Sums of Random Symmetric Matrices and Quadratic Optimization under Orthogonality Constraints

Arkadi Nemirovski*

Abstract

Let B_i be deterministic symmetric $m \times m$ matrices, and ξ_i be independent random scalars with zero mean and "of order of one" (e.g., $\xi_i \sim \mathcal{N}(0,1)$). We are interested in conditions for the "typical norm" of the random matrix $S_N = \sum_{i=1}^N \xi_i B_i$ to be of order of 1. An evident necessary condition is $\mathbf{E}\{S_N^2\} \preceq O(1)I$, which, essentially, translates to $\sum_{i=1}^N B_i^2 \preceq I$; a natural conjecture is that the latter condition is sufficient as well. In the paper, we prove a relaxed version of this conjecture, specifically, show that under the above condition the typical norm of S_N is $\leq O(1)m^{\frac{1}{6}}$: $\mathrm{Prob}\{\|S_N\| > \Omega m^{1/6}\} \leq O(1)\exp\{-O(1)\Omega^2\}$ for all $\Omega > 0$ We outline some applications of this result, primarily in investigating the quality of semidefinite relaxations of a general quadratic optimization problem with orthogonality constraints $\mathrm{Opt} = \max_{X_j \in \mathbf{R}^{m\times m}} \left\{F(X_1,...,X_k) : X_j X_j^T = I,\, j = 1,...,k\right\}$, where F is quadratic in $X = (X_1,...,X_k)$. We show that when F is convex in every one of X_j, a natural semidefinite relaxation of the problem is tight within a factor slowly growing with the size m of the matrices X_j: $\mathrm{Opt} \leq \mathrm{Opt}(SDP) \leq O(1)[m^{1/3} + \ln k]\mathrm{Opt}$.

AMS Subject Classification: 60F10, 90C22, 90C25, 90C59.

Key words: large deviations, random perturbations of linear matrix inequalities, semidefinite relaxations, orthogonality constraints, Procrustes problem.

1 Introduction

In this paper, we address the following question:

> (Q): *Let Ξ_i, $1 \leq i \leq N$, be independent random $m \times m$ symmetric matrices with zero mean and "light-tail" distributions, and let $S_N = \sum_{i=1}^N \Xi_i$. Under what conditions a "typical value" of $\|S_N\|$ is "of order of 1" so that the probability for $\|S_N\|$ to be $\geq \Omega$ goes to 0 exponentially fast as $\Omega > 1$ grows?*
> Here and in what follows $\|A\|$ denotes the standard spectral norm (the largest singular value) of a matrix A.

This informal question admits various formal settings; to motivate the one we focus on, we start with describing two applications we intend to consider: *tractable approximations of randomly perturbed Linear Matrix Inequalities* (LMI) and *semidefinite relaxations of nonconvex quadratic minimization under orthogonality constraints*.

*on leave from the Technion – Israel Institute of Technology at School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA; nemirovs@ie.technion.ac.il

Randomly perturbed LMI's. Consider a randomly perturbed LMI

$$A_0[x] - \sum_{i=1}^{N} \xi_i A_i[x] \succeq 0, \tag{1}$$

where $A_0[x], ..., A_N[x]$ are affine functions of the decision vector x taking values in the space $\mathbf{S}^m$ of symmetric $m \times m$ matrices, and ξ_i are independent of each other random perturbations (which w.l.o.g. can be assumed to have zero means). Constraints of this type arise in many applications, e.g., in various optimization and Control problems with randomly perturbed data. A natural way to treat a randomly perturbed constraint is to pass to its *chance* form, which in the case of constraint (1) is the deterministic constraint

$$\text{Prob}\left\{\xi = (\xi_1, ..., \xi_N) : A_0[x] - \sum_{i=1}^{N} \xi_i A_i[x] \succeq 0\right\} \geq 1 - \epsilon, \tag{2}$$

where $\epsilon > 0$ is a small tolerance. The resulting chance constraint, however, typically is "heavily computationally intractable" – usually, the probability in the left hand side cannot be computed efficiently, and its reliable estimation by Monte-Carlo techniques requires samples of order of ϵ^{-1}, which is prohibitively time-consuming when ϵ is small (like 1.e-6 or 1.e-8). In the rare cases when this difficulty can be circumvented (e.g., when ϵ is not too small), one still have a severe problem: chance constraint (2) defines, in general, a nonconvex set in the space of x-variables, and therefore it is absolutely unclear how to optimize under this constraint. A natural way to overcome this difficulty is to replace "intractable" chance constraint (1) with its "tractable approximation" – an explicit convex constraint on x such that its validity at a point x implies that x is feasible for (1). Now note that an evident necessary condition for x to be feasible for (1) is $A_0[x] \succeq 0$; strengthening this necessary condition to $A_0[x] \succ 0$, x is feasible for the chance constraint if and only if the random sum $S_N = \sum_{i=1}^{N} \underbrace{\xi_i A_0^{-1/2}[x] A_i[x] A_0^{-1/2}[x]}_{\Xi_i}$ is $\preceq I_m$ with probability $\geq 1 - \epsilon$. Assuming, as it is typically the case, that the distributions of ξ_i are symmetric, this condition is essentially the same as the condition $\|S_N\| \leq 1$ with probability $\geq 1 - \epsilon$. If we knew how to answer (Q), we could use this answer to build a "tractable" sufficient condition for $\|S_N\|$ to be ≤ 1 with probability close to 1 and thus could build a tractable approximation of (2).

Nonconvex quadratic optimization under orthogonality constraints. Here we present a single example – the *Procrustes problem*, postponing the in-depth considerations till section 4. In the Procrustes problem, one is given matrices $a[k]$, $k = 1, ..., K$, of the same size $m \times n$ and is looking for K orthogonal $n \times n$ matrices $x[k]$ minimizing the objective

$$\sum_{1 \leq k < k' \leq K} \|a[k]x[k] - a[k']x[k']\|_2^2,$$

where $\|a\|_2 = \sqrt{\text{Tr}(aa^T)}$ is the Frobenius norm of a matrix. Informally speaking, we are given K collections of points in $\mathbf{R}^n$ (s-th element of k-th collection is the s-th row of $a[k]$) and are seeking for rotations which make these collections as close to each other as possible, the closeness being quantified by the sum, over s, k, k', of squared Euclidean distances between s-th points of

k-th and k'-th collections. For various applications of this problem, see [2, 7, 8, 9]. The problem clearly is equivalent to the quadratic maximization problem

$$\max_{x[1],...,x[K]} \left\{ 2 \sum_{k<k'} \mathrm{Tr}(a[k]x[k]x^T[k']a^T[k']) : x[k] \in \mathbf{R}^{n\times n},\, x[k]x^T[k] = I_n,\, i = 1, ..., K \right\}. \qquad (P)$$

When $K > 2$, the problem is intractable (for $K = 2$, there is a closed form solution); it, however, allows for a straightforward semidefinite relaxation. Let $X = X[x[1], ..., x[K]]$ be the symmetric matrix defined as follows: the rows and the columns in X are indexed by triples (k, i, j), where k runs from 1 to K and i, j run from 1 to n; the entry $X_{kij,k'i'j'}$ in X is $x_{ij}[k]x_{i'j'}[k']$. Note that X is symmetric positive semidefinite matrix of rank 1. Further, the relation $x[k]x^T[k] = I_n$ is equivalent to a certain system $\mathcal{S}_k$ of linear equations on the entries of X, while the relation $x^T[k]x[k] = I_n$ (in fact equivalent to $x^T[k]x[k] = I_n$) is equivalent to another system $\mathcal{T}_k$ of linear equations on the entries of X. Finally, the objective in (P) is a linear function $\mathrm{Tr}(AX)$ of X, where A is an appropriate symmetric matrix of the same size $Kn^2 \times Kn^2$ as X. It is immediately seen that (P) is equivalent to the problem

$$\max_{X\in\mathbf{S}^{Kn^2}} \left\{ \mathrm{Tr}(AX) : X \succeq 0, X \text{ satisfies } \mathcal{S}_k, \mathcal{T}_k,\, k = 1, ..., K,\, \mathrm{Rank}(X) = 1 \right\};$$

removing the only troublemaking constraint $\mathrm{Rank}(X) = 1$, we end up with an explicit semidefinite program

$$\max_{X\in\mathbf{S}^{Kn^2}} \left\{ \mathrm{Tr}(AX) : X \succeq 0, X \text{ satisfies } \mathcal{S}_k, \mathcal{T}_k,\, k = 1, ..., K \right\} \qquad (\mathrm{SDP})$$

which is a relaxation of (P), so that $\mathrm{Opt(SDP)} \geq \mathrm{Opt}(P)$. We shall see in section 4 that an appropriate answer to (Q) allows to prove that

$$\mathrm{Opt(SDP)} \leq O(1)(n^{\frac{1}{3}} + \ln K)\mathrm{Opt}(P), \qquad (3)$$

and similarly for other problems of quadratic optimization under orthogonality constraints. To the best of our knowledge, (3) is the first nontrivial bound on the quality of semidefinite relaxation for problems of this type.

