COMPUTING RESTRICTED ISOMETRY CONSTANTS VIA MIXED-INTEGER SEMIDEFINITE PROGRAMMING

TRISTAN GALLY AND MARC E. PFETSCH

ABSTRACT. One of the fundamental tasks in compressed sensing is finding the sparsest solution to an underdetermined system of linear equations. It is well known that although this problem is NP-hard, under certain conditions it can be solved by using a linear program which minimizes the 1-norm. The restricted isometry property has been one of the key conditions in this context. However, computing the best constants for this condition is itself NP-hard. In this paper we propose a mixed-integer semidefinite programming approach for computing these optimal constants. This also subsumes sparse principal component analysis. Computational results with this approach allow to evaluate earlier semidefinite relaxations and show that the quality that can be obtained in reasonable time is limited.

Keywords. Compressed Sensing, Restricted Isometry Property, Mixed-Integer Semidefinite Programming

1. Introduction

Finding a sparsest solution of an underdetermined system of linear equations is one key problem in compressed sensing and has been thoroughly investigated in recent years. Given a matrix $A \in \mathbb{R}^{m \times n}$ with $m \leq n$ and right-hand side $b \in \mathbb{R}^m$, the core problem is

$$(P_0) \qquad \begin{aligned} \min \quad & \|x\|_0 \\ \text{s.t.} \quad & Ax = b, \end{aligned}$$

where $\|x\|_0$ denotes the number of nonzero components of x. This problem is NP-hard, see Garey and Johnson [16]. One key observation in compressed sensing is that a solution of (P_0) can be found by solving the following convex optimization problem, if certain conditions are fulfilled:

$$(P_1) \qquad \begin{aligned} \min \quad & \|x\|_1 \\ \text{s.t.} \quad & Ax = b. \end{aligned}$$

In fact, this so-called *basis pursuit* problem can be reformulated as a linear program (LP) and efficient specialized solution algorithms are available both in theory and practice: as an extremely selected list consider the articles [3, 7, 22, 28] and the practical comparison in [18].

2. The Restricted Isometry Property

Several conditions in compressed sensing imply that a solution of (P_0) can be *recovered* by solving (P_1), in which case one speaks of *ℓ_0-ℓ_1-equivalence*. First results in this direction were established by Chen, Donoho, and Saunders [7] and Donoho and Huo [11], with many articles following. We therefore refer the reader to the excellent book by Foucart and Rauhut [14] for further information.

One important condition is the *restricted isometry property* (RIP), which was first introduced by Candès and Tao [6].

Definition 1. *The* restricted isometry property *(RIP) of order k holds for a matrix $A \in \mathbb{R}^{m\times n}$ with constant $\delta \geq 0$, if*

$$(1) \qquad (1-\delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta)\|x\|_2^2$$

for all $x \in \Sigma_k \coloneqq \{x \in \mathbb{R}^n : \|x\|_0 \leq k\}$.

Since the following results depend on the constant δ in (1) being small enough, one is generally interested in the smallest constant satisfying (1).

Definition 2. *For a matrix $A \in \mathbb{R}^{m\times n}$ and order k, the smallest constant in (1),*

$$\delta_k \coloneqq \operatorname*{argmin}_{\delta\geq 0} \big\{(1-\delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta)\|x\|_2^2 \quad \textit{for all } x \in \Sigma_k\big\},$$

is called restricted isometry constant *(RIC) of order k.*

If the restricted isometry property holds for a sufficiently small restricted isometry constant, the optimal solutions of (P_0) and (P_1) coincide. The best-known result of this type is probably the following:

Theorem 1 (Candés [5])**.** *Let the RIC of order $2k$ of A satisfy the condition $\delta_{2k} < \sqrt{2}-1 \approx 0.4142$. If there exists $\tilde{x} \in \mathbb{R}^n$ with $\|\tilde{x}\|_0 \leq k$ and $A\tilde{x} = b$, then the optimal solutions of (P_0) and (P_1) coincide.*

Many results similar to Theorem 1 exist: For example, Foucart [12] relaxed the condition to $\delta_{2k} < 0.4652$ and Foucart and Rauhut [14] to $\delta_{2k} < 0.6246$. Moreover, Cai et al. [4] showed the condition $\delta_k < 0.307$ for Theorem 1 to hold.

For this paper, it is useful to take an asymmetric viewpoint and distinguish between the lower and upper bounds in (1), as proposed by Foucart and Lai [13]:

Definition 3. *The* lower *and* upper restricted isometry constant *are defined as*

$$(2) \qquad \alpha_k \coloneqq \operatorname*{argmax}_{\alpha\geq 0} \ \big\{\alpha^2\,\|x\|_2^2 \leq \|Ax\|_2^2, \ \ \textit{for all } x \in \Sigma_k\big\},$$

$$(3) \qquad \beta_k \coloneqq \operatorname*{argmin}_{\beta\geq 0} \ \big\{\|Ax\|_2^2 \leq \beta^2\,\|x\|_2^2, \ \ \textit{for all } x \in \Sigma_k\big\},$$

respectively. The restricted isometry ratio *(RIR) is defined as $\gamma_k \coloneqq \beta_k^2/\alpha_k^2$ for $\alpha_k \neq 0$.*

Definition 3 is a generalization of Definition 1 in the following sense: if the matrix A satisfies the RIP of order k with RIC δ_k, then it has a RIR of at most $(1+\delta_k)/(1-\delta_k)$. Conversely, if A has lower/upper RICs α_k and β_k, respectively, then this implies a restricted isometry constant of $\delta_k = \max\{1-\alpha_k^2,\ \beta_k^2-1\}$ in (1). So A has a small RIC if and only if it has a large lower and a small upper RIC. But since in (1) usually only one of the two inequalities will be sharp, the asymmetric version yields more information about A.

