Improved Approximation Algorithms for Low-Rank Problems Using Semidefinite Optimization

Ryan Cory-Wright

Department of Analytics, Marketing and Operations, Imperial College Business School, London, UK
ORCID: 0000-0002-4485-0619
r.cory-wright@imperial.ac.uk

Jean Pauphilet

Management Science and Operations, London Business School, London, UK
ORCID: 0000-0001-6352-0984
jpauphilet@london.edu

Inspired by the impact of the Goemans-Williamson algorithm on combinatorial optimization, we construct an analogous relax-then-round strategy for low-rank optimization problems. First, for orthogonally constrained quadratic optimization problems, we derive a semidefinite relaxation and a randomized rounding scheme that obtains provably near-optimal solutions, building on the blueprint from Goemans and Williamson for the Max-Cut problem. For a given $n \times m$ semi-orthogonal matrix, we derive a purely multiplicative approximation ratio for our algorithm, and show that it is never worse than $\max(2/(\pi m), 1/(\pi(\log(2m)+1)))$. We also show how to compute a tighter constant for a finite (n, m) by solving a univariate optimization problem. We then extend our approach to generic low-rank optimization problems by developing new semidefinite relaxations that are both tighter and more broadly applicable than those in prior works. Although our original proposal introduces large semidefinite matrices as decision variables, we show that most of the blocks in these matrices can be safely omitted without altering the optimal value, hence improving the scalability of our approach. Using several examples (including matrix completion, basis pursuit, and reduced-rank regression), we show how to reduce the size of our relaxation even further. Finally, we numerically illustrate the effectiveness and scalability of our relaxation and sampling scheme on orthogonally constrained quadratic optimization and matrix completion problems.

Key words: Low-rank; semidefinite relaxation; randomized rounding; approximation algorithm

1. Introduction

Many important optimization problems feature semi-orthogonal matrices, i.e., matrices $\boldsymbol{U} \in \mathbb{R}^{n \times m}$ such that $\boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_m$. Orthogonality constraints force the columns of $\boldsymbol{U}$ to be orthogonal and unit length, and are central to quadratic assignment (Gilman et al. 2022), quantum nonlocality (Briët et al. 2011), control theory (Ben-Tal and Nemirovski 2002), and sparse PCA (Cory-Wright and Pauphilet 2022) problems. The set of semi-orthogonal matrices is often called the Stiefel manifold (Gilman et al. 2022, Burer and Park 2024). Orthogonality constraints are also related to the rank of a matrix, which models a matrix's complexity in imputation (Bell and Koren 2007), factor analysis (Bertsimas et al. 2017), and multi-task regression (Negahban and Wainwright 2011) settings.

For any semi-orthogonal matrix $\boldsymbol{U} \in \mathbb{R}^{n \times m} : \boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_m$, the matrix $\boldsymbol{Y} := \boldsymbol{U} \boldsymbol{U}^\top$ is an orthogonal projection matrix of rank m, i.e., it satisfies $\boldsymbol{Y}^2 = \boldsymbol{Y}$. Moreover, for any symmetric orthogonal matrix $\boldsymbol{U} \in \mathcal{S}^n$, the matrix $\boldsymbol{Y} = \frac{1}{2}\boldsymbol{U} + \frac{1}{2}\boldsymbol{I}_n$ is a projection matrix. Building on the algebraic similarities between binary variables and projection matrices (which solve the polynomial equations $z^2 = z$ and $\boldsymbol{Y}^2 = \boldsymbol{Y}$), efficient approaches for mixed-integer optimization have been extended to rank-constrained optimization problems, including outer approximation (Bertsimas et al. 2022), perspective relaxations (Bertsimas et al. 2023c), and branch-and-bound (Bertsimas et al. 2023b).

In mixed-integer optimization, a major advance in the design of approximation algorithms occurred with the relax-and-round algorithm of Goemans and Williamson (1995). For Max-Cut problems, Goemans and Williamson (1995) propose a randomized rounding algorithm that achieves a constant factor approximation guarantee of 0.87856. The theoretical and computational success of Goemans and Williamson (1995)'s algorithm has had implications far beyond Max-Cut. Their algorithm provides a $2/\pi$-approximation for general binary quadratic optimization (BQO) problems (Nesterov 1998), and can be extended to linearly-constrained BQO problems (Bertsimas and Ye 1998). More recently, Dong et al. (2015) developed a sampling scheme *à la* Goemans and Williamson for a broad class of mixed-integer optimization problems with logical constraints. Conceptually, the Goemans-Williamson algorithm propelled semidefinite optimization and correlated rounding at the core of approximation algorithms for combinatorial optimization (see Wolkowicz et al. 1998, Williamson and Shmoys 2011).

The objective of this paper is to extend the core ideas underpinning the Goemans-Williamson algorithm to quadratic semi-orthogonal and rank-constrained optimization problems by leveraging the connection between binary and low-rank optimization as explored in (Bertsimas et al. 2022).

1.1. Binary Quadratic Optimization and the Goemans-Williamson Algorithm

Binary quadratic optimization (BQO) is a canonical optimization problem with numerous applications throughout machine learning, statistics, and quantum computing (see Luo et al. 2010, for a review). As we discuss in detail in Section EC.1.1, it also serves as an important building block for logically constrained optimization problems with quadratic objectives. Formally, given a matrix $\boldsymbol{Q} \succeq \boldsymbol{0}$, BQO selects a vector $\boldsymbol{z}$ in $\{-1,1\}^n$ that solves

$$\max_{\boldsymbol{z} \in \{-1,1\}^n} \sum_{i,j} Q_{i,j} z_i z_j = \max_{\boldsymbol{z} \in \{-1,1\}^n} \langle \boldsymbol{Q}, \boldsymbol{z}\boldsymbol{z}^\top \rangle, \tag{1}$$

where $\langle \cdot, \cdot \rangle$ denotes the Frobenius inner product between matrices. Problem (1) is NP-hard and often challenging to solve to certifiable optimality when $n \geq 100$ (Rehfeldt et al. 2023). Accordingly,

a popular approach for obtaining near-optimal solutions is to sample from a distribution parameterized by the solution of (1)'s convex relaxation. Specifically, we introduce a rank-one matrix $\boldsymbol{Z}$ to model the product $\boldsymbol{z}\boldsymbol{z}^\top$. Then, (1) is equivalent to

$$\max_{\boldsymbol{Z} \in \mathcal{S}_+^n} \quad \langle \boldsymbol{Q}, \boldsymbol{Z} \rangle \text{ s.t. } \mathrm{diag}(\boldsymbol{Z}) = \boldsymbol{e}, \, \mathrm{rank}(\boldsymbol{Z}) = 1.$$

We obtain a valid semidefinite relaxation of (1) by relaxing the rank constraint, as in Shor (1987):

$$\max_{\boldsymbol{Z} \in S_+^n} \quad \langle \boldsymbol{Q}, \boldsymbol{Z} \rangle \text{ s.t. } \mathrm{diag}(\boldsymbol{Z}) = \boldsymbol{e}. \tag{2}$$

Probabilistically speaking, (2) is a device for constructing a pseudodistribution over $\boldsymbol{z} \in \{-1, 1\}^n$, which aims to match the first two moments of the distribution of optimal solutions to the original binary quadratic problem (d'Aspremont and Boyd 2003, Barak et al. 2014). This suggests to sample from a distribution parameterized by the relaxed solution and round to restore feasibility, as proposed by Goemans and Williamson (1995) for Max-Cut and described in Algorithm 1.

Algorithm 1 The Goemans-Williamson rounding algorithm for Problem (1)

Require: Positive semidefinite matrix $\boldsymbol{Q} \in \mathcal{S}_+^n$

    Compute $\boldsymbol{Z}^\star$ a solution of (2)

    Sample $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Z}^\star)$

    Construct $\hat{\boldsymbol{z}} \in \{-1, 1\}^n : \hat{z}_i := \mathrm{sign}(y_i)$

    return $\hat{\boldsymbol{z}}$ a solution to Problem (1)

The overall idea of Algorithm 1 is that the projection step (i.e., taking the coordinate-wise sign of $\boldsymbol{y}$) aims to match the second moment of the distribution of $\boldsymbol{y}$, $\mathbb{E}[\boldsymbol{y}\boldsymbol{y}^\top] = \boldsymbol{Z}^\star$. Precisely, we have $\mathbb{E}[\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}^\top] \succeq \frac{2}{\pi}\boldsymbol{Z}^\star$ (see Nesterov 1998, Bertsimas and Ye 1998). This inequality implies a $2/\pi$-factor guarantee for BQO when $\boldsymbol{Q} \succeq \boldsymbol{0}$. By further assuming that $\boldsymbol{Q}$ is the Laplacian matrix of a graph, Goemans and Williamson (1995) obtain a tighter constant of $\frac{2}{\pi} \min_{0 \leq \theta \leq \pi} \left(\frac{\theta}{1 - \cos\theta}\right) = 0.87856\dots$.

1.2. Problem Setting

In this work, we generalize Algorithm 1 to address orthogonally and rank-constrained problems. We first consider a general family of orthogonally constrained quadratic problems that subsumes binary quadratic optimization. Formally, we find m orthogonal vectors $\boldsymbol{u}_i \in \mathbb{R}^n$ which solve

$$\max_{\boldsymbol{u}_i \in \mathbb{R}^n, \, i \in [m]} \quad \sum_{i,j=1}^m \boldsymbol{u}_i^\top \boldsymbol{A}^{(i,j)} \boldsymbol{u}_j \text{ s.t. } \boldsymbol{u}_i^\top \boldsymbol{u}_j = \delta_{i,j}, \quad \forall i, j \in [m], \tag{3}$$

where $\boldsymbol{A}^{(i,j)} \in \mathbb{R}^{n \times n}$ and $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise is the Kronecker delta indicator variable, $\boldsymbol{A} \in \mathcal{S}_+^{nm}$ is a semidefinite matrix with block matrices $\boldsymbol{A}^{(i,j)}$, and we require $n \geq m$ so that the

problem is feasible. By introducing the semi-orthogonal matrix $\boldsymbol{U} \in \mathbb{R}^{n \times m}$ whose columns are the vectors $\boldsymbol{u}_i$, we can write our problem as

$$\max_{\boldsymbol{U} \in \mathbb{R}^{n \times m}} \quad \langle \boldsymbol{A}, \mathrm{vec}(\boldsymbol{U})\mathrm{vec}(\boldsymbol{U})^\top \rangle \quad \text{s.t.} \quad \boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_m, \tag{4}$$

where the $\mathrm{vec}(\cdot)$ operator stacks the columns of $\boldsymbol{U}$ together into a single vector.

The similarities between Problems (4) and (1) are striking: For example, we can formulate any BQO instance (1) as a special case of Problem (4) by defining $\boldsymbol{u}_i = z_i \boldsymbol{e}_i$ and $\boldsymbol{A}^{(i,j)} = Q_{i,j} \boldsymbol{e}_i \boldsymbol{e}_j^\top$ (see Section EC.2 for a detailed reduction). From a worst-case complexity perspective, Problem (4) is thus NP-hard by reduction from Max-Cut, as formally proven by Lai et al. (2025, Theorem 3.1).

However, while Problem (4) arises in a wide variety of problem settings, including clustering, quantum non-locality, or generalized trust-region problems (see Burer and Park 2024, and references therein), our prime motivation for studying Problem (4) in this paper is that it appears as a relevant substructure for rank-constrained quadratic optimization problems of the form

$$\min_{\boldsymbol{X} \in \mathbb{R}^{n \times m}} \quad \lambda \cdot \mathrm{rank}(\boldsymbol{X}) + \langle \mathrm{vec}(\boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{X}^\top)^\top, \boldsymbol{H} \rangle + \langle \boldsymbol{D}, \boldsymbol{X} \rangle \quad \text{s.t.} \quad \mathrm{rank}(\boldsymbol{X}) \leq k, \tag{5}$$

in much the same way as BQO appears as a relevant substructure for logically constrained quadratic optimization problems (see Section EC.1). Problem (5) with $\lambda \geq 0$, $\boldsymbol{H} \in \mathcal{S}_+^{nm}$, and $\boldsymbol{D} \in \mathbb{R}^{n \times m}$ includes matrix completion (Candès and Recht 2009) and reduced rank regression (Negahban and Wainwright 2011) as special cases.

In this paper, inspired by the Goemans-Williamson algorithm for BQO, we develop a relax-then-round strategy with a multiplicative-factor performance guarantee for Problem (4). To our knowledge, our algorithm is the first rounding mechanism with an approximation guarantee for this problem. Then, we extend our semidefinite relaxations and rounding mechanism to rank-constrained problems of the form (5). Regarding our semidefinite relaxations, unlike prior work (e.g., Kim et al. 2022, Bertsimas et al. 2023c, Li and Xie 2024) they do not require the presence of a spectral term in the objective of (5) (i.e., a term that only depends on the singular values of $\boldsymbol{X}$) and are thus more broadly applicable, for instance to unregularized matrix completion.

1.3. Related Work

We now review the relevant literature on orthogonally constrained quadratic optimization.

Burer and Park (2024) develop a hierarchy of semidefinite relaxations for Problem (4). To evaluate the tightness of their relaxations numerically, they apply several 'feasible rounding procedures' to generate feasible solutions, but do not provide any theoretical performance guarantee for these heuristics. In contrast, we develop a randomized rounding procedure and show it achieves a multiplicative factor guarantee, which is independent of the ambient dimension n and only decreases

as $1/\log m$. As a non-convex quadratic optimization problem, Problem (4) can be solved to provable optimality via global solvers such as `Gurobi` or `BARON`, or custom branch-and-bound schemes (Bertsimas et al. 2023b). However, the scalability of these global solvers is currently limited by the size of computer chips.

Special Cases: A larger body of work considers a special case of Problem (4), where the matrix $\boldsymbol{A}$ is block-diagonal, namely $\boldsymbol{A}^{(i,j)} = \boldsymbol{0}$ if $i \neq j$. In this case, Problem (4) reduces to

$$\max_{\boldsymbol{U} \in \mathbb{R}^{n \times m}} \sum_{i \in [m]} \boldsymbol{u}_i^\top \boldsymbol{A}^{(i,i)} \boldsymbol{u}_i \quad \text{s.t.} \quad \boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_m, \tag{6}$$

which is referred to as the sum of heterogeneous quadratic forms or the heterogeneous PCA problem. Indeed, when all the matrices $\boldsymbol{A}^{(i,i)}$ are equal, we recover the Principal Component Analysis (PCA) problem. Bolla et al. (1998, section 5) solve Problem (6) in polynomial time via linear algebra techniques when the matrices $\boldsymbol{A}^{(i,i)}$ are diagonal or commute with each other. For general matrices, Gilman et al. (2022) further tailor the semidefinite relaxations of Burer and Park (2024). Although tighter, they numerically show that their relaxations are not always tight. Indeed, for some instances, they even obtain optimality gaps higher than 100%. We are not aware of any approximation algorithms with guarantees for general (non-diagonal) instances of Problem (6).

Approximation Algorithms: To our knowledge, existing approximation algorithms do not apply to Problem (4) exactly, but to optimization problems with different orthogonality structures. Briët et al. (2010) propose an approximation algorithm for problems of the form

$$\max_{\boldsymbol{U} \in \mathbb{R}^{n \times m}} \sum_{i,j \in [m]} A_{i,j} \boldsymbol{u}_i^\top \boldsymbol{u}_j \text{ s.t. } \boldsymbol{u}_i^\top \boldsymbol{u}_i = 1 \ \forall i \in [m], \tag{7}$$

which also subsumes BQO (for $n = 1$), but does not enforce orthogonality between the columns of $\boldsymbol{U}$. They devise a relax-and-round strategy analogous to Goemans-Williamson that achieves an approximation ratio of $2/\pi + \Theta(1/n)$. A second line of work (Nemirovski 2007, So 2009) proposes $O(1/\log(n+m))$-approximation algorithms for quadratic optimization problems over matrices $\boldsymbol{U}$ that satisfy $\boldsymbol{U}^\top \boldsymbol{U} \preceq \boldsymbol{I}_m$. In other words, to problems where the largest singular value of $\boldsymbol{U}$, $\sigma_{\max}(\boldsymbol{U})$, is at most one. This constraint does not ensure that the columns of $\boldsymbol{U}$ are orthogonal, nor that they are of norm 1. Nonetheless, Nemirovski (2007) shows that, in several cases such as the orthogonal Procrustes or quadratic assignment problems, orthogonality constraints $\boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_m$ can be relaxed into $\boldsymbol{U}^\top \boldsymbol{U} \preceq \boldsymbol{I}_m$ without loss of optimality. Finally, Bandeira et al. (2016) study approximation algorithms for problems of the form

$$\max_{\boldsymbol{U}_i \in \mathbb{R}^{n \times m}, i=1,\ldots,k} \sum_{i,j \in [k]} \langle \boldsymbol{A}^{(i,j)}, \boldsymbol{U}_i^\top \boldsymbol{U}_j \rangle \text{ s.t. } \boldsymbol{U}_i^\top \boldsymbol{U}_i = \boldsymbol{I}_m \ \forall i \in [k]. \tag{8}$$

Unfortunately, Problem (8) is not equivalent to (4) and the proof techniques in the aforementioned works do not extend to our case. There are two key differences in the objective function of (8).

Compared to (4), Problem (8) involves the *inner*-products between columns of *different* semi-orthogonal matrices $\boldsymbol{U}_i$, $\boldsymbol{U}_{i'}$ for $i \neq i'$. On the other hand, the objective in (4) depends on *outer*-products between columns of the *same* matrix $\boldsymbol{U}$. In particular, we can restore feasibility for each matrix $\boldsymbol{U}_i$, $i = 1, \ldots, k$ in (8) separately, while the columns of $\boldsymbol{U}$ in (4) need to be considered together. In particular, heterogeneous PCA is a special case of (4) but cannot be cast as a problem of the form (8).

1.4. Contributions and Structure

Our main contribution is the development of a Goemans-Williamson sampling algorithm for the class of semi-orthogonal problems (4) and its extension to rank-constrained optimization.

We begin by studying approximation algorithms for Problem (4) in Section 2. We derive a semidefinite relaxation and propose a sampling procedure to generate high-quality feasible solutions. We show that our algorithm achieves a purely multiplicative approximation guarantee (Theorems 1–2) for Problem (4), with a constant that scales as $O(1/\log m)$. We also identify a class of problem instances (Proposition 2) for which our algorithm cannot achieve a performance guarantee better than $O(1/\log m)$. In comparison, we show that sampling feasible solutions uniformly at random achieves a $1/nm$ approximation ratio. Notably, our approximation ratio does not depend on the ambient dimension n.

We then extend our approach to low-rank optimization problems in Section 3. To facilitate this extension, we first derive new Shor relaxations for low-rank optimization problems. Unlike prior works (Recht et al. 2010, Bertsimas et al. 2023c, Kim et al. 2022, Li and Xie 2024), our relaxations do not require a spectral or permutation-invariant term in the objective or constraints. Conscious that these relaxations involve a number of additional semidefinite variables that may be prohibitively large in practice, we show how to eliminate many of these variables in the relaxation without altering its optimal value (Theorem 3). Finally, we describe a sampling algorithm to generate high-quality solutions from this relaxation. As low-rank optimization is strongly NP-hard (Gillis and Glineur 2011), our theoretical polynomial-time approximation guarantees derived in Section 2 cannot be generalized to this broader class of problem.

To illustrate our approach, we apply our Shor relaxation to three prominent low-rank optimization problems in Section 4. In particular, we show how to exploit further problem structure and eliminate more variables from our relaxations, making our new relaxation more scalable.

Finally, in Section 5, we numerically benchmark our convex relaxations and randomized rounding schemes on quadratic semi-orthogonal and low-rank matrix completion problems.

1.5. Notation

We let nonbold face characters such as b denote scalars, lowercase boldfaced characters such as $\boldsymbol{x}$ denote vectors, uppercase boldfaced characters such as $\boldsymbol{X}$ denote matrices, and calligraphic uppercase characters such as $\mathcal{Z}$ denote sets. We let $[n]$ denote the set of running indices $\{1, ..., n\}$. The cone of $n \times n$ symmetric (resp. positive semidefinite) matrices is denoted by $\mathcal{S}^n$ (resp. $\mathcal{S}^n_+$). Inner products are denoted $\langle \cdot, \cdot \rangle$, and are associated with the Euclidean norm $\|\boldsymbol{x}\|$ for vectors and the Frobenius norm $\|\boldsymbol{X}\|_F$ for matrices. We also denote the spectral norm of a matrix $\boldsymbol{X}$ by $\|\boldsymbol{X}\|_\sigma$.

For a matrix $\boldsymbol{X} \in \mathbb{R}^{n \times m}$, we let $\boldsymbol{x}_i$ denote its ith column and $\boldsymbol{X}_{i,\cdot}$ denote a vector containing its ith row. We let $\mathrm{vec}(\boldsymbol{X}) : \mathbb{R}^{n \times m} \to \mathbb{R}^{nm}$ denote the vectorization operator which maps matrices to vectors by stacking columns. For a square matrix $\boldsymbol{X}$, $\mathrm{diag}(\boldsymbol{X})$ compiles the diagonal entries of $\boldsymbol{X}$ into a vector, while $\mathrm{Diag}(\boldsymbol{x})$ is a square matrix with diagonal equal to $\boldsymbol{x}$. For a matrix $\boldsymbol{W}$, we may find it convenient to describe it as a block matrix composed of equally sized blocks and denote the (i, i') block by $\boldsymbol{W}^{(i,i')}$. The dimension of each block will be clear from the context, given the size of the matrix $\boldsymbol{W}$ and the number of blocks. In particular, $\boldsymbol{I}_m \otimes \boldsymbol{\Sigma}$ with $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ denotes an $nm \times nm$ block-diagonal matrix whose m diagonal blocks are equal to $\boldsymbol{\Sigma}$ (see Gupta and Nagar 2018, Chapter 1.2, for an introduction to the Kronecker product $\otimes$). With this notation, $\mathrm{vec}(\boldsymbol{\Sigma X}) = (\boldsymbol{I}_m \otimes \boldsymbol{\Sigma}) \mathrm{vec}(\boldsymbol{X})$.

We let $\boldsymbol{X}^\dagger$ be the pseudoinverse of $\boldsymbol{X}$, which is used in the Schur complement lemma (Boyd et al. 1994, Eqn. 2.41). We let $\mathcal{Y}^k_n := \{\boldsymbol{Y} \in \mathcal{S}^n_+ : \boldsymbol{Y}^2 = \boldsymbol{Y}, \mathrm{tr}(\boldsymbol{Y}) \leq k\}$ denote the set of orthogonal projection matrices with rank at most k, whose convex hull is $\{\boldsymbol{P} \in \mathcal{S}^n_+ : \boldsymbol{P} \preceq \boldsymbol{I}_n, \mathrm{tr}(\boldsymbol{P}) \leq k\}$ (Overton and Womersley 1992, Theorem 3). Analogously, we let $\mathcal{Y}_n$ denote the set of $n \times n$ orthogonal projection matrices of any rank, with convex hull $\mathrm{Conv}(\mathcal{Y}_n) = \{\boldsymbol{P} \in \mathcal{S}^n_+ : \boldsymbol{P} \preceq \boldsymbol{I}_n\}$. In particular, we have $\mathrm{rank}(\boldsymbol{Y}) = \mathrm{tr}(\boldsymbol{Y})$ for any projection matrix $\boldsymbol{Y}$.

Finally, our sampling procedure invokes the multivariate Gaussian probability measure: we let $\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ denote a centered multivariate normal distribution with covariance matrix $\boldsymbol{\Sigma}$; see Grimmett and Stirzaker (2020) for an overview of the Gaussian distribution and Gupta and Nagar (2018) for an overview of its matrix extensions.

2. A Goemans-Williamson Approach for Orthogonality Constraints

In this section, we propose a new Goemans-Williamson type approach for semi-orthogonal quadratic optimization problems, mirroring the development of the Goemans-Williamson algorithm for BQO in Section 1.1. First, in Section 2.1, we review a semidefinite relaxation for semi-orthogonal quadratic optimization originally developed by Burer and Park (2024). Then, in Section 2.2, we propose a randomized rounding scheme to generate high-quality solutions, which largely follows the blueprint of the Goemans-Williamson strategy in BQO. We derive multiplicative performance

guarantees for our rounding mechanism in Section 2.3. In particular, we show that our algorithm achieves an $O(1/\log m)$-factor approximation for Problem (4) and that our analysis of our algorithm is tight. As a benchmark, we analyze the performance of uniform rounding (achieving $(1/(nm))$-factor approximation) in Section 2.4, and conclude the section by discussing potential variants of and improvements to our rounding mechanism in Section 2.5.

The rest of the paper extends the relax-then-round scheme developed in this section from low-rank orthogonal to low-rank quadratic optimization.

2.1. A Shor Relaxation

We study quadratic optimization over orthogonality constraints as described in Problem (4). As reviewed in Section 1.3, Burer and Park (2024, section 2.2) derive the following semidefinite relaxation for Problem (4):

$$\max_{\boldsymbol{W} \in \mathcal{S}_+^{mn}} \quad \langle \boldsymbol{A}, \boldsymbol{W} \rangle \quad \text{s.t.} \quad \text{tr}\left(\boldsymbol{W}^{(j,j')}\right) = \delta_{j,j'} \ \forall j, j' \in [m], \ \sum_{i \in [m]} \boldsymbol{W}^{(i,i)} \preceq \boldsymbol{I}_n, \tag{9}$$

where the matrix $\boldsymbol{W}$ encodes for the outer-product of $\text{vec}(\boldsymbol{U})$ with itself, and the trace constraints on the blocks of $\boldsymbol{W}$ stem from the columns of $\boldsymbol{U}$ having unit norm and being pairwise orthogonal.

Similarly to semidefinite relaxation of (1), imposing the constraint[1] that $\boldsymbol{W}$ is rank-one in (9) would result in an exact reformulation of (4). Accordingly, The relaxation (9) is tight whenever some optimal solution is rank-one. However, the optimal solutions to (9) are often all high-rank. Actually, it follows from manipulating the Barvinok-Pataki bound (Barvinok 2001, Pataki 1998) that there exists[2] some optimal solution to Problem (9) with rank at most $n + m$. However, not all optimal solutions obey this bound; thus, we do not use this observation in our analysis. An interesting question is how to generate a high-quality feasible solution to (4), with provable performance guarantee, by leveraging a solution of (9), which is the focus of the rest of the section.

Note that Problem (9) corresponds to the 'DiagSum' relaxation of Burer and Park (2024). Burer and Park (2024, section 2.3) derive an even stronger relaxation, which they call a 'Kronecker' relaxation. Thus, the approximation guarantees we derive for Problem (9) directly apply if we solve their Kronecker relaxation instead. However, we do not explicitly analyze the Kronecker relaxation here because it is significantly less tractable (as reported in Burer and Park 2024, Table 1), and it would not lead to a tighter order[3] of approximation guarantee than the $O(1/\log m)$ guarantee derived here, although it could improve the semidefinite relaxation for some specific instances (Burer and Park 2024).

2.2. A Sample-Then-Stochastically-Project Procedure

We propose a randomized rounding scheme to generate high-quality feasible solutions to (4) from an optimal solution to (9). In the next section, we show that it attains an $O(1/\log m)$ multiplicative approximation factor guarantee for Problem (4).

First, we solve (9) and obtain a semidefinite matrix $\boldsymbol{W}^\star$. Second, using $\boldsymbol{W}^\star$, we sample an $n \times m$ matrix $\boldsymbol{G}$ such that $\text{vec}(\boldsymbol{G})$ follows a normal distribution with mean $\boldsymbol{0}_{nm}$ and covariance matrix $\boldsymbol{W}^\star$. Third, from the matrix $\boldsymbol{G}$, we generate a feasible solution to (4). Specifically, we compute a singular value decomposition of $\boldsymbol{G}$, $\boldsymbol{G} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top$ and define $\boldsymbol{Q} := \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^\top$ where $\boldsymbol{D}$ is a diagonal matrix such that each diagonal entry is equal to ± 1, sampled independently such that $\mathbb{P}(D_{i,i} = 1) = (1 + \sigma_i/\sigma_{\max})/2$, where σ_i is the ith singular value of $\boldsymbol{G}$ and $\sigma_{\max}$ is the largest singular value of $\boldsymbol{G}$. We have $\boldsymbol{Q}^\top\boldsymbol{Q} = \boldsymbol{I}_m$ because $\boldsymbol{D}^2 = \boldsymbol{I}_m$ We summarize our procedure in Algorithm 2.

Algorithm 2 A Relax-then-Project Algorithm for Orthogonality Constrained Optimization

---

Require: Positive semidefinite matrix $\boldsymbol{A} \in \mathcal{S}_+^{nm}$

Compute $\boldsymbol{W}^\star$ a solution of (9)

Sample $\boldsymbol{G}$ according to $\text{vec}(\boldsymbol{G}) \sim \mathcal{N}(\boldsymbol{0}_{nm}, \boldsymbol{W}^\star)$

Compute the SVD of $\boldsymbol{G}$, $\boldsymbol{G} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top$

Sample $D_{i,i} = \pm 1$ independently such that $\mathbb{P}(D_{i,i} = 1) = (1 + \sigma_i/\sigma_{\max})/2$

Construct $\boldsymbol{Q} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^\top$

return A semi-orthogonal matrix $\boldsymbol{Q}$

We note that the normal distribution in our second step differs from the most widely used 'matrix normal distribution' (see, e.g., Gupta and Nagar 2018, Chapter 2) and, to the best of our knowledge, has only been studied by Barratt (2018). In contrast with other definitions of matrix Gaussian distributions, the entries of $\boldsymbol{G}$ in our sampling are neither independent nor identically distributed. In our implementation of Algorithm 2, we can sample $\text{vec}(\boldsymbol{G}) \sim \mathcal{N}(\boldsymbol{0}_{nm}, \boldsymbol{W}^\star)$ even when $\boldsymbol{W}^\star$ is rank-deficient via the following construction—which will also be relevant for the theoretical analysis in Section 2.3. Denoting $r = \text{rank}(\boldsymbol{W}^\star)$, we first construct a Cholesky decomposition of $\boldsymbol{W}$: $\boldsymbol{W} = \sum_{k \in [r]} \text{vec}(\boldsymbol{B}_k)\text{vec}(\boldsymbol{B}_k)^\top$ with $\boldsymbol{B}_k \in \mathbb{R}^{n \times m}$. Then, we sample $\text{vec}(\boldsymbol{G}) = \sum_{k \in [r]} \text{vec}(\boldsymbol{B}_k)z_k$ with $\boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}_r, \boldsymbol{I}_r)$. This procedure ensures that $\text{vec}(\boldsymbol{G}) \in \text{span}(\boldsymbol{W}^\star)$ almost surely, and that if the semidefinite relaxation is tight then $\boldsymbol{G}$ is optimal almost surely. In particular, if $r = 1$ then it is immediate that our semidefinite relaxation is tight, and thus our rounding is exact in this case.