The outlined applications motivate our specific approach to treating (Q). First, we are interested in the case when the size m of the random matrices in question can be large, and pay primary attention on how this size enters the results (as we shall see, this is the only way to get nontrivial bounds for our second application). In this respect, our goals are similar to those pursued in huge literature on large-scale random matrices inspired by applications in Physics. However, we cannot borrow much from this literature, since the assumptions which are traditional there (appropriate pattern of independence/weak dependence of entries in S_N) makes no sense for our applications. What we are interested in when answering (Q), are conditions expressed in terms of distributions of random terms Ξ_i in S_N. Let us try to understand what could be the "weakest possible" condition of this type. In the case when $\mathrm{Prob}\{\|S_N\| > \Omega\}$ goes rapidly to 0 as $\Omega > 1$ grows, we clearly should have $\mathbf{E}\{S_N^2\} \preceq O(1)I_m$ (since $S_N^2 \preceq \|S_N\|^2 I_m$). Thus, the condition

$$\left[\mathbf{E}\left\{S_N^2\right\} =\right] \quad \sum_{i=1}^N \mathbf{E}\left\{\Xi_i^2\right\} \preceq O(1)I_m \qquad (4)$$

is *necessary* for $\|S_N\|$ to be "of order of 1". A natural *guess* is that this necessary condition plus appropriate "light-tail" assumptions on the distributions of Ξ_i is sufficient for the property

in question; we shall see in a while that if this guess were true, it would provide us with all we need in our applications. Unfortunately, when interpreted literally, the guess fails to be true. First, it is immediately seen that in fact $O(1)I_m$ in the right hand side of (4) should be reduced to $O(1)\frac{1}{\ln m}I_m$. Indeed, let Ξ_i be diagonal matrices with independent (from position to position and for different t's) diagonal entries taking values $\pm\alpha N^{-1/2}$ with probabilities 1/2, so that

$$\Sigma \equiv \sum_{i=1}^{N} \mathbf{E}\left\{\Xi_i^2\right\} = \alpha^2 I_m.$$

Here S_N is a random diagonal matrix with i.i.d. diagonal entries; by Central Limit Theorem, the distribution of these entries approaches, as N grows, the Gaussian distribution $\mathcal{N}(0,\alpha^2)$. It follows that when N is large, the typical value of $\|S_N\|$ is the same as the typical value of $\max\limits_{i\le m}|\zeta_i|$, with independent $\zeta_i \sim \mathcal{N}(0,\alpha^2)$; in other words, for large N the typical value of $\|S_N\|$ is $\alpha\sqrt{2\ln m}$. In order for this quantity to be of order of 1, α should be of order of $(\ln m)^{-1/2}$, which corresponds to Σ of order of $(\ln m)^{-1}I_m$ rather than of order of I_m. In our context, the consequences of the outlined correction are not that dramatic, since $\ln m$, for all practical purposes, is a moderate constant. A less pleasant observation is that *the corrected guess still fails to be true, unless we impose further restrictions on the distributions of* Ξ_i. Indeed, consider the case when $m = 2k$ is even, $N = 1$, and the random matrix $\Xi_1 = S_N$ is $\left[\begin{array}{c|c} & \eta\xi^T \\ \hline \xi\eta^T & \end{array}\right]$, where η is uniformly distributed on the unit sphere in $\mathbf{R}^k$, $\xi \sim \mathcal{N}(0, I_k)$ and η, ξ are independent. In this case, direct computation demonstrates that $\mathbf{E}\left\{\Xi_1^2\right\} = I_m$, while $\|\Xi_1\| = \|S_N\| = \|\eta\|_2\|\xi\|_2$, so that the typical value of $\|S_N\|$ is as large as $O(\sqrt{m})$. It follows that in order to make our guess valid for the particular case we are considering, the right hand side in (4) should be reduced to $O(1)m^{-1}I_m$. After such a correction, our guess does become valid, but the correction itself turns out to be too bad for our tentative applications. What we intend to do is to try to save the "logarithmically corrected" guess at the cost of restricting Ξ_i to be *semi-scalar*, that is, to be random matrices of the form $\xi_i B_i$, where B_i are deterministic symmetric matrices and ξ_i are independent random scalars with zero mean and light-tail distributions. Specifically, we make a *conjecture* as follows:

Conjecture 1.1 *Let B_i, $i = 1, ..., N$, be deterministic symmetric $m \times m$ matrices such that*

$$\sum_{i=1}^{N} B_i^2 \preceq I_m, \tag{5}$$

and let ξ_i, $i = 1, ..., N$, be independent random scalars with zero mean and "of order of 1", e.g., such that (a) $|\xi_i| \le 1$, or (b) $\xi_i \sim \mathcal{N}(0,1)$, or (c) $\mathbf{E}\left\{\exp\{\xi_i^2\}\right\} \le \exp\{1\}$. Then

$$\Omega \ge O(1)\sqrt{\ln m} \Rightarrow \mathrm{Prob}\left\{\xi = (\xi_1, ..., \xi_N) : \|\sum_{i=1}^{N} \xi_i B_i\| \ge \Omega\right\} \le O(1)\exp\{-O(1)\Omega^2\} \tag{6}$$

with appropriate positive absolute constants $O(1)$.

It turns out that (6) would satisfy all the requirements posed by the applications we bear in mind. Unfortunately, for the time being we are unable to prove the conjecture "as it is". The primary goal of this paper is to prove a weaker statement – the one where $\sqrt{\ln m}$ in the premise of (6) is replaced with $m^{\frac{1}{6}}$, and to use this weaker fact in the applications we have mentioned.

In our opinion, question (Q) in general, and its specialization as presented in Conjecture 1.1, in particular are quite natural and deserve attention by their own right. Surprisingly, the only, to the best of our knowledge, result in this direction which makes no assumptions on how strong the entries in S_N depend on each other, is recent result announced in [5] (for proof, see [6]) as follows:

Proposition 1 *Let Ξ_i be independent symmetric $m \times m$ matrices with zero mean such that*

$$\mathbf{E}\left\{\exp\{\|\Xi_i\|^2\sigma_i^{-2}\}\right\} \le \exp\{1\}, \quad i = 1, ..., N$$

($\sigma_i > 0$ are deterministic scale factors). Then

$$\text{Prob}\left\{\|S_N\| \ge t\sqrt{\sum_{i=1}^N \sigma_i^2}\right\} \le O(1)\exp\{-O(1)\frac{t^2}{\ln m}\} \quad \forall t > 0. \tag{7}$$

with positive absolute constants $O(1)$.

From this Proposition it follows that when strengthening the premise (5) in Conjecture 1.1 to $\sum_{i=1}^N \|B_i\|^2 \le 1$, the conclusion becomes "nearly true":

$$\text{Prob}\left\{\xi = (\xi_1, ..., \xi_N) : \|\sum_{i=1}^T \xi_i B_i\| \ge \Omega\sqrt{\ln m}\right\} \le O(1)\exp\{-O(1)\Omega^2\}.$$

Unfortunately, in the applications we intend to consider strengthening the matrix inequality $\sum_i B_i^2 \preceq I$ to the scalar inequality $\sum_i \|B_i\|^2 \le 1$ is too costly to be of actual use.

The rest of the paper is organized as follows. In section 2, we prove that our conjecture, in its outlined weaker form, indeed is valid. In sections 3 and 4 we apply this result to approximating chance constraints associated with randomly perturbed LMI's, and to deriving bound on the quality of semidefinite relaxations of problems of quadratic approximation under orthogonality constraints.