For the asymmetric version of the RIP in Definition 3, one can again show ℓ_0-ℓ_1-equivalence:

Theorem 2 (Foucart and Lai [13]). *Let the RIR of order $2k$ of A satisfy $\gamma_{2k} \leq 4\sqrt{2} - 3 \approx 2.6569$. If there exists $\tilde{x} \in \mathbb{R}^n$ with $\|\tilde{x}\|_0 \leq k$ such that $A\tilde{x} = b$, then the optimal solutions of (P_0) and (P_1) coincide.*

One advantage of Definition 3 and Theorem 2 is that they are invariant under scaling of $Ax = b$. In fact, the condition of Theorem 1 holding for $Ax = b$ does not imply the same for the scaled equation $\lambda Ax = \lambda b$ with $\lambda \in \mathbb{R} \setminus \{0\}$. In Definition 3, however, we can use $\tilde{\alpha}_k = \lambda\alpha_k$, $\tilde{\beta}_k = \lambda\beta_k$, and therefore $\tilde{\gamma}_k = \lambda^2\alpha_k^2/(\lambda^2\beta_k^2) = \gamma_k$.

Furthermore, we note that the recovery results given in this section can also be derived for the *denoising* case, where instead of (P_0) one considers

$$(P_0^\delta) \qquad \begin{aligned} \min \quad & \|x\|_0 \\ \text{s.t.} \quad & \|Ax - b\|_2 \leq \delta, \end{aligned}$$

for some $\delta \geq 0$; the ℓ_1-problem (P_1) is changed similarly. For more details we refer the reader to [14].

3. Contribution of this Work

There are several types of matrices A for which it is possible to show analytically that the RIP holds, for instance random matrices with high probability, see, e.g., Baraniuk et al. [2]. However, it turns out that computing the restricted isometry constants for a given matrix A and k is NP-hard, see [27]. Note the unfortunate implication that it is hard to check for a concrete random matrix whether it satisfies the RIP, since the theoretical results hold asymptotically and with respect to high probability.

This motivated d'Aspremont et al. [8, 10] to propose several relaxations based on semidefinite programming (SDP) to compute upper bounds on the restricted isometry constants; small upper bounds suffice to guarantee ℓ_0-ℓ_1-equivalence between (P_0) and (P_1), but there exist cases in which the bounds are too weak to give a guarantee.

The relaxations in [8, 10] were originally derived for the connected *sparse principal component analysis* (SPCA). Principal component analysis (PCA) finds orthogonal directions maximizing the variance given in matrix A, solving problems

$$\max \{\|Ax\|_2^2 \,:\, \|x\|_2 \leq 1\}$$

along the way. For SPCA, one further adds the condition that $\|x\|_0 \leq k$ with the goal to reduce the dependency on small components. Note that this is then equivalent to computing the upper RIC and it is NP-hard in the strong sense [27]. SPCA was originally proposed by Zuo et al. [31]. Other authors like Moghaddam et al. [21] also proposed computing a sparse maximal eigenvalue for SPCA.

The goal of this paper is to compute optimal restricted isometry constants via mixed-integer semidefinite programming (MISDP). To this end, we will introduce an MISDP with a rank-1 constraint. In order to utilize this MISDP, we will show that the rank constraint can be relaxed without changing the optimal value. We can then apply an MISDP solver.

Our computational results show that the resulting problems are quite hard to solve, even for matrices of smaller size. However, this is true for the SDP relaxations of d'Aspremont et al. as well, and it allows for an evaluation of the quality of these bounds. In fact, these bounds turn out to be quite tight or weak, depending on the instance. This complements the theoretical approximation guarantees developed by d'Aspremont et al. in [9].

In the following, we start by reviewing the SDP-relaxations of [8] and [10] in Section 4. Afterwards, in Section 5, we construct a mixed-integer semidefinite program and show that it produces exact restricted isometry constants. Finally, in Section 6, we present numerical results and comparisons with the SDP relaxations in [8] and [10].

4. SDP Relaxations

In this section, we will introduce two known SDP relaxations to compute an upper bound on the restricted isometry constant δ_k in (1), since they have some ideas in common with the mixed-integer semidefinite program we want to propose and will also be checked against in the section on numerical results.