Second, we should comment on our procedure to obtain a feasible semi-orthogonal matrix $\boldsymbol{Q}$ from $\boldsymbol{G}$. Conditioned on $\boldsymbol{G}$, we have $\mathbb{E}[\boldsymbol{D}|\boldsymbol{G}] = \boldsymbol{\Sigma}/\sigma_{\max}$, so that $\mathbb{E}[\boldsymbol{Q}|\boldsymbol{G}] = \boldsymbol{G}/\sigma_{\max}$. This observation will be crucial in our theoretical analysis, enabling us to relate the second moment of the distribution of $\boldsymbol{Q}$ to that of $\boldsymbol{G}$, as formally stated below.

PROPOSITION 1. *Consider matrices $\boldsymbol{G} \in \mathbb{R}^{n \times m}$ and $\boldsymbol{Q} \in \mathbb{R}^{n \times m}$ generated according to Algorithm 2. The following holds:*

$$\mathbb{E}\left[\text{vec}(\boldsymbol{Q})\text{vec}(\boldsymbol{Q})^\top\right] \succeq \mathbb{E}\left[\frac{\text{vec}(\boldsymbol{G})\text{vec}(\boldsymbol{G})^\top}{\sigma_{\max}(\boldsymbol{G})^2}\right]. \tag{10}$$

Proof of Proposition 1   Observe that, conditioned on $\boldsymbol{G}$, we have

$$\mathrm{Cov}(\mathrm{vec}(\boldsymbol{Q})|\boldsymbol{G}) = \mathbb{E}\left[\mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top|\boldsymbol{G}\right] - \mathbb{E}\left[\mathrm{vec}(\boldsymbol{Q})|\boldsymbol{G}\right]\mathbb{E}\left[\mathrm{vec}(\boldsymbol{Q})|\boldsymbol{G}\right]^\top \succeq \boldsymbol{0},$$

leading to

$$\mathbb{E}\left[\mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top|\boldsymbol{G}\right] \succeq \frac{1}{\sigma_{\max}(\boldsymbol{G})^2}\mathrm{vec}(\boldsymbol{G})\mathrm{vec}(\boldsymbol{G})^\top,$$

since $\mathbb{E}\left[\mathrm{vec}(\boldsymbol{Q})|\boldsymbol{G}\right] = \boldsymbol{G}/\sigma_{\max}(\boldsymbol{G})$. Taking expectation with respect to $\boldsymbol{G}$ yields (10). $\square$

Similar to the original algorithm of Goemans and Williamson (1995), the intuition behind Algorithm 2 is that the sampled matrix $\boldsymbol{G}$ achieves an average performance equal to the relaxation value $(\mathbb{E}[\langle\boldsymbol{A},\mathrm{vec}(\boldsymbol{G})\mathrm{vec}(\boldsymbol{G})^\top\rangle] = \langle\boldsymbol{A},\boldsymbol{W}^\star\rangle)$ and is feasible on average $(\mathbb{E}[\boldsymbol{G}^\top\boldsymbol{G}] = \boldsymbol{I}_m)$. Therefore, the objective value of the feasible solution $\boldsymbol{Q}$ should not be too different from that of $\boldsymbol{G}$. We theoretically analyze the performance of our algorithm in Section 2.3.

2.3. Theoretical Analysis: Multiplicative Performance Guarantees

We now theoretically analyze the performance of Algorithm 2 in the case where the objective matrix $\boldsymbol{A}$ in (4) is positive semidefinite.

Solutions $\boldsymbol{Q}$ generated by Algorithm 2 achieve an average performance of $\mathbb{E}\left[\mathrm{vec}(\boldsymbol{Q})^\top\boldsymbol{A}\,\mathrm{vec}(\boldsymbol{Q})\right] = \langle\boldsymbol{A},\mathbb{E}\left[\mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top\right]\rangle$. By Proposition 1 and the fact that $\boldsymbol{A} \succeq \boldsymbol{0}$, we have $\langle\boldsymbol{A},\mathbb{E}\left[\mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top\right]\rangle \geq \langle\boldsymbol{A},\mathbb{E}\left[\mathrm{vec}(\boldsymbol{G})\mathrm{vec}(\boldsymbol{G})^\top/\sigma_{\max}(\boldsymbol{G})^2\right]\rangle$. Hence, to obtain a β-multiplicative guarantee for our algorithm, it suffices to show that $\mathbb{E}\left[\mathrm{vec}(\boldsymbol{G})\mathrm{vec}(\boldsymbol{G})^\top/\sigma_{\max}(\boldsymbol{G})^2\right] \succeq \beta\boldsymbol{W}^\star$, which is the main focus of this section.

Our first result is an analytical multiplicative performance guarantee, which asymptotically scales as $O(1/\log m)$. It arises as a consequence of a Cauchy-Schwarz inequality and concentration bounds on the largest singular value of $\boldsymbol{G}$. Indeed, $\sigma_{\max}(\boldsymbol{G})$ satisfies the following technical lemma (proof deferred to Section EC.3.1):

LEMMA 1. *Consider a random matrix* $\boldsymbol{G} \in \mathbb{R}^{n\times m}$ *sampled according to* $\mathrm{vec}(\boldsymbol{G}) \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{W})$*, where the matrix* $\boldsymbol{W}$ *is a feasible solution to* (9)*. Then, the following inequality holds*

$$\mathbb{E}[\sigma_{\max}(\boldsymbol{G})^2] \leq \min(m, 2\log(2m) + 2). \tag{11}$$

Furthermore, $\sigma_{\max}(\boldsymbol{G})$ *satisfies the tail bounds: for any* $t > 0$*,* $\mathbb{P}(\sigma_{\max}(\boldsymbol{G}) > t) \leq 2m\,e^{-t^2/2}$.

From this technical lemma, we derive the following semidefinite relationship:

THEOREM 1. *The matrix* $\boldsymbol{G} \in \mathbb{R}^{n\times m}$ *generated by Algorithm 2 satisfies the inequality*

$$\mathbb{E}\left[\frac{\mathrm{vec}(\boldsymbol{G})\mathrm{vec}(\boldsymbol{G})^\top}{\sigma_{\max}(\boldsymbol{G})^2}\right] \succeq \max\left(\frac{2}{\pi m},\ \frac{1}{\pi(\log(2m)+1)}\right)\boldsymbol{W}^\star. \tag{12}$$

Theorem 1 leads to a purely multiplicative performance guarantee for Algorithm 2, $\mathbb{E}\left[\langle \boldsymbol{A}, \text{vec}(\boldsymbol{Q})\text{vec}(\boldsymbol{Q})^\top\rangle\right] \geq \beta \ \langle \boldsymbol{A}, \boldsymbol{W}^\star\rangle$, with $\beta = \max\left(\frac{2}{\pi m}, \ \frac{1}{\pi(\log(2m)+1)}\right)$. Interestingly, the multiplicative constant is independent of the ambient dimension n, but only depends on the number of vectors m. For small values of m, the $2/(\pi m)$ term dominates and drives the value of β (e.g., it equals 0.636 for $m = 1$). Asymptotically, however, our bound scales as $1/\log m$ and exhibits a very mild dependency on m. Actually, we can show that our analysis of Algorithm 2 is essentially tight (proof of Proposition 2 deferred to Section EC.3.2)

PROPOSITION 2. *There exists a family of matrices $\boldsymbol{W}^\star \in \mathcal{S}_+^{nm}$ for Algorithm 2 for which, for any $\beta > 0$ such that $\mathbb{E}[\text{vec}(\boldsymbol{Q})\,\text{vec}(\boldsymbol{Q})^\top] \succeq \beta \boldsymbol{W}^\star$, we must have $\beta = O\left(1/\log m\right)$.*

Proof of Theorem 1 Consider an arbitrary unit vector $\boldsymbol{v} \in \mathbb{R}^{nm}$. From Cauchy-Schwarz, we have that for any random variables $A \geq 0, B > 0$ a.s. that $\mathbb{E}[A/B] \geq \mathbb{E}[\sqrt{A}]^2/\mathbb{E}[B]$. Thus, applying this inequality to $A = (\boldsymbol{v}^\top\text{vec}(\boldsymbol{G}))^2$ and $B = \sigma_{\max}(\boldsymbol{G})^2$ yields

$$\mathbb{E}[(\boldsymbol{v}^\top\text{vec}(\boldsymbol{G}))^2/\sigma_{\max}^2(\boldsymbol{G})] \geq \frac{\mathbb{E}[|\boldsymbol{v}^\top\text{vec}(\boldsymbol{G}))|]^2}{\mathbb{E}[\sigma_{\max}(\boldsymbol{G})^2]}.$$

For the numerator, $\boldsymbol{v}^\top\text{vec}(\boldsymbol{G}) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{v}^\top\boldsymbol{W}^\star\boldsymbol{v})$, so, by definition of the half-normal distribution (e.g., Grimmett and Stirzaker 2020)

$$\mathbb{E}[|\boldsymbol{v}^\top\text{vec}(\boldsymbol{G})|] = \sqrt{\frac{2}{\pi}}\sqrt{\boldsymbol{v}^\top\boldsymbol{W}^\star\boldsymbol{v}}.$$

For the denominator, we refer to Lemma 1, where we show that $\mathbb{E}[\sigma_{\max}(\boldsymbol{G})^2] \leq \min(m, 2\log(2m) + 2)$. Thus, we have the inequality

$$\mathbb{E}[(\boldsymbol{v}^\top\text{vec}(\boldsymbol{G}))^2/\sigma_{\max}^2(\boldsymbol{G})] \geq \frac{2}{\pi\min(m, 2\log(2m)+2)}\boldsymbol{v}^\top\boldsymbol{W}^\star\boldsymbol{v}.$$

Combining this inequality with (10), we obtain the desired result. □

Despite its strong asymptotic behavior, for a fixed value of m, the constant in Theorem 1 can be weak (largely because of the use of the Cauchy-Schwarz inequality). To get a more accurate estimate of the performance of Algorithm 2, we now derive tighter bounds that can be computed numerically.

THEOREM 2. *Let $\boldsymbol{G} \in \mathbb{R}^{n\times m}$ be a Gaussian matrix generated by Algorithm 2. Then, the matrix $\boldsymbol{G}$ satisfies the inequality:*

$$\mathbb{E}\left[\frac{\text{vec}(\boldsymbol{G})\text{vec}(\boldsymbol{G})^\top}{\sigma_{\max}(\boldsymbol{G})^2}\right] \succeq \beta_{n,m}\boldsymbol{W}^\star, \tag{13}$$

with

$$\beta_{n,m} := \min_{\lambda\in[0,1]} \int_0^\infty \left(1 + 2t\,m\frac{1-\lambda}{nm-1}\right)^{-(nm-1)/2} (1 + 2tm\lambda)^{-3/2}dt. \tag{14}$$

In particular, the constant $\beta_{n,m}$ satisfies the following properties:

Table 1   Values of the approximation factor from Theorems 1 and 2 for some values of n and m.

	Theorem 1	$\beta_{n,m}$ (Theorem 2)			
m		$n=5$	$n=10$	$n=15$	$n=\infty$
1	0.636620	0.735264	0.706972	0.697920	0.680415
2	0.318310	0.353486	0.346734	0.344533	0.340208
3	0.212207	0.232640	0.229689	0.228720	0.226805
4	0.159155	0.173367	0.171721	0.171179	0.170104
5	0.127324	0.138164	0.137116	0.136770	0.136083
10	0.079662	—	0.068299	0.068213	0.068042
15	0.072323	—	—	0.045437	0.045361

(a) *For any integer m, $\beta_{n,m}$ is non-increasing in n. Moreover, it is non-increasing in m for any fixed n wherever $\beta_{n,m}$ exists ($n \geq m$)*

(b) *For any integer m, we have $\beta_{n,m} \to \beta_{\infty,m}$ as $n \to \infty$ with*

$$\beta_{\infty,m} := \min_{\lambda \in [0,1]} \int_0^\infty e^{-t\,m(1-\lambda)} (1+2tm\lambda)^{-3/2} dt \quad = \min_{\lambda \in [0,1]} \mathbb{E}_{X \sim \chi_1^2} \left[\frac{X}{m(1-\lambda) + m\lambda X} \right].$$

(c) *For $m = 1$, $\beta_{n,1}$ is optimal, since there exists a covariance matrix $\boldsymbol{W}^\star$ satisfying (13) at equality.*

Compared with Theorem 1, the value of the constant in Theorem 2 is primarily computational. By solving numerically the one-dimensional minimization problem in λ, it provides tighter estimates of the performance of our algorithm, especially for small values of m, as reported in Table 1. While the guarantee from Theorem 1 is independent of n, the constant $\beta_{n,m}$ in Theorem 2 is monotonically decreasing with n, obtaining stronger performance guarantees for finite values of n.

However, we should acknowledge that $\beta_{\infty,m}$ does not scale as $\Theta(1/\log m)$ for large values of m, and thus is asymptotically weaker than Theorem 1 (actually, Remark EC.3 identifies a class of matrices $\boldsymbol{W}^\star$ for which $\beta_{n,m} \leq 1/m$). We can further strengthen Theorem 2 and view $\beta_{n,m}$ as a special case of an even tighter bound (Theorem EC.1), which recovers the asymptotic scaling of Theorem 1. For the sake of exposition, we only present Theorem 2 in the main paper. We present and prove the more general result (Theorem EC.1) in the Electronic Companion. Theorem 2 follows immediately as a special case.

2.4. Benchmark: Uniform Sampling

To evaluate the performance of Algorithm 2, it is interesting to compare its performance to a naive baseline where we draw $\boldsymbol{Q}$ uniformly from the set of semi-orthogonal matrices i.e., sample $\boldsymbol{Q}$ from the Haar measure (Meckes 2019). Note that this is analogous to generating i.i.d. Bernoulli vectors in binary quadratic optimization, which achieves a $1/2$ approximation ratio in the Max-Cut case:

PROPOSITION 3. *Let $\boldsymbol{Q} \in \mathbb{R}^{n \times m}$ be distributed uniformly over $\{\boldsymbol{U} \in \mathbb{R}^{n \times m} : \boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_m\}$. We have*

$$\mathbb{E}[\langle \boldsymbol{A}, \mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top \rangle] \leq \max_{\boldsymbol{U} \in \mathbb{R}^{n \times m} : \boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_m} \langle \boldsymbol{A}, \mathrm{vec}(\boldsymbol{U})\mathrm{vec}(\boldsymbol{U})^\top \rangle \leq nm\,\mathbb{E}[\langle \boldsymbol{A}, \mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top \rangle].$$

Proposition 3 implies that taking $\boldsymbol{Q}$ to be uniformly distributed gives a $1/(nm)$-factor approximation algorithm for Problem (4). This is a significantly worse multiplicative term than Theorems 1 and 2 for Algorithm 2.

Proof of Proposition 3   By optimality, $\boldsymbol{Q}$ being feasible for (4),

$$\langle \boldsymbol{A}, \mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top \rangle \leq \max_{\boldsymbol{U} \in \mathbb{R}^{n \times m}: \boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_m} \langle \boldsymbol{A}, \mathrm{vec}(\boldsymbol{U})\mathrm{vec}(\boldsymbol{U})^\top \rangle,$$

which leads to the first inequality.

Furthermore,

$$\max_{\boldsymbol{U} \in \mathbb{R}^{n \times m}: \boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_m} \langle \boldsymbol{A}, \mathrm{vec}(\boldsymbol{U})\mathrm{vec}(\boldsymbol{U})^\top \rangle \leq \max_{\boldsymbol{u} \in \mathbb{R}^{nm}: \|\boldsymbol{u}\|^2 = m} \langle \boldsymbol{A}, \boldsymbol{u}\boldsymbol{u}^\top \rangle = m\lambda_{\max}(\boldsymbol{A}) \leq m\,\mathrm{tr}(\boldsymbol{A}).$$

To conclude, observe that since $\boldsymbol{Q}$ is distributed according to the Haar measure, we have $\mathbb{E}[\boldsymbol{q}_i\boldsymbol{q}_i^\top] = \frac{1}{n}\boldsymbol{I}_n$ and $\mathbb{E}[\boldsymbol{q}_i\boldsymbol{q}_j^\top] = \boldsymbol{0}$ for $i \neq j$ (cf. Meckes 2019). Therefore, we have $\mathbb{E}[\mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top] = \frac{1}{n}\boldsymbol{I}_{nm}$ and $\mathbb{E}[\langle \boldsymbol{A}, \mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top \rangle] = \frac{1}{n}\mathrm{tr}(\boldsymbol{A})$. $\qquad\square$

REMARK 1. Proposition 3's upper bound is tight for uniform rounding. Indeed, if $\boldsymbol{A}$ is an identity matrix, then any uniformly sampled $\boldsymbol{Q}$ is optimal and the left inequality is tight. Moreover, if $\boldsymbol{A}$ is a matrix such that $\boldsymbol{A}_{i,j}^{(i,j)} = 1$ for $i, j \in [m]$ and $A_{l_1,l_2}^{(i,j)} = 0$ for $l_1 \neq i$ or $l_2 \neq j$ otherwise, then $\mathrm{tr}(\boldsymbol{A}) = m$ and an optimal choice of $\boldsymbol{U}$ is $\boldsymbol{U}_i = \boldsymbol{e}_i$, giving both $\max_{\boldsymbol{U} \in \mathbb{R}^{n \times m}: \boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_m} \langle \boldsymbol{A}, \mathrm{vec}(\boldsymbol{U})\mathrm{vec}(\boldsymbol{U})^\top \rangle = m^2$ and $nm\,\mathbb{E}[\langle \boldsymbol{A}, \mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top \rangle] = m^2$. This corresponds to a family of instances of increasing size for which our bound on uniform rounding is tight.

In Section EC.5 of the online supplement, we augment the discussion in this section by deriving a second benchmark algorithm inspired by deflation algorithms in the sparse principal component analysis literature, and show that it attains a $1/m^2$ guarantee, which improves on the $1/nm$ guarantee of uniform rounding for $n \gg m$, but is worse than Algorithm 2. As deflation is more complicated to design and analyze than uniform rounding but simpler than Algorithm 2, we have demonstrated a clear trade-off between the difficulty of designing and analyzing an algorithm and the quality of the worst-case guarantee.

2.5. Discussion: Algorithm Variants

We conclude this section by discussing alternative rounding strategies from our relaxation (9).

Algorithm 2 can be interpreted as a two-step generalization of the randomized rounding scheme of Goemans and Williamson (1995), where we sample a large multivariate normal vector $\mathrm{vec}(\boldsymbol{G}) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{W}^\star)$ and generate a feasible semi-orthogonal matrix $\boldsymbol{Q}$ from $\boldsymbol{G}$. Interestingly, our algorithm introduces an additional source of randomness in the generation of $\boldsymbol{Q}$ (hence, the qualification 'two-step'), which is key for guaranteeing a relationship between the second moment of $\mathrm{vec}(\boldsymbol{Q})$ and $\mathrm{vec}(\boldsymbol{G})$ as derived in Equation (10).

Alternatively, we could have taken $\boldsymbol{D} = \boldsymbol{I}_m$ in Algorithm 2, i.e., define $\boldsymbol{Q}$ as the projection (with respect to the Frobenius norm) of $\boldsymbol{G}$ onto the space of semi-orthogonal matrices. However, with this deterministic construction, Equation (10) may not hold. Indeed, consider the counterexample with $n = 4, m = 2$ and

$$(\boldsymbol{W}^\star)^{(1,1)} = \mathrm{Diag}\begin{pmatrix} 0.025 \\ 0.177 \\ 0.263 \\ 0.535 \end{pmatrix}, (\boldsymbol{W}^\star)^{(1,2)} = \mathrm{Diag}\begin{pmatrix} -0.042979 \\ 0.229513 \\ 0.201629 \\ -0.388163 \end{pmatrix}, (\boldsymbol{W}^\star)^{(2,2)} = \mathrm{Diag}\begin{pmatrix} 0.076 \\ 0.300 \\ 0.159 \\ 0.465 \end{pmatrix},$$

Averaging over 200 repetitions with 25000 Gaussian samples per repetition, we obtain a 95% confidence interval on $\lambda_{\min}$ of the form $\lambda_{\min}\left(\mathbb{E}\left[\mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top - \frac{\mathrm{vec}(\boldsymbol{G})\mathrm{vec}(\boldsymbol{G})^\top}{\sigma_{\max}(\boldsymbol{G})^2}\right]\right) = -0.0154 \pm 0.000025 < 0$

Our rounding procedures sample $D_{i,i} \in \{\pm 1\}$ at random, in particular, without taking into account the downstream objective $\mathrm{vec}(\boldsymbol{Q})^\top \boldsymbol{A} \, \mathrm{vec}(\boldsymbol{Q})$. Instead, we could also optimize the diagonal entries of $\boldsymbol{D}$ to explicitly maximize the objective, by solving a binary quadratic optimization problem. Doing so would give a solution at least as good as the one obtained via a random sampling, at the expense of solving a BQO problem with m variables, which might be practically feasible for moderate values of m.

3. New Relaxations and Sampling for Low-Rank Optimization Problems

In this section, we generalize our Goemans-Williamson algorithm for semi-orthogonal quadratic optimization (Algorithm 2 in Section 2) to generic rank-constrained optimization. For readers familiar with the mixed-integer literature, our overall approach mirrors the extension of the Goemans-Williamson rounding for BQO to logically constrained optimization, as reviewed in Section EC.1.

We proceed in three steps: First, we derive new Shor relaxations for rank-constrained optimization problems (§3.1). Unlike prior work (Recht et al. 2010, Bertsimas et al. 2023c, Kim et al. 2022, Li and Xie 2024), our relaxations do not require the presence of a spectral or permutation-invariant term in the objective or constraints. Interestingly, we show that many of the variables in our Shor relaxations can be omitted without altering the objective value, leading to a more compact and tractable formulation. Compared with Bertsimas et al. (2023c), we show that our new relaxations are stronger and more broadly applicable. Second, we discuss how our common ideas in logically constrained optimization, such as RLT, can be generalized to our context and further strengthen our relaxation (§3.2). Finally, we describe a sampling algorithm for these problems in §3.3.

3.1. A New Shor Relaxation and Its Compact Formulation

We study a quadratic low-rank optimization problem with linear constraints, which encompasses low-rank matrix completion (Candès and Recht 2009), and reduced rank regression (Negahban

and Wainwright 2011) problems among others; see Bertsimas et al. (2022) for a review of low-rank optimization. Formally, we study the problem:

$$\min_{\boldsymbol{Y} \in \mathcal{Y}_n^k} \min_{\boldsymbol{X} \in \mathbb{R}^{n \times m}} \quad \langle \boldsymbol{C}, \boldsymbol{Y} \rangle + \left\langle \mathrm{vec}(\boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{X}^\top)^\top, \boldsymbol{H} \right\rangle + \langle \boldsymbol{D}, \boldsymbol{X} \rangle \qquad (15)$$
$$\text{s.t.} \quad \langle \boldsymbol{A}_i, \boldsymbol{X} \rangle \leq b_i \ \forall i \in \mathcal{I}, \ \boldsymbol{X} = \boldsymbol{Y}\boldsymbol{X},$$

where $\boldsymbol{H} \in \mathcal{S}_+^{nm}, \boldsymbol{C} \in \mathcal{S}_+^n$ are positive semidefinite matrices, $\boldsymbol{D} \in \mathbb{R}^{n \times m}$ is a rectangular matrix, and $\mathcal{I}$ denotes the index set of constraints. As demonstrated in Bertsimas et al. (2022), any rank-constrained optimization problem of the form (5) can be formulated as an optimization over $(\boldsymbol{X}, \boldsymbol{Y})$ of the form (15), where the additional decision variable $\boldsymbol{Y}$ is a projection matrix which encodes the span of $\boldsymbol{X}$ and whose trace bounds $\mathrm{rank}(\boldsymbol{X})$. Here, we write $\mathrm{vec}(\boldsymbol{X}^\top)$ rather than the mathematically equivalent $\mathrm{vec}(\boldsymbol{X})$ to simplify the notation in our relaxations. Since there exists a permutation matrix $\boldsymbol{K}_{n,m} \in \mathbb{R}^{nm \times nm}$ such that $\mathrm{vec}(\boldsymbol{A}^\top) = \boldsymbol{K}_{n,m} \mathrm{vec}(\boldsymbol{A})$ for any $\boldsymbol{A} \in \mathbb{R}^{n \times m}$ ($\boldsymbol{K}_{n,m}$ is also called a commutation matrix, see Magnus and Neudecker 1979), both formulations are equivalent.

Problem (15) is quite a general formulation. It models matrix completion objectives like $\sum_{(i,j) \in \Omega} (X_{i,j} - A_{i,j})^2$ (as we detail in Section 4.1) and optimal power flow terms like $X_{i,j}X_{k,l}$. As a result of this generality, it is also challenging to solve.

We now develop a convex relaxation of (15). We remark that previous works on developing low-rank relaxations like Bertsimas et al. (2023c), Kim et al. (2022) require a spectral or permutation invariant term in the objective to develop a valid convex relaxation, hence do not apply to (15). Thus, designing a computationally tractable convex relaxation for (15) is arguably an open problem. Following the Shor relaxation blueprint, we introduce matrices $\boldsymbol{W}_{x,x}$, $\boldsymbol{W}_{x,y}$, $\boldsymbol{W}_{y,y}$ to model the outer products $\mathrm{vec}(\boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{X}^\top)^\top$, $\mathrm{vec}(\boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{Y})^\top$, and $\mathrm{vec}(\boldsymbol{Y})\mathrm{vec}(\boldsymbol{Y})^\top$ respectively.

PROPOSITION 4. *The convex semidefinite optimization problem*

$$\min_{\substack{\boldsymbol{Y} \in \mathcal{S}_+^n : \boldsymbol{Y} \preceq \boldsymbol{I}, \mathrm{tr}(\boldsymbol{Y}) \leq k \\ \boldsymbol{W}_{y,y} \in \mathcal{S}_+^{n^2}}} \min_{\substack{\boldsymbol{X} \in \mathbb{R}^{n \times m} : \langle \boldsymbol{A}_i, \boldsymbol{X} \rangle \leq b_i, i \in \mathcal{I} \\ \boldsymbol{W}_{x,x} \in \mathcal{S}_+^{nm}, \boldsymbol{W}_{x,y} \in \mathbb{R}^{nm \times n^2}}} \quad \langle \boldsymbol{C}, \boldsymbol{Y} \rangle + \langle \boldsymbol{W}_{x,x}, \boldsymbol{H} \rangle + \langle \boldsymbol{D}, \boldsymbol{X} \rangle$$

$$s.t. \quad \begin{pmatrix} 1 & \mathrm{vec}(\boldsymbol{X}^\top)^\top & \mathrm{vec}(\boldsymbol{Y})^\top \\ \mathrm{vec}(\boldsymbol{X}^\top) & \boldsymbol{W}_{x,x} & \boldsymbol{W}_{x,y} \\ \mathrm{vec}(\boldsymbol{Y}) & \boldsymbol{W}_{x,y}^\top & \boldsymbol{W}_{y,y} \end{pmatrix} \succeq \boldsymbol{0}, \qquad (16)$$

$$\sum_{i=1}^n \boldsymbol{W}_{y,y}^{(i,i)} = \boldsymbol{Y}, \ \sum_{i=1}^n \boldsymbol{W}_{x,y}^{(i,i)} = \boldsymbol{X}^\top$$

is a valid convex relaxation of Problem (15).

REMARK 2. If an optimal solution to (16) is such that $\boldsymbol{W}_{x,x}$ is a rank-one matrix then $\boldsymbol{W}_{x,x} = \mathrm{vec}(\boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{X}^\top)^\top$ and the optimal values of (16) and (15) coincide.

Proof of Proposition 4 Fix $(\boldsymbol{X}, \boldsymbol{Y})$ in (15) and set

$$(\boldsymbol{W}_{x,x}, \boldsymbol{W}_{x,y}, \boldsymbol{W}_{y,y}) := (\mathrm{vec}(\boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{X}^\top)^\top, \mathrm{vec}(\boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{Y})^\top, \mathrm{vec}(\boldsymbol{Y})\mathrm{vec}(\boldsymbol{Y})^\top).$$

It is sufficient to verify that $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{W}_{x,x}, \boldsymbol{W}_{x,y}, \boldsymbol{W}_{y,y})$ is feasible for (16)—it obviously attains the same objective value. First, by construction, the semidefinite constraint is satisfied (at equality). Moreover, we have

$$
\boldsymbol{Y}\boldsymbol{Y}^\top = \boldsymbol{Y} \implies \sum_{i=1}^{n} \boldsymbol{W}_{y,y}^{(i,i)} = \boldsymbol{Y},
$$
$$
\boldsymbol{X}^\top \boldsymbol{Y}^\top = \boldsymbol{X}^\top \implies \sum_{i\in[n]} \boldsymbol{W}_{x,y}^{(i,i)} = \boldsymbol{X}^\top.
$$

$\square$

Unfortunately, (16) is not compact and involves $n^2 \times n^2$ and $nm \times nm$ matrices. Therefore, a natural research question is whether it is possible to eliminate any variables from (16) without altering its optimal objective value. We answer this question affirmatively.