2 Main result

2.1 Preliminaries: Talagrand's Inequality

We start with the following instrumental fact:

Theorem 2.1 [Talagrand's Inequality] *Let $(E_i, \|\cdot\|_i)$, $i = 1, ..., N$, be finite-dimensional normed spaces and μ_i, $i = 1, ..., N$, be Borel probability measures on the balls $V_i = \{x_i \in T_i : \|x_i\|_i \le 1/2\}$. Let us equip the space $E = E_1 \times ... \times E_N$ with the norm $\|(x_1, ..., x_N)\| = \sqrt{\sum_{i=1}^N \|x_i\|_i^2}$ and with the probability distribution $\mu = \mu_1 \times ... \times \mu_N$, and let A be a closed convex set in E such that $\mu(A) > 0$. Then*

$$\int_E \exp\{\frac{\text{dist}^2_{\|\cdot\|}(x, A)}{4}\}\mu(dx) \le \frac{1}{\mu(A)}, \tag{8}$$

where $\text{dist}_{\|\cdot\|}(x, A) = \min_{z \in A} \|x - z\|$.

In this form, the Talagrand Inequality is proved in [4], up to the only difference that in [4], the supports of μ_i are assumed to be finite subsets of V_i. However, finiteness of the supports is of no importance, since a Borel probability measure on V_i can be weakly approximated by probability measures with finite supports contained in V_i.

2.2 Main result

Our main result related to question (Q) is as follows:

Theorem 2.2 *Let $\xi_1, ..., \xi_N$ be independent random variables with zero mean and zero third moment taking values in $[-1, 1]$, B_i, $i = 1, ..., N$, be deterministic symmetric $m \times m$ matrices, and $\Theta > 0$ be a real such that*

$$\sum_i B_i^2 \preceq \Theta^2 I. \tag{9}$$

Then

$$\begin{array}{l} \Omega \geq 7m^{1/4} \Rightarrow \mathrm{Prob}\{\| \sum\limits_{i=1}^N \xi_i B_i\| \geq \Omega\Theta\} \leq \frac{5}{4}\exp\{-\frac{\Omega^2}{32}\} \quad (a) \\ \Omega \geq 7m^{1/6} \Rightarrow \mathrm{Prob}\{\| \sum\limits_{i=1}^N \xi_i B_i\| \geq \Omega\Theta\} \leq 22\exp\{-\frac{\Omega^2}{32}\} \quad (b) \end{array} \tag{10}$$

Proof. 1^0. Our first observation is as follows:

Lemma 1 *Let $\Theta > 0$, let $B_i \in \mathbf{S}^m$ be deterministic matrices satisfying (9) and ζ_i be independent random scalar variables such that*

$$\mathbf{E}\{\zeta_i\} = 0,\ \mathbf{E}\left\{\zeta_i^2\right\} \leq \sigma^2,\ \mathbf{E}\left\{\zeta_i^4\right\} - \left(\mathbf{E}\{\zeta_i^2\}\right)^2 \leq 2\sigma^4.$$

Let, finally, $S_k = \sum\limits_{i=1}^k \zeta_i B_i$, $1 \leq k \leq N$. Then

$$1 \leq k \leq N \Rightarrow \mathbf{E}\left\{S_k^4\right\} \preceq 3\sigma^4\Theta^4 I. \tag{11}$$

Proof. Setting $S_0 = 0$, $E_i = \mathbf{E}\{S_i^4\}$, $\sigma_i = \left(\mathbf{E}\{\zeta_i^2\}\right)^{1/2}$, $\omega_i = \left(\mathbf{E}\{\zeta_i^4\}\right)^{1/4}$ and taking into account that ζ_i and S_{i-1} are independent with zero mean, we have

$$\begin{array}{rcl} E_i = & \mathbf{E}\left\{[S_{i-1} + \zeta_i B_i]^4\right\} & \\ = & \mathbf{E}\Big\{ S_{i-1}^4 + \sigma_i^2 \Big[\underbrace{S_{i-1}B_iS_{i-1}B_i + B_iS_{i-1}B_iS_{i-1}}_{\preceq S_{i-1}B_i^2S_{i-1} + B_iS_{i-1}^2B_i \text{ due to } xy^T+yx^T \preceq xx^T+yy^T} + S_{i-1}^2B_i^2 + B_i^2S_{i-1}^2 + S_{i-1}B_i^2S_{i-1} + B_iS_{i-1}^2B_i \Big] & \\ & + \omega_i^4 B_i^4 \Big\} & \\ \preceq & \mathbf{E}\left\{S_{i-1}^4 + 2\sigma_i^2 S_{i-1}B_i^2S_{i-1} + 2\sigma_i^2 B_iS_{i-1}^2B_i + S_{i-1}^2(\sigma_i^2B_i^2) + (\sigma_i^2B_i^2)S_{i-1}^2 + \sigma_i^4B_i^4 + (\omega_i^4 - \sigma_i^4)B_i^2\right\} & \\ = & E_{i-1} + 2\sum\limits_{j=1}^{i-1}\sigma_j^2\sigma_i^2B_jB_i^2B_j + 2\sum\limits_{j=1}^{i-1}\sigma_i^2\sigma_j^2B_iB_j^2B_i + \underbrace{\sum\limits_{j=1}^{i-1}\sigma_i^2\sigma_j^2B_j^2B_i^2 + \sum\limits_{j=1}^{i-1}\sigma_i^2\sigma_j^2B_i^2B_j^2 + \sigma_i^4B_i^4}_{= \left(\sum\limits_{j=1}^{i}\sigma_j^2B_j^2\right)^2 - \left(\sum\limits_{j=1}^{i-1}\sigma_j^2B_j^2\right)^2} & \\ & + [\omega_i^4 - \sigma_i^4]B_i^4 & \end{array}$$

whence

$$\begin{aligned}
E_k &\preceq 2\sum_{i=1}^{k}\sum_{j=1}^{i-1}\sigma_i^2\sigma_j^2 B_j B_i^2 B_j + 2\sum_{i=1}^{k}\sum_{j=1}^{i-1}\sigma_i^2\sigma_j^2 B_i B_j^2 B_i + \left(\sum_{j=1}^{k}\sigma_j^2 B_j^2\right)^2 + \sum_{i=1}^{k}\omega_i^4 B_i^4 \\
&= 2\sum_{\substack{1\le i,j\le k\\ i\ne j}}\sigma_i^2\sigma_j^2 B_i B_j^2 B_i + \left(\sum_{j=1}^{k}\sigma_j^2 B_j^2\right)^2 + \sum_{i=1}^{k}\left[\omega_i^4-\sigma_i^4\right]B_i^4 \\
&\preceq 2\sum_{\substack{1\le i,j\le k\\ i\ne j}}\sigma^4 B_i B_j^2 B_i + \left(\sum_{j=1}^{k}\sigma_j^2 B_j^2\right)^2 + \sum_{i=1}^{k}2\sigma^4 B_i^4 \preceq 2\sigma^4\sum_{i,j=1}^{k} B_i B_j^2 B_i + \left(\sum_{j=1}^{k}\sigma_j^2 B_j^2\right)^2 \\
&= 2\sigma^4\sum_{i=1}^{k} B_i\left[\sum_{j=1}^{k}B_j^2\right]B_i + \left(\sum_{j=1}^{k}\sigma_j^2 B_j^2\right)^2 \preceq 2\sigma^4\Theta^2\sum_{i=1}^{k}B_i^2 + \Big(\underbrace{\sum_{j=1}^{k}\sigma_j^2 B_j^2}_{=A,\,0\preceq A\preceq\sigma^2\Theta^2 I}\Big)^2 \\
&\preceq (2\sigma^4\Theta^4+\sigma^4\Theta^4)I,
\end{aligned}$$

as claimed. □

2^0. Now we are ready to prove (10.a). For $x\in\mathbf{R}^N$, let $S(x)=\sum_{i=1}^{N}x_iB_i$, and let $Q=\{x\in\mathbf{R}^N : \|S(x)\|\le\Theta\}$, so that Q is a closed convex set in $\mathbf{R}^N$ symmetric w.r.t. the origin. We claim that Q contains the centered at the origin unit Euclidean ball. Indeed, all we should verify is that if $\|x\|_2\le 1$, then $|y^T(\sum_i x_iB_i)y|\le\Theta$ for all $y\in\mathbf{R}^m$ with $y^Ty\le 1$. We have

$$|y^T(\sum_i x_iB_i)y|\le\sum_i|x_i|\|B_iy\|_2\le\left(\sum_i x_i^2\right)^{1/2}\left(\sum_i y^TB_i^2y\right)^{1/2}\le\Theta,$$

as claimed.