The idea is that the optimal constant α_k^2 in (2) can be computed by the non-convex quadratic optimization problem with a cardinality constraint

$$\text{(QP)} \qquad \begin{aligned} \min \quad & \|Ax\|_2^2 \\ \text{s.t.} \quad & \|x\|_2^2 = 1, \\ & \|x\|_0 \leq k. \end{aligned}$$

The corresponding maximization problem allows to compute the constant β_k^2 in (3). D'Aspremont et al. [10] tackled the non-convex quadratic equality constraint $\|x\|_2^2 = 1$ by the technique of semidefinite lifting, see, e.g., Goemans and Williamson [17]. The idea is to use new matrix variables $X = xx^\top$ for the quadratic terms, where the condition $X = xx^\top$ is enforced by the equivalent constraints $X \succeq 0$ and $\text{Rank}(X) = 1$, where, as usual, $X \succeq 0$ means that X is symmetric (shortly written $X \in S_n$) and positive semidefinite. The objective and the normalization-constraint can then be expressed as $\|Ax\|_2^2 = x^\top A^\top Ax = \text{Tr}(A^\top AX)$ and $\|x\|_2^2 = \sum_{i=1}^n x_i^2 = \text{Tr}(X) = 1$.

In the next step, the inequality $\|x\|_0 \leq k$ is substituted by the weaker constraint $\mathbb{1}^\top |X| \mathbb{1} \leq k$, which follows from the fact that $\mathbb{1}^\top |X| \mathbb{1} = \|\text{vec}(X)\|_1 \leq \sqrt{\|\text{vec}(X)\|_0}\, \|\text{vec}(X)\|_2$ and $\|\text{vec}(X)\|_0 = \|x\|_0^2$ in the rank-1 case, where $\text{vec}(X)$ is a vector in $\mathbb{R}^{n^2}$ consisting of all entries of X. Relaxing the non-convex rank constraint finally leads to the following SDP:

$$\text{(A1)} \qquad \begin{aligned} \min \quad & \text{Tr}(A^\top AX) \\ \text{s.t.} \quad & \text{Tr}(X) = 1, \\ & \mathbb{1}^\top |X| \mathbb{1} \leq k, \\ & X \succeq 0. \end{aligned}$$

Another SDP relaxation was proposed by d'Aspremont et al. in [8]. Here,

$$\phi(\Sigma, \rho) = \max\left\{x^\top \Sigma\, x - \rho \|x\|_0 \,:\, \|x\|_2 \le 1\right\}, \tag{4}$$

originally stemming from sparse principal component analysis, is approximately solved for a symmetric positive semidefinite matrix Σ. For $\Sigma = A^\top A$ we can use this to compute an upper bound on β_k via Lagrangian Relaxation as

$$\begin{aligned}\beta_k^2 &= \max\left\{x^\top A^\top A x \,:\, \|x\|_2 \le 1,\ \|x\|_0 \le k\right\}\\ &\le \inf_{\rho \ge 0} \max\left\{x^\top A^\top A x - \rho(\|x\|_0 - k) \,:\, \|x\|_2 \le 1\right\}\\ &= \inf_{\rho\ge 0} \phi(A^\top A, \rho) + \rho\, k.\end{aligned}$$

As one way to solve (4), the semidefinite lifting technique is used again to compute an upper bound on ϕ via

$$\text{(A2-Primal)} \qquad \begin{aligned} \max \quad & \sum_{i=1}^n \operatorname{Tr}(P_i\, B_i)\\ \text{s.t.} \quad & \operatorname{Tr}(X) = 1,\\ & X \succeq P_i \succeq 0 \quad \text{for } i = 1, \dots, n,\end{aligned}$$

for $B_i = b_i\, b_i^\top - \rho\, I$, where b_i is the i-th column of the square-root of Σ. For details, we refer the reader to [8]. The corresponding dual problem can then be written as

$$\text{(A2-Dual)} \qquad \begin{aligned} \min \quad & \lambda_{\max}\Big(\sum_{i=1}^n Y_i\Big)\\ \text{s.t.} \quad & Y_i \succeq B_i \quad \text{for } i = 1, \dots, n,\\ & Y_i \succeq 0 \quad\ \ \text{for } i = 1, \dots, n,\end{aligned}$$

which has only half as many variables as (A2-Primal) and one fewer linear and semidefinite constraint. Then (A2-Dual) can be solved for nonnegative ρ, which should be smaller than Σ_{11}, because otherwise the optimal solution for (4) will always be $x = 0$.

As shown in [8], a lower bound for α_k can be computed via

$$\alpha_k^2 \ge \sup_{\rho\ge 0}\ \psi(A^\top A, \rho) - \rho\, k, \tag{A2}$$

with

$$\psi(\Sigma, \rho) = \min\left\{x^\top \Sigma\, x + \rho \|x\|_0 \,:\, \|x\|_2 \le 1\right\}.$$

In this case problem (A2-Primal) or equivalently (A2-Dual) has to be solved for $\Sigma = MI - A^\top A$, where $M \in \mathbb{R}$ has to be big enough to make Σ positive semidefinite. This implies the following lower bound:

$$\alpha_k^2 \ge \sup_{\rho \ge 0}\ (M - \psi(MI - A^\top A, \rho)) - \rho\, k.$$

In [9] explicit bounds on the gap between (A2-Primal) and the penalized problem (4) were derived using randomization arguments.