THEOREM 3. *Problem* (16) *is equivalent to*

$$
\min_{\substack{\boldsymbol{Y}\in\mathcal{S}_+^n\,:\,\boldsymbol{Y}\preceq\boldsymbol{I},\ \mathrm{tr}(\boldsymbol{Y})\leq k}} \quad \min_{\substack{\boldsymbol{X}\in\mathbb{R}^{n\times m}\,:\,\langle\boldsymbol{A}_i,\boldsymbol{X}\rangle\leq b_i,\ i\in\mathcal{I} \\ \boldsymbol{W}_{x,x}\in\mathcal{S}_+^{nm}}} \quad \langle\boldsymbol{C},\boldsymbol{Y}\rangle + \langle\boldsymbol{W}_{x,x},\boldsymbol{H}\rangle + \langle\boldsymbol{D},\boldsymbol{X}\rangle
$$
$$
s.t. \quad \boldsymbol{W}_{x,x} \succeq \mathrm{vec}(\boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{X}^\top)^\top, \tag{17}
$$
$$
\begin{pmatrix} \sum_{i\in[n]}\boldsymbol{W}_{x,x}^{(i,i)} & \boldsymbol{X}^\top \\ \boldsymbol{X} & \boldsymbol{Y} \end{pmatrix} \succeq \boldsymbol{0}.
$$

Proof of Theorem 3 We show that given a feasible solution to either problem, we can generate an optimal solution to the other problem with an equal or lower objective value.

Suppose that $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{W}_{x,x}, \boldsymbol{W}_{x,y}, \boldsymbol{W}_{y,y})$ is feasible in (16). Then, by summing appropriate semidefinite submatrices of the overall PSD matrix, we have that

$$
\begin{pmatrix} \sum_{i\in[n]}\boldsymbol{W}_{x,x}^{(i,i)} & \sum_{i\in[n]}\boldsymbol{W}_{x,y}^{(i,i)} \\ \sum_{i\in[n]}\boldsymbol{W}_{x,y}^{(i,i)\top} & \sum_{i\in[n]}\boldsymbol{W}_{y,y}^{(i,i)} \end{pmatrix} \succeq \boldsymbol{0}.
$$

Moreover, from (16) we have that $\sum_{i\in[n]} \boldsymbol{W}_{x,y}^{(i,i)} = \boldsymbol{X}^\top$ and $\sum_{i\in[n]} \boldsymbol{W}_{y,y}^{(i,i)} = \boldsymbol{Y}$. Thus, $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{W}_{x,x})$ is feasible in (17) and attains the same objective value.

Next, suppose that $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{W}_{x,x})$ is feasible in (17). By the Schur complement lemma, we must have $\boldsymbol{Y} \succeq \boldsymbol{X}(\sum_i \boldsymbol{W}_{x,x}^{(i,i)})^\dagger \boldsymbol{X}^\top$. Since $\boldsymbol{C} \succeq \boldsymbol{0}$, we can set $\boldsymbol{Y} := \boldsymbol{X}(\sum_i \boldsymbol{W}_{x,x}^{(i,i)})^\dagger \boldsymbol{X}^\top$ without loss of optimality—doing so cannot increase the objective value. To construct admissible matrices $\boldsymbol{W}_{x,y}$ and $\boldsymbol{W}_{y,y}$, let us first define the auxiliary matrix

$$
\boldsymbol{U} := \left(\sum_{i\in[n]} \boldsymbol{W}_{x,x}^{(i,i)} \right)^\dagger \boldsymbol{X}^\top \in \mathbb{R}^{m\times n},
$$

and observe that $\boldsymbol{Y} = \boldsymbol{U}^\top \boldsymbol{X}^\top = \boldsymbol{X}\boldsymbol{U}$. Then, we define $\boldsymbol{W}_{x,y}$, $\boldsymbol{W}_{y,y}$ as the blocks of the matrix

$$
\boldsymbol{M} := \begin{pmatrix} 1 & \mathrm{vec}(\boldsymbol{X}^\top)^\top & \mathrm{vec}(\boldsymbol{Y})^\top \\ \mathrm{vec}(\boldsymbol{X}^\top) & \boldsymbol{W}_{x,x} & \boldsymbol{W}_{x,y} \\ \mathrm{vec}(\boldsymbol{Y}) & \boldsymbol{W}_{x,y}^\top & \boldsymbol{W}_{y,y} \end{pmatrix}
$$

defined as

$$M := \begin{pmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{nm} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I}_n \otimes \boldsymbol{U} \end{pmatrix}^{\top} \begin{pmatrix} 1 & \text{vec}(\boldsymbol{X}^{\top})^{\top} & \text{vec}(\boldsymbol{X}^{\top})^{\top} \\ \text{vec}(\boldsymbol{X}^{\top}) & \boldsymbol{W}_{x,x} & \boldsymbol{W}_{x,x} \\ \text{vec}(\boldsymbol{X}^{\top}) & \boldsymbol{W}_{x,x} & \boldsymbol{W}_{x,x} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{nm} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I}_n \otimes \boldsymbol{U} \end{pmatrix}.$$

Since $\boldsymbol{Y} = \boldsymbol{U}^{\top} \boldsymbol{X}^{\top}$, we have $\text{vec}(\boldsymbol{Y}) = \text{vec}(\boldsymbol{U}^{\top} \boldsymbol{X}^{\top}) = (\boldsymbol{I}_n \otimes \boldsymbol{U}^{\top})\text{vec}(\boldsymbol{X}^{\top})$ and thus our construction is consistent with the existing value of $\boldsymbol{Y}$. We now verify that $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{W}_{x,x}, \boldsymbol{W}_{x,y}, \boldsymbol{W}_{y,y})$ is feasible for (16). By construction, $\boldsymbol{M} \succeq \mathbf{0}$. Thus, $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{W}_{x,x}, \boldsymbol{W}_{x,y}, \boldsymbol{W}_{y,y})$ satisfies the semidefinite constraint in (16). Next, by construction, $\boldsymbol{W}_{x,y}$ and $\boldsymbol{W}_{y,y}$ can be decomposed into $n \times n$ blocks satisfying:

$$\boldsymbol{W}_{x,y}^{(i,j)} = \boldsymbol{W}_{x,x}^{(i,j)} \boldsymbol{U}, \ \boldsymbol{W}_{y,y}^{(i,j)} = \boldsymbol{U}^{\top} \boldsymbol{W}_{x,x}^{(i,j)} \boldsymbol{U}.$$

Summing the on-diagonal blocks of these matrices then reveals that

$$\sum_{i \in [n]} \boldsymbol{W}_{x,y}^{(i,i)} = \sum_{i \in [n]} \boldsymbol{W}_{x,x}^{(i,i)} \boldsymbol{U} = \left(\sum_{i \in [n]} \boldsymbol{W}_{x,x}^{(i,i)} \right) \left(\sum_{j \in [n]} \boldsymbol{W}_{x,x}^{(j,j)} \right)^{\dagger} \boldsymbol{X}^{\top} = \boldsymbol{X}^{\top},$$

$$\sum_{i \in [n]} \boldsymbol{W}_{y,y}^{(i,i)} = \sum_{i \in [n]} \boldsymbol{U}^{\top} \boldsymbol{W}_{x,x}^{(i,i)} \boldsymbol{U} = \boldsymbol{U}^{\top} \left(\sum_{i \in [n]} \boldsymbol{W}_{x,x}^{(i,i)} \boldsymbol{U} \right) = \boldsymbol{U}^{\top} \boldsymbol{X}^{\top} = \boldsymbol{Y}.$$

Therefore, we conclude that $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{W}_{x,x}, \boldsymbol{W}_{x,y}, \boldsymbol{W}_{y,y})$ is feasible in (16) and attains an equal or lower objective value. Thus, both relaxations are equivalent. $\square$

Problem (17) is much more compact than (16), as it does not require introducing the variables $\boldsymbol{W}_{y,y} \in \mathcal{S}_+^{n^2}$ or $\boldsymbol{W}_{x,y} \in \mathbb{R}^{nm \times n}$. The proof of Theorem 3 provides a recipe for reconstructing an optimal $\boldsymbol{W}_{y,y}$ given an optimal solution $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{W}_{x,x})$ to (17). Namely, compute the auxiliary matrix $\boldsymbol{U} := \left(\sum_{i \in [n]} \boldsymbol{W}_{x,x}^{(i,i)} \right)^{\dagger} \boldsymbol{X}^{\top}$ and set $\boldsymbol{W}_{y,y} := (\boldsymbol{I}_n \otimes \boldsymbol{U})^{\top} \boldsymbol{W}_{x,x} (\boldsymbol{I}_n \otimes \boldsymbol{U})$. With this observation, one can implement the Goemans-Williamson sampling scheme for $\boldsymbol{Y}$ we propose in Section 3.3, even without solving a relaxation that explicitly involves $\boldsymbol{W}_{y,y}$.

Finally, it is interesting to consider whether the relaxation developed here is at least as strong as the matrix perspective relaxation developed by Bertsimas et al. (2023c). We now prove this is indeed the case. Bertsimas et al. (2023c) only applies to partially separable objectives. Hence, we first need to impose more structure on the objective of (15) to compare relaxations.

PROPOSITION 5. *Assume that the term* $\langle \boldsymbol{H}, \text{vec}(\boldsymbol{X}^{\top})\text{vec}(\boldsymbol{X}^{\top})^{\top} \rangle + \langle \boldsymbol{D}, \boldsymbol{X} \rangle$ *in Problem* (15) *can be rewritten as the partially separable term* $\frac{1}{2\gamma}\|\boldsymbol{X}\|_F^2 + h(\boldsymbol{X})$, *where h is convex in $\boldsymbol{X}$. Then, the optimal value of Problem* (16) *is at least as large as the relaxation of Bertsimas et al.* (2023c)

$$\min_{\boldsymbol{Y} \in \text{Conv}\left(\mathcal{Y}_n^k\right)} \min_{\boldsymbol{X} \in \mathbb{R}^{n \times m}, \boldsymbol{\theta} \in \mathcal{S}_+^m} \quad \langle \boldsymbol{C}, \boldsymbol{Y} \rangle + \frac{1}{2\gamma}\text{tr}(\boldsymbol{\theta}) + h(\boldsymbol{X}) \tag{18}$$

$$s.t. \quad \langle \boldsymbol{A}_i, \boldsymbol{X} \rangle \le b_i \ \forall i \in \mathcal{I}, \ \begin{pmatrix} \boldsymbol{\theta} & \boldsymbol{X}^{\top} \\ \boldsymbol{X} & \boldsymbol{Y} \end{pmatrix} \succeq \mathbf{0},$$

Proof of Proposition 5 Given the equivalence between Problems (16)–(17) proven in Theorem 3, it suffices to show that the constraints in (17) imply the constraints in (18). Letting $\boldsymbol{\theta} :=$

$\sum_{i\in[n]} \boldsymbol{W}_{x,x}^{(i,i)}$, we observe that $\boldsymbol{\theta}$ is feasible for (18). In addition, given the additional assumption that the objective involves $\|\boldsymbol{X}\|_F^2 = \mathrm{tr}(\boldsymbol{X}^\top \boldsymbol{X})$, the objective in the relaxation is

$$\langle \boldsymbol{H}, \boldsymbol{W}_{x,x} \rangle = \mathrm{tr}\left(\sum_{i\in[n]} \boldsymbol{W}_{x,x}^{(i,i)} \right) = \mathrm{tr}(\boldsymbol{\theta}),$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

The proof of Proposition 5 reveals that our Shor relaxation (17) can be perceived as decomposing the variable $\boldsymbol{\theta}$ in (18), and strengthening the relaxation by imposing additional constraints on the elements of this decomposition.

3.2. Strategies for Strengthening the Shor Relaxation

Theorem 3 might give the unfair impression that Problem (16) is not a useful relaxation, because it is equivalent to the much more compact optimization problem (17). However, explicit decision variables $\boldsymbol{W}_{y,y}, \boldsymbol{W}_{x,y}$ allow us to express additional valid inequalities to strengthen the relaxation:

- The matrix $\boldsymbol{Y}$ being symmetric, $\mathrm{vec}(\boldsymbol{Y}) = \mathrm{vec}(\boldsymbol{Y}^\top) = \boldsymbol{K}_{n,n}\mathrm{vec}(\boldsymbol{Y})$, which leads to the constraints

$$\begin{aligned} \mathrm{vec}(\boldsymbol{Y})\mathrm{vec}(\boldsymbol{Y})^\top = \boldsymbol{K}_{n,n}\mathrm{vec}(\boldsymbol{Y})\mathrm{vec}(\boldsymbol{Y})^\top \boldsymbol{K}_{n,n}^\top &\implies \boldsymbol{W}_{y,y} = \boldsymbol{K}_{n,n}\boldsymbol{W}_{y,y}\boldsymbol{K}_{n,n}^\top, \\ \mathrm{vec}(\boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{Y})^\top = \mathrm{vec}(\boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{Y})^\top \boldsymbol{K}_{n,n}^\top &\implies \boldsymbol{W}_{x,y} = \boldsymbol{W}_{x,y}\boldsymbol{K}_{n,n}^\top. \end{aligned} \tag{19}$$

- If we further require the matrix $\boldsymbol{X}$ to be symmetric (implying $n = m$), then we can impose the additional linear equalities $\boldsymbol{W}_{x,x} = \boldsymbol{K}_{n,n}\boldsymbol{W}_{x,x}\boldsymbol{K}_{n,n}^\top$ and $\boldsymbol{W}_{x,y} = \boldsymbol{K}_{n,n}\boldsymbol{W}_{x,y}$.

- As in binary optimization, we can impose triangle inequalities on $\boldsymbol{Y}$ and $\boldsymbol{W}_{yy}$. Indeed, from the fact that $0 \leq Y_{i,i} \leq 1$, we have that any triplet (i,j,ℓ) satisfies

$$(1 - Y_{i,i})(1 - Y_{j,j})(1 - Y_{\ell,\ell}) \geq 0$$

$$\iff 1 - Y_{i,i} - Y_{j,j} - Y_{\ell,\ell} + Y_{i,i}Y_{j,j} + Y_{i,i}Y_{\ell,\ell} + Y_{j,j}Y_{\ell,\ell} - Y_{i,i}Y_{j,j}Y_{\ell,\ell} \geq 0$$

$$\implies 1 - Y_{i,i} - Y_{j,j} - Y_{\ell,\ell} + Y_{i,i}Y_{j,j} + Y_{i,i}Y_{\ell,\ell} + Y_{j,j}Y_{\ell,\ell} \geq 0,$$

which can be expressed as a linear constraint in $(\boldsymbol{Y}, \boldsymbol{W}_{yy})$ after replacing each bilinear term with the appropriate entry of $\boldsymbol{W}_{yy}$. We can derive additional triangle inequalities by starting from the fact that $Y_{i,i}(1 - Y_{j,j})(1 - Y_{\ell,\ell}) \geq 0$ or $Y_{i,i}Y_{j,j}(1 - Y_{\ell,\ell}) \geq 0$. Triangle inequalities involving $Y_{i,j} \in [-1,1]$ rather than $Y_{i,i}$ follow similarly.

Finally, similarly to BQO, one can tighten Problem (16) and Problem (17) by applying RLT. Any constraint of the form $\boldsymbol{A}\mathrm{vec}(\boldsymbol{X}) \leq \boldsymbol{b}$ leads to the valid inequalities $\boldsymbol{A}\boldsymbol{W}_{x,x}\boldsymbol{A}^\top + \boldsymbol{b}\boldsymbol{b}^\top \geq \boldsymbol{b}\mathrm{vec}(\boldsymbol{X})^\top\boldsymbol{A} + \boldsymbol{A}\mathrm{vec}(\boldsymbol{X})\boldsymbol{b}^\top$, as reviewed by Bao et al. (2011).[4]

3.3. Generalization of Goemans-Williamson Rounding to Low-Rank Optimization

Mirroring the extension of Goemans-Williamson to mixed-integer optimization problems with logical constraints (see Section EC.1.2), we now extend Algorithm 2 to low-rank optimization. First, we observe that under the constraint $\boldsymbol{X} = \boldsymbol{Y}\boldsymbol{X}$, the term

$$\langle \mathrm{vec}(\boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{X}^\top)^\top, \boldsymbol{H}\rangle + \langle \boldsymbol{D}, \boldsymbol{X}\rangle$$

in the objective function of (15) is, through the identity $\mathrm{vec}(\boldsymbol{X}^\top) = (\boldsymbol{I}_n \otimes \boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{Y})$, equal to

$$\langle \mathrm{vec}(\boldsymbol{Y})\mathrm{vec}(\boldsymbol{Y})^\top, (\boldsymbol{I}_n \otimes \boldsymbol{X}^\top)\boldsymbol{H}(\boldsymbol{I}_n \otimes \boldsymbol{X}^\top)\rangle + \langle \boldsymbol{Y}, \boldsymbol{X}^\top \boldsymbol{D}\rangle.$$

Thus, Problem (15) can be rewritten as the following optimization problem

$$
\min_{\boldsymbol{X} \in \mathbb{R}^{n \times m}} \min_{\boldsymbol{Y} \in \mathcal{Y}_n^k} \quad \langle \boldsymbol{C} + \boldsymbol{X}^\top \boldsymbol{D}, \boldsymbol{Y}\rangle + \langle \mathrm{vec}(\boldsymbol{Y})\mathrm{vec}(\boldsymbol{Y})^\top, (\boldsymbol{I}_n \otimes \boldsymbol{X}^\top)\boldsymbol{H}(\boldsymbol{I}_n \otimes \boldsymbol{X}^\top)\rangle
$$
$$
\text{s.t.} \quad \langle \boldsymbol{A}_i, \boldsymbol{X}\rangle \le b_i \ \forall i \in \mathcal{I}, \ \boldsymbol{X} = \boldsymbol{Y}\boldsymbol{X},
$$

(20)

where the lower-level optimization problem is quadratic in $\boldsymbol{Y}$ and very much reminiscent of the orthogonally constrained problem studied in Section 2. This suggests that Algorithm 2 is a good candidate for generating feasible solutions to (20). We formalize our algorithm for rank-constrained optimization in Algorithm 3.

Algorithm 3 A Goemans-Williamson Rounding Method for Low-Rank Optimization

---

Generate $(\boldsymbol{Y}^\star, \boldsymbol{W}_{y,y}^\star)$ a solution to the semidefinite relaxation (16)

Compute $\hat{\boldsymbol{Y}} : \mathrm{vec}(\hat{\boldsymbol{Y}}) \sim \mathcal{N}\left(\mathrm{vec}(\boldsymbol{Y}^\star), \boldsymbol{W}_{y,y}^\star - \mathrm{vec}(\boldsymbol{Y}^\star)\mathrm{vec}(\boldsymbol{Y}^\star)^\top\right)$

Construct $\bar{\boldsymbol{Y}} \in \mathcal{Y}_n^k$ which solves $\min_{\boldsymbol{Y} \in \mathcal{Y}_n^k} \|\boldsymbol{Y} - \hat{\boldsymbol{Y}}\|_F$ (by performing an eigendecomposition)

Compute $\bar{\boldsymbol{X}}(\bar{\boldsymbol{Y}})$, an optimal $\boldsymbol{X}$ given $\bar{\boldsymbol{Y}}$ by solving

$$
\min_{\boldsymbol{X} \in \mathbb{R}^{n \times m}} \quad \langle \boldsymbol{C}, \boldsymbol{Y}\rangle + \langle \mathrm{vec}(\boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{X}^\top)^\top, \boldsymbol{H}\rangle + \langle \boldsymbol{D}, \boldsymbol{X}\rangle
$$
$$
\text{s.t.} \quad \langle \boldsymbol{A}_i, \boldsymbol{X}\rangle \le b_i \ \forall i \in \mathcal{I}, \ \boldsymbol{X} = \bar{\boldsymbol{Y}}\boldsymbol{X}
$$

return $\bar{\boldsymbol{Y}}, \bar{\boldsymbol{X}}(\bar{\boldsymbol{Y}})$ a feasible solution to (20)

We make the following remarks on our implementation of Algorithm 3

- It is challenging to produce a constant factor approximation on the performance of Algorithm 3. Indeed, Problem (15) models sparse regression as a special case, and it is strongly NP-hard (Chen et al. 2019, Theorem 1) to find a $O(n^{c_1} d^{c_2})$-approximation of sparse regression, where n is the number of data samples, d is the number of features, and $c_1 + c_2 < 1$. Moreover, in our numerical experiments, we uncover matrix completion instances where the optimal objective

value is non-zero but our Shor relaxation returns an objective value of zero, which suggests that Algorithm 3 does not provide a purely multiplicative constant factor approximation. Nonetheless, as we observe in numerical experiments (Section 5), it sometimes returns solutions within an optimality gap of 0–5% and thus is of interest.

- To obtain a solution to our Shor relaxation, $(\boldsymbol{Y}^\star, \boldsymbol{W}_{y,y}^\star)$, we can either solve (16), or solve the equivalent compact relaxation (17) and reconstruct $\boldsymbol{W}_{y,y}^\star$ as $\boldsymbol{W}_{y,y}^\star := (\boldsymbol{I}_n \otimes \boldsymbol{U})^\top \boldsymbol{W}_{x,x}(\boldsymbol{I}_n \otimes \boldsymbol{U})$ where $\boldsymbol{U} := \left(\sum_{i \in [n]} \boldsymbol{W}_{x,x}^{(i,i)} \right)^\dagger \boldsymbol{X}^\top$.

- We take the second moment of our Gaussian distribution to be $\boldsymbol{W}_{y,y}^\star - \mathrm{vec}(\boldsymbol{Y}^\star)\mathrm{vec}(\boldsymbol{Y}^\star)^\top$ so that $\mathbb{E}[\mathrm{vec}(\hat{\boldsymbol{Y}})\mathrm{vec}(\hat{\boldsymbol{Y}})^\top] = \boldsymbol{W}_{y,y}^\star$.

- The sampled matrices $\hat{\boldsymbol{Y}}$ are not necessarily symmetric in general. However, if $\boldsymbol{W}_{y,y}^\star$ satisfies (19), $\hat{\boldsymbol{Y}}$ is symmetric almost surely. Consequently, it could be beneficial to sample using a moment matrix that satisfies the permutation-invariance constraints (19). In numerical experiments, we investigate the benefits of projecting $\boldsymbol{W}_{y,y}^\star$ onto the set of matrices satisfying (19) before sampling.

- In practice, we randomly round multiple times from the solution to the Shor relaxation and return the best solution $\bar{\boldsymbol{Y}}$ found, rather than only rounding once. This repetition improves the quality of the returned solution significantly, and comes at a low increase in computational cost because solving the Shor relaxation is more expensive than sampling $\bar{\boldsymbol{Y}}$ and computing $\bar{\boldsymbol{X}}(\bar{\boldsymbol{Y}})$.

4. Examples of Low-Rank Relaxations

This section applies the Shor relaxation technique proposed in §3 to several important problems from the low-rank literature. By exploiting problem structure, we demonstrate that it is often possible to reduce our Shor relaxation to a relaxation that does not involve any $n^2 \times n^2$ matrices.

4.1. Matrix Completion

Given a random sample $\{A_{i,j} : (i,j) \in \Omega \subseteq [n] \times [m]\}$ of a matrix $\boldsymbol{A} \in \mathbb{R}^{n \times m}$, the goal of the low-rank matrix completion problem is to reconstruct the matrix $\boldsymbol{A}$, by assuming it is approximately low-rank (Candès and Recht 2009). This problem admits the formulation:

$$\min_{\boldsymbol{Y} \in \mathcal{Y}_n} \min_{\boldsymbol{X} \in \mathbb{R}^{n \times m}} \quad \|\mathcal{P}(\boldsymbol{A}) - \mathcal{P}(\boldsymbol{X})\|_F^2 + \lambda \cdot \mathrm{tr}(\boldsymbol{Y}) \text{ s.t.} \quad \boldsymbol{X} = \boldsymbol{Y}\boldsymbol{X}, \tag{21}$$

where $\lambda > 0$ is a penalty multiplier on the rank of $\boldsymbol{X}$ through the trace of $\boldsymbol{Y}$, and

$$\mathcal{P}(\boldsymbol{A})_{i,j} = \begin{cases} A_{i,j} & \text{if } (i,j) \in \Omega \\ 0 & \text{otherwise} \end{cases}$$

is a linear map which masks the hidden entries of $\boldsymbol{A}$. By expanding the quadratic $\|\mathcal{P}(\boldsymbol{A}) - \mathcal{P}(\boldsymbol{X})\|_F^2$, and invoking Theorem 3, we obtain the following relaxation of (21)

$$
\min_{\boldsymbol{Y} \in \text{Conv}(\mathcal{Y}_n)} \min_{\boldsymbol{X} \in \mathbb{R}^{n \times m}, \boldsymbol{W} \in \mathcal{S}_+^{nm}} \sum_{i \in [n]} \langle \boldsymbol{W}^{(i,i)}, \boldsymbol{H}^i \rangle - 2 \langle \mathcal{P}(\boldsymbol{X}), \mathcal{P}(\boldsymbol{A}) \rangle + \langle \mathcal{P}(\boldsymbol{A}), \mathcal{P}(\boldsymbol{A}) \rangle + \lambda \cdot \text{tr}(\boldsymbol{Y})
$$
$$
\text{s.t.} \quad \boldsymbol{W} \succeq \text{vec}(\boldsymbol{X}^\top)\text{vec}(\boldsymbol{X}^\top)^\top, \quad \begin{pmatrix} \sum_{i \in [n]} \boldsymbol{W}_{x,x}^{(i,i)} & \boldsymbol{X}^\top \\ \boldsymbol{X} & \boldsymbol{Y} \end{pmatrix} \succeq \boldsymbol{0},
\tag{22}
$$

where $\boldsymbol{H}^i$ is a diagonal matrix which takes entries $\boldsymbol{H}_j^i = 1$ if $(i,j) \in \Omega$ and $\boldsymbol{H}_j^i = 0$ otherwise.

Compared with the matrix perspective relaxation of Bertsimas et al. (2023c), our relaxation is directly applicable to (21), while Bertsimas et al. (2023c) requires the presence of an additional Frobenius regularization term $+\frac{1}{2\gamma}\|\boldsymbol{X}\|_F^2$ in the objective. With this additional term, our approach leads to relaxations of the form (22) after redefining $\boldsymbol{H}_i \leftarrow \boldsymbol{H}_i + \frac{1}{2\gamma}\boldsymbol{I}_m$, which are at least as strong as the relaxation of Bertsimas et al. (2023c) per Proposition 5.

We observe that the off-diagonal blocks of $\boldsymbol{W}$ do not appear in either the objective of (22) or any constraints other than $\boldsymbol{W} \succeq \text{vec}(\boldsymbol{X}^\top)\text{vec}(\boldsymbol{X}^\top)^\top$. For this reason, we can omit them entirely:

PROPOSITION 6. *Problem* (22) *attains the same optimal objective value as*

$$
\min_{\boldsymbol{Y} \in \text{Conv}(\mathcal{Y}_n)} \min_{\boldsymbol{X} \in \mathbb{R}^{n \times m}, \boldsymbol{S}^i \in \mathcal{S}_+^m} \sum_{i \in [n]} \langle \boldsymbol{S}^i, \boldsymbol{H}^i \rangle - 2 \langle \mathcal{P}(\boldsymbol{X}), \mathcal{P}(\boldsymbol{A}) \rangle + \langle \mathcal{P}(\boldsymbol{A}), \mathcal{P}(\boldsymbol{A}) \rangle + \lambda \cdot \text{tr}(\boldsymbol{Y})
$$
$$
s.t. \quad \boldsymbol{S}^i \succeq \boldsymbol{X}_{i,.}\boldsymbol{X}_{i,.}^\top, \quad \begin{pmatrix} \sum_{i \in [n]} \boldsymbol{S}^i & \boldsymbol{X}^\top \\ \boldsymbol{X} & \boldsymbol{Y} \end{pmatrix} \succeq \boldsymbol{0}.
\tag{23}
$$

Proof of Proposition 6 It suffices to show that given any feasible solution to (23) we can construct a feasible solution to (22) with the same objective value; the converse is immediate. Let $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{S}^i)$ be feasible in (23). Define the block matrix $\boldsymbol{W}$ by setting $\boldsymbol{W}^{(i,i)} = \boldsymbol{S}^i$ and $\boldsymbol{W}^{(i,j)} = (\boldsymbol{X}^\top)_i (\boldsymbol{X}^\top)_j^\top$. Then, it is not hard to see that $\boldsymbol{W} - \text{vec}(\boldsymbol{X}^\top)\text{vec}(\boldsymbol{X}^\top)^\top$ is a block matrix with zero off-diagonal blocks and on-diagonal blocks $\boldsymbol{S}^i - \boldsymbol{X}_{i,.}\boldsymbol{X}_{i,.}^\top \succeq \boldsymbol{0}$. Thus, $\boldsymbol{W} - \text{vec}(\boldsymbol{X}^\top)\text{vec}(\boldsymbol{X}^\top)^\top$ is a positive semidefinite matrix, and $\boldsymbol{W} \succeq \text{vec}(\boldsymbol{X}^\top)\text{vec}(\boldsymbol{X}^\top)^\top$. Moreover, $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{W})$ is feasible in (22) and attains the same objective value. $\square$

REMARK 3. Suppose that two rows of $\boldsymbol{A}$ have an identical sparsity pattern with respect to the known entries Ω, i.e., $\boldsymbol{H}^i = \boldsymbol{H}^j$. Then, we can replace the matrices $\boldsymbol{S}^i, \boldsymbol{S}^j$ with their sum $\tilde{\boldsymbol{S}}^{i,j} := \boldsymbol{S}^i + \boldsymbol{S}^j$ and rewrite (23) even more compactly, by omitting the matrices $\boldsymbol{S}^i, \boldsymbol{S}^j$, substituting $\tilde{\boldsymbol{S}}^{i,j}$ for $\boldsymbol{S}^i + \boldsymbol{S}^j$ in the objective/constraints, and requiring that $\tilde{\boldsymbol{S}}^{i,j} \succeq \boldsymbol{X}_{i,.}\boldsymbol{X}_{i,.}^\top + \boldsymbol{X}_{j,.}\boldsymbol{X}_{j,.}^\top$. This observation also applies if $k \geq 2$ rows share the same sparsity pattern. Moreover, if two rows of $\boldsymbol{A}$ have a similar sparsity pattern but with minor differences, we can obtain a computationally cheaper yet looser relaxation by masking all entries that do not appear in both rows, i.e., setting $\boldsymbol{H}^i, \boldsymbol{H}^j = \boldsymbol{H}^i \times \boldsymbol{H}^j$ and proceeding as before.