Applying Lemma 1 to $\zeta_i=\xi_i$ (which allows to take $\sigma=1$), we get

$$\mathbf{E}\left\{\|S(\xi)\|^4\right\}\le\mathbf{E}\left\{\mathrm{Tr}(S^4(\xi))\right\}\le 3m\Theta^4,$$

whence by Tschebyshev inequality

$$\gamma>0\Rightarrow\mathrm{Prob}\left\{\|S(\xi)\|>2\gamma\Theta\right\}<\frac{3m}{16\gamma^4}. \tag{12}$$

Now let $\gamma=m^{1/4}$ and $A=\gamma Q$. The set A is closed and convex and contains the centered at the origin ball of the radius γ. It follows that if $s>1$ and $x\notin sA=s\gamma Q$ (or, which is the same, $\|S(x)\|>\gamma s\Theta$), then $\mathrm{dist}_{\|\cdot\|}(x,A)\ge(s-1)\gamma=(s-1)m^{1/4}$. Applying Talagrand Inequality to the distribution of the random vector $\zeta=\xi/2$ (Theorem 2.1), we get

$$\begin{aligned}
&\mathrm{Prob}\left\{\|S(\xi)\|>2sm^{1/4}\Theta\right\}=\mathrm{Prob}\left\{\xi/2\notin sA\right\}\le\frac{1}{\mathrm{Prob}\{\xi\in 2A\}}\exp\{-\tfrac{(s-1)^2m^{1/2}}{4}\}\\
&\le\tfrac{5}{4}\exp\{-\tfrac{(s-1)^2m^{1/2}}{4}\},
\end{aligned} \tag{13}$$

where the concluding inequality is given by (12). The resulting inequality is valid for all $s>1$, and (10.a) follows.

3^0. Now let us prove (10.b). We start with the following weak analogy to Lemma 1:

Lemma 2 *Let B_i, $i=1,...,N$, be deterministic symmetric matrices satisfying* (9)*, and ζ_i, $i=1,...,N$, be independent scalar random variables with zero mean and zero third moment*

such that $\sigma_i^2 \equiv \mathbf{E}\{\zeta_i^2\} \le \sigma^2$, $\omega_i^4 \equiv \mathbf{E}\{\zeta_i^4\} \le \min\left[\sigma_i^4 + 2\sigma^4, \omega^4\right]$, $\chi_i^6 \equiv \mathbf{E}\{\zeta_i^6\} \le \chi^6$, *and let* $S_k = \sum_{i=1}^{k} \zeta_i B_i$. *Then*

$$\mathbf{E}\left\{\mathrm{Tr}(S_k^6)\right\} \le [45\sigma^6 + 15\omega^4\sigma^2 + \chi^6]\Theta^6 m. \tag{14}$$

Proof. Let $E_i = \mathbf{E}\left\{S_i^4\right\}$, $\phi_i = \mathbf{E}\left\{\mathrm{Tr}(S_i^6)\right\}$. Given a multi-index $\iota = (\iota_1, ..., \iota_n)$ with entries $0, 1$ and two symmetric matrices P, Q, let $[P, Q]^\iota$ stand for the product of n matrices, with ℓ-th factor being P or Q depending on whether $\iota_\ell = 1$ or $\iota_\ell = 0$ (e.g., $[P, Q]^{(0,1,1)} = QP^2$). Let I, J be the sets of 6-dimensional multi-indices ι with entries 0,1 such that exactly 4, respectively, 2 of the entries are equal to 1 (so that both I and J contain 15 multi-indices each). Taking into account that S_{i-1} has zero mean and is independent of $\zeta_i B_i$, and that ζ_i has zero first and third moments, we have

$$\mathbf{E}\left\{S_i^6\right\} = \mathbf{E}\left\{S_{i-1}^6\right\} + \sigma_i^2 \sum_{\iota\in I} \mathbf{E}\left\{[S_{i-1}, B_i]^\iota\right\} + \omega_i^4 \sum_{\iota\in J} \mathbf{E}\left\{[S_{i-1}, B_i]^\iota\right\} + \chi_i^6 B_i^6,$$

whence

$$\phi_i \le \phi_{i-1} + \sigma_i^2 \sum_{\iota\in I} \mathbf{E}\left\{\mathrm{Tr}([S_{i-1}, B_i]^\iota)\right\} + \omega_i^4 \sum_{\iota\in J} \mathbf{E}\left\{\mathrm{Tr}([S_{i-1}, B_i]^\iota)\right\} + \chi_i^6 \mathrm{Tr}(B_i^6). \tag{15}$$

Now let us list all 15 products $[S_{i-1}, B_i]^\iota$, $\iota \in I$; we split these products into groups, all members of the same group being of equal trace in view of the identities $\mathrm{Tr}(A) = \mathrm{Tr}(A^T)$ and $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. Here are the groups (to simplify notation, we skip indices of S_{i-1} and B_i)

$$\begin{array}{ll} BS^4B, S^2B^2S^2, (S^4B^2, B^2S^4), (S^3B^2S, SB^2S^3) & (a)\\ SBS^2BS, (S^2BS^2B, BS^2BS^2) & (b)\\ (S^3BSB, BSBS^3), (SBS^3B, BS^3BS), (S^2BSBS, SBSBS^2) & (c)\end{array}$$

Let the traces of products in the respective groups be $T_a = T_{a,i}(\zeta_1, ..., \zeta_{i-1})$, $T_b = T_{b,i}(\zeta_1, ..., \zeta_{i-1})$, $T_c = T_{c,i}(\zeta_1, ..., \zeta_{i-1})$. We have

$$\underbrace{BS^2}_{X}\underbrace{BS^2}_{Y^T} + \underbrace{S^2BS^2B}_{YX^T} \preceq \underbrace{BS^4B}_{XX^T} + \underbrace{S^2B^2S^2}_{YY^T},$$

whence $T_b \le T_a$, and similarly

$$\underbrace{S^2B}_{X}\underbrace{SBS}_{Y^T} + \underbrace{SBSBS^2}_{YX^T} \preceq \underbrace{S^2B^2S^2}_{XX^T} + \underbrace{SBS^2BS}_{YY^T},$$

whence $2T_c \le T_a + T_b \le 2T_a$. The conclusion is that the sum $\sum_{\iota\in I}$ in (15) does not exceed the quantity

$$\mathcal{I}_i = 15\mathbf{E}\left\{\mathrm{Tr}(B_i S_{i-1}^4 B_i)\right\} = 15\mathrm{Tr}(B_i \mathbf{E}\{S_{i-1}^4\} B_i) = 15\mathrm{Tr}(B_i E_{i-1} B_i).$$

Invoking Lemma 1, we get

$$\mathcal{I}_i \le 45\sigma^4\Theta^4\mathrm{Tr}(B_i^2).$$

Completely similar reasoning as applied to the sum $\sum_{\iota\in J}$ in (15) implies that this sum does not exceed the quantity

$$\begin{aligned}&\mathcal{J}_i = 15\mathbf{E}\left\{\mathrm{Tr}(S_{i-1}B_i^4 S_{i-1})\right\} = 15\mathbf{E}\left\{\mathrm{Tr}(B_i^2 S_{i-1}^2 B_i^2)\right\} = 15\mathrm{Tr}\left(B_i^2\left[\sum_{j=1}^{i-1}\sigma_j^2 B_j^2\right]B_i^2\right)\\ &\le 15\Theta^2\sigma^2\mathrm{Tr}(B_i^4)\end{aligned}$$