5. An MISDP Formulation for the RIP

In this section we propose an MISDP formulation to compute the optimal constants α_k and β_k in (2) and (3), respectively. We will again start with the problem (QP). For the non-convex quadratic equality constraint $\|x\|_2^2 = 1$, we again use semidefinite lifting, introducing the new variable $X = xx^\top$ like in the last section. The cardinality-constraint in (QP) can equivalently be written using binary variables z_i as $-z_i \leq x_i \leq z_i$ and $\sum_{i=1}^n z_i \leq k$. We can thus equivalently rewrite (QP) as

$$\begin{aligned} \min \quad & \mathrm{Tr}(A^\top AX) \\ \text{s.t.} \quad & \mathrm{Tr}(X) = 1, \\ & -z_j \leq X_{ij} \leq z_j \quad \text{for } j = 1, \dots, n, \\ & \sum_{i=1}^n z_i \leq k, \\ & \mathrm{Rank}(X) = 1, \\ & X \succeq 0, \\ & z \in \{0,1\}^n. \end{aligned} \tag{5}$$

We then relax the non-convex rank constraint to arrive at the following MISDP formulation:

$$\begin{aligned} \min \quad & \mathrm{Tr}(A^\top AX) \\ \text{s.t.} \quad & \mathrm{Tr}(X) = 1, \\ & -z_j \leq X_{ij} \leq z_j \quad \text{for } j = 1, \dots, n, \\ & \sum_{i=1}^n z_i \leq k, \\ & X \succeq 0, \\ & z \in \{0,1\}^n. \end{aligned} \tag{MISDP}$$

In the following, we will show that (MISDP) actually produces the exact value of α_k. We will use the following result about the relationship between the rank of solutions of semidefinite programs and the dimension of corresponding faces of the feasible set S, i.e., subsets $F \subseteq S$ such that $\frac{1}{2}(p+q) \in F$ for $p, q \in S$ implies $p, q \in F$:

Theorem 3 (Pataki [24]). *Let $X \in F$, where F is a face of*

$$S := \{X \in S_n \,:\, X \succeq 0,\ \mathit{Tr}(A_i X) = b_i \ \text{ for } i = 1, \dots, \tilde{m}\}$$

for symmetric matrices $A_i \in S_n$, $i = 1, \dots, \tilde{m}$. Then

$$\tfrac{1}{2}\mathit{Rank}(X) \cdot (\mathit{Rank}(X) + 1) \leq \tilde{m} + \mathit{dim}(F).$$

Using this result on the projection of (MISDP) onto the X variables, yields the following.

Theorem 4. *For every $A \in \mathbb{R}^{m \times n}$ and $k \in \mathbb{Z}_{>0}$, there exists an optimal solution $(X^\star, z^\star)$ of (MISDP) with $\mathit{Rank}(X^\star) = 1$.*

Proof. First note that because of the constraint $\mathrm{Tr}(X) = 1$, (MISDP) is bounded. Moreover, $X_{11} = 1$, $X_{ij} = 0$ for all $i, j \neq 1$, $z_1 = 1$, $z_i = 0$ for all $i = 2, \dots, n$ is feasible for $k > 0$. Thus, an optimal solution to (MISDP) exists.

Let $(X^\star, z^\star)$ be an optimal solution to (MISDP) with $\ell := \sum_{i=1}^n z_i^\star \leq k$. Let $T := \{i_1, i_2, \dots, i_\ell\}$ be the support of $z^\star$, and $\check{A} = (A_{i_1}, A_{i_2}, \dots, A_{i_\ell})$ the submatrix of A formed by the columns indexed by T. Then consider

$$\begin{aligned} \min \quad & \mathrm{Tr}(\check{A}^\top \check{A} \check{X}), \\ \text{s.t.} \quad & \mathrm{Tr}(\check{X}) = 1, \\ & \check{X} \succeq 0, \; \check{X} \in S_\ell. \end{aligned} \tag{6}$$

Let $\check{X}$ be the $\ell \times \ell$ submatrix of $X^\star$ with all rows and columns outside of T removed. Then $\check{X}$ is feasible for (6) with the same objective value, since we only removed zero rows and columns of $X^\star$ and $(A^\top A)X^\star$, which do not influence the trace or the positive semidefiniteness.

On the other hand, we can lift any solution $\check{X}$ of (6) to a point $\hat{X} \in S_n$ by extending with zeros. Defining $\hat{z}_i$ to be 1 if and only if $i \in T$, yields a feasible point $(\hat{X}, \hat{z})$ for (MISDP) with identical objective value, since $\mathrm{Tr}(\check{X}) = 1$ and $\check{X} \in S_\ell$ imply $0 \leq \check{X}_{jj} = \hat{X}_{i_j i_j} \leq 1$ for $j = 1, \dots, \ell$. Therefore, $\hat{X}_{ij} \leq \hat{z}_j$ by diagonal dominance of semidefinite matrices.

Thus, the optimal objective values of (6) and (MISDP) agree, and, furthermore, $\mathrm{Rank}(\hat{X}) = \mathrm{Rank}(\check{X})$. Therefore, it suffices to show that (6) has an optimal solution of rank one.

