty} \exp(-k^2/(2p)). \tag{A.2.2}$$

Since $\sum_{k=1}^{\infty} \exp(-k^2/(2p))$ is dominated by a geometric series, the summation in (A.2.2) is a finite number depending on p. Therefore, once h is chosen, it is possible to choose B, depending only on p and h, such that each of the two terms in (A.2.2) is at most $1/3$.

Therefore, we can choose h and B so that $f_j \leq 2/3$ for all $j \in I_1$. It follows that

$$P\left(\sum_{i\in I_1} \frac{|n_i - p_2|^\alpha}{r_2 - n_i} \geq B\frac{r_1}{r_2^{1-\alpha/2}} \wedge \tilde{\Psi}\right) \leq (2/3)^{r_1} \tag{A.2.3}$$

To obtain a bound on the second term in (A.2.1), notice that the probability that $n_j > tr_2$ is at most $v_p^{r_2}$ where $v_p = (e^\delta/(1+\delta)^{(1+\delta)})^p$ by Theorem 2.6.1, where $\delta = t/p - 1 = \min\{p, \sqrt{p} - p\}$. Applying the union bound shows

$$P\left(\sum_{i\in I_1}\sum_{i\in I_1} \frac{|n_i - p_2|^\alpha}{r_2 - n_i} \geq B\frac{r_1}{r_2^{1-\alpha/2}}\right) \leq (2/3)^{r_1} + r_1 v_p^{\hat{r_2}}$$

as required.

A.3 Proof of Theorem 5.4.3

Denote by n_i the summation $\sum_{j\in I_2} X_{ij}$ for each $i \in I_1$. We define Ψ to be the event that at least one n_i differs significantly from its mean; that is, we define Ψ to be the event that there exists $i \in I_1$ such that $n_i > tr_2$ where $t := 1 + \bar{x}$. We let $\tilde{\Psi}$ denote the complement of Ψ, and let $\tilde{\psi}(n_i)$ be the indicator of this complement. That is, $\tilde{\psi}(n_i) = 1$ if $n_i \leq tr_2$ and is equal to 0 otherwise. We let B be a positive scalar to be determined later. Note that

$$P\left(\sum_{i\in I_1} \frac{|n_i - \bar{x}r_2|^\delta}{r_2} \geq B\frac{r_1}{r_2^{1-\delta/2}}\right) \leq P\left(\sum_{i\in I_1} \frac{|n_i - \bar{x}r_2|^\delta}{r_2} \geq B\frac{r_1}{r_2^{1-\delta/2}} \wedge \tilde{\Psi}\right) + P(\Psi). \tag{A.3.1}$$

We analyze the two terms separately. To obtain an upper bound on the first term we use the technique of Bernstein used in the proof of Theorem 5.3.5. Let ϕ be the indicator function of the nonnegative reals. Then

$$\begin{aligned} &P\left(\sum_{i\in I_1} \frac{|n_i - \bar{x}r_2|^\delta}{r_2} \geq B\frac{r_1}{r_2^{1-\delta/2}} \wedge \tilde{\Psi}\right) \\ &\qquad = P\left(\sum_{i\in I_1} \frac{|n_i - \bar{x}r_2|^\delta}{r_2} - B\frac{r_1}{r_2^{1-\delta/2}} \geq 0 \wedge \tilde{\psi}(n_i)\ \forall i \in I_1\right) \end{aligned}$$

$$= P\left(\sum_{i\in I_1} \frac{|n_i - \bar{x}r_2|^\delta}{r_2^{\delta/2}} - Br_1 \geq 0 \wedge \tilde{\psi}(n_i)\ \forall i \in I_1\right)$$
$$= E\left(\phi\left(\sum_{i\in I_1} \frac{|n_i - \bar{x}r_2|^\delta}{r_2^{\delta/2}} - Br_1\right) \cdot \prod_{i\in I_1} \tilde{\psi}(n_i)\right).$$