Thus, (15) implies that

$$\phi_i \le \phi_{i-1} + 45\sigma^6\Theta^4\mathrm{Tr}(B_i^2) + 15\omega_i^4\Theta^2\sigma^2\mathrm{Tr}(B_i^4) + \chi_i^6\mathrm{Tr}(B_i^6);$$

since $\sum_i B_i^2 \preceq \Theta^2 I$, we have $\mathrm{Tr}(B_i^4) \le \Theta^2\mathrm{Tr}(B_i^2)$ and $\mathrm{Tr}(B_i^6) \le \Theta^4\mathrm{Tr}(B_i^2)$. We arrive at the relation

$$\phi_i \le \phi_{i-1} + \Theta^4\left[45\sigma^6 + 15\omega_i^4\sigma^2 + \chi_i^6\right]\mathrm{Tr}(B_i^2).$$

Taking into account that $\sum_i B_i^2 \preceq \Theta^2 I$, whence of course $\sum_i \mathrm{Tr}(B_i^2) \le \Theta^2 m$, we conclude that

$$\mathbf{E}\left\{\mathrm{Tr}(S_N^6)\right\} \le [45\sigma^6 + 15\omega^4\sigma^2 + \chi^6]\Theta^6 m,$$

as claimed. □

4^0. Now we can derive (10.*b*) in the same fashion as (10.*a*). Let $S(\cdot)$ and Q be defined as in 2^0. Applying Lemma 2 to $\zeta_i = \xi_i$ (which allows to take $\sigma = \omega = \chi = 1$), we get

$$\mathbf{E}\left\{\|S(\xi)\|^6\right\} \le \mathbf{E}\left\{\mathrm{Tr}(S^6(\xi))\right\} \le 61m\Theta^6,$$

whence by Tschebyshev inequality

$$\gamma > 0 \Rightarrow \mathrm{Prob}\left\{\|S(\xi)\| > 2\gamma\Theta\right\} < \frac{61m}{64\gamma^6}. \tag{16}$$

Now let $\gamma = m^{1/6}$ and $A = \gamma Q$. The set A is closed and convex and contains the centered at the origin ball of the radius $\gamma = m^{1/6}$. It follows that if $s > 1$ and $x \notin sA = s\gamma Q$ (or, which is the same, $\|S(x)\| > \gamma s\Theta$), then $\mathrm{dist}_{\|\cdot\|}(x, A) \ge (s-1)\gamma = (s-1)m^{1/6}$. Applying Talagrand Inequality to the distribution of the random vector $\zeta = \xi/2$ (Theorem 2.1), we get

$$s > 1 \Rightarrow \mathrm{Prob}\left\{\|S(\xi)\| > 2sm^{1/6}\Theta\right\} = \mathrm{Prob}\left\{\xi/2 \notin sA\right\} \le \frac{1}{\mathrm{Prob}\{\xi \in 2A\}}\exp\{-\tfrac{(s-1)^2m^{1/3}}{4}\}$$
$$\le 22\exp\{-\tfrac{(s-1)^2m^{1/3}}{4}\},$$

where the concluding inequality is given by (16). The resulting inequality is valid for all $s > 1$, and (10.*b*) follows. ∎

Corollary 1 *Let Ξ_1,...,Ξ_n be independent Gaussian symmetric $m \times m$ random matrices with zero means and $\Theta > 0$ be such that*

$$\sum_{i=t}^{n} \mathbf{E}\left\{\Xi_i^2\right\} \preceq \Theta^2 I. \tag{17}$$

Then relations (10) *hold true.*

Proof. By evident reasons every Gaussian symmetric random matrix Ξ_i can be represented as $\sum_{i=1}^{M} \eta_{it}B^{it}$ with independent $\eta_{it} \sim \mathcal{N}(0,1)$ and deterministic symmetric matrices B^{it}; observe that $\mathbf{E}\{\Xi_i^2\} = \sum_i (B^{it})^2$. Representing in this way every one of the matrices $\Xi_1, ..., \Xi_n$ and taking into account that the resulting Gaussian random variables $\{\eta_{it}\}$ are mutually independent, we conclude that

$$S_n \equiv \sum_{i=1}^{n} \Xi_i = \sum_{i=1}^{N} \xi_i B_i$$

with independent $\xi_i \sim \mathcal{N}(0,1)$ and deterministic symmetric matrices B_i satisfying the relation $\sum_i B_i^2 \preceq \Theta^2 I$. Now let $\{\zeta_{ij}\}_{\substack{1\le i\le N\\ j=1,2,...}}$ be a collection of independent random variables taking values ± 1 with probabilities $1/2$, and let

$$S_{n,\nu} = \sum_{i=1}^{N}\sum_{j=1}^{\nu} \zeta_{ij}\frac{1}{\sqrt{\nu}} B_i.$$

By Theorem 2.2, we have

$$\begin{array}{ll} \Omega \ge 7m^{1/4} \Rightarrow \mathrm{Prob}\{\|S_{n,\nu}\| > \Omega\Theta\} \le \frac{5}{4}\exp\{-\frac{\Omega^2}{64}\}, & (a)\\ \Omega \ge 7m^{1/6} \Rightarrow \mathrm{Prob}\{\|S_{n,\nu}\| > \Omega\Theta\} \le 22\exp\{-\frac{\Omega^2}{64}\}. & (b)\end{array}$$

As $\nu \to \infty$, the distribution of $S_{n,\nu}$, by Central Limit Theorem, converges weakly to the distribution of S_n, and (10) follows. ∎

2.3 Non-symmetric case

Question (Q) makes sense for non-symmetric (and even non-square) random matrices. In this case validity of Conjecture 1.1 would imply the following statement:

(!) *Let C_i be deterministic $m \times n$ matrices such that*

$$\sum_{i=1}^{N} C_i C_i^T \preceq \Theta^2 I_m,\ \sum_{i=1}^{N} C_i^T C_i \preceq \Theta^2 I_n \tag{18}$$

and ξ_i be independent random scalars with zero mean and of order of 1. Then

$$\Omega \ge O(1)\sqrt{\ln(m+n)} \Rightarrow \mathrm{Prob}\left\{\xi = (\xi_1,...,\xi_N) : \|\sum_{i=1}^{N}\xi_i C_i\| \ge \Omega\Theta\right\} \le O(1)\exp\{-O(1)\Omega^2\}. \tag{19}$$

Indeed, to in order to extract (!) from the assertion proposed by Conjecture 1.1, it suffices to apply the assertion to our ξ_i's and the deterministic symmetric $(m+n)\times(m+n)$ matrices

$$B_i = \left[\begin{array}{c|c} & C_i^T \\ \hline C_i & \end{array}\right]. \tag{20}$$

Utilizing in exactly the same fashion Theorem 2.2 and Corollary 1, we arrive at the following

Proposition 2 *Let deterministic $m \times n$ matrices C_i satisfy* (18)*, and let ξ_i be independent random scalars with zero first and third moment and such that either $|\xi_i| \le 1$ for all $i \le N$, or $\xi_i \sim \mathcal{N}(0,1)$ for all $i \le N$.*

$$\begin{array}{ll} \Omega \ge 7(m+n)^{1/4} \Rightarrow \mathrm{Prob}\{\|\sum\limits_{i=1}^{N}\xi_i C_i\| \ge \Omega\Theta\} \le \frac{5}{4}\exp\{-\frac{\Omega^2}{32}\}, & (a)\\ \Omega \ge 7(m+n)^{1/6} \Rightarrow \mathrm{Prob}\{\|\sum\limits_{i=1}^{N}\xi_i C_i\| \ge \Omega\Theta\} \le 22\exp\{-\frac{\Omega^2}{32}\}. & (b)\end{array} \tag{21}$$

We are about to add to Proposition 2 a simple additional statement, which allows to strengthen the result in the case when one of the sizes m, n is much smaller than another:

Proposition 3 *Let C_i, ξ_i be as in Proposition 2. Then*

$$\Omega \geq 4\sqrt{\min[m,n]} \Rightarrow \mathrm{Prob}\{\|\sum_{i=1}^{N} \xi_i C_i\| \geq \Omega\Theta\} \leq \frac{4}{3}\exp\{-\frac{\Omega^2}{16}\} \tag{22}$$

Proof. It suffices to consider the case when $|\xi_i| \leq 1$; the Gaussian version of the statement can be derived from the one with $|\xi_i| \leq 1$ in exactly the same fashion as in the proof of Corollary 1.