In Section EC.6 we support our discussion on low-rank matrix completion by demonstrating that analogous reductions hold for low-rank basis pursuit.

4.2. Reduced Rank Regression

Given a response matrix $\boldsymbol{B} \in \mathbb{R}^{n \times m}$ and a predictor matrix $\boldsymbol{A} \in \mathbb{R}^{n \times p}$, an important problem in high-dimensional statistics is to recover a low-complexity model which relates the matrices $\boldsymbol{B}$ and $\boldsymbol{A}$. A popular choice for doing so is to assume that $\boldsymbol{B}, \boldsymbol{A}$ are related via $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{E}$, where $\boldsymbol{X} \in \mathbb{R}^{p \times m}$ is a coefficient matrix, $\boldsymbol{E}$ is a matrix of noise, and we require that the rank of $\boldsymbol{X}$ is small so that the linear model is parsimonious Negahban and Wainwright (2011). This gives:

$$\min_{\boldsymbol{X} \in \mathbb{R}^{p \times m}} \quad \|\boldsymbol{B} - \boldsymbol{A}\boldsymbol{X}\|_F^2 + \mu \cdot \operatorname{rank}(\boldsymbol{X}), \tag{24}$$

where $\mu > 0$ controls the complexity of the estimator. For this problem, our Shor relaxation (17) is equivalent to the (improved) matrix perspective relaxation of Bertsimas et al. (2023c).

Indeed, by invoking Theorem 3, we obtain (24)'s Shor relaxation

$$\min_{\boldsymbol{Y} \in \operatorname{Conv}(\mathcal{Y}_m)} \min_{\boldsymbol{X} \in \mathbb{R}^{p \times m}, \boldsymbol{W} \in \mathcal{S}_+^{pm}} \quad \left\langle \boldsymbol{A}^\top \boldsymbol{A}, \sum_{i \in [m]} \boldsymbol{W}^{(i,i)} \right\rangle + \langle \boldsymbol{B}, \boldsymbol{B} \rangle - 2\langle \boldsymbol{A}\boldsymbol{X}, \boldsymbol{B} \rangle + \mu \cdot \operatorname{tr}(\boldsymbol{Y})$$

$$\text{s.t.} \quad \boldsymbol{W} \succeq \operatorname{vec}(\boldsymbol{X})\operatorname{vec}(\boldsymbol{X})^\top, \; \begin{pmatrix} \sum_{i \in [m]} \boldsymbol{W}^{(i,i)} & \boldsymbol{X} \\ \boldsymbol{X}^\top & \boldsymbol{Y} \end{pmatrix} \succeq \boldsymbol{0}, \tag{25}$$

for which we show the following equivalence result:

PROPOSITION 7. *Problem* (25) *attains the same objective value as*

$$\min_{\boldsymbol{Y} \in \operatorname{Conv}(\mathcal{Y}_m)} \min_{\boldsymbol{X} \in \mathbb{R}^{p \times m}, \boldsymbol{\theta} \in \mathcal{S}_+^p} \quad \langle \boldsymbol{A}^\top \boldsymbol{A}, \boldsymbol{\theta} \rangle + \langle \boldsymbol{B}, \boldsymbol{B} \rangle - 2\langle \boldsymbol{A}\boldsymbol{X}, \boldsymbol{B} \rangle + \mu \cdot \operatorname{tr}(\boldsymbol{Y})$$

$$s.t. \quad \begin{pmatrix} \boldsymbol{\theta} & \boldsymbol{X} \\ \boldsymbol{X}^\top & \boldsymbol{Y} \end{pmatrix} \succeq \boldsymbol{0}, \tag{26}$$

which corresponds to the improved relaxation of Bertsimas et al. (2023c, Equation 7)

Proof of Proposition 7   We show that for any solution to (26) one can construct a solution to (25) with the same objective value or vice versa. Indeed, for any feasible solution $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{W})$ to (25), $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{\theta} = \sum_{i \in [m]} \boldsymbol{W}^{(i,i)})$ is feasible for (26) with the same objective value. Conversely, let us consider $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{\theta})$ a feasible solution to (26). Then,

$$\begin{pmatrix} \boldsymbol{\theta} & \boldsymbol{X}^\top \\ \boldsymbol{X} & \boldsymbol{I}_m \end{pmatrix} = \begin{pmatrix} \boldsymbol{\theta} & \boldsymbol{X}^\top \\ \boldsymbol{X} & \boldsymbol{Y} \end{pmatrix} + \begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0}^\top & \boldsymbol{I} - \boldsymbol{Y} \end{pmatrix} \succeq \boldsymbol{0},$$

because both matrices are PSD given that $\boldsymbol{Y} \preceq \boldsymbol{I}$. Therefore, it follows from the Schur complement lemma that $\boldsymbol{\theta} \succeq \boldsymbol{X}\boldsymbol{X}^\top = \sum_{i \in [m]} \boldsymbol{X}_i \boldsymbol{X}_i^\top$. Thus, there exists a decomposition $\boldsymbol{\theta} = \sum_{i \in [m]} \boldsymbol{S}^i$ with $\boldsymbol{S}_i \succeq \boldsymbol{X}_i \boldsymbol{X}_i^\top$ for each i. In particular, by assigning $\boldsymbol{X}_i \boldsymbol{X}_i^\top$ to each $\boldsymbol{S}^i$, plus the remaining $\boldsymbol{\theta} - \sum_i \boldsymbol{X}_i \boldsymbol{X}_i^\top$ arbitrarily between the $\boldsymbol{S}^i$'s. Finally, let us define the matrix $\boldsymbol{W}$ such that $\boldsymbol{W}^{(i,i)} = \boldsymbol{S}^i$ and $\boldsymbol{W}^{(i,j)} = \boldsymbol{X}_i \boldsymbol{X}_j^\top$ for $i \neq j$. Then, $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{W})$ is feasible for (25) and attains the same objective value. The relaxation (26) is precisely the relaxation developed in Bertsimas et al. (2023c). $\qquad\square$

Proposition 7's proof technique uses the fact that $\boldsymbol{X}$ enters the objective quadratically via $\boldsymbol{X}\boldsymbol{X}^\top$, rather than properties specific to reduced rank regression. This suggests other low-rank problems

which are quadratic through $\boldsymbol{X}\boldsymbol{X}^\top$ (or $\boldsymbol{X}^\top\boldsymbol{X}$), e.g., low-rank factor analysis (Bertsimas et al. 2017), sparse plus low-rank matrix decompositions (Bertsimas et al. 2023a) and quadratically constrained programming (Wang and Kılınç-Karzan 2022) admit similarly compact Shor relaxations.

We have shown in this section that for two quadratic low-rank problems, it is possible to eliminate enough variables in the Shor relaxation that no matrices of size $n^2 \times n^2$ remain. Further, we observe that the same reduction holds for a third problem, namely low-rank basis pursuit, in Section EC.6 of the Electronic Complement. Thus, we argue that our proof technique is very general, and can likely be applied to other low-rank problems of practical interest (e.g., sensor location). This suggests that while Shor relaxations involving $n^2 \times n^2$ matrices may appear to be too large to be useful in practice, they can often be reduced to forms that are useful.

5. Numerical Results

In this section, we benchmark our relax-then-round schemes on synthetic semi-orthogonal quadratic and low-rank matrix completion problems. We also compare the performance of our schemes with the matrix perspective relaxation proposed by Bertsimas et al. (2023c). We emphasize that we introduce $\|\cdot\|_F^2$ regularization to perform this comparison, as the relaxation of Bertsimas et al. (2023c) is not applicable to generic low-rank quadratic optimization.

All experiments are conducted on a MacBook Pro laptop with a 36 GB Apple M3 CPU, using MOSEK version 10.1, Julia version 1.9, and JuMP.jl version 1.13.0. All solver parameters are set to their default values. We divide our discussion into two parts. First, in §5.1, we study the quality of our relax-and-round scheme for semi-orthogonal quadratic problems. Second, in §5.2 we investigate the quality of our relax-and-round scheme for low-rank matrix completion problems and compare with prior literature.

5.1. Semi-Orthogonal Quadratic Optimization

We evaluate the performance of our Shor relaxation and Algorithm 2 for semi-orthogonal quadratic optimization problems (4). For fixed (n, m), we generate a random semidefinite matrix $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{B}^\top \in \mathcal{S}_+^{nm}$ where the entries of $\boldsymbol{B} \in \mathbb{R}^{nm \times 10}$ are standard independent random variables. We solve the Shor relaxation (9) and sample $N = 100$ feasible solutions from Algorithm 2. For comparison, we also implement the following benchmarks:

1. We sample N solutions uniformly at random (Uniform, as analyzed in Section 2.4).
2. We sample N solutions using Algorithm 2 but generate $\boldsymbol{Q}$ by projecting the rectangular matrix $\boldsymbol{G}$ onto the set of semi-orthogonal matrices directly (Algorithm 2 with projection).
3. We sample N solutions by applying a deflation heuristic (Deflation, described in Section EC.5).
4. We follow the heuristic in Burer and Park (2024), namely, we project the reshaped leading eigenvector of $\boldsymbol{W}^\star$ (Burer and Park).
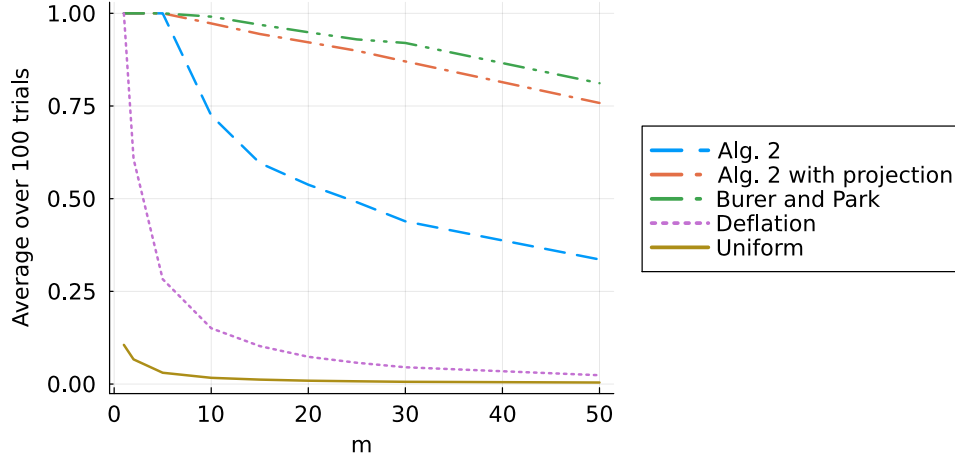
Figure 1   Average performance ratio $\langle \boldsymbol{A}, \mathrm{vec}(\boldsymbol{Q})\,\mathrm{vec}(\boldsymbol{Q})^{\top}\rangle / \langle \boldsymbol{A}, \boldsymbol{W}^{\star}\rangle$ over $N = 100$ generated solutions for different feasibility heuristics. Note that the method of Burer and Park (2024) is deterministic (always returns the same solution for a given instance). For each value of m, results are averaged over 5 instances.

We consider $n = 50$ and $m \in \{1, 2, 5, 10, 15, 20, 25, 30, 50\}$. We generate five instances for each (n, m).

Figure 1 compares the average performance ratio $\langle \boldsymbol{A}, \hat{\mathbb{E}}^{N}[\mathrm{vec}(\boldsymbol{Q})\,\mathrm{vec}(\boldsymbol{Q})^{\top}]\rangle / \langle \boldsymbol{A}, \boldsymbol{W}^{\star}\rangle$, for these five algorithms. Confirming our theoretical analysis, we observe that the average performance ratio degrades as m increases. We also observe that our Algorithm 2 strongly outperforms Deflation and Uniform —theoretically, Uniform and Deflation achieve a $1/nm$- and $1/m^2$-performance guarantee respectively (Proposition 3 and EC.4). A crucial step in the theoretical analysis of Algorithm 2 is the fact that we generate a feasible matrix $\boldsymbol{Q}$ from a randomly generated matrix $\boldsymbol{G}$, by randomly switching the singular values of $\boldsymbol{G}$ to ± 1. Instead, we find that using a deterministic projection (Alg. 2 with projection) leads to much stronger performance, comparable to that of the heuristic in Burer and Park (2024). However, as observed in Section 2.5, our analysis cannot be easily generalized to such deterministic projection schemes.

In practice, one might be interested in the performance of the best solution found, rather than the average performance over N solutions. Figure 2 reports the best performance of each method, and shows that the relative ordering of the methods remains unchanged, although the difference in performance between methods shrinks.

5.2. Low-Rank Matrix Completion

In this section, we evaluate the performance of Algorithm 3 on synthetic low-rank matrix completion instances. We use the data generation process of Candès and Recht (2009): We construct a matrix of observations, $\boldsymbol{A}_{\mathrm{full}} \in \mathbb{R}^{n \times m}$, from a rank-$r$ model: $\boldsymbol{A}_{\mathrm{full}} = \boldsymbol{UV} + \epsilon \boldsymbol{Z}$, where the entries of $\boldsymbol{U} \in \mathbb{R}^{n \times r}, \boldsymbol{V} \in \mathbb{R}^{r \times m}$, and $\boldsymbol{Z} \in \mathbb{R}^{n \times m}$ are drawn independently from a standard normal distribution, and
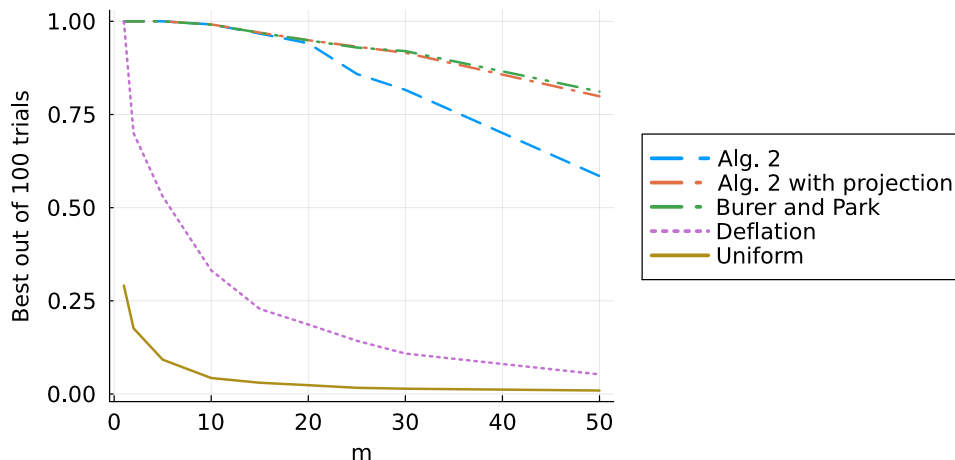
Figure 2 Performance ratio $\langle \boldsymbol{A}, \text{vec}(\boldsymbol{Q})\,\text{vec}(\boldsymbol{Q})^\top \rangle / \langle \boldsymbol{A}, \boldsymbol{W}^\star \rangle$ of the best out of $N = 100$ solutions for different feasibility heuristics. Note that the method of Burer and Park (2024) only returns one solution. For each value of m, results are averaged over 5 instances.

$\epsilon \geq 0$ models the degree of noise. We fix $\epsilon = 0.1, m = n$ and $r = 2$ for all experiments. We sample a random subset $\Omega \subseteq [n] \times [m]$, of predefined size (see also Candès and Recht 2009, section 1.1.2). Each result reported in this section is averaged over 10 random seeds.

We first evaluate the quality of our new relaxations, compared with the matrix perspective relaxation of Bertsimas et al. (2023c, MPRT). Unfortunately, MPRT does not apply to (21) as it requires a Frobenius regularization term in the objective. Hence, instead of (21), we consider

$$\min_{\boldsymbol{X} \in \mathbb{R}^{n \times m}} \quad \frac{1}{2\gamma} \|\boldsymbol{X}\|_F^2 + \frac{1}{2} \sum_{(i,j) \in \Omega} (A_{i,j} - X_{i,j})^2 \text{ s.t. } \text{rank}(\boldsymbol{X}) \leq r.$$

for some regularization parameter $\gamma > 0$. As $\gamma \to \infty$, we recover the solution of (21). We compare the (lower) bounds obtained by three different approaches: MPRT, our full Shor relaxation (16) with the permutation equalities (19), hereafter denoted "Shor-Perm", and our compact Shor relaxation (23) ("Shor-Red"). Figure 3 reports the lower bounds achieved by each approach —in relative terms compared with an upper bound achieved by the alternating minimization method of Burer and Monteiro (2003) initialized with a truncated SVD of $\mathcal{P}(\boldsymbol{A})$ (absolute values are reported in Figures EC.1–EC.2)— as γ increases, for different proportion of entries sampled $p = |\Omega|/mn$ ($n = 8$ being fixed). Supporting Proposition 5, we observe that Shor-Perm and Shor-Red obtain smaller optimality gaps than MPRT, for all values of γ, and that the benefit increases as the fraction of sampled entries p increases. In particular, when $p = 0.95$, there is a regime of values of γ (around 10^2) where both Shor relaxations are tight (as evidenced by a gap of 0%), while MPRT is not. In addition, as γ increases, MPRT achieves an uninformative gap of 100% (by returning a trivial lower bound of 0, see Figure EC.1), while our Shor relaxations provide non-trivial bounds (and

(a) $p = 0.5$ (b) $p = 0.95$
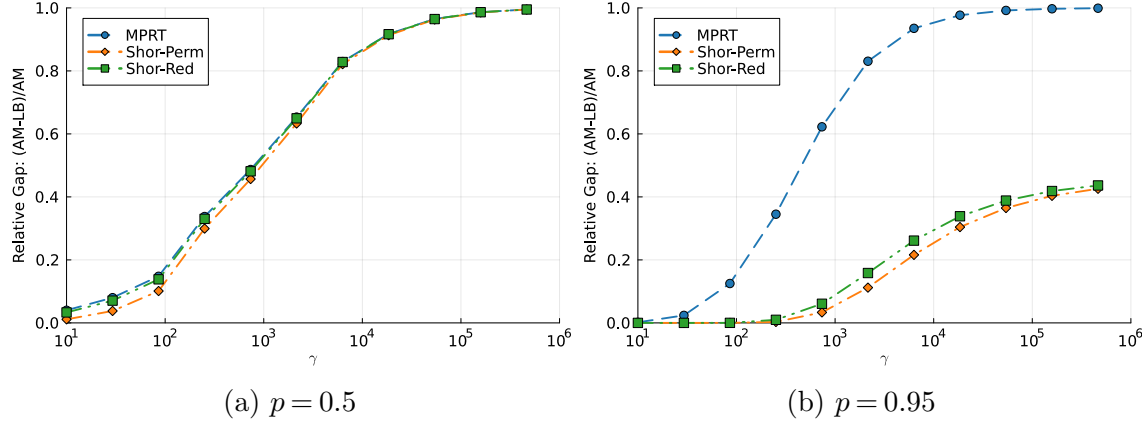
Figure 3   Relative gap obtained with different relaxations of the regularized matrix completion problem as we
vary γ. We fix $n = 8$. Results are averaged over 10 replications.

gaps). From this experiment, it seems that imposing the permutation equalities (19) on $\boldsymbol{W}_{y,y}$ in our Shor relaxation (Shor-Perm vs. Shor-Red) does not lead to significantly tighter bounds, while being computationally much more expensive (see Figure EC.3 for computational times).

Our second experiment investigates the performance of our rounding strategy for the Shor relaxations, on the same instances. The relaxation Shor-Perm provides a matrix $\boldsymbol{W}_{y,y}$ directly. From a solution to the compact relaxation Shor-Red, we can reconstruct a matrix $\boldsymbol{W}_{y,y}$ using the reconstruction strategy discussed in Section 3.3. Figure 4 reports the best upper bound found from 1,000 sampled solutions. Interestingly, we observe that while the lower bounds from both relaxations in Figure 3 are rather similar, Shor-Perm provides a substantial improvement in the quality of the upper bound obtained, especially for higher values of p. Intuitively, this can be explained by the fact that the constraints (19) ensure that the sampled solution $\hat{\boldsymbol{Y}}$ is symmetric almost surely, hence is closer to being feasible. However, the matrix $\boldsymbol{W}_{y,y}$ recovered from Shor-Red does not satisfy these constraints. To support this intuition, we consider a third approach where we project the matrix $\boldsymbol{W}_{y,y}$ recovered from (19) onto the set $\{\boldsymbol{W} \in \mathcal{S}_+^{nm} : \boldsymbol{W} = \boldsymbol{K}_{n,m} \boldsymbol{W} \boldsymbol{K}_{n,m}^\top\}$ before sampling ("Shor-Red-Proj"). As displayed on the right panel of Figure 4, this additional projection step improves the quality of the solutions sampled from Shor-Red further, without significant additional computational cost, thus we use this projection technique for the rest of our numerics.

On the same instances, our third experiment compares Goemans-Williamson rounding with two other methods for generating feasible solutions: taking a truncated SVD of the MPRT relaxation (as advocated in Bertsimas et al. 2023c, "MPRT + Greedy") and alternating minimization initialized with a truncated SVD of $\mathcal{P}(\boldsymbol{A})$ ("AM"). Figure 5 depicts the upper bounds achieved by each method. Among the rounding-based schemes, we observe that Goemans-Williamson rounding on Shor-Perm performs significantly better than MPRT + Greedy when $p = 0.5$ and comparably
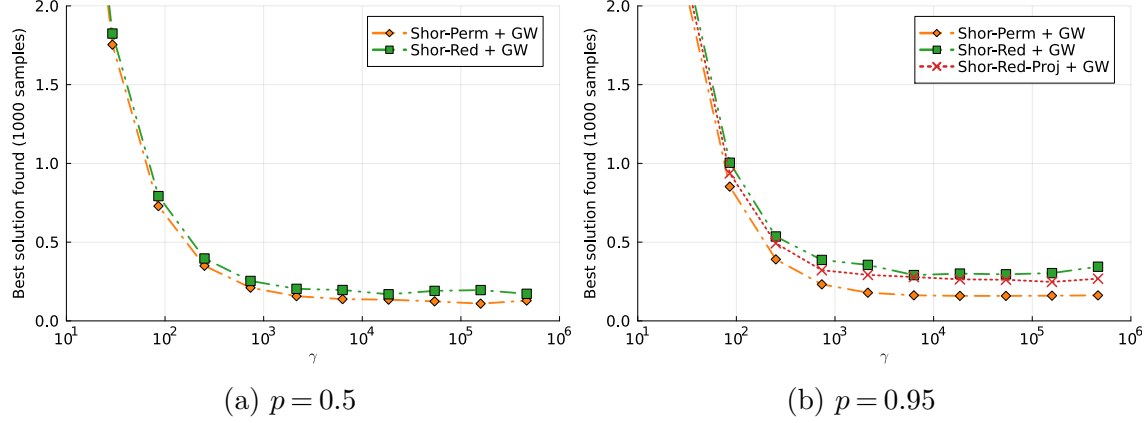
Figure 4   Average quality of GW rounding as we vary γ for rounding the full Shor relaxation ("GW-Full") and the reduced relaxation with and without projecting $\boldsymbol{W}_{y,y}$ ("GW-Red-Proj", "GW-Red-NoProj").
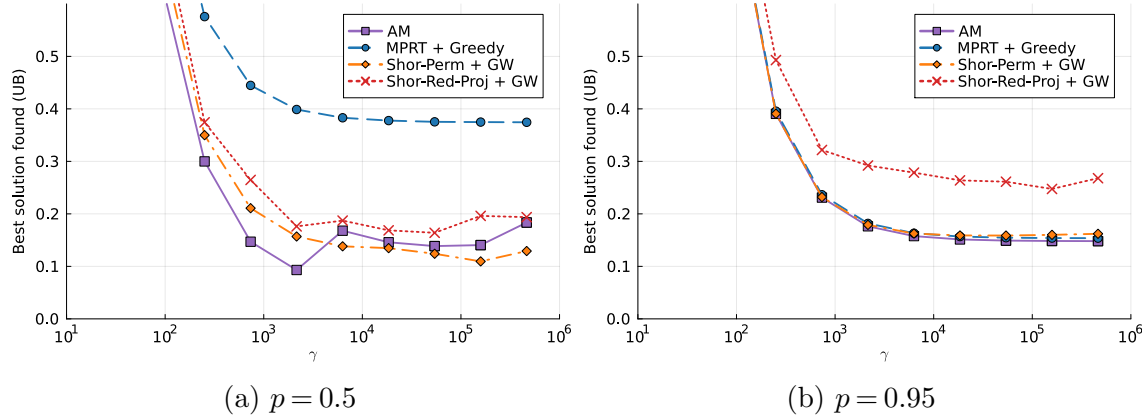
**Figure 5**   Average quality of feasible methods as we vary γ, for GW rounding on the full Shor relaxation ("GW-Full"), on the reduced relaxation with projecting $\boldsymbol{W}_{y,y}$ ("GW-Red-Proj"), greedily rounding the matrix perspective relaxation ("MPRT-GD"), and alternating minimization ("AM").

when $p = 0.95$. The alternating minimization method of Burer and Monteiro (2003) is generally the best-performing method, except for instances with $p = 0.5$ and $\gamma \geq 10^4$. For these particularly challenging instances, which have many local optima, our Goemans-Williamson rounding could serve as an alternative or the initialization of the AM algorithm.

Our final experiment benchmarks the scalability of our reduced Shor relaxation and Goemans-Williamson rounding as we vary $n = m$ with the proportion of entries fixed at $p = 0.5$. We set $\gamma = 10^4/n^2$. We report the average upper and lower bound (divided by n^2 so that quantities have the same meaning as we vary n; left) and the average computational time (right) in Figure 6. We also report the average objective value obtained by alternating minimization as a baseline. Note that we do not consider the full Shor relaxation in this experiment, as it requires more RAM than is available for these experiments when $n = 10$. For any $n \in \{4, \ldots, 42\}$, the Shor relaxation can be

solved in seconds, while when $n > 44$, Mosek runs out of RAM. Moreover, the lower bound from the Shor relaxation is tight for $n \geq 18$, although only alternating minimization matches the bound.

(a) Objective values   (b) Runtimes
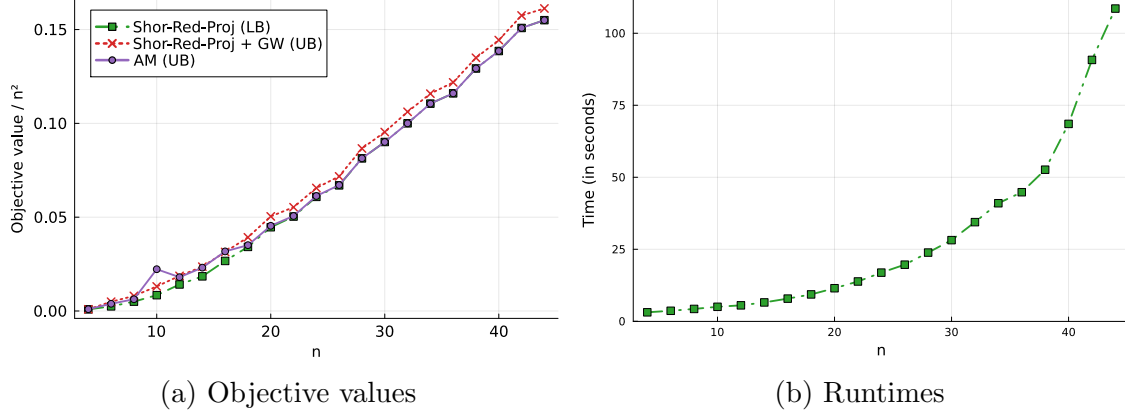
Figure 6   Objective value (left panel) and runtime for Shor-Red-Proj (right panel) as we vary $n = m$ with $p = 0.5$ for our reduced Shor relaxation followed by Goemans-Williamson rounding. Results are averaged over 10 replications.

6. Conclusion

This paper proposes a new technique for relaxing and rounding quadratic optimization problems over semi-orthogonal matrices, and generalizes it to a broader class of low-rank optimization problems. We obtain a new semidefinite relaxation by vectorizing the matrices and modeling the outer product of this vectorization with itself. By exploiting problem structure to eliminate most of the variables in our semidefinite relaxations, we show how to solve our relaxation efficiently. By interpreting the new decision variables in these relaxations as the second moment of a multivariate Gaussian distribution, we propose a sampling procedure, reminiscent of the Goemans-Williamson algorithm for BQO, which, as demonstrated throughout our numerical experiments, obtains high-quality solutions to low-rank problems in polynomial time.

Endnotes

1. Imposing the constraint $\text{rank}(\sum_{i \in [m]} \boldsymbol{W}^{(i,i)}) \leq m$, which is equivalent to a rank-one constraint on $\boldsymbol{W}$ under trace and $\preceq \boldsymbol{I}_n$ constraints, would also suffice.

2. Explicitly, this follows from Pataki (1998, theorem 2.2). After introducing a slack matrix $\boldsymbol{S}$ for the semidefinite inequality $\sum_{i \in [m]} \boldsymbol{W}^{(i,i)} \preceq \boldsymbol{I}_n$, i.e., writing $\boldsymbol{S} = \boldsymbol{I}_n - \sum_{i \in [m]} \boldsymbol{W}^{(i,i)}$, we have $m(m+1)/2 + n(n+1)/2$ scalar inequalities. Thus, there exists some optimal solution $(\boldsymbol{W}, \boldsymbol{S})$ with $\text{rank}(\boldsymbol{W})(\text{rank}(\boldsymbol{W}) + 1)/2 + \text{rank}(\boldsymbol{S})(\text{rank}(\boldsymbol{S}) + 1)/2 \leq m(m+1)/2 + n(n+1)/2$, which implies

$\mathrm{rank}(\boldsymbol{W})(\mathrm{rank}(\boldsymbol{W})+1)/2 \le m(m+1)/2+n(n+1)/2$. Since $m(m+1)/2+n(n+1)/2 \le (n+m)(n+m+1)/2$, this implies $\mathrm{rank}(\boldsymbol{W}) \le n+m$.