Let h be a positive scalar satisfying $h \leq 1/4$. For any $h > 0$ and all $x \in \mathbf{R}$, $\phi(x) \leq \exp(hx)$. By the independence of the n_i's, we have

$$\begin{aligned}
&P\left(\sum_{i\in I_1} \frac{|n_i - \bar{x}r_2|^\delta}{r_2} \geq B\frac{r_1}{r_2^{1-\delta/2}} \wedge \tilde{\Psi}\right) \\
&\quad = E\left(\phi\left(\sum_{i\in I_1} \frac{|n_i - \bar{x}r_2|^\delta}{r_2^{\delta/2}} - Br_1\right) \prod_{i\in I_1} \tilde{\psi}(n_i)\right) \\
&\quad \leq E\left(\exp\left(h\left(\sum_{i\in I_1} \frac{|n_i - \bar{x}r_2|^\delta}{r_2^{\delta/2}} - Br_1\right)\right) \cdot \prod_{i\in I_1} \tilde{\psi}(n_i)\right) \\
&\quad = \prod_{i\in I_1} E\left(\exp\left(h\left(\frac{|n_i - \bar{x}r_2|^\delta}{r_2^{\delta/2}} - B\right)\right)\tilde{\psi}(n_i)\right) \\
&\quad = \prod_{i\in I_1} f_i,
\end{aligned}$$

where

$$\begin{aligned}
f_i &:= E\left(\exp\left(h\left(\frac{|n_i - \bar{x}r_2|^\delta}{r_2^{\delta/2}} - B\right)\right)\tilde{\psi}(n_i)\right) \\
&= \int_0^{trs} \exp\left(h\left(\frac{|x - \bar{x}r_2|^\delta}{r_2^{\delta/2}} - B\right)\right) g(x)dx \qquad \text{(A.3.2)}
\end{aligned}$$

where g is the probability density function of n_i for all $i \in I_1$.

We analyze each f_i individually. Fix $i \in I_1$. For each integer, $k = 0, 1, \ldots, \lfloor tr_2 \rfloor$, and x such that

$$k\sqrt{r_2} \leq |x - \bar{x}r_2| \leq (k+1)\sqrt{r_2},$$

we overestimate the integrand in (A.3.2) as

$$\exp\left(h\left(\frac{|x-\bar{x}r_2|^\delta}{r_2^{\delta/2}}-B\right)\right)P(|x-\bar{x}r_2|\geq k\sqrt{r_2})\leq 2\exp(h(k+1)^\delta-hB)\exp(-2k^2) \tag{A.3.3}$$

by (2.6.4). It follows that

$$f_i\leq\sum_{k=0}^{\lfloor t\sqrt{r_2}\rfloor}2\exp(h(i+1)^\delta-hB-2k^2). \tag{A.3.4}$$

Overestimating the finite sum in (A.3.4) as an infinite sum and observing that the choice of $h\leq 1/4$ ensures that $h(k+1)^\delta-2k^2\leq -k^2$ yields

$$f_i\leq 2\exp(h(1-B))+2\exp(-hB)\sum_{k=1}^{\infty}\exp(-k^2). \tag{A.3.5}$$

Note that $\sum_{k=1}^{\infty}\exp(-k^2)$ is dominated by a geometric series. Therefore, for fixed $h\leq 1/4$, it is possible to choose B, depending only on h, so that each term in (A.3.5) is at most $1/3$ and, hence $f_i\leq 2/3$ for all $i\in I_1$. It follows immediately that

$$P\left(\sum_{i\in I_1}\frac{|n_i-\bar{x}r_2|^\delta}{r_2}\geq B\frac{r_1}{r_2^{1-\delta/2}}\wedge\tilde{\Psi}\right)\leq(2/3)^{r_1}. \tag{A.3.6}$$

We next obtain a bound on the second term in (A.3.1). For any $i\in I_1$, the probability that $|n_i-\bar{x}r_2|>r_2$ is at most $2\exp(-2r_2)$ by (2.6.3). Applying the union bound shows that the probability that any $i\in I_1$ satisfies $|n_i-\bar{x}r_2|>r_2$ is at most

$$P(\Psi)\leq 2r_1\exp(-2r_2). \tag{A.3.7}$$

Substituting (A.3.6) and (A.3.7) in (A.3.1) completes the proof.