Let B_i be given by (20). Same as in item 2^0 of the proof of Theorem 2.2, setting $Q = \{x \in \mathbf{R}^N : \|\sum_{i=1}^{N} x_i B_i\| \leq \Theta\}$, we conclude from (18) that the closed convex set Q contains the unit Euclidean ball centered at the origin, and that for every $\gamma > 0$ one has

$$s > 1 \Rightarrow \mathrm{Prob}\left\{\|\sum_{i=1}^{N} \xi_i B_i\| > 2s\gamma\Theta\right\} \leq \frac{1}{\mathrm{Prob}\{\xi \in 2\gamma Q\}}\exp\{-\frac{(s-1)^2\gamma^2}{4}\}, \tag{23}$$

Assume w.l.o.g. that $\min[m,n] = n$. We have $\sum_{i=1}^{N} C_i^T C_i \preceq \Theta I_n$, whence, taking traces, $\sum_{i=1}^{N} \|C_i\|_2^2 \leq n\Theta^2$. It follows that

$$\mathbf{E}\left\{\|\sum_i \xi_i C_i\|_2^2\right\} = \sum_i \mathbf{E}\left\{\xi_i^2\right\}\|C_i\|_2^2 \leq \sum_i \|C_i\|_2^2 \leq n\Theta^2,$$

whence by Tschebyshev inequality and due to $\|C\| \leq \|C\|_2$

$$\forall t > 0 : \mathrm{Prob}\left\{\|\sum_i \xi_i C_i\| \geq tn^{1/2}\Theta\right\} \leq \mathrm{Prob}\left\{\|\sum_i \xi_i C_i\|_2 \geq tn^{1/2}\Theta\right\} \leq t^{-2}.$$

Setting $\gamma = n^{1/2}$, we conclude from the latter inequality that $\mathrm{Prob}\left\{\|\sum_i \xi_i C_i\| \geq 2\gamma\Theta\right\} \leq 1/4$, whence, in view of $\|\sum_i \xi_i B_i\| = \|\sum_i \xi_i C_i\|$,

$$\mathrm{Prob}\left\{\|\sum_i \xi_i B_i\| > 2\gamma\Theta\right\} = \mathrm{Prob}\{\xi \notin 2\gamma Q\} \leq 1/4.$$

Thus, (23) with $\gamma = n^{1/2}$ implies that

$$s > 1 \Rightarrow \mathrm{Prob}\left\{\|\sum_{i=1}^{N} \xi_i B_i\| > 2sn^{1/2}\Theta\right\} \leq \frac{4}{3}\exp\{-\frac{(s-1)^2 n}{4}\},$$

and (22) follows (recall that $\|\sum_i \xi_i C_i\| \equiv \|\sum_i \xi_i B_i\|$). ∎

3 Application: Randomly perturbed Linear Matrix Inequality

Consider a randomly perturbed Linear Matrix Inequality (LMI)

$$A_0[x] - \rho\sum_{i=1}^{N} \xi_i A_i[x] \succeq 0, \tag{24}$$

where $A_0[x], ..., A_N[x]$ are symmetric matrices affinely depending on decision vector x, ξ_i, $i = 1, ..., N$, are random real perturbations which we assume to be independent with zero means "of order of 1" and with "light tails" (precise formulations of these two assumptions will be given later). We are interested to describe those x for which the randomly perturbed LMI (24) holds true with probability $\geq 1 - \epsilon$, where $\epsilon << 1$. Clearly, for such an x one should have $A_0[x] \succeq 0$. We will simplify a little bit our task and focus on points x with $A_0[x] \succ 0$. For such an x, setting $B_i[x] = A_0^{-1/2}[x]A_i[x]A_0^{-1/2}[x]$, the question becomes to describe those x for which

$$\text{Prob}\left\{\sum_{i=1}^{N} \xi_i B_i[x] \preceq I\right\} \geq 1 - \epsilon. \tag{25}$$

Precise description seems to be completely intractable; what we are about to present are verifiable *sufficient conditions* for (25) to hold true.

3.1 Condition based on Proposition 1

Proposition 4 *Let $m \geq 2$, let perturbations ξ_i be independent with zero means and such that $\mathbf{E}\{\exp\{\xi_i^2\}\} \leq \exp\{1\}$. Then the condition*

$$A_0[x] \succ 0 \ \& \ \rho^2 \sum_{i=1}^{N} \|A_0^{-1/2}[x]A_i[x]A_0^{-1/2}[x]\|^2 \leq \frac{1}{450 \exp\{1\}(\ln \frac{3}{\epsilon})(\ln m)} \tag{26}$$

is sufficient for (24) to be valid with probability $\geq 1 - \epsilon$.

This is a straightforward corollary of Proposition 1 (we use the actual values of absolute constants in (7) presented in [6]).

A severe shortcoming of (26) is that this condition, although verifiable, in general defines a nonconvex set in the space of decision variables x, which makes it problematic to optimize in x under the conditions. There are, however, two simple cases when the conditions are free of this shortcoming. The first is when $A_i[x]$ are independent of x ("perturbations in the constant term of LMI"); here the "problematic" part of the conditions – the inequality

$$\sum_{i=1}^{N} \|A_0^{-1/2}[x]A_i[x]A_0^{-1/2}[x]\|^2 \leq \tau \tag{$*$}$$

on x, τ – can be represented by the system of convex inequalities

$$-A_0[x] \preceq \mu_i A_i \preceq A_0[x],\ \mu_i > 0,\ i = 1, ..., N, \sum_{i=1}^{N} \mu_i^{-2} \leq \tau.$$

in variables x, μ_i, τ. The second "good" case is the one when $A_0[x] \equiv A$ is constant. Here $(*)$ can be represented by system of convex constraints

$$-\lambda_i A \preceq A_i[x] \preceq \lambda_i A,\ i = 1, ..., N, \sum_i \lambda_i^2 \leq \tau$$

in variables x, λ_i, τ.

3.2 Conditions based on Theorem 2.2 and Corollary 1

With these statements in the role of Proposition 1, we arrive at the following statement:

Proposition 5 *Let perturbations ξ_i be independent with zero means and zero third moments and either such that $|\xi_i| \le 1$, $i = 1, ..., N$, or such that $\xi_i \sim \mathcal{N}(0,1)$, $i = 1, ..., N$. Let, further, $\epsilon \in (0,1)$ be such that one of the following two conditions is satisfied:*

$$\begin{array}{ll} (a) & \ln\left(\frac{5}{4\epsilon}\right) \ge \frac{49m^{1/2}}{32} \\ (b) & \ln\left(\frac{22}{\epsilon}\right) \ge \frac{49m^{1/3}}{32} \end{array} \tag{27}$$

Then the condition

$$A_0[x] \succ 0 \ \& \ \rho^2 \Big\| \sum_{i=1}^N (A_0^{-1/2}[x] A_i[x] A_0^{-1/2}[x])^2 \Big\| \le \begin{cases} \frac{1}{32\ln\left(\frac{5}{4\epsilon}\right)}, & \text{case of (27.}a\text{)} \\ \frac{1}{32\ln\left(\frac{22}{\epsilon}\right)}, & \text{case of (27.}b\text{)} \end{cases} \tag{28}$$

is sufficient for (24) *to be valid with probability $\ge 1 - \epsilon$.*

Note that condition (28), in contrast to (26), defines a convex domain in the space of design variables. Indeed, this condition is of the form

$$A_0[x] \succ 0 \ \& \ \rho^2 \sum_{i=1}^N A_i[x] A_0^{-1}[x] A_i[x] \preceq c(\epsilon) A_0[x],$$

which can be represented by system of LMI's

$$A_0[x] \succ 0 \ \& \ \left[\begin{array}{cc} Y_i & A_i[x] \\ A_i[x] & A_0[x] \end{array} \right] \succeq 0, \ i = 1, ..., N \ \& \ \rho^2 \sum_{i=1}^N Y_i \preceq c(\epsilon) A_0[x] \tag{29}$$

in variables x, Y_i. Note also that in Control applications (which are of primary importance for randomly perturbed LMI) m does not exceed few tens, and in this range of values of m the only advantage of (26) as compared with (28), that is, $\ln(m)$ in the right hand side of (26) vs. $m^{1/2}$ and $m^{1/3}$ in the right hand side of (27), becomes unimportant (in fact, (26), because of large constant factors, in a reasonable range of values of m leads to much more conservative conclusions than (27)).