3. First, as we show in Section EC.4.2, the Kronecker constraints of Burer and Park (2024) do not rule out any feasible matrices $\boldsymbol{W}$ that are diagonal. We show in Theorem 2 that for $m=1$ there exists a worst-case $\boldsymbol{W}^\star$ which is diagonal for each n. Thus, Kronecker constraints cannot improve the worst-case approximation ratio when $m=1$. Second, Kronecker constraints would also not rule out the family of worst-case instances identified in Proposition 2, because they do not rule out any matrices $\boldsymbol{W}$ where each block matrix $\boldsymbol{W}^{(i,j)}$ is rank-one, as occurs for that family of instances (see Section EC.4.2). Thus, Kronecker constraints cannot improve the order of our approximation guarantee.

4. One may also be tempted to impose the inequalities $\boldsymbol{A}\boldsymbol{W}_{x,x}\boldsymbol{A}^\top + \boldsymbol{b}\boldsymbol{b}^\top \succeq \boldsymbol{b}\mathrm{vec}(\boldsymbol{X})^\top\boldsymbol{A} + \boldsymbol{A}\mathrm{vec}(\boldsymbol{X})\boldsymbol{b}^\top$ but they are actually implied by $\boldsymbol{W}_{x,x} \succeq \mathrm{vec}(\boldsymbol{X}^\top)\mathrm{vec}(\boldsymbol{X}^\top)^\top$ and $\boldsymbol{A}\mathrm{vec}(\boldsymbol{X}) \le \boldsymbol{b}$.

References

Bandeira AS, Kennedy C, Singer A (2016) Approximating the little Grothendieck problem over the orthogonal and unitary groups. *Mathematical Programming* 160:433–475.

Bao X, Sahinidis NV, Tawarmalani M (2011) Semidefinite relaxations for quadratically constrained quadratic programming: A review and comparisons. *Mathematical Programming* 129:129–157.

Barak B, Kelner JA, Steurer D (2014) Rounding sum-of-squares relaxations. *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, 31–40.

Barratt S (2018) A matrix Gaussian distribution. *arXiv preprint arXiv:1804.11010* .

Barvinok A (2001) A remark on the rank of positive semidefinite matrices subject to affine constraints. *Discrete & Computational Geometry* 25(1):23–31.

Bell RM, Koren Y (2007) Lessons from the Netflix prize challenge. *ACM SIGKDD Explorations Newsletter* 9(2):75–79.

Ben-Tal A, Nemirovski A (2002) On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty. *SIAM Journal on Optimization* 12(3):811–833.

Bertsimas D, Copenhaver MS, Mazumder R (2017) Certifiably optimal low rank factor analysis. *Journal of Machine Learning Research* 18(29):1–53.

Bertsimas D, Cory-Wright R, Johnson NA (2023a) Sparse plus low rank matrix decomposition: A discrete optimization approach. *Journal of Machine Learning Research* 24(1):12478–12528.

Bertsimas D, Cory-Wright R, Lo S, Pauphilet J (2023b) Optimal low-rank matrix completion: Semidefinite relaxations and eigenvector disjunctions. *arXiv preprint arXiv:2305.12292* .

Bertsimas D, Cory-Wright R, Pauphilet J (2022) Mixed-projection conic optimization: A new paradigm for modeling rank constraints. *Operations Research* 70(6):3321–3344.

Bertsimas D, Cory-Wright R, Pauphilet J (2023c) A new perspective on low-rank optimization. *Mathematical Programming* 202:47–92.

Bertsimas D, Ye Y (1998) Semidefinite relaxations, multivariate normal distributions, and order statistics. *Handbook of Combinatorial Optimization*, 1473–1491 (Springer).

Bolla M, Michaletzky G, Tusnády G, Ziermann M (1998) Extrema of sums of heterogeneous quadratic forms. *Linear Algebra and its Applications* 269(1-3):331–365.

Boyd S, El Ghaoui L, Feron E, Balakrishnan V (1994) *Linear Matrix Inequalities in System and Control Theory* (SIAM).

Briët J, Buhrman H, Toner B (2011) A generalized Grothendieck inequality and nonlocal correlations that require high entanglement. *Communications in Mathematical Physics* 305(3):827–843.

Briët J, de Oliveira Filho FM, Vallentin F (2010) The positive semidefinite Grothendieck problem with rank constraint. *International Colloquium on Automata, Languages, and Programming*, 31–42 (Springer).

Burer S, Monteiro RD (2003) A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming* 95(2):329–357.

Burer S, Park K (2024) A strengthened SDP relaxation for quadratic optimization over the Stiefel manifold. *Journal of Optimization Theory and Applications* 202(1):320–339.

Candès EJ, Recht B (2009) Exact matrix completion via convex optimization. *Foundations of Computational Mathematics* 9(6):717–772.

Chen Y, Ye Y, Wang M (2019) Approximation hardness for a class of sparse optimization problems. *Journal of Machine Learning Research* 20(38):1–27.

Cory-Wright R, Pauphilet J (2022) Sparse PCA with multiple components. *arXiv preprint arXiv:2209.14790* .

Dong H, Chen K, Linderoth J (2015) Regularization vs. relaxation: A conic optimization perspective of statistical variable selection. *arXiv preprint arXiv:1510.06083* .

d'Aspremont A, Boyd S (2003) Relaxations and randomized methods for nonconvex QCQPs. *EE392o Class Notes, Stanford University* 1:1–16.

Gillis N, Glineur F (2011) Low-rank matrix approximation with weights or missing data is NP-hard. *SIAM Journal on Matrix Analysis and Applications* 32(4):1149–1165.

Gilman K, Burer S, Balzano L (2022) A semidefinite relaxation for sums of heterogeneous quadratics on the Stiefel manifold. *arXiv preprint arXiv:2205.13653* .

Goemans MX, Williamson DP (1995) Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)* 42(6):1115–1145.

Grimmett G, Stirzaker D (2020) *Probability and Random Processes* (Oxford University Press).

Gupta AK, Nagar DK (2018) *Matrix Variate Distributions* (Chapman and Hall/CRC).

Kim J, Tawarmalani M, Richard JPP (2022) Convexification of permutation-invariant sets and an application to sparse principal component analysis. *Mathematics of Operations Research* 74(4):2547–2584.

Lai Z, Lim LH, Tang T (2025) Stiefel optimization is NP-hard. *arXiv preprint arXiv:2507.02839* .

Li Y, Xie W (2024) On the partial convexification of the low-rank spectral optimization: Rank bounds and algorithms. *International Conference on Integer Programming and Combinatorial Optimization*, 265–279 (Springer).

Luo ZQ, Ma WK, So AMC, Ye Y, Zhang S (2010) Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Processing Magazine* 27(3):20–34.

Magnus JR, Neudecker H (1979) The commutation matrix: some properties and applications. *The Annals of Statistics* 7(2):381–394.

Meckes ES (2019) *The Random Matrix Theory of the Classical Compact Groups*, volume 218 (Cambridge University Press).

Negahban S, Wainwright MJ (2011) Estimation of (near) low-rank matrices with noise and high-dimensional scaling. *Annals of Statistics* 39:1069–1097.

Nemirovski A (2007) Sums of random symmetric matrices and quadratic optimization under orthogonality constraints. *Mathematical Programming* 109(2-3):283–317.

Nesterov Y (1998) Semidefinite relaxation and nonconvex quadratic optimization. *Optimization Methods and Software* 9(1-3):141–160.

Overton ML, Womersley RS (1992) On the sum of the largest eigenvalues of a symmetric matrix. *SIAM Journal on Matrix Analysis and Applications* 13(1):41–45.

Pataki G (1998) On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Mathematics of Operations Research* 23(2):339–358.

Recht B, Fazel M, Parrilo PA (2010) Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Review* 52(3):471–501.

Rehfeldt D, Koch T, Shinano Y (2023) Faster exact solution of sparse MaxCut and QUBO problems. *Mathematical Programming Computation* 15(3):445–470.

Shor NZ (1987) Quadratic optimization problems. *Soviet Journal of Computer and Systems Sciences* 25:1–11.

So AMC (2009) Improved approximation bound for quadratic optimization problems with orthogonality constraints. *Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, 1201–1209 (SIAM).

Wang AL, Kılınç-Karzan F (2022) On the tightness of SDP relaxations of QCQPs. *Mathematical Programming* 193(1):33–73.

Williamson DP, Shmoys DB (2011) *The Design of Approximation Algorithms* (Cambridge University Press).

Wolkowicz H, Saigal R, Vandenberghe L (1998) *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, volume 27 (Springer Science & Business Media).

Supplementary Material

EC.1. Goemans-Williamson and Logically Constrained Optimization

To build intuition for readers familiar with the mixed-integer optimization literature, we review, in this section, how the classical Goemans-Williamson algorithm for BQO has been generalized to mixed-integer optimization problems. Precisely, we review a semidefinite relaxation and randomized rounding scheme for logically constrained problems, which prepares the ground for the extension of the semidefinite relaxation and randomized rounding scheme from Section 2 to rank-constrained optimization in Section 3. Our algorithmics and some of our proof techniques in the main paper follow analogously to the techniques derived in the mixed-integer optimization case. However, with the exception of some of the proof techniques, the results in this section can be found in the mixed-integer optimization literature.

EC.1.1. A Shor Relaxation and Its Compact Version

We consider a quadratic optimization problem that unfolds over two stages, as occurs in sparse regression, portfolio selection, and network design problems; see Bertsimas et al. (2021) for a review. In the first stage, a decision-maker activates binary variables subject to resource budget constraints and activation costs. Subsequently, in the second stage, the decision-maker optimizes over the continuous variables. Formally, we consider the problem

$$\min_{\boldsymbol{z}\in\mathcal{Z}_n^k}\min_{\boldsymbol{x}\in\mathbb{R}^n}\quad \boldsymbol{c}^\top\boldsymbol{z}+\frac{1}{2}\boldsymbol{x}^\top\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{d}^\top\boldsymbol{x}\ \text{s.t.}\ \boldsymbol{A}\boldsymbol{x}\leq\boldsymbol{b},\ x_i=z_i x_i\ \forall i\in[n], \tag{EC.1}$$

where $\mathcal{Z}_n^k:=\{\boldsymbol{z}\in\{0,1\}^n:\boldsymbol{e}^\top\boldsymbol{z}\leq k\}$, $\boldsymbol{Q}\succeq\boldsymbol{0}$ is a positive semidefinite matrix, and $\boldsymbol{c}\in\mathbb{R}_+^n$. Note that the bilinear constraints $x_i=z_i x_i$ for $z_i\in\{0,1\}$ enforces the logical relationships '$x_i=0$ if $z_i=0$'.

Problem (EC.1) has a convex quadratic objective function. Therefore, a viable technique for obtaining a strong convex relaxation is introducing semidefinite matrices to model products of variables. This technique was first proposed by Shor (1987) in the context of non-convex quadratic optimization and has since been studied by many other authors; see Han et al. (2022) for a review. In particular, we introduce the block matrix $\boldsymbol{W}\in\mathcal{S}_+^{2n}$ to represent the outer product of $\begin{pmatrix}\boldsymbol{x}\\\boldsymbol{z}\end{pmatrix}$ with itself. Specifically, we partition $\boldsymbol{W}$ into four blocks: $\boldsymbol{W}_{x,x}$, $\boldsymbol{W}_{z,z}$, $\boldsymbol{W}_{x,z}$, and $\boldsymbol{W}_{x,z}^\top$, which model $\boldsymbol{x}\boldsymbol{x}^\top$, $\boldsymbol{z}\boldsymbol{z}^\top$, $\boldsymbol{x}\boldsymbol{z}^\top$, and $\boldsymbol{z}\boldsymbol{x}^\top$, respectively. With these additional variables, we have the following semidefinite relaxation for Problem (EC.1):

PROPOSITION EC.1. *The optimization problem*

$$\min_{\boldsymbol{z}\in[0,1]^n:\boldsymbol{e}^\top\boldsymbol{z}\leq k}\quad \min_{\substack{\boldsymbol{x}\in\mathbb{R}^n:\boldsymbol{A}\boldsymbol{x}\leq\boldsymbol{b}\\\boldsymbol{W}\in\mathcal{S}_+^{2n}}}\quad \boldsymbol{c}^\top\boldsymbol{z}+\frac{1}{2}\langle\boldsymbol{Q},\boldsymbol{W}_{x,x}\rangle+\boldsymbol{d}^\top\boldsymbol{x}$$

$$\text{s.t.}\quad \boldsymbol{W}\succeq\begin{pmatrix}\boldsymbol{x}\\\boldsymbol{z}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}\\\boldsymbol{z}\end{pmatrix}^\top,\ \text{diag}(\boldsymbol{W}_{z,z})=\boldsymbol{z},\ \text{diag}(\boldsymbol{W}_{x,z})=\boldsymbol{x}, \tag{EC.2}$$

is a valid convex relaxation of Problem (EC.1).

Proof of Proposition EC.1   It suffices to show that any feasible solution to (EC.1) corresponds to a feasible solution in (EC.2) with the same objective value. To see this, fix $\boldsymbol{z}, \boldsymbol{x}$ in (EC.1), and set

$$\boldsymbol{W} := \begin{pmatrix} \boldsymbol{W}_{x,x} & \boldsymbol{W}_{x,z} \\ \boldsymbol{W}_{x,z}^\top & \boldsymbol{W}_{z,z} \end{pmatrix} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{z} \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{z} \end{pmatrix}^\top.$$

Furthermore, $(\boldsymbol{W}_{z,z})_{i,i} = z_i^2 = z_i$ because z_i is binary, and $(\boldsymbol{W}_{x,z})_{i,i} = x_i z_i = x_i$. Hence, the solution $(\boldsymbol{z}, \boldsymbol{x}, \boldsymbol{W})$ is feasible in (EC.2) and attains the same objective value.   □

REMARK EC.1. Problem (EC.2) is a relaxation of Problem (EC.1) by allowing $\boldsymbol{z} \in [0,1]^n$ and by omitting the rank-1 constraint on $\boldsymbol{W}$. Reimposing the rank-one constraint obtains an equivalent reformulation of Problem (EC.1).

REMARK EC.2. We can strengthen Problem (EC.2) by applying the Reformulation-Linearization Technique (RLT; e.g., Bao et al. 2011) to the linear constraints on $\boldsymbol{x}$, $\boldsymbol{Ax} \leq \boldsymbol{b}$, leading to $\boldsymbol{A W}_{x,x} \boldsymbol{A}^\top + \boldsymbol{b b}^\top \geq \boldsymbol{b x}^\top \boldsymbol{A} + \boldsymbol{A x b}^\top$. All results follow identically with RLT constraints on $(\boldsymbol{x}, \boldsymbol{W}_{x,x})$.

While Problem (EC.2) is a valid convex relaxation, it may be expensive to solve, because it involves large semidefinite matrices. Surprisingly, Han et al. (2022) demonstrated that Problem (EC.2) is equivalent to the so-called "optimal perspective relaxation" originally proposed by Zheng et al. (2014), Dong et al. (2015), which is much more compact. We now recall this compact relaxation and prove its equivalence. We acknowledge that this result has been proven previously in Han et al. (2022, Theorem 6), in a non-constructive way. Here, we develop a new, constructive proof for it, which we will be able to extend to rank-constrained optimization.

PROPOSITION EC.2. *Problem* (EC.2) *is equivalent to*

$$\min_{\boldsymbol{z} \in [0,1]^n : \boldsymbol{e}^\top \boldsymbol{z} \leq k} \quad \min_{\substack{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{Ax} \leq \boldsymbol{b} \\ \boldsymbol{X} \in \mathcal{S}^n}} \quad \boldsymbol{c}^\top \boldsymbol{z} + \frac{1}{2} \langle \boldsymbol{Q}, \boldsymbol{X} \rangle + \boldsymbol{d}^\top \boldsymbol{x}$$

$$\text{s.t.} \quad \boldsymbol{X} \succeq \boldsymbol{x x}^\top, \ x_i^2 \leq X_{i,i} z_i, \forall i \in [n], \tag{EC.3}$$

Proposition EC.2 leverages our assumption that $\boldsymbol{c} \geq \boldsymbol{0}$ in our statement of (EC.1)–(EC.2) to show that we can solve the semidefinite relaxation (EC.2) by solving the much smaller semidefinite optimization problem (EC.3), which only involves one semidefinite variable $\boldsymbol{X} \in \mathcal{S}_+^n$, and reconstruct an optimal solution involving $\boldsymbol{W} \in \mathcal{S}_+^{2n}$ to (EC.2). Our proof of Proposition EC.2 makes this reconstruction step explicit.

Proof of Proposition EC.2   We show that any feasible solution to Problem (EC.2) generates a feasible solution to (EC.3) with an equal or lower objective and vice versa.

First, we consider a feasible solution to (EC.2), $(\boldsymbol{z}, \boldsymbol{x}, \boldsymbol{W})$, and show that the solution $(\boldsymbol{z}, \boldsymbol{x}, \boldsymbol{X}) = (\boldsymbol{z}, \boldsymbol{x}, \boldsymbol{W}_{x,x})$ is a feasible solution to (EC.3), with the same objective value. To establish feasibility,

we only need to verify that $x_i^2 \leq X_{i,i} z_i$, since the remaining constraints in (EC.3) are present in (EC.2). From the non-negativity of the 2×2 minors of the semidefinite matrix $\boldsymbol{W}$, we have $(\boldsymbol{W}_{x,x})_{i,i}(\boldsymbol{W}_{z,z})_{i,i} \geq (\boldsymbol{W}_{x,z})_{i,i}^2$. Substituting the identities $(\boldsymbol{W}_{x,z})_{i,i} = x_i$ and $(\boldsymbol{W}_{z,z})_{i,i} = z_i$ yields the result.

Next, consider a feasible solution $(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{X})$ to (EC.3). Observe that the constraint $x_i^2 \leq X_{i,i} z_i$ imposes $x_i = 0$ if $X_{i,i} = 0$. Since $\boldsymbol{c} \geq \boldsymbol{0}$, it follows that, if the constraint $z_i X_{i,i} \geq x_i^2$ is not binding for some index i, we can decrease z_i without impacting feasibility or worsening the objective value. Accordingly, we can assume $z_i = x_i^2 / X_{i,i}$ without loss of generality (with the convention $0/0 = 0$ so that $z_i = 0$ if $X_{i,i} = 0$). We now define a matrix $\boldsymbol{W}$ such that $(\boldsymbol{z}, \boldsymbol{x}, \boldsymbol{W})$ is feasible for (EC.2) and achieves the same objective value. Observe that the matrix $\boldsymbol{M}$

$$\underbrace{\begin{pmatrix} 1 & \boldsymbol{x}^\top & \boldsymbol{z}^\top \\ \boldsymbol{x} & \boldsymbol{X} & \boldsymbol{W}_{x,z} \\ \boldsymbol{z} & \boldsymbol{W}_{x,z}^\top & \boldsymbol{W}_{z,z} \end{pmatrix}}_{\boldsymbol{M}} := \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_n \\ \boldsymbol{0} & \mathrm{Diag}(\boldsymbol{u}) \end{pmatrix} \begin{pmatrix} 1 & \boldsymbol{x}^\top \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_n \\ \boldsymbol{0} & \mathrm{Diag}(\boldsymbol{u}) \end{pmatrix}^\top$$

with $u_i = \frac{x_i}{X_{i,i}}$ if $X_{i,i} > 0$ and 0 otherwise, is positive semidefinite as a positive semidefinite matrix, left and right multiplied by a matrix and the same matrix transposed. Hence, we consider the matrices $\boldsymbol{W}_{z,z}, \boldsymbol{W}_{x,z}$ as defined above. Moreover, we note that the vector $\boldsymbol{z}$ defined as a block of the matrix $\boldsymbol{M}$ is equal to our original $\boldsymbol{z}$. Indeed, $(\mathrm{Diag}(\boldsymbol{u})\boldsymbol{x})_i = (\boldsymbol{x} \circ \boldsymbol{u})_i = \frac{x_i^2}{X_{i,i}} = z_i$.

To complete the proof, we verify that $\boldsymbol{M}$ gives a feasible solution to (EC.2). First, by the Schur complement lemma, we have

$$\boldsymbol{M} \succeq \boldsymbol{0} \text{ if and only if } \boldsymbol{W} = \begin{pmatrix} \boldsymbol{X} & \boldsymbol{W}_{x,z} \\ \boldsymbol{W}_{x,z}^\top & \boldsymbol{W}_{z,z} \end{pmatrix} \succeq \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{z} \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{z} \end{pmatrix}^\top.$$

Second, by the definition of $\boldsymbol{W}_{z,z}$, we have

$$(\boldsymbol{W}_{z,z})_{ii} = X_{i,i} u_i^2 = \begin{cases} \frac{x_i^2}{X_{i,i}} & \text{if } X_{i,i} > 0 \\ 0 & \text{if } X_{i,i} = 0 \end{cases} = z_i,$$

because $x_i^2 / X_{i,i} = z_i$. Finally, by the definition of $\boldsymbol{W}_{x,z}$, we have

$$(\boldsymbol{W}_{x,z})_{ii} = X_{i,i} u_i = \begin{cases} x_i & \text{if } X_{i,i} > 0 \\ 0 & \text{if } X_{i,i} = 0 \end{cases} = x_i.$$

Therefore, $(\boldsymbol{z}, \boldsymbol{x}, \boldsymbol{W})$ is feasible in (EC.3) and attains an equal objective value. $\square$

We close this section by pointing out that Proposition EC.2 does *not* imply that Problem (EC.2) cannot be useful in practice. Indeed, the variables $\boldsymbol{W}_{z,z}$ and $\boldsymbol{W}_{x,z}$ enable to express constraints that further tighten the relaxation. For example, one can tighten Problem (EC.2)'s relaxation by imposing the so-called triangle inequalities on $(\boldsymbol{z}, \boldsymbol{W}_{z,z})$, as derived by Padberg (1989). As we demonstrate via a simple sparse linear regression example in Section EC.1.3, the equivalence demonstrated in Proposition EC.2 does not hold in the presence of these triangle inequalities.

EC.1.2. Goemans-Williamson Rounding for Logically Constrained Optimization

The equivalence result in Proposition EC.2 reveals that it is possible to reconstruct an optimal $\boldsymbol{W}_{z,z}$ given an optimal solution to the semidefinite relaxation (EC.3) that involves $\boldsymbol{z}, \boldsymbol{x}, \boldsymbol{X} = \boldsymbol{W}_{x,x}$ only. This raises the following research question: how to use the reconstructed solution $\boldsymbol{W}_{z,z}$ as part of a rounding scheme for constructing a high-quality solution to (EC.1). To answer this question, Dong et al. (2015) observe, in the context of sparse regression, that the variable $\boldsymbol{x}$ being fixed, the objective function in Problem (EC.1) is quadratic in $\boldsymbol{z}$, given that $x_i = z_i x_i$. This observation suggests that the rounding mechanism of Goemans and Williamson (1995) is a good candidate for generating high-quality feasible solutions $\boldsymbol{z}$ to (EC.1). In particular, rounding for a binary $\boldsymbol{z}$ using a Goemans-Williamson scheme, then solving for $\boldsymbol{x}$ with $\boldsymbol{z}$ being fixed to $\bar{\boldsymbol{z}}$. Accordingly, we now describe a Goemans-Williamson rounding to logically constrained quadratic optimization problems, in Algorithm EC.1 (see also Dong et al. (2015)).

Algorithm EC.1 Goemans-Williamson Rounding for Logically Constrained Optimization

---

Compute solution $\boldsymbol{z}^\star, \boldsymbol{W}_{z,z}^\star$ either by solving (EC.2), or solving(EC.3) and reconstructing $\boldsymbol{W}_{z,z}^\star$.

Sample $\hat{\boldsymbol{z}} \sim \mathcal{M}(\boldsymbol{z}^\star, \boldsymbol{W}_{z,z}^\star - \boldsymbol{z}^\star \boldsymbol{z}^{\star\top})$

Construct $\bar{\boldsymbol{z}} \in \{0,1\}^n : \bar{z}_i = \mathrm{Round}(\hat{z}_i)$.

Compute $\bar{\boldsymbol{x}}(\bar{\boldsymbol{z}})$, an optimal $\boldsymbol{x}$ given $\bar{\boldsymbol{z}}$ by solving

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} \quad \frac{1}{2} \boldsymbol{x}^\top \boldsymbol{Q} \boldsymbol{x} + \boldsymbol{d}^\top \boldsymbol{x} \text{ s.t. } \boldsymbol{A} \boldsymbol{x} \leq \boldsymbol{b}, \ x_i = 0 \text{ if } \bar{z}_i = 0, \forall i \in [n]$$

return $\bar{\boldsymbol{z}}, \bar{\boldsymbol{x}}(\bar{\boldsymbol{z}})$

We remark that $\hat{\boldsymbol{z}}$ is sampled according to a normal distribution with covariance matrix $\boldsymbol{W}_{z,z}^\star - \boldsymbol{z}^\star \boldsymbol{z}^{\star\top}$ in Algorithm EC.1 to ensure that $\mathbb{E}[\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}^\top] = \boldsymbol{W}_{z,z}^\star$, and thus the random solution $\hat{\boldsymbol{z}}$ is feasible and has an objective value equal to the optimal value of the semidefinite relaxation in expectation.

Unfortunately, it is challenging to produce a constant-factor approximation guarantee for Algorithm EC.1, as discussed for the case of sparse linear regression by Dong et al. (2015). This is perhaps unsurprising, indeed, solving logically constrained quadratic optimization problems is strongly NP-hard (Chen et al. 2019). Nonetheless, the Goemans-Williamson algorithm is useful in practice even in settings where we cannot obtain constant-factor theoretical guarantees.

In the main paper, we extend the Goemans-Williamson approach to logically constrained quadratic optimization, mirroring the extension to further generalize our Shor relaxation and Goemans-Williamson sampling scheme for semi-orthogonal quadratic optimization to low-rank quadratic optimization problems.

EC.1.3. Non-Equivalence of Shor and Optimal Perspective Relaxations

Consider a sparse linear regression problem setting of the form

$$\min_{\boldsymbol{\beta} \in \mathbb{R}^p} \quad \|\boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{y}\|_2^2 \text{ s.t. } \|\boldsymbol{\beta}\|_0 \leq k,$$

and its two semidefinite relaxations: (a) Problem (EC.2) reinforced with the triangle inequalities

$$z_i + z_j + z_l \leq Z_{i,j} + Z_{i,k} + Z_{j,k} + 1 \qquad\qquad \forall i,j,k \in [n],$$

$$Z_{i,j} + Z_{i,k} \leq z_i + Z_{j,k} \qquad\qquad \forall i,j,k \in [n],$$

and (b) the more compact semidefinite relaxation (EC.3), which as proven in Proposition EC.2 is equivalent to Problem (EC.2) (without the triangle inequalities).

Let the problem data be $p = 6, n = 8, k = 3$ and

$$\boldsymbol{X} = \begin{pmatrix} 1.04 & 0.97 & 0.35 & 0.34 & 0.04 & 0.62 \\ 1.13 & 1.08 & 0.66 & 0.78 & 0.85 & 0.45 \\ 1.50 & 2.54 & 1.73 & 0.11 & -1.06 & -0.41 \\ 0.65 & -1.42 & -1.52 & -1.03 & -0.11 & 0.81 \\ 0.49 & -1.17 & -1.58 & 0.60 & 0.70 & 1.53 \\ 0.51 & -1.34 & -1.53 & 0.07 & -0.10 & 0.17 \\ 0.81 & 2.63 & -0.90 & 1.73 & 1.36 & 1.73 \\ 0.76 & 0.71 & 0.08 & -0.20 & -0.57 & -0.13 \end{pmatrix}, \quad \boldsymbol{y} = \begin{pmatrix} 0.43 \\ 0.84 \\ 1.15 \\ -2.22 \\ -1.44 \\ -1.94 \\ -3.18 \\ -2.44 \end{pmatrix}.$$

Then, using Mosek version 10.2 to solve all relaxations and Gurobi version 10.0.2 to solve the mixed-integer problem:

- Relaxation (a) has an optimal objective value of 1.45886.

- Relaxation (b) has an optimal objective value of 1.4118.

- The original MINLO has an optimal objective value of 1.5336.

Thus, the Shor relaxation with triangle inequalities and the more compact semidefinite relaxation are not equivalent.

EC.2. Connection with Binary Quadratic Optimization and the Original Goemans-Williamson Algorithm

We now connect our quadratic semi-orthogonal optimization problem (4), its semidefinite relaxation (9), and our rounding algorithm (Algorithm 2) to the canonical binary quadratic optimization problem.

Consider a binary quadratic optimization problem (1). For each $i = 1, \ldots, n$, define $\boldsymbol{u}_i = z_i \boldsymbol{e}_i$. By construction, we have $\boldsymbol{u}_i^\top \boldsymbol{u}_j = 0$ if $i \neq j$ and $\boldsymbol{u}_i^\top \boldsymbol{u}_i = z_i^2 = 1$. With this notation,

$$\boldsymbol{z}^\top \boldsymbol{Q} \boldsymbol{z} = \sum_{i,j} Q_{i,j} z_i z_j = \sum_{i,j} \boldsymbol{u}_i^\top \boldsymbol{e}_i Q_{i,j} \boldsymbol{e}_j^\top \boldsymbol{u}_j = \sum_{i,j} \boldsymbol{u}_i^\top \boldsymbol{A}^{(i,j)} \boldsymbol{u}_j,$$

with $\boldsymbol{A}^{(i,j)} := Q_{i,j} \boldsymbol{e}_i \boldsymbol{e}_j^\top$. In particular, we have that $\boldsymbol{Q} \succeq \boldsymbol{0} \iff \boldsymbol{A} \succeq \boldsymbol{0}$. Hence, in the variable $\boldsymbol{U} = [\boldsymbol{u}_1 \cdots \boldsymbol{u}_n] \in \mathbb{R}^{n \times n}$ we have a quadratic semi-orthogonal problem of the form (4). We have the

additional constraints that each vector $\boldsymbol{u}_i$ is colinear to $\boldsymbol{e}_i$, i.e., constraints of the form $U_{j,i} = 0$ for $j \neq i$.