Let $\check{X}$ be an extreme point of (6) (since the set of optimal points of (6) is nonempty, convex, and compact, such a point exists, see [25, Corollary 18.5.1]). Defining the face $F = \{\check{X}\}$ and applying Theorem 3 on (6) gives

$$\tfrac{1}{2}\mathrm{Rank}(\check{X}) \cdot (\mathrm{Rank}(\check{X}) + 1) \leq \tilde{m} + dim(F) = 1 + 0 = 1,$$

since there are $\tilde{m} = 1$ linear constraints. Since $\check{X} = 0$ is infeasible for $\mathrm{Tr}(X) = 1$, it follows that $\mathrm{Rank}(\check{X}) = 1$, which can then be lifted to an optimal solution of (MISDP) of rank one. □

Theorem 4 guarantees that (MISDP) always has a solution of rank one. Therefore, the optimal values of (MISDP) and (QP) agree. We can therefore compute the optimal constant α_k for the RIP by solving (MISDP). Since the same argumentation also holds for the maximization problem for β_k, this allows us to compute the restricted isometry constant and the restricted isometry ratio by mixed-integer semidefinite programming.

Theorem 4 does not guarantee that any found solution will have rank one, only that the optimal objective values agree. This suffices for our application of computing the restricted isometry constants α_k and β_k. For other applications like SPCA, an extreme solution can be generated with an algorithm proposed by Pataki [23].

For efficient application of branch-and-bound to (MISDP), it is important to use strong semidefinite relaxations for fractional z variables. To this end, we can use the following strengthening of the constraints $-z_j \leq X_{ij} \leq z_j$.

Lemma 1. *The inequalities*

$$-\tfrac{1}{2}z_j \le X_{ij} \le \tfrac{1}{2}z_j, \tag{7}$$

for $i \neq j \in \{1,\dots,n\}$*, are valid for* (MISDP).

Proof. Let $X \succeq 0$ satisfy $\mathrm{Tr}(X) = 1$. Then $X_{ii} + X_{jj} \le 1$ and nonnegativity of the diagonal entries imply $X_{ii}X_{jj} \le \frac{1}{4}$ for $i \neq j$. Since all minors of X need to be nonnegative, it follows that

$$X_{ij}^2 \le X_{ii}X_{jj} \le \tfrac{1}{4}.$$

Therefore $-\frac{1}{2} \le X_{ij} \le \frac{1}{2}$, which shows validity of (7). □

6. Numerical Results

In this section we demonstrate the applicability of the proposed MISDP-formulation and use it to assess the quality of the relaxations proposed in [8] and [10]. For solving the mixed-integer semidefinite programs, we use the MISDP solver SCIP-SDP [26, 15], originally developed in [19] and [20]. It combines the branch-and-bound framework SCIP 3.2.1 [1] with interior-point SDP-solvers, in our case SDPA 7.3.8 [29, 30].

For solving the continuous SDPs, the command-line version of SDPA 7.3.8 is used. We compute the left- and right-hand restricted isometry constant α_k and β_k either exactly using (MISDP) or produce corresponding bounds via (A1) and (A2). In the latter case, we solve the SDP (A2-Dual) 15 times for ρ between 0 and the maximum diagonal entry of $A^\top A$. This is the same approach as in [8], except that we reduced the number from 25 to 15, as this did not significantly seem to influence the quality of the solutions, but reduced the solving times.

We use a testset consisting of 63 matrices of the following seven different types: band matrices, with band size three or five with entries within the band chosen uniformly in $\{0,1\}$, binary matrices with all entries chosen uniformly in $\{0,1\}$, Gaussian-distributed matrices, and rank one matrices $A = aa^\top$ with $N(0,1)$-distributed entries of a. In addition to these four types, we also use three types of matrices with scaled parameters. These include matrices with $N(0,1/m)$-distributed entries, with entries chosen uniformly in $\pm 1/\sqrt{m}$, and finally with distribution

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

For each type we used MATLAB 8.3.0 to randomly generate nine matrices of three different sizes with the number of rows ranging from 15 to 30, number of columns from 25 to 40, and order k between three and five, see Table 1.

For the latter three kinds of matrices, it was shown by Baraniuk et al. [2] that for $n \to \infty$, sufficiently small k, and given δ, they satisfy the restricted isometry property of order k and constant δ with probability converging exponentially to one. The needed proportion between n and k is so small, however, that we cannot expect the generated matrices in small dimensions to satisfy the RIP.

Table 1. Sizes of the used matrices; three matrices were randomly generated for each size

	A			B			C		
type	m	n	k	m	n	k	m	n	k
$N(0,1)$	15	30	5	25	35	4	30	40	3
binary	15	30	5	25	35	4	30	40	3
band matrix	30	30	5	35	35	4	40	40	3
rank 1	30	30	5	35	35	4	40	40	3
$N(0,1/m)$	15	30	5	25	35	4	30	40	3
$\pm 1/\sqrt{m}$	15	30	5	25	35	4	30	40	3
$0, \pm\sqrt{3/m}$	15	30	5	25	35	4	30	40	3

The tests were performed on a Linux cluster with Intel i3 CPUs with 3.2GHz, 4MB cache, and 8GB memory running Linux. Each computation was performed single-threaded with a single process running on each computer and with a time limit of four hours. The code was compiled with gcc 4.4.5 with `-O3` optimization. For 39 matrices, the evaluation of (A2-Dual) was aborted for all 15 choices of ρ, because of the memory limit of 8GB. We also failed to solve (MISDP) for the left-hand side of the three medium sized rank one matrices with default parameters, because SDPA failed to solve the root node relaxation. With different settings or a different SDP-solver, however, we could verify that α_k is zero for these instances. These are counted with the maximum allowed time of four hours.