4 Application: semidefinite relaxation of quadratic minimization under orthogonality constraints

4.1 Problem of interest

Consider the following optimization problem:

$$\max_{x \in \mathbf{M}^{m,n}} \left\{ \langle x, \mathcal{A}x \rangle : \begin{array}{ll} \langle x, \mathcal{B}x \rangle \le 1 & (a) \\ \langle x, \mathcal{B}_\ell x \rangle \le 1, \ \ell = 1, ..., L & (b) \\ \mathcal{C}x = 0 & (c) \\ \|x\| \le 1 & (d) \end{array} \right\} \tag{P}$$

where

- $\mathbf{M}^{m,n}$ is the space of $m \times n$ matrices equipped with the Frobenius inner product $\langle x, y\rangle = \mathrm{Tr}(xy^T)$,
- the mappings $\mathcal{A}$, $\mathcal{B}$, $\mathcal{B}_\ell$ are symmetric linear mappings from $\mathbf{M}^{m,n}$ into itself,
- $\mathcal{B}$ is positive semidefinite of rank 1,
- $\mathcal{B}_\ell$, $\ell = 1, ..., L$, are positive semidefinite,
- $\mathcal{C}$ is a linear mapping from $\mathbf{M}^{m,n}$ into $\mathbf{R}^M$.

Note that (P) covers a number of problems of quadratic optimization under orthogonality constraints, e.g.

1. Inhomogeneous modification

$$\max_{x\in\mathbf{M}^{m,n}} \left\{ \langle x, \mathcal{A}x\rangle + 2\langle b, x\rangle : \begin{array}{ll} \langle x, \mathcal{B}x\rangle \le 1 & (a) \\ \langle x, \mathcal{B}_\ell x\rangle \le 1,\ \ell = 1, ..., L & (b) \\ \mathcal{C}x = 0 & (c) \\ \|x\| \le 1 & (d) \end{array} \right\} \qquad (P_+)$$

   of (P). Indeed, partitioning a matrix $y \in \mathbf{M}^{n+1,m+1}$ as $\left[\begin{array}{c|c} y_{00} & y_{01} \\ \hline y_{10} & y_{11} \end{array}\right]$ with scalar y_{00}, (P_+) is equivalent to the problem

$$\max_y \left\{ \langle y, \mathcal{A}^+ y\rangle \equiv \langle y_{11}, \mathcal{A}y_{11}\rangle + 2y_{00}\langle b, x_{11}\rangle : \begin{array}{ll} \langle y_{11}, \mathcal{B}y_{11}\rangle \le 1 & (a) \\ \langle y_{11}, \mathcal{B}_\ell y_{11}\rangle \le 1,\ \ell = 1, ..., L & (b) \\ \mathcal{C}y_{11} = 0, y_{01} = 0, y_{10} = 0 & (c) \\ \|y\| \le 1 & (d) \end{array} \right\}$$

   of the form of (P);

2. Orthogonal relaxation of the quadratic assignment problem (see [10, 11, 12] and references therein)

$$\max_X \left\{ \mathrm{Tr}(BXAX^T) - 2\mathrm{Tr}(CX) : X \in \mathbf{M}^{m,m}, XX^T = I_m \right\} \qquad (\mathrm{QA})$$

   with symmetric $m \times m$ matrices A, B. Indeed, the transformation $B \leftarrow B + bI_m$ converts (QA) into an equivalent problem, thus we can assume that $B \succ 0$. Similarly, the transformation $A \leftarrow A + aI_m$ converts (QA) into equivalent problem, thus we can assume that $A \succ 0$. In the case when $B \succ 0$, $A \succ 0$, representing $B = D^2$ and $A = E^2$ with symmetric D, E, we see that the objective in (QA) is $f(X) = \mathrm{Tr}([DXE][DXE]^T) + 2\mathrm{Tr}(CX)$, which is a convex quadratic form of X. Consequently, the maximum of f over the set $\{X \in \mathbf{M}^{m,m} : XX^T = I_m\}$ is exactly the same as the maximum of f over the set $\{X \in \mathbf{M}^{m,m} : \|X\| \le 1\}$, since the former set is exactly the set of extreme points of the latter one. Thus, (QA) is equivalent to the problem of the form

$$\max_X \left\{ \mathrm{Tr}(\bar{B}X\bar{A}X^T) - 2\mathrm{Tr}(CX) : \|X\| \le 1 \right\},$$

   which is of the form of (P_+);

3. Procrustes problem which can be posed as (see Introduction)

$$\max_{X[1],...,X[K]}\left\{\sum_{1\le\ell<\ell'\le K}\mathrm{Tr}(A[\ell]X[\ell]X^T[\ell']A^T[\ell']) : X[\ell]X^T[\ell]=I_n,\ \ell=1,...,K\right\} \qquad (\mathrm{Pr})$$

Indeed, the objective in (Pr) is linear in every one of $X[\ell]$; thus, we do not affect the problem by relaxing the orthogonality constraints $X[\ell]X^T[\ell]=I_n$ to $\|X[\ell]\|\le 1$. Indeed, such a relaxation could only increase the optimal value. This, however, does not happen, since given a feasible solution to the problem

$$\max_{X[1],...,X[K]}\left\{\sum_{1\le\ell<\ell'\le K}\mathrm{Tr}(A[\ell]X[\ell]X^T[\ell']A^T[\ell']) : \|X[\ell]\|\le 1,\ \ell=1,...,K\right\} \qquad (\mathrm{Pr}_+)$$

we can easily convert it into a feasible solution to (Pr) with the same or larger value of the objective.

> Indeed, keeping $X[2],...,X[K]$ intact, we can straightforwardly replace $X[1]$ by an orthogonal matrix without spoiling the objective value[1)]. After $X[1]$ is made orthogonal, we can repeat the outlined procedure with $X[2]$ in the role of $X[1]$, and so on. After K steps we end up with a feasible solution to both (Pr_+) and (Pr) which is at least as good as the solution we have started with.

It remains to note that problem (Pr_+) is of the form of (P) – we can arrange all matrices $X[\ell]$, $\ell=1,...,K$, in a large block-diagonal matrix $x=\begin{bmatrix} X[1] & & \\ & \ddots & \\ & & X[K]\end{bmatrix}$, thus converting ($\mathrm{Pr}_+$) into the equivalent problem

$$\max_x\left\{F(x)\equiv\sum_{\ell<\ell'}\mathrm{Tr}(A[\ell]X[\ell]X^T[\ell']A^T[\ell']) : \mathcal{C}x=0, \|x\|\le 1\right\}$$

where the homogeneous equations $\mathcal{C}x=0$ express the fact that x is of the outlined block-diagonal form; the resulting problem is in the form of (P);

4. The problem

$$\max_{X[1],...,X[K]}\left\{\sum_{\ell<\ell'}\|A[\ell]X[\ell]-A[\ell']X[\ell']\|_2^2 : X[\ell]X^T[\ell]=I_n,\ \ell=1,...,k\right\}$$

"opposite" to the Procrustes problem. Indeed, since the objective is convex in every one of $X[\ell]$, we, same as above, lose nothing when relaxing the constraints $X[\ell]X^T[\ell]=I_n$ to $\|X[\ell]\|\le 1$. The resulting problem can be converted to the form of (P) in exactly the same manner as in the previous example. The same argument applies to a general-type problem of quadratic maximization under orthogonality constraints, provided that the objective is convex in every one of the corresponding variable matrices.

[1)] since the objective is linear in $X[1]$, the remaining variable matrices being fixed, and thus attains its maximum in $X[1]$ varying in the set $\{X : \|X\|\le 1\}$ at an extreme point of the set, which is an orthogonal matrix; this matrix is easily computable, given $X[2],...,X[K]$.

Note that in some of the outlined examples we end up with a particular case of problem (P) where the homogeneous linear constraints (c) in (P) imply that x is a block-diagonal matrix $\left[\begin{array}{c|c|c} x_1 & & \\ \hline & \ddots & \\ \hline & & x_K \end{array}\right]$ with $m_k \times n_k$ diagonal blocks x_k, $k = 1, .., K$. We shall refer to $\mathbf{\Delta} = \{(m_k, n_k)\}_{k=1}^K$ as to the *structure* of (P), with the case of no nontrivial block-diagonal structure in x corresponding to the trivial structure $\mathbf{\Delta} = (m, n)$ with $K = 1$.