We introduce the variable $\boldsymbol{W} = \text{vec}(\boldsymbol{U})\,\text{vec}(\boldsymbol{U})^{\top}$ and write the semidefinite relaxation (9). The constraints on the support of $\boldsymbol{U}$ further impose that each block of $\boldsymbol{W}$ is of the form $\boldsymbol{W}^{(i,j)} = Z_{i,j}\boldsymbol{e}_i\boldsymbol{e}_j^{\top}$. The objective of (9) can thus be written as

$$\langle \boldsymbol{A}, \boldsymbol{W} \rangle = \sum_{i,j} \langle \boldsymbol{A}^{(i,j)}, \boldsymbol{W}^{(i,j)} \rangle = \sum_{i,j} Q_{i,j} Z_{i,j},$$

and the constraints on the matrix $\boldsymbol{W}$ are equivalent to:

$$\boldsymbol{W} \succeq \boldsymbol{0}: \qquad\qquad\qquad\qquad\qquad\qquad \boldsymbol{Z} \succeq \boldsymbol{0},$$

$$\text{tr}(\boldsymbol{W}^{(j,j')}) = \delta_{j,j'}: \qquad\qquad\qquad\qquad\qquad Z_{j,j} = 1,$$

$$\sum_{i \in [n]} \boldsymbol{W}^{(i,i)} \preceq \boldsymbol{I}_n: \qquad\qquad \text{Diag}(Z_{1,1}, \ldots, Z_{n,n}) \preceq \boldsymbol{I}_n.$$

So, we recover the semidefinite relaxation of BQO, (2), exactly.

Consider a solution to the semidefinite relaxation (9), $\boldsymbol{W}^{\star}$, and $\boldsymbol{Z}^{\star}$ such that $\boldsymbol{W}^{\star(i,j)} = Z_{i,j}^{\star}\boldsymbol{e}_i\boldsymbol{e}_j^{\top}$. By sampling $\text{vec}(\boldsymbol{G}) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{W}^{\star})$, the sparsity pattern of $\boldsymbol{W}^{\star}$ implies that each column of $\boldsymbol{G}$, $\boldsymbol{g}_i$, is of the form $\boldsymbol{g}_i = y_i\boldsymbol{e}_i$, with $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Z}^{\star})$. In this case, the matrix $\boldsymbol{G}$ is diagonal and its SVD can be written

$$\boldsymbol{G} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\top} := \boldsymbol{I}_n \begin{pmatrix} |y_1| & & \\ & \ddots & \\ & & |y_n| \end{pmatrix} \begin{pmatrix} \text{sign}(y_1) & & \\ & \ddots & \\ & & \text{sign}(y_n) \end{pmatrix}.$$

We then generate $\boldsymbol{Q} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^{\top}$ for some random diagonal matrix $\boldsymbol{D}$. Each column of $\boldsymbol{Q}$, $\boldsymbol{q}_i$, can be expressed as $\boldsymbol{q}_i = D_{i,i}\,\text{sign}(y_i)\boldsymbol{e}_i$ and we can identify the feasible solution to the BQO problem as $\hat{z}_i = D_{i,i}\,\text{sign}(y_i)$. If we had taken $\boldsymbol{D} = \boldsymbol{I}_n$ in Algorithm 2, then we would get $\boldsymbol{q}_i = \text{sign}(y_i)\boldsymbol{e}_i$, i.e., $\hat{z}_i = \text{sign}(y_i)$, which is precisely the original Goemans-Williamson algorithm (Algorithm 1). Instead, $D_{i,i} \in \{\pm 1\}$ is sampled at random with

$$\mathbb{P}(D_{i,i} = 1) = \frac{1}{2}\left(1 + \frac{\sigma_i}{\sigma_{\max}}\right) = \frac{1}{2}\left(1 + \frac{|y_i|}{\max_j |y_j|}\right).$$

We can interpret our algorithm as a regularization of the Goemans-Williamson procedure. If $|y_i|$ is very large, then we would get $D_{i,i} = 1$ with high probability, and we would follow the Goemans-Williamson rounding rule $\hat{z}_i = \text{sign}(y_i)$ closely. On the other hand, if $|y_i|$ is close to 0, we disregard the sign of y_i and instead sample $\hat{z}_i = \pm 1$ with probability 0.5

EC.3. Technical Appendix to Section 2

EC.3.1. Bounding the Largest Singular Values of the Stochastic Matrix G

In this section, we prove concentration results on $\sigma_{\max}(\boldsymbol{G})$ (Lemma 1).

As described in Section 2.2, in our implementation of Algorithm 2, we sample $\text{vec}(\boldsymbol{G}) \sim \mathcal{N}(\boldsymbol{0}_{nm}, \boldsymbol{W}^\star)$ as $\text{vec}(\boldsymbol{G}) = \sum_{k \in [r]} \text{vec}(\boldsymbol{B}_k) z_k$ with $\boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}_r, \boldsymbol{I}_r)$ and $\boldsymbol{W}^\star = \sum_{k \in [r]} \text{vec}(\boldsymbol{B}_k) \text{vec}(\boldsymbol{B}_k)^\top$ a Cholesky decomposition of $\boldsymbol{W}^\star$ This construction interprets $\boldsymbol{G}$ as a matrix series, $\boldsymbol{G} = \sum_{k \in [r]} \boldsymbol{B}_k z_k$, as studied in the statistics literature (see, e.g., Tropp 2015).

To analyze the behavior of $\sigma_{\max}(\boldsymbol{G})$, it is important to understand the spectral behavior of $\sum_k \boldsymbol{B}_k^\top \boldsymbol{B}_k$ and $\sum_k \boldsymbol{B}_k \boldsymbol{B}_k^\top$.

LEMMA EC.1. *Let $\boldsymbol{W}$ be a feasible solution of* (9) *and consider a Cholesky decomposition of $\boldsymbol{W}$, $\boldsymbol{W} = \sum_{k \in [r]} \text{vec}(\boldsymbol{B}_k) \text{vec}(\boldsymbol{B}_k)^\top$ with $r = \text{rank}(\boldsymbol{W})$ and $\boldsymbol{B}_k \in \mathbb{R}^{n \times m}$. Then, we have*

$$\sum_k \boldsymbol{B}_k^\top \boldsymbol{B}_k = \boldsymbol{I}_m, \quad and \quad \sum_{k \in [r]} \boldsymbol{B}_k \boldsymbol{B}_k^\top \preceq \boldsymbol{I}_m.$$

Proof of Lemma EC.1 Noting that $\boldsymbol{W}^{(i,j)} = \sum_{k \in [r]} \boldsymbol{B}_k \boldsymbol{e}_i \boldsymbol{e}_j^\top \boldsymbol{B}_k^\top$, we have

$$\left(\sum_k \boldsymbol{B}_k^\top \boldsymbol{B}_k \right)_{i,j} = \sum_{k \in [r]} \boldsymbol{e}_i^\top \boldsymbol{B}_k^\top \boldsymbol{B}_k \boldsymbol{e}_j = \text{tr}(\boldsymbol{W}^{(i,j)}),$$

$$\text{and } \sum_{k \in [r]} \boldsymbol{B}_k \boldsymbol{B}_k^\top = \sum_{k \in [r]} \sum_{i \in [m]} \boldsymbol{B}_k \boldsymbol{e}_i \boldsymbol{e}_i^\top \boldsymbol{B}_k^\top = \sum_{i \in [m]} \boldsymbol{W}^{(i,i)}.$$

The fact that $\boldsymbol{W}$ satisfies the constraints in (9) concludes the proof.   □

We can now prove Lemma 1.

Proof of Lemma 1 For the first inequality, we use the simple bound $\sigma_{\max}(\boldsymbol{G})^2 \leq \|\boldsymbol{G}\|_F^2$. Then, we have $\mathbb{E}[\|\boldsymbol{G}\|_F^2] = \text{tr}(\mathbb{E}[\boldsymbol{G}^\top \boldsymbol{G}]) = \text{tr}\left(\sum_k \boldsymbol{B}_k^\top \boldsymbol{B}_k \right) = m$. Hence, $\mathbb{E}[\sigma_{\max}(\boldsymbol{G})^2] \leq m$.

For the second bound, this is a consequence of tail bounds for Gaussian matrix series. While typical results provide a logarithmic dependency in $(n + m)$ (see equation (4.1.7) in Tropp 2015), we can obtain bounds that only depend on m by leveraging the fact that the matrix $\sum_k \boldsymbol{B}_k \boldsymbol{B}_k^\top$, although $n \times n$, has trace $m \leq n$. Specifically, to apply theorem 1 of Gao et al. (2024), we first define the following *Hermitian* Gaussian series

$$\boldsymbol{Y} := \sum_k z_k \boldsymbol{A}_k \quad \text{with} \quad \boldsymbol{A}_k := \begin{pmatrix} \boldsymbol{0} & \boldsymbol{B}_k \\ \boldsymbol{B}_k^\top & \boldsymbol{0} \end{pmatrix}.$$

By construction, $\sigma_{\max}(\boldsymbol{G}) = \lambda_{\max}(\boldsymbol{Y})$ and

$$\sum_k \boldsymbol{A}_k^2 = \begin{pmatrix} \sum_k \boldsymbol{B}_k \boldsymbol{B}_k^\top & \boldsymbol{0} \\ \boldsymbol{0} & \sum_k \boldsymbol{B}_k^\top \boldsymbol{B}_k. \end{pmatrix}$$

From Lemma EC.1, we have $\sum_k \boldsymbol{A}_k^2 \preceq \boldsymbol{I}_{n+m}$ so $\lambda_{\max}(\sum_k \boldsymbol{A}_k^2) \leq 1$. Furthermore, $\mathrm{tr}(\sum_k \boldsymbol{A}_k^2) = 2\,\mathrm{tr}(\sum_k \boldsymbol{B}_k^\top \boldsymbol{B}_k) = 2m$. Then, according to the proof of theorem 1 in Gao et al. (2024), for any $t, \theta > 0$, we have (equation 22),

$$\mathbb{P}\left(\sigma_{\max}(\boldsymbol{G}) > t\right) = \mathbb{P}\left(\lambda_{\max}(\boldsymbol{Y}) > t\right) \leq (2m)\frac{e^{\theta^2/2} - 1}{e^{\theta t} - 1}.$$

Taking $\theta = t$ and using the fact that $\dfrac{x-1}{x^2-1} = \dfrac{1}{x+1} \leq \dfrac{1}{x}$, we finally get

$$\mathbb{P}\left(\sigma_{\max}(\boldsymbol{G}) > t\right) \leq (2m)e^{-t^2/2},$$

as claimed. Finally, to convert this tail bound into a bound on $\mathbb{E}[\sigma_{\max}(\boldsymbol{G})^2]$ we use the characterization of the expected value for non-negative random variables:

$$\mathbb{E}[\sigma_{\max}(\boldsymbol{G})^2] = \int_0^\infty \mathbb{P}\left(\sigma_{\max}(\boldsymbol{G})^2 \geq t\right) dt = \int_0^\infty \mathbb{P}\left(\sigma_{\max}(\boldsymbol{G}) \geq \sqrt{t}\right) dt$$
$$= \int_0^\tau \mathbb{P}\left(\sigma_{\max}(\boldsymbol{G}) \geq \sqrt{t}\right) dt + \int_\tau^\infty \mathbb{P}\left(\sigma_{\max}(\boldsymbol{G}) \geq \sqrt{t}\right) dt,$$

with $\tau := 2\log(2m)$ (such that $2m\, e^{-\tau/2} = 1$). We bound the probability in the first integral by 1. For the second integral, we have from our tail bound

$$\int_\tau^\infty \mathbb{P}\left(\sigma_{\max}(\boldsymbol{G}) \geq \sqrt{t}\right) dt \leq (2m) \int_\tau^\infty e^{-t/2} dt = (2m)\left[-2e^{-t/2}\right]_\tau^\infty = 2$$

All together, we get $\mathbb{E}[\sigma_{\max}(\boldsymbol{G})^2] \leq \tau + 2 = 2\log(2m) + 2$. $\qquad\square$

EC.3.2. Proof of Proposition 2

In this section, we construct an example of a matrix $\boldsymbol{W}$ for which Algorithm 2 cannot achieve a performance guarantee that scales better than $1/\log m$.

Consider m orthonormal vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ and apply Algorithm 2 with a covariance matrix $\boldsymbol{W}$ defined as

$$\boldsymbol{W}^{(i,i)} = \boldsymbol{u}_i \boldsymbol{u}_i^\top, \quad \text{and } \boldsymbol{W}^{(i,j)} = \alpha \boldsymbol{u}_i \boldsymbol{u}_j^\top,$$

for some $\alpha \in (0,1)$. This matrix satisfies all the constraints of the semidefinite relaxation (9). The columns of the matrix $\boldsymbol{G}$ generated by Algorithm 2 are of the form

$$\boldsymbol{g}_i = z_i \boldsymbol{u}_i, \quad \text{with } \boldsymbol{z} \sim \mathcal{N}\left(\boldsymbol{0}, (1-\alpha)\boldsymbol{I}_m + \alpha \boldsymbol{e}\boldsymbol{e}^\top\right).$$

The SVD of $\boldsymbol{G}$ is precisely

$$\boldsymbol{G} = \boldsymbol{U} \begin{pmatrix} |z_1| & & \\ & \ddots & \\ & & |z_m| \end{pmatrix} \begin{pmatrix} \mathrm{sign}(z_1) & & \\ & \ddots & \\ & & \mathrm{sign}(z_m) \end{pmatrix}.$$

Hence, $\sigma_{\max} = \max_i |z_i|^2$ and the columns of the matrix $\boldsymbol{Q}$ are of the form

$$\boldsymbol{q}_i = d_i \boldsymbol{u}_i, \quad \text{with } \mathbb{P}(d_i = 1) = 1 - \mathbb{P}(d_i = -1) = \frac{1 + |z_i|/\sigma_{\max}}{2}.$$

In particular, $\boldsymbol{q}_i \boldsymbol{q}_i^\top = \boldsymbol{u}_i \boldsymbol{u}_i^\top$ a.s., and conditioned on $\boldsymbol{z}$, we have

$$\mathbb{E}\left[\boldsymbol{q}_i \boldsymbol{q}_j^\top | \boldsymbol{z}\right] = \mathbb{E}\left[d_i | \boldsymbol{z}\right] \mathbb{E}\left[d_j | \boldsymbol{z}\right] \boldsymbol{u}_i \boldsymbol{u}_j^\top = \frac{|z_i||z_j|}{\max_k |z_k|^2} \boldsymbol{u}_i \boldsymbol{u}_j^\top.$$

Let us denote $b_m(\alpha) := \mathbb{E}\left[\dfrac{|z_i||z_j|}{\max_k |z_k|^2}\right]$. If there exists a constant $\beta > 0$ such that $\mathbb{E}[\mathrm{vec}(\boldsymbol{Q}) \mathrm{vec}(\boldsymbol{Q})^\top] \succeq \beta \boldsymbol{W}$, then, applying it to the vector $\mathrm{vec}(\boldsymbol{U})$, we get

$$(1 + b_m(\alpha)(m-1))m \geq \beta(1 + \alpha(m-1))m, \quad \text{i.e.,} \quad \frac{1 + b_m(\alpha)(m-1)}{1 + \alpha(m-1)} \geq \beta.$$

In other words, for large values of m, $\beta = O\left(\dfrac{b_m(\alpha)}{\alpha}\right)$. Let us assume for now that there exists a constant $C_\alpha > 0$ such that

$$b_m(\alpha) \leq \frac{C_\alpha}{\log m}, \tag{EC.4}$$

then it rules out the existence of a constant β that vanishes to 0 as $m \to +\infty$ slower than $1/\log m$, thus ensuring that our analysis of Algorithm 2 is tight (in terms of dependency on m).

Proof of Equation (EC.4)   For any $a > 0$,

$$\mathbb{E}\left[\frac{|z_1||z_2|}{\max_k |z_k|^2}\right] = \mathbb{E}\left[\frac{|z_1||z_2|}{\max_k |z_k|^2} \mathbf{1}(\max_k |z_k|^2 < a)\right] + \mathbb{E}\left[\frac{|z_1||z_2|}{\max_k |z_k|^2} \mathbf{1}(\max_k |z_k|^2 \geq a)\right]$$

$$\leq \mathbb{P}(\max_k |z_k|^2 < a) + \frac{\mathbb{E}\left[|z_1||z_2|\right]}{a},$$

where the inequality follows from the fact that $|z_1||z_2|/\max_k |z_k|^2 \leq 1$. We control each term separately.

For the first term, let us write each random variable z_k as $z_k = \sqrt{1-\alpha}y_k + \sqrt{\alpha}g$ with $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_m)$ and $g \sim \mathcal{N}(0,1)$ independent from $\boldsymbol{y}$. Conditioned on g, we have $\boldsymbol{z}|g \sim \mathcal{N}(\sqrt{\alpha}g, (1-\alpha)\boldsymbol{I}_m)$, i.e., the z_k's are i.i.d. Hence, conditioning on g, we have for the tail probability:

$$\mathbb{P}\left(\max_k |z_k|^2 < a \Big| g\right) = \mathbb{P}\left(|z_k|^2 < a, \forall k | g\right) = \prod_k \mathbb{P}\left(|z_k| < \sqrt{a} | g\right) = \mathbb{P}\left(|z| < \sqrt{a}\right)^m$$

with $z \sim \mathcal{N}(\sqrt{\alpha}g, 1-\alpha)$. Denote $p(g) := \mathbb{P}\left(|z| < \sqrt{a}\right) = \mathbb{P}(|\mathcal{N}(\sqrt{\alpha}g, 1-\alpha)| < \sqrt{a})$. Observe that, for any α and any $a > 0$, the function $g \mapsto p(g)$ is maximized at $g = 0$. Hence, we have

$$\mathbb{P}(|z| < \sqrt{a})^m = p(g)^m \leq p(0)^m.$$

Furthermore,

$$p(0) = \mathbb{P}\left(|\mathcal{N}(0, 1-\alpha)| < \sqrt{a}\right) = 1 - 2\mathbb{P}\left(\mathcal{N}(0, 1-\alpha) < -\sqrt{a}\right)$$

and, denoting $t = a/\sqrt{1-\alpha}$, we have

$$\mathbb{P}\left(\mathcal{N}(0, 1-\alpha) < -\sqrt{a}\right) \quad = \mathbb{P}\left(\mathcal{N}(0,1) < -\sqrt{t}\right) \quad \geq \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{t} + 1/\sqrt{t}} e^{-t/2} \quad \geq \frac{1}{\sqrt{2\pi}} \frac{1}{2\sqrt{t}} e^{-t/2};$$

see (Vershynin 2018, Proposition 2.1.2) for a proof of the first inequality. Taking $a = 2\sqrt{1-\alpha}\,(1-\epsilon)\log m$ for some $\epsilon \in (0,1)$, i.e., $t = 2\,(1-\epsilon)\log m$ yields

$$\mathbb{P}\left(\mathcal{N}(0,1-\alpha) < -a\right) \geq \frac{1}{4\sqrt{\pi}}\frac{1}{\sqrt{(1-\epsilon)\log m}}m^{-(1-\epsilon)},$$

and finally, we get

$$\mathbb{P}\left(\max_k |z_k|^2 < a|g\right) \leq \left(1 - 2\mathbb{P}\left(\mathcal{N}(0,1-\alpha) < -a\right)\right)^m$$

$$\leq \exp\left(-2m\mathbb{P}\left(\mathcal{N}(0,1-\alpha) < -a\right)\right)$$

$$\leq \exp\left(-\frac{1}{2\sqrt{\pi}}\frac{1}{\sqrt{(1-\epsilon)\log m}}m^\epsilon\right).$$

Taking the expectation over g, we obtain

$$\mathbb{P}\left(\max_k |z_k|^2 < a\right) = \mathbb{E}\left[\mathbb{P}\left(\max_k |z_k|^2 < a|g\right)\right] \leq \exp\left(-\frac{1}{2\sqrt{\pi}}\frac{1}{\sqrt{(1-\epsilon)\log m}}m^\epsilon\right).$$

For the second term, (z_1, z_2) is a two-dimensional Gaussian vector with unit variance and correlation α. We have $\mathbb{E}[|z_1 z_2|] = \mathbb{E}[|z_1||z_2|] \leq \sqrt{\mathbb{E}[z_1^2]}\sqrt{\mathbb{E}[z_2^2]} = 1$.

All together, we have

$$\mathbb{E}\left[\frac{|z_1||z_2|}{\max_k |z_k|^2}\right] \leq \mathbb{P}(\max_k |z_k|^2 < a) + \frac{\mathbb{E}\left[|z_1||z_2|\right]}{a}$$

$$\leq \exp\left(-\frac{1}{2\sqrt{\pi}}\frac{1}{\sqrt{(1-\epsilon)\log m}}m^\epsilon\right) + \frac{1}{2\sqrt{1-\alpha}(1-\epsilon)\log m}.$$

For any value of ϵ, we must have $m^\epsilon \geq 2\sqrt{\pi(1-\epsilon)\log m}\,(\log\log m)$ for sufficiently large m, in which case we have

$$\mathbb{E}\left[\frac{y_i^2}{\max_k |y_k|^2}\right] \leq \frac{1}{\log m} + \frac{1}{2\sqrt{1-\alpha}(1-\epsilon)\log m},$$

which proves Equation (EC.4). $\qquad\square$

EC.3.3. Proof of Theorem 2

In this section, we prove Theorem 2. Actually, we will obtain Theorem 2 as a special case of a more general performance guarantee for Algorithm 2, which we now formally state and prove.

THEOREM EC.1. *Let $\boldsymbol{G} \in \mathbb{R}^{n \times m}$ be a Gaussian matrix generated by Algorithm 2. Then, for any $T \geq 0$ and $\delta \in (0,1)$, $\boldsymbol{G}$ satisfies the inequality:*

$$\mathbb{E}\left[\frac{\text{vec}(\boldsymbol{G})\text{vec}(\boldsymbol{G})^\top}{\sigma_{\max}(\boldsymbol{G})^2}\right] \succeq \left(\beta_{n,m}(T) + \frac{e^{-2T\log(6m/\delta)}}{2\log(6m/\delta)}\left(1 - \sqrt{\delta}\right)\right)\boldsymbol{W}^\star, \qquad (\text{EC.5})$$

with

$$\beta_{n,m}(T) := \min_{\lambda \in [0,1]}\int_0^T \left(1 + 2t\,m\frac{1-\lambda}{nm-1}\right)^{-(nm-1)/2}(1 + 2tm\lambda)^{-3/2}dt.$$

We can numerically optimize for $T \geq, \delta \in (0, 1)$ to compute the tightest constant and better evaluate the performance of our algorithm. We recover Theorem 2 by setting $T = \infty$. Qualitatively, we recover the same $\Theta(1/\log m)$ asymptotic regime as Theorem 1 by taking $T = 0$:

$$\mathbb{E}\left[\frac{\text{vec}(\boldsymbol{G})\text{vec}(\boldsymbol{G})^\top}{\sigma_{\max}(\boldsymbol{G})^2}\right] \succeq \frac{1 - \sqrt{\delta}}{2\log(m) + 2\log(6/\delta)}\boldsymbol{W}^\star,$$

which scales like $1/(2\log m)$.

Proof of Theorem EC.1   For any $T \geq 0$, we can write

$$\frac{1}{\sigma_{\max}(\boldsymbol{G})^2} = \int_{t=0}^T e^{-t\sigma_{\max}(\boldsymbol{G})^2}dt + \int_{t=T}^\infty e^{-t\sigma_{\max}(\boldsymbol{G})^2}dt. \tag{EC.6}$$

By Tonelli's theorem (e.g., Grimmett and Stirzaker 2020), and using non-negativity of each term in the integral, this leads to

$$\mathbb{E}\left[\frac{\text{vec}(\boldsymbol{G})\text{vec}(\boldsymbol{G})^\top}{\sigma_{\max}(\boldsymbol{G})^2}\right] = \underbrace{\int_{t=0}^T \mathbb{E}[\text{vec}(\boldsymbol{G})\text{vec}(\boldsymbol{G})^\top e^{-t\sigma_{\max}(\boldsymbol{G})^2}]dt}_{\boldsymbol{J}_1(T)} + \underbrace{\int_{t=T}^\infty \mathbb{E}[\text{vec}(\boldsymbol{G})\text{vec}(\boldsymbol{G})^\top e^{-t\sigma_{\max}(\boldsymbol{G})^2}]dt}_{\boldsymbol{J}_2(T)}.$$

The rest of the proof follows by deriving lower bounds for each integral $\boldsymbol{J}_1(T)$ and $\boldsymbol{J}_2(T)$, and combining them. We obtain a bound that is valid for any $T \geq 0$, and thus can be optimized with respect to T to obtain the tightest possible lower bound.

Lower bound on $\boldsymbol{J}_1(T)$: We use the operator bound $\sigma_{\max}(\boldsymbol{G})^2 \leq \|\boldsymbol{G}\|_F^2$ to obtain

$$\boldsymbol{J}_1(T) \succeq \int_{t=0}^T \mathbb{E}[\text{vec}(\boldsymbol{G})\text{vec}(\boldsymbol{G})^\top e^{-t\|\boldsymbol{G}\|_F^2}]dt.$$

The inner expectation can be computed analytically: Denote $r = \text{rank}(\boldsymbol{W}^*) \leq nm$ and consider an eigenvalue decomposition of $\boldsymbol{W}^\star$, $\boldsymbol{W}^\star = \boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}^\top$. We have the multivariate normal identity $\text{vec}(\boldsymbol{G}) = \boldsymbol{H}\boldsymbol{\Lambda}^{1/2}\boldsymbol{z}$, with $\boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}_r, \boldsymbol{I}_r)$ and thus

$$\mathbb{E}[\text{vec}(\boldsymbol{G})\text{vec}(\boldsymbol{G})^\top \exp(-t\|\boldsymbol{G}\|_F^2)] = \boldsymbol{H}\boldsymbol{\Lambda}^{1/2}\mathbb{E}[\boldsymbol{z}\boldsymbol{z}^\top \exp(-t\boldsymbol{z}^\top\boldsymbol{\Lambda}\boldsymbol{z})]\boldsymbol{\Lambda}^{1/2}\boldsymbol{H}^\top.$$

Furthermore,

$$\mathbb{E}[\boldsymbol{z}\boldsymbol{z}^\top \exp(-t\boldsymbol{z}^\top\boldsymbol{\Lambda}\boldsymbol{z})] = \frac{1}{(2\pi)^{r/2}}\int \boldsymbol{z}\boldsymbol{z}^\top e^{-t\boldsymbol{z}^\top\boldsymbol{\Lambda}\boldsymbol{z}}e^{-\frac{1}{2}\boldsymbol{z}^\top\boldsymbol{z}}d\boldsymbol{z} = \frac{1}{\sqrt{\det(\boldsymbol{I}_{nm} + 2t\boldsymbol{\Lambda})}}(\boldsymbol{I}_r + 2t\boldsymbol{\Lambda})^{-1},$$

by completing the square. So, the first integral is lower bounded by

$$\boldsymbol{H}\boldsymbol{\Lambda}^{1/2}\boldsymbol{B}\boldsymbol{\Lambda}^{1/2}\boldsymbol{H}^\top \text{ with } \boldsymbol{B} := \int_0^T \frac{1}{\sqrt{\det(\boldsymbol{I}_{nm} + 2t\boldsymbol{\Lambda})}}(\boldsymbol{I}_r + 2t\boldsymbol{\Lambda})^{-1}dt.$$

Indeed, if there existed a scalar $\beta > 0$ such that $\boldsymbol{B} \succeq \beta\boldsymbol{I}_r$, then we could conclude that $\boldsymbol{J}_1(T) \succeq \beta\boldsymbol{W}^\star$.

To find such β, observe that $\boldsymbol{B}$ is a diagonal matrix with diagonal entries

$$\int_0^T \prod_{i'=1}^r (1 + 2t\Lambda_{i'})^{-1/2}(1 + 2t\Lambda_i)^{-1}dt.$$

Hence, it is sufficient to find a lower bound on the diagonal entries of $\boldsymbol{B}$. Given the constraints on $\boldsymbol{W}^\star$, the eigenvalues $\boldsymbol{\Lambda}$ must satisfy: $\Lambda_i \geq 0$ (from $\boldsymbol{W}^\star \succeq 0$), $\sum_{i=1}^r \Lambda_i = m$ (from $\mathrm{tr}(\boldsymbol{W}^\star) = \sum_{i=1}^m \mathrm{tr}(\boldsymbol{W}^{\star(i,i)}) = m$). Hence, we can take

$$\beta = \min_{\boldsymbol{\Lambda} \in [0,m]^r : \sum_{i=1}^r \Lambda_i = m} \int_0^T \prod_{i=1}^r (1 + 2t\Lambda_{i'})^{-1/2} (1 + 2t\Lambda_1)^{-1} dt.$$

For any $t \geq 0$, the function $\boldsymbol{\Lambda} \mapsto \int_0^T \prod_{i=1}^r (1 + 2t\Lambda_{i'})^{-1/2} (1 + 2t\Lambda_1)^{-1}$ is log-convex, hence is convex (Boyd and Vandenberghe 2004, Section 3.5.1). By integration over t, the function $\boldsymbol{\Lambda} \mapsto \int_0^T \prod_{i=1}^r (1 + 2t\Lambda_{i'})^{-1/2} (1 + 2t\Lambda_1)^{-1} dt$ is also convex. In addition, we observe that this function is invariant by any permutation of the Λ_i, $i > 1$. So, by Jensen's inequality, we can restrict our attention to minimizers of the form $\Lambda_1 = \lambda$, $\Lambda_i = \dfrac{m - \lambda}{r - 1}, i > 1$ without loss of optimality, and

$$\beta = \min_{\lambda \in [0,m]} \int_0^T \left(1 + 2t\frac{m - \lambda}{r - 1}\right)^{-(r-1)/2} (1 + 2t\lambda)^{-3/2} dt. \tag{EC.7}$$

Recall that for any scalar x, the sequence $(1 + x/k)^{-k}$ is monotonically decreasing and converges to e^{-x}. As a result, for a fixed value of t and λ, the integrand is decreasing in $r = \mathrm{rank}(\boldsymbol{W}^\star)$. Looking at the worst case, we have

$$\beta \geq \beta_{n,m}(T) := \min_{\lambda \in [0,m]} \int_0^T \left(1 + 2t\frac{m - \lambda}{nm - 1}\right)^{-(nm-1)/2} (1 + 2t\lambda)^{-3/2} dt$$

$$= \min_{\lambda \in [0,1]} \int_0^T \left(1 + 2t\,m\frac{1 - \lambda}{nm - 1}\right)^{-(nm-1)/2} (1 + 2tm\lambda)^{-3/2} dt,$$

where we relabel $\lambda \leftarrow \lambda/m$ to normalize the optimization problem.