In Table 2, we compare the relative gap between the bounds on α_k and β_k produced by the two heuristics and the exact values given by (MISDP), computed as their difference divided by the exact value. We present arithmetic means over all of the nine matrices for the given type that could be solved within the time- and memory-limit. Note that for the left-hand side we also omitted all instances with $\alpha_k = 0$, since we cannot compute a relative gap in that case. This included all rank one matrices. The number of instances used to compute the average is given in parentheses.

For the left-hand-side α_k of the RIP, the quality of both relaxations is bad: Especially (A2) regularly returned worse solutions than the trivial lower bound of zero. Relaxation (A1) performed slightly better, but the bounds were still not nearly good enough to have any chance of proving ℓ_0-ℓ_1-equivalence. The relatively bad performance of both relaxations for the left-hand side might not be too surprising though, since both relaxations were originally designed for sparse principal component analysis, which corresponds to the right-hand side of the RIP.

In fact, for the right-hand side β_k, both relaxations perform much better. Problem (A1) constantly produces bounds within 10 to 30 % of the optimal solution, with only slightly worse results for binary and rank one matrices. The formulation (A2) shows bigger fluctuations, but produces better results for five of the seven matrix types, while performing even worse for binary and rank one matrices. For band matrices, however, the average gap was below 2 %, the best result for any matrix type among both relaxations.

The restricted isometry constants of all matrices can be found in the appendix. Note that for the unscaled matrices (the first 36 instances until

Table 2. Average gap of relaxations for RICs

matrices	α_k (A1)	α_k (A2)	β_k (A1)	β_k (A2)
$N(0,1)$	90.57 % (9)	384.11 % (6)	26.78 % (9)	12.26 % (6)
binary	87.68 % (8)	958.80 % (5)	55.31 % (9)	143.00 % (6)
band matrix	100.00 % (1)	435.18 % (1)	12.90 % (9)	1.90 % (6)
rank 1	– (0)	– (0)	47.14 % (9)	85.07 % (6)
$N(0,1/m)$	90.04 % (9)	1107.69 % (6)	23.78 % (9)	12.68 % (6)
$\pm 1/\sqrt{m}$	88.39 % (9)	876.21 % (6)	29.20 % (9)	19.67 % (6)
$0, \pm\sqrt{3/m}$	89.30 % (8)	1678.80 % (5)	25.49 % (9)	14.61 % (6)
total	89.48 % (44)	959.70 % (29)	31.51 % (63)	41.31 % (42)

Table 3. Shifted geometric mean of solving times for RICs

matrices	α_k (MISDP)	α_k (A1)	α_k (A2)	β_k (MISDP)	β_k (A1)	β_k (A2)
$N(0,1)$	470.3	2.6	5163.8	137.7	2.5	6568.0
binary	428.1	2.4	5464.8	658.2	3.1	5614.9
band matrix	13.9	2.4	4885.3	12.2	2.8	5380.0
rank 1	–	–	–	102.7	1.5	4145.7
$N(0,1/m)$	324.2	3.8	5667.4	96.3	3.2	7013.3
$\pm 1/\sqrt{m}$	319.3	3.2	5763.8	171.4	2.9	6322.9
$0, \pm\sqrt{3/m}$	229.0	3.6	5534.0	106.3	3.1	6942.2
total	215.7	3.0	5404.6	117.7	2.7	5914.8

instance `rnk140403C`), we expect the restricted isometry constants to be much larger than for the scaled ones. Even for the scaled matrices, the restricted isometry property never holds for the matrices in the testset, since the instances are too small to satisfy the needed relation between n and k.

For reporting the solving times, we will use the *shifted geometric mean* to decrease the influence of easy instances, see Achterberg [1] for more details. The shifted geometric mean of values $x_1, \ldots, x_n$ is computed as

$$\Big(\prod_{i=1}^{n}(x_i + s)\Big)^{1/n} - s,$$

where we used a shift of $s = 10$.

The solving times in seconds for the left- and right-hand side of the RIP are given in Table 3, again excluding the left-hand sides for the rank one matrices, but including all others. The relaxation (A1) is very fast, taking only a few seconds for all instances. The mixed-integer semidefinite program obviously cannot compete with these solving times, although for band-matrices, the difference is relatively small. The Lagrangian Relaxation (A2), however, takes more time to produce bounds than is needed by (MISDP) to compute exact solutions. This is mainly caused by the fact that (A2-Dual) involves $n+1$ matrix variables of dimension $n\times n$ with corresponding SDP-constraints instead of a single one each in (MISDP) and (A1). Furthermore, the big difference is also caused by the instances running into the memory limit, but even without those, the algorithm still takes slightly longer than (MISDP).

7. Conclusion

This paper shows that exact values of restricted isometry constants can be computed via mixed-integer semidefinite programming. The computed bounds significantly improve the bounds computed by the semidefinite relaxations (A1) and (A2) proposed in [10] and [8], respectively. In conclusion, the relaxation (A1) is fast to compute, but produces a gap of around 90 % for α_k and 30 % for β_k. If successfully solved, (A2) provides better bounds for β_k for most types of matrices, but its running time is too large to be used in practice. Moreover, directly solving (MISDP) is often faster and produces the exact value!