Lower bound on $\boldsymbol{J}_2(T)$: For the second integral, we leverage tail bounds on $\sigma_{\max}(\boldsymbol{G})^2$; see Lemma 1. For any $\theta > 0$, we have

$$\int_{t=T}^\infty e^{-t\sigma_{\max}(\boldsymbol{G})^2} dt = \frac{e^{-T\sigma_{\max}(\boldsymbol{G})^2}}{\sigma_{\max}(\boldsymbol{G})^2} \geq \frac{e^{-T\theta}}{\theta}(1 - \mathbf{1}(\sigma_{\max}(\boldsymbol{G})^2 > \theta)).$$

So for any unit vector $\boldsymbol{u}$,

$$\boldsymbol{u}^\top \left(\int_{t=T}^\infty \mathbb{E}[\mathrm{vec}(\boldsymbol{G})\mathrm{vec}(\boldsymbol{G})^\top e^{-t\sigma_{\max}(\boldsymbol{G})^2}] dt \right) \boldsymbol{u} \geq \frac{e^{-T\theta}}{\theta} \left(\boldsymbol{u}^\top \boldsymbol{W}^* \boldsymbol{u} - \mathbb{E}\left[(\mathrm{vec}(\boldsymbol{G})^\top \boldsymbol{u})^2 (\mathbf{1}(\sigma_{\max}(\boldsymbol{G})^2 > \theta)) \right] \right),$$

For the last term, we apply Cauchy-Schwarz to get

$$\mathbb{E}\left[(\mathrm{vec}(\boldsymbol{G})^\top \boldsymbol{u})^2 (\mathbf{1}(\sigma_{\max}(\boldsymbol{G})^2 > \theta)) \right] \leq \sqrt{\mathbb{E}\left[(\mathrm{vec}(\boldsymbol{G})^\top \boldsymbol{u})^4 \right] \mathbb{E}\left[(\mathbf{1}(\sigma_{\max}(\boldsymbol{G})^2 > \theta)) \right]}$$

$$\leq \sqrt{3}\,(\boldsymbol{u}^\top \boldsymbol{W}^* \boldsymbol{u})\,\sqrt{2m}\,e^{-\theta/4},$$

where the last inequality follows from 4th moment formula for multivariate Gaussian variables $(E[Z^4] = 3\sigma^4)$ applied to $Z := \mathrm{vec}(\boldsymbol{G})^\top \boldsymbol{u} \sim \mathcal{N}(0, \boldsymbol{u}^\top \boldsymbol{W}^\star \boldsymbol{u})$; and the tail bound $\mathbb{P}(\sigma_{\max}(\boldsymbol{G})^2 > \theta) \leq 2m e^{-\theta/2}$ (Lemma 1). All together,

$$\boldsymbol{J}_2(T) \succeq \frac{e^{-T\theta}}{\theta} \left(1 - \sqrt{6m}\,e^{-\theta/4} \right) \boldsymbol{W}^*.$$

Taking $\theta = 2\log(6m/\delta)$ for $\delta \in (0,1)$, we get

$$\boldsymbol{J}_2(T) \succeq \frac{e^{-2T\log(6m/\delta)}}{2\log(6m/\delta)} \left(1 - \sqrt{\delta}\right)\boldsymbol{W}^*. \tag{EC.8}$$

Combining the bounds for $\boldsymbol{J}_1(T)$ and $\boldsymbol{J}_2(T)$ concludes the proof. □

REMARK EC.3. We observe that the first part of the bound, $\boldsymbol{J}_1(T)$, is obtained by looking at the worst-case instance over all covariance matrices $\boldsymbol{W}^\star$. In particular, Equation (EC.7) provides a tighter value of $\beta_{n,m}(T)$ that depends explicitly on the rank of $\boldsymbol{W}^\star$, r, instead of the ambient dimension n. For instance, if $r = 1$, we get

$$\beta = \min_{\lambda \in [0,m]} \int_0^T (1 + 2t\lambda)^{-3/2} dt = \min_{\lambda \in [0,m]} \frac{1 - (1 + 2\lambda T)^{-1/2}}{\lambda} = \frac{1 - (1 + 2mT)^{-1/2}}{m}.$$

Alternatively, by the Barvinok-Pataki bound, we know there exists some optimal solution $\boldsymbol{W}^\star$ with rank at most $n + m$ and we could use this bound to refine our constant. Furthermore, if $\boldsymbol{W}^\star$ has additional structure (e.g., $\boldsymbol{W}^\star$ is block diagonal), we can derive additional constraints on the eigenvalues $\boldsymbol{\Lambda}$, hence tighter constants $\beta_{n,m}(T)$.

We now provide some interesting qualitative features of the bound in Theorem EC.1.

PROPOSITION EC.3. *The constant*

$$\beta_{n,m}(T) = \min_{\lambda \in [0,1]} \int_0^T \left(1 + 2t\,m\frac{1-\lambda}{nm-1}\right)^{-(nm-1)/2} (1 + 2tm\lambda)^{-3/2} dt.$$

in Theorem EC.1 satisfies the following properties:

(a) For any $T \geq 0$ and any integer m, the constant $\beta_{n,m}(T)$ is non-increasing in n.

(b) For any $T \geq 0$ and any integer n, the constant $\beta_{n,m}(T)$ is non-increasing in m wherever it exists $(n \geq m)$.

(c) For any $T \geq 0$, any m, we have

$$\beta_{n,m}(T) \underset{n \to +\infty}{\to} \beta_{\infty,m}(T) = \min_{\lambda \in [0,1]} \int_0^T e^{-t\,m(1-\lambda)} (1 + 2tm\lambda)^{-3/2} dt.$$

(d) For any $T \geq 0$ and any integer n,m (with $n \geq m$), we can also express $\beta_{n,m}(T)$ as

$$\beta_{n,m}(T) = \min_{\lambda \in [0,1]} \mathbb{E}_{X \sim \chi_1^2, Y \sim \chi_{nm-1}^2} \left[\frac{X}{\frac{m(1-\lambda)}{nm-1}Y + m\lambda X} \left(1 - e^{-\frac{Tm(1-\lambda)}{nm-1}Y - Tm\lambda X}\right) \right].$$

(e) For any $T \geq 0$ and any integer m, we can also write $\beta_{\infty,m}(T)$ as

$$\beta_{\infty,m}(T) = \min_{\lambda \in [0,1]} \mathbb{E}_{X \sim \chi_1^2} \left[\frac{X}{m(1-\lambda) + m\lambda X} \left(1 - e^{-T(m(1-\lambda) + m\lambda X)}\right) \right].$$

(f) For $m = 1$, taking $T = \infty$ in (EC.5) is optimal and the inequality

$$\mathbb{E}\left[\frac{\mathrm{vec}(\boldsymbol{G})\mathrm{vec}(\boldsymbol{G})^\top}{\sigma_{\max}(\boldsymbol{G})^2} \right] \succeq \beta_{n,m}(\infty)\,\boldsymbol{W}^\star$$

is tight (in the sense that there exists a covariance matrix $\boldsymbol{W}^\star$ satisfying it at equality).

Proof of Proposition EC.3 We prove each claim separately.

Claim (a) For any $\lambda \in [0,1]$ and $t \in [0,T]$, the integrand

$$\left(1 + 2t\,m\frac{1-\lambda}{nm-1}\right)^{-(nm-1)/2} (1+2tm\lambda)^{-3/2}$$

is decreasing in n. Integrating over t and minimizing over λ gives $\beta_{n+1,m}(T) \le \beta_{n,m}(T)$.

Claim (b) For any $\lambda \in [0,1]$ and $t \in [0,T]$, we claim that the integrand

$$\left(1 + 2t\,m\frac{1-\lambda}{nm-1}\right)^{-(nm-1)/2} (1+2tm\lambda)^{-3/2}$$

is decreasing in m. To see this, first observe that $(1+2tm\lambda)^{-3/2}$ is obviously decreasing in m. Next, consider the quantity $\left(1 + 2t\,m\frac{1-\lambda}{nm-1}\right)^{-(nm-1)/2}$. By letting $u = nm-1$, $a = 2t(1-\lambda)/n \ge 0$ and $c = a/(1+a)$, we obtain the relationships

$$\frac{m}{nm-1} = \frac{(u+1)/n}{u} = \frac{1}{n}\left(1 + \frac{1}{u}\right),$$

and thus we get

$$1 + 2tm\frac{1-\lambda}{nm-1} = 1 + a\left(1 + \frac{1}{u}\right) = (1+a)\left(1 + \frac{c}{u}\right), \qquad c := \frac{a}{1+a} \in [0,1).$$

In particular, we have the equivalent polynomial $(1+a)^{-u/2}(1+c/u)^{-u/2}$. This polynomial is decreasing in u, since $h(u) := u\log(1+c/u)$ has derivative

$$h'(u) = \log\left(1 + \frac{c}{u}\right) - \frac{c}{u+c} \ge 0,$$

because $\log(1+x) \ge \frac{x}{1+x}$ for $x > 0$. Thus, the polynomial is also decreasing in m. Thus, the integrand is decreasing in m, and integrating with respect to t gives the result.

Claim (c) The series $(1+x/k)^{-k}$ converging to e^{-x}, we have that, for any $\lambda \in [0,1]$ and $t \in [0,T]$, the integrand monotonically converges to $e^{-tm(1-\lambda)}(1+2tm\lambda)^{-3/2}$, as $n \to \infty$. By the dominated convergence theorem, the functions

$$f_n(\lambda) := \int_0^T \left(1 + 2t\,m\frac{1-\lambda}{nm-1}\right)^{-(nm-1)/2} (1+2tm\lambda)^{-3/2} dt$$

are continuous and converge monotonically to $f_\infty(\lambda) := \int_0^T e^{-tm(1-\lambda)}(1+2tm\lambda)^{-3/2} dt$. From $f_n(\lambda) \ge f_\infty(\lambda)$, we get $\beta_{n,m}(T) \ge \beta_{\infty,n}(T)$. Taking $\lambda^\star$ the minimizer of the continuous function $f_\infty(\lambda)$ over the compact set $[0,1]$, we have $f_n(\lambda^\star) \ge \beta_{n,m}(T) \ge \beta_{\infty,n}(T)$. In the limit, $f_n(\lambda^\star) \to f_\infty(\lambda^\star) = \beta_{\infty,m}(T)$ by continuity, so, by sandwiching, $\beta_{n,m}(T) \to \beta_{\infty,m}(T)$.

Claim (d) Take $Z \sim N(0,1)$ and observe that for any scalar $a > 0$, $\mathbb{E}\left[e^{-aZ^2}\right] = (1+2a)^{-1/2}$, which implies (by differentiation w.r.t. a) $\mathbb{E}\left[Z^2 e^{-aZ^2}\right] = (1+2a)^{-3/2}$. Introducing $nm-1$ additional independent, standard normal random variables $Z_1, \ldots, Z_{nm-1}$, we have

$$\left(1 + 2t\,m\frac{1-\lambda}{nm-1}\right)^{-(nm-1)/2} (1+2tm\lambda)^{-3/2} = \prod_{i=1}^{nm-1} \mathbb{E}\left[e^{-\frac{tm(1-\lambda)}{nm-1}Z_i^2}\right] \mathbb{E}\left[Z^2 e^{-tm\lambda Z^2}\right]$$

$$= \mathbb{E}\left[Z^2 e^{-\frac{tm(1-\lambda)}{nm-1}\sum_i Z_i^2 - tm\lambda Z^2}\right]$$

$$= \mathbb{E}\left[Xe^{-\frac{tm(1-\lambda)}{nm-1}Y-tm\lambda X}\right],$$

where $X \sim \chi_1^2$ and $Y \sim \chi_{nm-1}^2$. Integrating over $t \in [0,T]$ and invoking Tonelli's theorem to exchange the order of the integral and the expectation, we get

$$\int_0^T \left(1 + 2t\,m\frac{1-\lambda}{nm-1}\right)^{-(nm-1)/2} (1+2tm\lambda)^{-3/2}dt = \mathbb{E}\left[\int_0^T Xe^{-\frac{tm(1-\lambda)}{nm-1}Y-tm\lambda X}dt\right]$$

$$= \mathbb{E}\left[\frac{X}{\frac{m(1-\lambda)}{nm-1}Y+m\lambda X}\left(1 - e^{-\frac{Tm(1-\lambda)}{nm-1}Y-Tm\lambda X}\right)\right],$$

as claimed.

Claim (e) Taking $Z \sim N(0,1)$ and making the same observations as in the proof of Claim (c), we get

$$e^{-t\,m(1-\lambda)}(1+2tm\lambda)^{-3/2} = \mathbb{E}\left[Z^2 e^{-tm\lambda Z^2}e^{-t\,m(1-\lambda)}\right],$$

$$\int_{t=0}^T e^{-t\,m(1-\lambda)}(1+2tm\lambda)^{-3/2}dt = \mathbb{E}\left[Z^2 \int_{t=0}^T e^{-tm\lambda Z^2}e^{-t\,m(1-\lambda)}dt\right] \quad = \mathbb{E}\left[Z^2 \frac{1 - e^{-Tm\lambda Z^2-T\,m(1-\lambda)}}{m\lambda Z^2 + m(1-\lambda)}\right],$$

where the second equality permutes the order of the integral and the expectation, according to Tonelli's theorem. Defining $X := Z^2 \sim \chi_1^2$ leads to the expression of Claim (d).

Claim (f) Observe that for $m = 1$, $\sigma_{\max}(\boldsymbol{G})^2 = \|\boldsymbol{G}\|_F^2$ and $\beta_{n,m}(\infty)$ is the tightest constant (over all possible covariance matrices $\boldsymbol{W}^\star$) such that

$$\mathbb{E}\left[\frac{\mathrm{vec}(\boldsymbol{G})\mathrm{vec}(\boldsymbol{G})^\top}{\|\boldsymbol{G}\|_F^2}\right] \succeq \beta\boldsymbol{W}^\star.$$

$\square$

EC.3.4. Computing the Approximation Constant for Finite n, m

We report the value of the constant $\beta_{n,m} = \beta_{n,m}(\infty)$ from Theorem (2) in Table EC.1 and the constant from Theorem EC.1 in Table EC.2 for some values of n, m.

To compute these constants numerically in `Julia`, we model all integrals using Gauss-Kronrod quadrature in t with a relative tolerance of 10^{-8} and an absolute tolerance of 10^{-10}. We identify an approximately optimal T using a grid of 1000 values distributed uniformly in log space over $[10^{-6}, 10^6]$, in addition to explicitly considering 0 and $+\infty$. For each value of T, in our outer maximization problem, we use golden section search with a tolerance of 10^{-8} to maximize for δ. Given a value of T and δ, we minimize with respect to λ via an inner golden section search with a tolerance of 10^{-8}. To improve stability when t is large, we evaluate the two factors in the integrand in the log domain and exponentiate at the end. The edge case $nm = 1$ is handled separately via its analytic limit.

As a sanity check, we tightened all our tolerances by two orders of magnitude and increased the grid resolution for T by an order of magnitude, and then recomputed our constants. We found that

none of them changed to within the first six decimal places, which indicates that the aggregate numerical error is below 10^{-6}.

We observe that for $m \leq 13$ and all n considered, the optimal constant is attained by setting $T = +\infty$. However, once $m > 13$, we obtain a strictly better constant by optimizing for T.

Finally, we made the observation in Remark EC.3 that one can marginally improve the constants by leveraging the Barvinok-Pataki bound to bound the rank of $\boldsymbol{W}^{\star}$. For instance, when $n = m = 2$, we find that it improves the approximation constant from 0.375 to 0.3877. However, for larger n, m, the effects of this observation are negligible.

n\m	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1.000005	—	—	—	—	—	—	—	—	—	—	—	—	—	—
2	0.828427	0.375000	—	—	—	—	—	—	—	—	—	—	—	—	—
3	0.775334	0.362826	0.236678	—	—	—	—	—	—	—	—	—	—	—	—
4	0.750000	0.356945	0.234140	0.174200	—	—	—	—	—	—	—	—	—	—	—
5	0.735264	0.353486	0.232640	0.173367	0.138164	—	—	—	—	—	—	—	—	—	—
6	0.725652	0.351211	0.231649	0.172815	0.137813	0.114602	—	—	—	—	—	—	—	—	—
7	0.718895	0.349600	0.230945	0.172423	0.137564	0.114429	0.097955	—	—	—	—	—	—	—	—
8	0.713889	0.348401	0.230420	0.172130	0.137377	0.114300	0.097861	0.085556	—	—	—	—	—	—	—
9	0.710033	0.347473	0.230013	0.171902	0.137232	0.114199	0.097787	0.085499	0.075955	—	—	—	—	—	—
10	0.706972	0.346734	0.229689	0.171721	0.137116	0.114119	0.097728	0.085454	0.075920	0.068299	—	—	—	—	—
11	0.704484	0.346131	0.229424	0.171573	0.137022	0.114054	0.097680	0.085418	0.075891	0.068275	0.062049	—	—	—	—
12	0.702421	0.345630	0.229203	0.171449	0.136943	0.113999	0.097640	0.085387	0.075866	0.068256	0.062033	0.056850	—	—	—
13	0.700684	0.345207	0.229017	0.171345	0.136876	0.113953	0.097606	0.085361	0.075846	0.068239	0.062019	0.056839	0.052457	—	—
14	0.699201	0.344846	0.228858	0.171256	0.136820	0.113914	0.097577	0.085339	0.075828	0.068225	0.062008	0.056829	0.052448	0.048695	—
15	0.697920	0.344533	0.228720	0.171179	0.136770	0.113879	0.097552	0.085320	0.075813	0.068213	0.061998	0.056820	0.052441	0.048688	0.045437
∞	0.680415	0.340208	0.226805	0.170104	0.136083	0.113403	0.097202	0.085052	0.075602	0.068042	0.061856	0.056701	0.052340	0.048601	0.045361

Table EC.1: $\alpha_{n,m}(T{=}\infty) = \beta_{n,m}(\infty)$ (tail term vanishes). Entries shown only for $n \geq m$; last row is $n \to \infty$ limit of β.

n\m	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1.000005	—	—	—	—	—	—	—	—	—	—	—	—	—	—
2	0.828427	0.375000	—	—	—	—	—	—	—	—	—	—	—	—	—
3	0.775334	0.362826	0.236678	—	—	—	—	—	—	—	—	—	—	—	—
4	0.750000	0.356945	0.234140	0.174200	—	—	—	—	—	—	—	—	—	—	—
5	0.735264	0.353486	0.232640	0.173367	0.138164	—	—	—	—	—	—	—	—	—	—
6	0.725652	0.351211	0.231649	0.172815	0.137813	0.114602	—	—	—	—	—	—	—	—	—
7	0.718895	0.349600	0.230945	0.172423	0.137564	0.114429	0.097955	—	—	—	—	—	—	—	—
8	0.713889	0.348401	0.230420	0.172130	0.137377	0.114300	0.097861	0.085556	—	—	—	—	—	—	—
9	0.710033	0.347473	0.230013	0.171902	0.137232	0.114199	0.097787	0.085499	0.075955	—	—	—	—	—	—
10	0.706972	0.346734	0.229689	0.171721	0.137116	0.114119	0.097728	0.085454	0.075920	0.068299	—	—	—	—	—
11	0.704484	0.346131	0.229424	0.171573	0.137022	0.114054	0.097680	0.085418	0.075891	0.068275	0.062049	—	—	—	—
12	0.702421	0.345630	0.229203	0.171449	0.136943	0.113999	0.097640	0.085387	0.075866	0.068256	0.062033	0.056850	—	—	—
13	0.700684	0.345207	0.229017	0.171345	0.136876	0.113953	0.097606	0.085361	0.075846	0.068239	0.062019	0.056839	0.054157	—	—
14	0.699201	0.344846	0.228858	0.171256	0.136820	0.113914	0.097577	0.085339	0.075828	0.068225	0.062008	0.056829	0.054157	0.053438	—
15	0.697920	0.344533	0.228720	0.171179	0.136770	0.113879	0.097552	0.085320	0.075813	0.068213	0.061998	0.056820	0.054157	0.053438	0.052814
∞	0.680415	0.340208	0.226805	0.170104	0.136083	0.113403	0.097202	0.085052	0.075602	0.068042	0.061856	0.056701	0.054157	0.053438	0.052814

Table EC.2: Best lower-bound constant $\alpha_{n,m}^*$ from Theorem EC.1, optimized over $T \geq 0$ and $\delta \in (0,1)$. Entries shown only for $n \geq m$; last row is $n \to \infty$ limit.

EC.4. Counterexamples

EC.4.1. Non-Equivalence of Reduced Shor Relaxation and Shor Relaxation in Presence of Permutation Equalities

Consider a low-rank matrix completion problem of the form

$$\min_{\boldsymbol{X}\in\mathbb{R}^{n\times m}}\quad \frac{1}{2\gamma}\|\boldsymbol{X}\|_F^2 + \frac{1}{2}\sum_{(i,j)\in\Omega}(X_{i,j}-A_{i,j})^2 \text{ s.t. } \mathrm{rank}(\boldsymbol{X})\le k,$$

and three semidefinite relaxations: (a) the matrix perspective relaxation as introduced in the paper Bertsimas et al. (2023), (b) the semidefinite relaxation (16) with the inequalities on $\boldsymbol{W}_{x,y}$ and $\boldsymbol{W}_{y,y}$, (c) the semidefinite relaxation (17).

Let the problem data be $\gamma=100, k=2, n=7, m=5$, and suppose we are trying to impute the following matrix, where $*$ denotes a missing entry:

$$\boldsymbol{A}=\begin{pmatrix} -2 & * & -1 & 1 & -1 \\ * & 4 & -4 & -5 & -4 \\ * & -3 & 1 & 4 & 3 \\ 3 & 5 & -5 & -5 & -1 \\ 7 & 8 & -10 & -8 & 1 \\ 3 & 1 & -2 & * & 5 \\ 7 & 7 & -13 & -8 & * \end{pmatrix}.$$

Then (using Mosek version 10.2 to solve all semidefinite relaxations):

- The matrix perspective relaxation as introduced in the paper Bertsimas et al. (2023) has an optimal objective value of 3.9275.

- The semidefinite relaxation (16) has an optimal objective value of 5.1387.

- The more compact semidefinite relaxation (17) has an objective value of 4.314.

- The method of Burer and Monteiro (2003) finds a feasible solution with objective value 9.495.

Thus, we conclude that the permutation inequalities in (16) are not redundant, and the reduction in Theorem 3 does not hold in the presence of these inequalities. Nonetheless, the reduction is useful because it produces a non-trivial lower bound after solving a smaller semidefinite problem.

EC.4.2. Partial Redundancy of Kronecker Constraints of Burer and Park (2024)

In this section, we support our discussion in Endnote 3 by demonstrating that if $\boldsymbol{W}^\star$ is a matrix which is either (a) diagonal or (b) of the form $\boldsymbol{W}^{(i,i)}=\boldsymbol{u}_i\boldsymbol{u}_i^\top$, $\boldsymbol{W}^{(i,j)}=\alpha\boldsymbol{u}_i\boldsymbol{u}_j^\top$ for $\alpha\in[0,1]$, then the Kronecker constraints introduced by Burer and Park (2024) are redundant. This shows that the order of our approximation guarantee for Algorithm EC.1 would not be improved from $O(1/\log m)$ by imposing the Kronecker constraints of Burer and Park (2024), since it does not rule out the family of worst-case matrices $\boldsymbol{W}^\star$ studied in Proposition 2.

To show our claim holds, we first remind the reader of the form of the Kronecker product constraints proposed in Burer and Park (2024). Note that, as there is no linear term in the objective

function of (4), the term $\boldsymbol{U}$ in (Burer and Park 2024, Section 2.3)'s constraint can be set to $\boldsymbol{0}$ without loss of generality. With this simplification, for each $i \in [m]$ and $p \in [n]$, define the matrices

$$\boldsymbol{K}_{ip} := \boldsymbol{e}_i \boldsymbol{e}_{m+p}^\top + \boldsymbol{e}_{m+p} \boldsymbol{e}_i^\top \in \mathcal{S}^{m+n}.$$

Then, for any matrix $\boldsymbol{W} \in \mathcal{S}_+^{mn}$, (Burer and Park 2024)'s Kronecker constraint is equivalent to

$$M(\boldsymbol{W}) := I_{(m+n)^2} + \sum_{i,j=1}^{m} \sum_{p,q=1}^{n} [\boldsymbol{W}^{(i,j)}]_{pq} \boldsymbol{K}_{ip} \otimes \boldsymbol{K}_{jq} \succeq 0. \tag{EC.9}$$

We now compare (EC.9) against the constraints

$$\boldsymbol{W} \succeq \boldsymbol{0}, \qquad \text{tr}(\boldsymbol{W}^{(j,j)}) = 1 \; \forall j \in [m], \; \sum_{j=1}^{m} \boldsymbol{W}^{(j,j)} \preceq \boldsymbol{I}_n. \tag{EC.10}$$

via the following results:

LEMMA EC.2 (**Partial redundancy of Kronecker constraints**). *Under* (EC.10) *the following properties hold:*

(i)   *If* $\boldsymbol{W} \in \mathcal{S}_+^{mn}$ *is a diagonal matrix then* $M(\boldsymbol{W}) \succeq 0$.

(ii) *Let* $\boldsymbol{W} \in \mathcal{S}_+^{mn}$ *be a matrix of the form*

$$\boldsymbol{W}^{(i,i)} = \boldsymbol{u}_i \boldsymbol{u}_i^\top, \qquad \boldsymbol{W}^{(i,j)} = \alpha \, \boldsymbol{u}_i \boldsymbol{u}_j^\top \quad (i \neq j).$$

for a set of orthonormal vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m \in \mathbb{R}^n$ *and* $\alpha \in [0, 1]$. *Then,* $M(\boldsymbol{W}) \succeq 0$.

Proof of Lemma EC.2   We invoke an invariant subspace decomposition of $\mathbb{R}^{m+n} \otimes \mathbb{R}^{m+n}$:

(i) Define $\boldsymbol{W}^{(i,i)} = \text{Diag}(\alpha_{i1}, \ldots, \alpha_{in})$. From (EC.10) we have

$$\sum_{p=1}^{n} \alpha_{ip} = 1 \; (\forall i), \qquad 0 \leq \alpha_{ip} \leq 1 \; (\forall i, p), \qquad \sum_{i=1}^{m} \alpha_{ip} \leq 1 \; (\forall p). \tag{EC.11}$$

In the Kronecker constraint (EC.9), K_{ip} acts non-trivially only on the two–dimensional subspace $\mathcal{S}_{ip} := \text{span}\{e_i, \; e_{m+p}\}$, where it is the exchange matrix $J := \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. Hence, on the $\mathcal{S}_{ip} \otimes \mathcal{S}_{ip}$ the restriction of $I + \alpha_{ip}(K_{ip} \otimes K_{ip})$ equals $I_4 + \alpha_{ip}(J \otimes J)$ and has eigenvalues $1 \pm \alpha_{ip} \in [0, 2]$.

The remaining coupling occurs on

$$\mathcal{T} := \text{span}\{e_i \otimes e_i : 1 \leq i \leq m\} \; \oplus \; \text{span}\{e_{m+p} \otimes e_{m+p} : 1 \leq p \leq n\}.$$

Summing over (i, p), the restriction of $M(\boldsymbol{W})$ to $\mathcal{T}$ is

$$\begin{bmatrix} \boldsymbol{I}_m & \boldsymbol{A} \\ \boldsymbol{A}^\top & \boldsymbol{I}_n \end{bmatrix}, \qquad \boldsymbol{A} := (\alpha_{ip})_{i \leq m, \; p \leq n}. \tag{EC.12}$$

By the Schur complement lemma, (EC.12) is positive semidefinite when the spectral norm of $\boldsymbol{A}$ does not exceed 1, i.e., $\|\boldsymbol{A}\|_\sigma \leq 1$, which holds for any diagonal $\boldsymbol{W}$ that satisfies (EC.10). Moreover, on the orthogonal complement of $\mathcal{T}$, $M(\boldsymbol{W}) = \boldsymbol{I}_{(m+n)^2} \succeq \boldsymbol{0}$. Therefore, $M(\boldsymbol{W}) \succeq 0$.