Still, the matrices currently handable by (MISDP) are small. For the future, the hope is that with the advancement of solving techniques for mixed-integer semidefinite programs, the sizes of the matrices can be significantly increased. The components to be improved include cutting planes, primal heuristics, and branching rules.

Acknowledgments

The authors would like to thank the German Research Foundation (DFG) for funding, since parts of this research have been carried out within the Collaborative Research Center 805.

References

[1] T. Achterberg. SCIP: solving constraint integer programs. *Mathematical Programming Computation*, 1(1):1–41, 2009.

[2] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3):253–263, 2008.

[3] S. Becker, J. Bobin, and E. J. Candès. NESTA: A fast and accurate first-order method for sparse recovery. *SIAM Journal on Imaging Sciences*, 4(1):1–39, 2011.

[4] T. Cai, L. Wang, and G. Xu. New bounds for restricted isometry constants. *IEEE Transactions on Information Theory*, 56(9):4388–4394, September 2010.

[5] E. J. Candès. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathematique*, 346(9):589–592, 2008.

[6] E. J. Candès and T. Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 59:1207–1223, 2005.

[7] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1999.

[8] A. d'Aspremont, F. Bach, and L. El Ghaoui. Optimal solutions for sparse principal component analysis. *Journal of Machine Learning Research*, 9:1269–1294, 2008.

[9] A. d'Aspremont, F. Bach, and L. El Ghaoui. Approximation bounds for sparse principal component analysis. *Mathematical Programming B*, 148(1-2):89–110, 2014.

[10] A. d'Aspremont, L. El Ghaoui, M. I. Jordan, and G. R. G. Lanckriet. A direct formulation for sparse PCA using semidefinite programming. *SIAM Review*, 49(3):434–448, 2007.

[11] D. L. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory*, 47(7):2845–2862, 2001.

[12] S. Foucart. A note on guaranteed sparse recovery via ℓ_1-minimization. *Applied and Computational Harmonic Analysis*, 29(1):97–103, 2010.

[13] S. Foucart and M.-J. Lai. Sparsest solutions of underdetermined linear systems via ℓ_q-minimization for $0 < q \leq 1$. *Applied and Computational Harmonic Analysis*, 26(3):395–407, 2009.

[14] S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing*. Birkhäuser, Basel, 2013.

[15] T. Gally, M. E. Pfetsch, and S. Ulbrich. A framework for solving mixed-integer semidefinite programs. Technical report, Optimization Online, 2016.

[16] M. Garey and D. Johnson. *Computers and intractability. A guide to the theory of NP-completeness*. W. H. Freeman and Company, 1979.

[17] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the Association for Computing Machinery*, 42(6):1115–1145, 1995.

[18] D. A. Lorenz, M. E. Pfetsch, and A. M. Tillmann. Solving basis pursuit: Heuristic optimality check and solver comparison. *ACM Transactions on Mathematical Software*, 41(2):Article 8, 29 pages, 2015.

[19] S. Mars. *Mixed-Integer Semidefinite Programming with an Application to Truss Topology Design*. PhD thesis, FAU Erlangen-Nürnberg, 2013.

[20] S. Mars and L. Schewe. An SDP-package for SCIP. Technical report, TU Darmstadt and FAU Erlangen-Nürnberg, August 2012.

[21] B. Moghaddam, Y. Weiss, and S. Avidan. Spectral bounds for sparse PCA: Exact and greedy algorithms. *Advances in Neural Information Processing Systems*, 18:915–922, 2006.

[22] M. Osbourne, B. Presnell, and B. Turlach. A new approach to variable selection in least squares problems. *IMA Journal of Numerical Analysis*, 20(3):389–402, 2000.

[23] G. Pataki. Cone-LP's and semidefinite programs: Geometry and a simplex-type method. In W. H. Cunningham, S. T. McCormick, and M. Queyranne, editors, *Proceedings of the 5th International IPCO Conference, Vancouver, Canada*, volume 1084 of *LNCS*, pages 162–174. Springer, Berlin, Heidelberg, 1996.

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

[25] R. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[26] SCIP-SDP. Available at http://www.opt.tu-darmstadt.de/scipsdp/, visited 03/2016.

[27] A. M. Tillmann and M. E. Pfetsch. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing. *IEEE Transactions on Information Theory*, 60(2):1248–1259, 2014.

[28] E. van den Berg and M. P. Friedlander. Probing the Pareto frontier for basis pursuit solutions. *SIAM Journal on Scientific Computing*, 31(2):890–912, 2008.

[29] M. Yamashita, K. Fujisawa, and M. Kojima. Implementation and evaluation of SDPA 6.0 (SemiDefinite Programming Algorithm 6.0). *Optimization Methods and Software*, 18:491–505, 2003.

[30] M. Yamashita, K. Fujisawa, K. Nakata, M. Nakata, M. Fukuda, K. Kobayashi, and K. Goto. A high-performance software package for semidefinite programs: SDPA 7. Technical Report Research Report B-460, Dept. of Mathematical and Computing Science, Tokyo Institute of Technology, September 2010.