(ii) Define the auxiliary vectors $\boldsymbol{w}_i := \sum_{p=1}^{n} (\boldsymbol{u}_i)_p \, \boldsymbol{e}_{m+p} \in \mathbb{R}^{m+n}$, so that $\langle \boldsymbol{w}_i, \boldsymbol{w}_j \rangle = \langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = \delta_{ij}$, and further define

$$\boldsymbol{K}_i := \sum_{p=1}^{n} (\boldsymbol{u}_i)_p \, \boldsymbol{K}_{ip} = \boldsymbol{e}_i \boldsymbol{w}_i^\top + \boldsymbol{w}_i \boldsymbol{e}_i^\top.$$

Then, each $\boldsymbol{K}_i$ acts as J on the two–dimensional subspace $\mathcal{S}_i := \operatorname{span}\{e_i, w_i\}$, and vanishes on $\mathcal{S}_i^\perp$. Moreover, the subspaces $\{\mathcal{S}_i\}_{i=1}^{m}$ are pairwise orthogonal. Plugging $\boldsymbol{W}^{(i,j)}$ into (EC.9) yields

$$M(\boldsymbol{W}) = \boldsymbol{I} + \alpha \, \boldsymbol{S} \otimes \boldsymbol{S} + (1-\alpha) \sum_{i=1}^{m} \boldsymbol{K}_i \otimes \boldsymbol{K}_i, \quad \boldsymbol{S} := \sum_{i=1}^{m} \boldsymbol{K}_i. \tag{EC.13}$$

Since $\mathcal{S}_i \perp \mathcal{S}_j$ for $i \neq j$, $\mathbb{R}^{m+n} \otimes \mathbb{R}^{m+n}$ decomposes orthogonally as

$$\bigoplus_{i,j=1}^{m} \left(\mathcal{S}_i \otimes \mathcal{S}_j \right) \; \oplus \; \mathcal{S}^\perp,$$

and $M(\boldsymbol{W})$ is block–diagonal with respect to this splitting. On each block:

- If $i \neq j$, then $\boldsymbol{S} \otimes \boldsymbol{S}$ restricts to $\boldsymbol{K}_i \otimes \boldsymbol{K}_j = \boldsymbol{J} \otimes \boldsymbol{J}$, while $\sum_k \boldsymbol{K}_k \otimes \boldsymbol{K}_k$ vanishes; thus $M(\boldsymbol{W})\big|_{\mathcal{S}_i \otimes \mathcal{S}_j} = \boldsymbol{I}_4 + \alpha(\boldsymbol{J} \otimes \boldsymbol{J})$, whose eigenvalues are $1 \pm \alpha$ (each with multiplicity two) and hence nonnegative for $\alpha \in [0,1]$.

- If $i = j$, then (EC.13) restricts to $M(\boldsymbol{W})\big|_{\mathcal{S}_i \otimes \mathcal{S}_i} = \boldsymbol{I}_4 + (\boldsymbol{J} \otimes \boldsymbol{J})$, whose spectrum is $\{2, 2, 0, 0\}$.

- On $\mathcal{S}^\perp$, $M(\boldsymbol{W}) = \boldsymbol{I}$.

Therefore, $M(\boldsymbol{W}) \succeq 0$ for all $\alpha \in [0,1]$. $\quad\square$

We caution that Lemma EC.2 does not imply that the Kronecker constraints of Burer and Park (2024) are always redundant. In fact, the numerical results in Burer and Park (2024) demonstrate that Kronecker constraints substantially tighten our semidefinite relaxations for some instances, albeit at the price of compromising numerical tractability. This situation is quite similar to the use of triangle inequalities for Max-Cut: triangle inequalities do not improve the worst-case approximation ratio for max-cut beyond the Goemans-Williamson ratio of 0.87856 (O'Donnell and Wu 2008), but they often improve the quality of semidefinite relaxations in practice, sometimes substantially (Rendl et al. 2010).

EC.5. A Stronger Benchmark via Linear Algebra Techniques

To further evaluate the performance of Algorithm 2, we propose a benchmark for constructing a feasible solution to (4). The procedure is inspired by the deflation procedure for PCA and considers eigenvectors of the on-diagonal blocks $\boldsymbol{A}$. Because it ignores the off-diagonal blocks of $\boldsymbol{A}$ in the design of $\boldsymbol{U}$, we show that it leads to a $1/m^2$–approximation guarantee, which is better than uniform sampling for $m \ll n$ but weaker than Algorithm 2.

Algorithm (EC.2) constructs the columns of $\boldsymbol{U}$: To compute $\boldsymbol{u}_i$, it projects the diagonal block $\boldsymbol{A}^{(i,i)}$ onto the subspace orthogonal to the columns already constructed, $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{i-1}$, and takes its leading eigenvector. In practice, we can process the diagonal blocks in any order, with each

ordering leading to a different candidate solution. In our implementation, we consider N random permutations of $\{1, \ldots, m\}$ and generate N feasible solutions to allow for a fair comparison with our sampling-based approach.

Algorithm EC.2 A Deflation-Inspired Benchmark for Orthogonality Constrained Optimization

Require: Positive semidefinite matrix $\boldsymbol{A} \in \mathcal{S}_+^{nm}$

    Initialize $\boldsymbol{U} = \boldsymbol{0}$

    for $i = 1, \ldots, m$ **do**

        Define $\boldsymbol{B} = (\boldsymbol{I}_n - \boldsymbol{u}_{i-1}\boldsymbol{u}_{i-1}^\top) \cdots (\boldsymbol{I}_n - \boldsymbol{u}_1\boldsymbol{u}_1^\top)\boldsymbol{A}^{(i,i)}(\boldsymbol{I}_n - \boldsymbol{u}_1\boldsymbol{u}_1^\top) \cdots (\boldsymbol{I}_n - \boldsymbol{u}_{i-1}\boldsymbol{u}_{i-1}^\top)$

        Compute $\boldsymbol{v}_i \in \arg\max_{\boldsymbol{x}\,\|\boldsymbol{x}\|_2=1} \boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x}$

        Define $\boldsymbol{u}_i = z_i \boldsymbol{v}_i$ with $\mathbb{P}(z_i = 1) = 1 - \mathbb{P}(z_i = -1) = 1/2$.

    end for

    return Semi-orthogonal matrix $\boldsymbol{U}$

We can show the following guarantee for this deflation procedure:

PROPOSITION EC.4. *Let $\boldsymbol{Q}$ be generated according to Algorithm EC.2 with a random ordering of the blocks. Then, we have*

$$\mathbb{E}\left[\langle\boldsymbol{A}, \mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top\rangle\right] \le \max_{\boldsymbol{U}\in\mathbb{R}^{n\times m}:\boldsymbol{U}^\top\boldsymbol{U}=\boldsymbol{I}_m} \left[\langle\boldsymbol{A}, \mathrm{vec}(\boldsymbol{U})\mathrm{vec}(\boldsymbol{U})^\top\rangle\right] \le m^2\mathbb{E}\left[\langle\boldsymbol{A}, \mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top\rangle\right].$$

REMARK EC.4. Observe that if $\boldsymbol{A}$ is a block diagonal matrix with identical on-diagonal blocks $\boldsymbol{A}^{(i,i)} = \boldsymbol{\Sigma}$ and zero off diagonal blocks $\boldsymbol{A}^{(i,j)} = \boldsymbol{0}$ for $i \ne j$, as in principal component analysis, then the proposed algorithm corresponds to deflation in PCA and is thus exact.

Proof of Proposition EC.4 First, at iteration i, since $\boldsymbol{q}_i$ is colinear to the leading eigenvector of the matrix obtained by $\boldsymbol{A}^{(i,i)}$ onto a space orthogonal to $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_{i-1}$, we have that $\boldsymbol{q}_i^\top\boldsymbol{q}_j = 0$ for each $i > j$ and thus $\boldsymbol{Q}$ is feasible, leading the left inequality holds.

Second, since z_i, z_j are i.i.d. with mean 0, in expectation, we have that $\mathbb{E}\left[\boldsymbol{q}_i^\top\boldsymbol{A}^{(i,j)}\boldsymbol{q}_j\right] = 0$ for $i \ne j$, and thus the expected objective value attained by $\boldsymbol{Q}$ is $\mathbb{E}[\langle\mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top, \boldsymbol{A}\rangle] = \sum_{i\in[m]}\mathbb{E}[\boldsymbol{q}_i^\top\boldsymbol{A}^{(i,i)}\boldsymbol{q}_i] = \sum_{i\in[m]}\boldsymbol{v}_i^\top\boldsymbol{A}^{(i,i)}\boldsymbol{v}_i$. When treating the blocks in the order $1, \ldots, m$, we have $\mathbb{E}[\langle\mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top, \boldsymbol{A}\rangle] \ge \boldsymbol{v}_1^\top\boldsymbol{A}^{(1,1)}\boldsymbol{v}_1 = \lambda_{\max}(\boldsymbol{A}^{(1,1)})$. By taking the average over random permutations of $\{1, \ldots, m\}$ as well, we get $\mathbb{E}[\langle\mathrm{vec}(\boldsymbol{Q})\mathrm{vec}(\boldsymbol{Q})^\top, \boldsymbol{A}\rangle] \ge \frac{1}{m}\sum_{i\in[m]}\lambda_{\max}(\boldsymbol{A}^{(i,i)})$.

On the other hand, for any 2×2 block of $\boldsymbol{A}$ we have

$$\begin{pmatrix} \boldsymbol{A}^{(i,i)} & \boldsymbol{A}^{(i,j)} \\ \boldsymbol{A}^{(j,i)} & \boldsymbol{A}^{(j,j)} \end{pmatrix} \succeq \boldsymbol{0},$$

so for any orthogonal matrix $\boldsymbol{U}$,

$$\boldsymbol{u}_i^\top \boldsymbol{A}^{(i,j)} \boldsymbol{u}_j + \boldsymbol{u}_j^\top \boldsymbol{A}^{(j,i)} \boldsymbol{u}_i \leq \boldsymbol{u}_i^\top \boldsymbol{A}^{(i,i)} \boldsymbol{u}_i + \boldsymbol{u}_j^\top \boldsymbol{A}^{(j,j)} \boldsymbol{u}_j,$$

$$\text{and} \quad \sum_{i,j \in [m]} \boldsymbol{u}_i^\top \boldsymbol{A}^{(i,j)} \boldsymbol{u}_j \leq m \sum_{i \in [m]} \boldsymbol{u}_i^\top \boldsymbol{A}^{(i,i)} \boldsymbol{u}_i \leq m \sum_{i \in [m]} \lambda_{\max}(\boldsymbol{A}^{(i,i)}) \leq m^2 \mathbb{E}[\langle \text{vec}(\boldsymbol{Q})\text{vec}(\boldsymbol{Q})^\top, \boldsymbol{A} \rangle].$$

$$\square$$

REMARK EC.5. Our proof technique shows that $\mathbb{E}[\langle \text{vec}(\boldsymbol{Q})\text{vec}(\boldsymbol{Q})^\top, \boldsymbol{A} \rangle] \geq \frac{1}{m} \sum_{i \in [m]} \lambda_{\max}(\boldsymbol{A}^{(i,i)})$. Thus, Algorithm (EC.2) yields a $1/m$-factor approximation when $\boldsymbol{A}$ is a block diagonal matrix and a $1/(2m)$-factor approximation when $\boldsymbol{A}$ is block diagonally dominant. In general, however, the block diagonal objective bounds the full objective within a factor of m, hence the overall $1/m^2$ guarantee. It is also worth noting that the semidefinite

$$\begin{pmatrix} \boldsymbol{A}^{(1,1)} & \boldsymbol{A}^{(1,2)} & \dots \boldsymbol{A}^{(1,m)} \\ \boldsymbol{A}^{(2,1)} & \boldsymbol{A}^{(2,2)} & \dots \boldsymbol{A}^{(2,m)} \\ \boldsymbol{A}^{(3,1)} & \vdots & \ddots \boldsymbol{A}^{(3,m)} \\ \boldsymbol{A}^{(m,1)} & \boldsymbol{A}^{(m,2)} & \dots \boldsymbol{A}^{(m,m)} \end{pmatrix} \preceq m \begin{pmatrix} \boldsymbol{A}^{(1,1)} & \boldsymbol{0} & \dots \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}^{(2,2)} & \dots \boldsymbol{0} \\ \boldsymbol{0} & \vdots & \ddots \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \dots \boldsymbol{A}^{(m,m)} \end{pmatrix}$$

which we implicitly prove as part of our approximation guarantee, is actually a special case of the pinching inequality from quantum information theory (see Mosonyi and Ogawa 2015, Lemma II.2).

EC.6. Basis Pursuit Discussion

In this section, we support our discussion of compact relaxations for low-rank matrix completion problems (Section 4.1) by demonstrating analogous results hold in the low-rank basis pursuit case.

Given a sample $\{A_{i,j}, (i,j) \in \Omega \subseteq [n] \times [m]\}$ of an *exactly* low-rank matrix $\boldsymbol{A} \in \mathbb{R}^{n \times m}$, the goal of the low-rank basis pursuit problem is to recover the lowest rank matrix $\boldsymbol{X}$ that exactly matches all observed entries of $\boldsymbol{A}$ (Candès and Recht 2009). This problem admits the formulation:

$$\min_{\boldsymbol{Y} \in \mathcal{Y}_n} \min_{\boldsymbol{X} \in \mathbb{R}^{n \times m}} \text{tr}(\boldsymbol{Y}) \quad \text{s.t.} \quad \mathcal{P}(\boldsymbol{A}) = \mathcal{P}(\boldsymbol{X}), \boldsymbol{X} = \boldsymbol{Y}\boldsymbol{X}, \tag{EC.14}$$

where $\mathcal{P}(\boldsymbol{A})$ denotes a linear map that masks the hidden entries of $\boldsymbol{A}, \boldsymbol{X}$ such that $\mathcal{P}(\boldsymbol{A})_{i,j} = A_{i,j}$ if $(i,j) \in \Omega$ and 0 otherwise. Following Theorem 3 and applying RLT to the constraints $A_{i,j} - X_{i,j} = 0, \forall (i,j) \in \Omega$ leads to the following relaxation

$$\min_{\boldsymbol{Y} \in \text{Conv}(\mathcal{Y}_n)} \min_{\boldsymbol{X} \in \mathbb{R}^{n \times m}, \boldsymbol{W} \in \mathcal{S}_+^{nm}} \text{tr}(\boldsymbol{Y})$$

$$\text{s.t.} \quad A_{i,j} A_{k,\ell} - A_{k,\ell} X_{i,j} - A_{i,j} X_{k,\ell} + (\boldsymbol{W}^{(i,k)})_{j,\ell} = 0, \forall (i,j), (k,\ell) \in \Omega \times \Omega$$

$$A_{i,j} = X_{i,j}, \forall (i,j) \in \Omega$$

$$\boldsymbol{W} \succeq \text{vec}(\boldsymbol{X}^\top)\text{vec}(\boldsymbol{X}^\top)^\top, \quad \begin{pmatrix} \sum_{i \in [n]} \boldsymbol{W}^{(i,i)} & \boldsymbol{X}^\top \\ \boldsymbol{X} & \boldsymbol{Y} \end{pmatrix} \succeq \boldsymbol{0},$$

$$\tag{EC.15}$$

Similarly to the low-rank matrix completion case, the structure of the compact Shor relaxation means that the off-diagonal blocks of $\boldsymbol{W}$ do not appear in either the objective nor any constraint

involving $\boldsymbol{Y}$. As we prove below, the off-diagonal blocks can, therefore, be eliminated from the relaxation without impacting its optimal value:

PROPOSITION EC.5. *Problem* (EC.15) *attains the same objective value as*

$$
\min_{\boldsymbol{Y} \in \mathrm{Conv}(\mathcal{Y}_n)} \min_{\boldsymbol{X} \in \mathbb{R}^{n \times m}, \ \boldsymbol{S}^i \in \mathcal{S}_+^m, i \in [n]} \mathrm{tr}(\boldsymbol{Y})
$$
$$
s.t. \quad A_{i,j} A_{i,\ell} - A_{i,\ell} X_{i,j} - A_{i,j} X_{i,\ell} + (\boldsymbol{S}^i)_{j,\ell} = 0, \forall (i,j), (i,\ell) \in \Omega \times \Omega
$$
$$
A_{i,j} = X_{i,j}, \forall (i,j) \in \Omega
$$
$$
\boldsymbol{S}^i \succeq \boldsymbol{X}_{i,\cdot} \boldsymbol{X}_{i,\cdot}^\top, \ \begin{pmatrix} \sum_{i \in [n]} \boldsymbol{S}^i & \boldsymbol{X}^\top \\ \boldsymbol{X} & \boldsymbol{Y} \end{pmatrix} \succeq \boldsymbol{0},
$$
$$
\text{(EC.16)}
$$

where $\boldsymbol{X}_{i,\cdot}$ *denotes a column vector containing the ith row of* $\boldsymbol{X}$.

Proof of Proposition EC.5   From a solution to (EC.15), defining $\boldsymbol{S}^i := \boldsymbol{W}^{(i,i)}$ yields a feasible solution to (EC.16) with same objective value. In turn, let us consider a feasible solution to (EC.15), $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{S}^i)$. Define the block matrix $\boldsymbol{W} \in \mathcal{S}^{nm}$ by setting $\boldsymbol{W}^{(i,i)} = \boldsymbol{S}^i$ and $\boldsymbol{W}^{(i,k)} = \boldsymbol{X}_{i,\cdot} \boldsymbol{X}_{k,\cdot}^\top$. Then, it is not hard to see that $\boldsymbol{W} - \mathrm{vec}(\boldsymbol{X}^\top) \mathrm{vec}(\boldsymbol{X}^\top)^\top$ is a block diagonal matrix with on-diagonal blocks $\boldsymbol{S}^i - \boldsymbol{X}_{i,\cdot} \boldsymbol{X}_{i,\cdot}^\top \succeq \boldsymbol{0}$. Thus, $\boldsymbol{W} - \mathrm{vec}(\boldsymbol{X}^\top) \mathrm{vec}(\boldsymbol{X}^\top)^\top \succeq \boldsymbol{0}$. Moreover,

$$
(\boldsymbol{W}^{(i,k)})_{j,\ell} = \begin{cases} (\boldsymbol{S}^i)_{j,\ell} & \text{if } i = k, \\ X_{i,k} X_{k,\ell} & \text{otherwise.} \end{cases}
$$

So the linear constraints indexed by $(i,j), (i,\ell) \in \Omega \times \Omega$ are all satisfied. Thus, $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{W})$ is feasible in (EC.15) and attains the same objective value. $\square$

The preprocessing techniques proposed here also apply directly to phase retrieval problems (cf. Candès and Li 2014). Indeed, phase retrieval is essentially basis pursuit, except we replace the linear constraint $\mathcal{P}(\boldsymbol{A} - \boldsymbol{X}) = \boldsymbol{0}$ with other constraints $\langle \boldsymbol{g}_i \boldsymbol{g}_i^\top, \boldsymbol{X} \rangle = b_i \ \forall i \in \mathcal{I}$. However, the unstructured nature of the linear constraints implies that eliminating as many variables may not be possible.

EC.7. Additional Numerical Results

This section complements Section 5.

EC.7.1. Additional Results for Semi-Orthogonal Quadratic Optimization

Table EC.3 reports the time required to solve our semidefinite relaxation (9) for different values of m, using Mosek as the semidefinite optimization solver. Table EC.4 reports the time required by each feasibility heuristic (excluding time to solve the relaxation when needed).

Table EC.3   Computational time (average and standard deviation) for solving (9) for $n = 50$ and various values of m. Results are aggregated over 5 instances.

m	Average Time (s)	Std Dev (s)
1	0.16	0.01
2	0.31	0.01
5	1.44	0.064
10	5.25	0.22
15	8.10	0.36
20	16.00	0.53
25	25.45	0.62
30	41.33	1.01
50	213.41	22.63

Table EC.4   Computational time (average and standard deviation) for different feasibility heuristics for $n = 50$ and various values of m. For methods that require solving the relaxation (9) (Alg. 2, Alg. 2 with projection, Burer and Park (2024)), time for solving the relaxation is not included but reported in Table EC.3. Results are aggregated over 5 instances.

| m	Alg. 2	Alg. 2 with projection	Burer and Park (2024)	Deflation	Uniform
2	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.02 (0.02)	0.0 (0.0)
5	0.02 (0.0)	0.02 (0.0)	0.01 (0.0)	0.11 (0.02)	0.02 (0.0)
10	0.07 (0.01)	0.07 (0.01)	0.06 (0.0)	0.39 (0.12)	0.07 (0.01)
15	0.15 (0.0)	0.15 (0.0)	0.14 (0.0)	0.62 (0.02)	0.15 (0.0)
20	0.29 (0.03)	0.29 (0.03)	0.29 (0.06)	0.95 (0.02)	0.29 (0.03)
25	0.47 (0.04)	0.47 (0.04)	0.43 (0.0)	1.4 (0.01)	0.47 (0.04)
30	0.82 (0.28)	0.82 (0.28)	0.68 (0.01)	1.99 (0.01)	0.82 (0.28)
50	3.36 (1.62)	3.36 (1.62)	2.35 (0.06)	5.39 (0.01)	3.35 (1.63)

EC.7.2. Additional Results for Low-Rank Matrix Completion

Figure 3 compares the quality of different relaxations for low-rank matrix completion by returning the optimality gap achieved, defined as the relative difference between the lower bound (obtained by each relaxation) and one upper bound (obtained by alternating minimization, AM). Figure EC.1 and EC.2 report the lower and upper bounds separately.
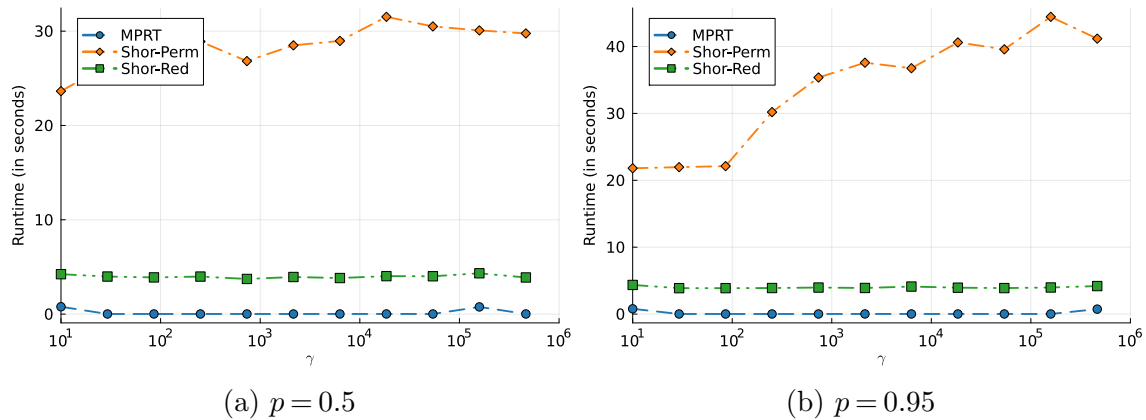
| (a) $p = 0.5$ | (b) $p = 0.95$ |
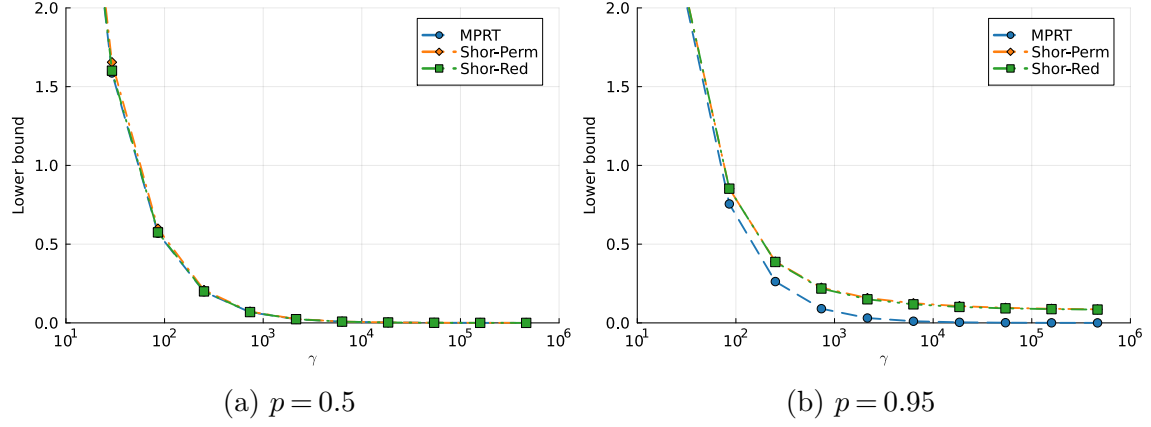
Figure EC.1   Absolute lower bounds as we vary γ for (a) a matrix perspective relaxation ("MPRT"), (b) our Shor relaxation with permutation equalities ("Shor-Perm"), (c) our compact Shor relaxation with no permutation equalities ("GW-Red"), for $p \in \{0.5, 0.95\}$ and $n = 8$.

| (a) $p = 0.5$ | (b) $p = 0.95$ |
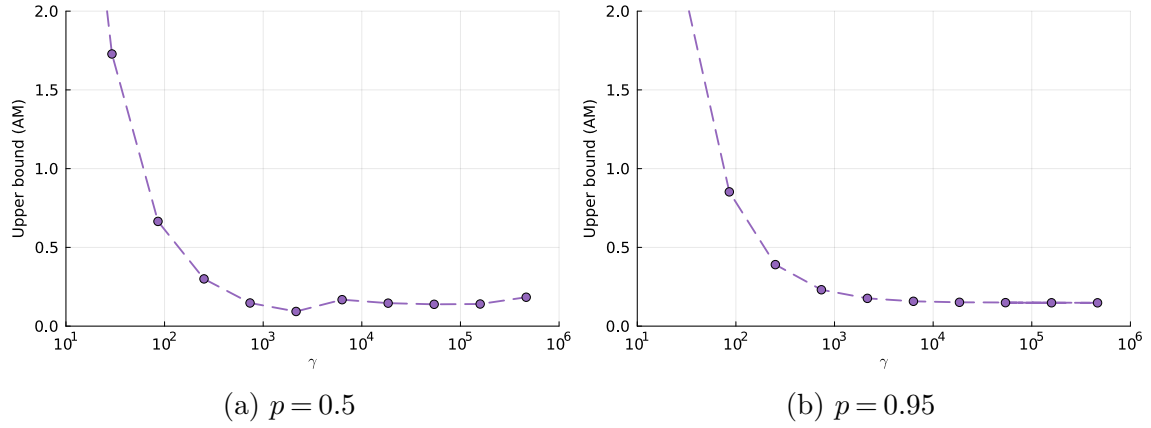
Figure EC.2   Absolute upper bounds as we vary γ for the alternating minimization method of Burer and Monteiro (2003) initialized at a rank-r SVD of $\mathcal{P}(\boldsymbol{A})$ for $p \in \{0.5, 0.95\}$ and $n = 8$.

Figure EC.3 compares the same three relaxations in terms of computational time.

(a) $p = 0.5$ (b) $p = 0.95$

Figure EC.3   Runtimes for (a) a matrix perspective relaxation ("MPRT"), (b) our Shor relaxation with permutation equalities ("Shor-Perm"), (c) our Shor relaxation with no permutation equalities ("GW-Red"), for $p \in \{0.5, 0.95\}$, $n = 8$, and increasing γ.

References

Bao X, Sahinidis NV, Tawarmalani M (2011) Semidefinite relaxations for quadratically constrained quadratic programming: A review and comparisons. *Mathematical Programming* 129:129–157.

Bertsimas D, Cory-Wright R, Pauphilet J (2021) A unified approach to mixed-integer optimization problems with logical constraints. *SIAM Journal on Optimization* 31(3):2340–2367.

Bertsimas D, Cory-Wright R, Pauphilet J (2023) A new perspective on low-rank optimization. *Mathematical Programming* 202:47–92.

Boyd SP, Vandenberghe L (2004) *Convex optimization* (Cambridge University Press).

Burer S, Monteiro RD (2003) A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming* 95(2):329–357.

Burer S, Park K (2024) A strengthened SDP relaxation for quadratic optimization over the Stiefel manifold. *Journal of Optimization Theory and Applications* 202(1):320–339.

Candès EJ, Li X (2014) Solving quadratic equations via phaselift when there are about as many equations as unknowns. *Foundations of Computational Mathematics* 14:1017–1026.

Candès EJ, Recht B (2009) Exact matrix completion via convex optimization. *Foundations of Computational Mathematics* 9(6):717–772.

Chen Y, Ye Y, Wang M (2019) Approximation hardness for a class of sparse optimization problems. *Journal of Machine Learning Research* 20(38):1–27.

Dong H, Chen K, Linderoth J (2015) Regularization vs. relaxation: A conic optimization perspective of statistical variable selection. *arXiv preprint arXiv:1510.06083* .

Gao X, Zhang M, Luo J (2024) Optimized tail bounds for random matrix series. *Entropy* 26(8):633.

Goemans MX, Williamson DP (1995) Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)* 42(6):1115–1145.

Grimmett G, Stirzaker D (2020) *Probability and Random Processes* (Oxford University Press).

Han S, Gómez A, Atamtürk A (2022) The equivalence of optimal perspective formulation and Shor's SDP for quadratic programs with indicator variables. *Operations Research Letters* 50(2):195–198.

Mosonyi M, Ogawa T (2015) Quantum hypothesis testing and the operational interpretation of the quantum rényi relative entropies. *Communications in Mathematical Physics* 334(3):1617–1648.

O'Donnell R, Wu Y (2008) An optimal sdp algorithm for max-cut, and equally optimal long code tests. *Proceedings of the fortieth annual ACM symposium on theory of computing*, 335–344.

Padberg M (1989) The Boolean quadric polytope: some characteristics, facets and relatives. *Mathematical Programming* 45:139–172.

Rendl F, Rinaldi G, Wiegele A (2010) Solving max-cut to optimality by intersecting semidefinite and polyhedral relaxations. *Mathematical Programming* 121:307–335.

Shor NZ (1987) Quadratic optimization problems. *Soviet Journal of Computer and Systems Sciences* 25:1–11.

Tropp JA (2015) An introduction to matrix concentration inequalities. *Foundations and Trends® in Machine Learning* 8(1-2):1–230.

Vershynin R (2018) *High-Dimensional Probability: An Introduction with Applications in Data Science*, volume 47 (Cambridge University Press).

Zheng X, Sun X, Li D (2014) Improving the performance of MIQP solvers for quadratic programs with cardinality and minimum threshold constraints: A semidefinite program approach. *INFORMS Journal on Computing* 26(4):690–703.