[31] H. Zou, T. Hastie, and R. Tibshirani. Sparse principal component analysis. *Journal of Computational & Graphical Statistics*, 15(2):265–286, 2006.

Tristan Gally, Technische Universität Darmstadt, Department of Mathematics, Dolivostrasse 15, 64293 Darmstadt, Germany
E-mail address: gally@mathematik.tu-darmstadt.de

Marc E. Pfetsch, Technische Universität Darmstadt, Department of Mathematics, Dolivostrasse 15, 64293 Darmstadt, Germany
E-mail address: pfetsch@mathematik.tu-darmstadt.de

Appendix

Table 4. List of all matrices in the testset

matrix	m	n	k	α_k^2	β_k^2	γ_k^2	δ_k^2
norm15305A	30	30	5	0.32	59.29	182.61	58.29
norm15305B	30	30	5	0.45	53.97	121.03	52.97
norm15305C	30	30	5	0.44	58.88	133.86	57.88
norm25354A	35	35	4	3.16	71.25	22.54	70.25
norm25354B	35	35	4	3.54	65.82	18.57	64.82
norm25354C	35	35	4	3.82	79.48	20.78	78.48
norm30403A	40	40	3	8.11	66.99	8.26	65.99
norm30403B	40	40	3	6.46	71.20	11.03	70.20
norm30403C	40	40	3	5.63	67.78	12.05	66.78
bina15305A	15	30	5	0.00	41.90	–	40.90
bina15305B	15	30	5	0.16	39.43	252.02	38.43
bina15305C	15	30	5	0.15	40.73	277.19	39.73
bina25354A	25	35	4	0.89	48.21	54.06	47.21
bina25354B	25	35	4	0.94	55.17	58.97	54.17
bina25354C	25	35	4	1.04	45.62	44.07	44.62
bina30403A	30	40	3	2.29	46.70	20.41	45.70
bina30403B	30	40	3	2.85	43.74	15.34	42.74
bina30403C	30	40	3	1.84	49.40	26.83	48.40
band30305A	15	30	5	0.00	5.19	–	4.19
band30305B	15	30	5	0.00	4.46	–	3.46
band30305C	15	30	5	0.00	6.12	–	5.12
band35354A	25	35	4	0.00	8.53	–	7.53
band35354B	25	35	4	0.00	9.58	–	8.58
band35354C	25	35	4	0.09	10.88	125.97	9.88
band40403A	30	40	3	0.00	13.06	–	12.06
band40403B	30	40	3	0.00	14.15	–	13.15
band40403C	30	40	3	0.00	12.07	–	11.07
rnk130305A	15	30	5	0.00	408.40	–	407.40
rnk130305B	15	30	5	0.00	515.61	–	514.61
rnk130305C	15	30	5	0.00	260.21	–	259.21
rnk135354A	25	35	4	0.00	1095.61	–	1094.61
rnk135354B	25	35	4	0.00	566.16	–	565.16
rnk135354C	25	35	4	0.00	678.58	–	677.58
rnk140403A	30	40	3	0.00	508.41	–	507.41
rnk140403B	30	40	3	0.00	652.89	–	651.89
rnk140403C	30	40	3	0.00	718.17	–	717.17
wish15305A	15	30	5	0.03	3.66	141.77	2.66
wish15305B	15	30	5	0.03	3.11	123.06	2.11
wish15305C	15	30	5	0.04	3.29	87.16	2.29
wish25354A	25	35	4	0.15	2.25	14.74	1.25
wish25354B	25	35	4	0.15	2.43	16.06	1.43
wish25354C	25	35	4	0.18	2.87	15.80	1.87
wish30403A	30	40	3	0.19	2.27	12.21	1.27
wish30403B	30	40	3	0.22	2.33	10.84	1.33
wish30403C	30	40	3	0.21	2.37	11.24	1.37
bern15305A	30	30	5	0.03	3.20	96.84	2.20
bern15305B	30	30	5	0.04	2.79	73.25	1.79
bern15305C	30	30	5	0.04	3.22	83.90	2.22
bern25354A	35	35	4	0.17	2.41	14.11	1.41
bern25354B	35	35	4	0.17	2.36	14.25	1.36
bern25354C	35	35	4	0.13	2.55	20.10	1.55
bern30403A	40	40	3	0.29	1.94	6.66	0.94
bern30403B	40	40	3	0.26	2.12	8.23	1.12
bern30403C	40	40	3	0.30	1.85	6.16	0.85
0+-115305A	15	30	5	0.00	4.16	–	3.16
0+-115305B	15	30	5	0.01	3.25	218.58	2.25
0+-115305C	15	30	5	0.01	2.95	198.23	1.95
0+-125354A	25	35	4	0.14	2.92	20.91	1.92
0+-125354B	25	35	4	0.10	2.58	25.94	1.58
0+-125354C	25	35	4	0.16	2.52	15.49	1.52
0+-130403A	30	40	3	0.26	2.24	8.46	1.24
0+-130403B	30	40	3	0.25	2.34	9.40	1.34
0+-130403C	30	40	3	0.07	2.54	35.50	1.54