Semidefinite relaxation of (P). Problem (P), in general, is NP-hard (this is the case already for the generic inhomogeneous ($C \neq 0$) orthogonal relaxation of the quadratic assignment problem, see [10]). At the same time, (P) admits a straightforward semidefinite relaxation as follows. We can identify $\mathcal{A}$ in (P) with a symmetric $mn \times mn$ matrix $A = [A_{ij,k\ell}]$ with rows and columns indexed by pairs (i, j), $1 \leq i \leq m$, $1 \leq j \leq n$ satisfying the relation

$$[\mathcal{A}x]_{ij} = \sum_{k,\ell} A_{ij,k\ell} x_{k\ell}$$

(from now on, if opposite is not stated, in a sum $\sum_{p,q}$, p runs from 1 to m, and q runs from 1 to n). Similarly, $\mathcal{B}$, $\mathcal{B}_\ell$ can be identified with symmetric positive semidefinite $mn \times mn$ matrices B, B_ℓ, with B of rank 1. Finally, $\mathcal{C}$ can be identified with a $M \times mn$ matrix $C = [C_{\mu,ij}]$:

$$(\mathcal{C}x)_\mu = \sum_{i,j} C_{\mu,ij} x_{ij}.$$

For $x \in \mathbf{M}^{m,n}$, let $\mathrm{Vec}(x)$ be the mn-dimensional vector obtained from the matrix x by arranging its columns into a single column, and let $X(x) \in \mathbf{S}_+^{mn}$ be the matrix $\mathrm{Vec}(x)\mathrm{Vec}^T(x)$, that is, the $mn \times mn$ matrix $[x_{ij}x_{k\ell}]$. Observe that $X(x) \succeq 0$, and that $\sum_{i,j} c_{ij}x_{ij} = 0$ if and only if $0 = \left(\sum_{i,j} c_{ij}x_{ij}\right)^2 \equiv \sum_{ij,k\ell} c_{ij}x_{ij}c_{k\ell}x_{k\ell} = \mathrm{Tr}(X(c)X(x))$. Further,

$$\langle x, \mathcal{A}x\rangle = \sum_{ij,k\ell} A_{ij,k\ell} x_{ij} x_{k\ell} = \mathrm{Tr}(AX(x)),$$

and similarly

$$\langle x, \mathcal{B}\rangle = \mathrm{Tr}(BX(x)), \quad \langle x, \mathcal{B}_\ell\rangle = \mathrm{Tr}(B_\ell X(x)).$$

Finally, $\|x\| \leq 1$ if and only if $xx^T \preceq I_m$. The entries in the matrix xx^T are linear combinations of the entries in $X(x)$, so that

$$xx^T \preceq I_m \Leftrightarrow \mathcal{S}(X(x)) \preceq I_n,$$

where $\mathcal{S}$ is an appropriate linear mapping from $\mathbf{S}^{mn}$ to $\mathbf{S}^m$. Similarly, $\|x\| \leq 1$ if and only if $x^Tx \preceq I_n$, which again is a linear restriction on $X(x)$:

$$x^Tx \preceq I_n \Leftrightarrow \mathcal{T}(X(x)) \preceq I_n,$$

where $\mathcal{T}$ is an appropriate linear mapping from $\mathbf{S}^{mn}$ to $\mathbf{S}^n$. With the above observations, (P) can be rewritten as the problem

$$\min_{x \in \mathbf{M}^{m,n}} \left\{ \mathrm{Tr}(AX(x)) : \begin{array}{ll} \mathrm{Tr}(BX(x)) \leq 1 & (a) \\ \mathrm{Tr}(B_\ell X(x)) \leq 1, \ell = 1, ..., L & (b) \\ \mathrm{Tr}(C^\mu X(x)) = 0,\ \mu = 1, ..., M & (c) \\ \mathcal{S}(X(x)) \preceq I_m, \mathcal{T}(X(x)) \preceq I_n & (d) \end{array} \right\},$$

where $C^\mu \in \mathbf{S}^{mn}_+$ is given by $C^\mu_{ij,k\ell} = C_{\mu,ij}C_{\mu,k\ell}$. Since $X(x) \succeq 0$ for all x, the problem

$$\min_{X\in\mathbf{S}^{mn}} \left\{ \mathrm{Tr}(AX) : \begin{array}{ll} \mathrm{Tr}(BX) \le 1 & (a) \\ \mathrm{Tr}(B_\ell X) \le 1, \ell = 1,...,L & (b) \\ \mathrm{Tr}(C^\mu X) = 0,\ \mu = 1,...,M & (c) \\ \mathcal{S}(X) \preceq I_m, \mathcal{T}(X) \preceq I_n & (d) \\ X \succeq 0 & (e) \end{array} \right\} \qquad \text{(SDP)}$$

is a relaxation of (P), so that $\mathrm{Opt}(P) \le \mathrm{Opt}(\mathrm{SDP})$. Observe that problem (SDP) is a semidefinite program and as such is computationally tractable.

Remark 4.1 When (P) possesses a nontrivial structure, the design dimension of relaxation (SDP) can be reduced. Indeed, in this case, as it is immediately seen, (SDP.c) imply that $X_{ij,k\ell}$ should be zero unless both the cells (i,j), (k,ℓ) belong to diagonal blocks in x. Consequently, in fact the decision matrix X in (SDP) can be thought of as a symmetric matrix of the row size $\sum_{k=1}^{K} m_k n_k$ rather than of the size mn.

4.2 Quality of the relaxation

Our goal is to prove the following

Proposition 6 (i) *There exists $\bar{x} \in \mathbf{M}^{m,n}$ such that*

$$\begin{array}{llll} (*) & \langle \bar{x}, \mathcal{B}\bar{x}\rangle = \mathrm{Opt}(\mathrm{SDP}) & (a) & \langle \bar{x}, \mathcal{B}\bar{x}\rangle \le 1 \\ (b) & \langle \bar{x}, \mathcal{B}_\ell\bar{x}\rangle \le \Omega^2,\ \ell = 1,...,L & (c) & \mathcal{C}\bar{x} = 0 \\ (d) & \|\bar{x}\| \le \Omega & & \end{array} \qquad (30)$$

where

$$\begin{array}{rcl} \Omega & = & \max\left[\max\limits_{1\le k\le K} \mu_k + \sqrt{32\ln(132K)}, \sqrt{32\ln(12(L+1))}\right], \\ \mu_k & = & \min\left[7(m_k+n_k)^{\frac{1}{6}}, 4\sqrt{\min[m_k,n_k]}\right] \end{array} \qquad (31)$$

(ii) *In particular, one has*

$$\mathrm{Opt}(P) \le \mathrm{Opt}(\mathrm{SDP}) \le \Omega^2 \mathrm{Opt}(P). \qquad (32)$$

Proof. 0^0. (ii) is an immediate consequence of (i). Indeed, with $\bar{x}$ satisfying (30), the matrix $\tilde{x} = \Omega^{-1}x$ clearly is a feasible solution to (P), and the value of the objective at this solution is $\Omega^{-2}\mathrm{Opt}(\mathrm{SDP})$ by (30.$*$), which gives the right inequality in (32); the left inequality is readily given by the origin of (SDP).

It remains to prove (i). Let Y be an optimal solution to (SDP); then $Y \succeq 0$, so that the matrix $S = Y^{1/2}$ is well defined. Let us set

$$SAS = U^T \Lambda U,$$

where Λ is a diagonal $mn \times mn$ matrix, and U is an orthogonal $mn \times mn$ matrix. Let ξ be a random mn-dimensional vector with independent entries ξ_{ij}, $1 \le i \le m$, $1 \le j \le n$, taking values ± 1 with probabilities $1/2$, and let random $m \times n$ matrix ζ be given by

$$\mathrm{Vec}(\zeta) = SU^T\xi, \qquad (33)$$

so that $\zeta = \zeta(\xi)$ is a deterministic function of ξ.

1^0. Observe that

$$\mathbf{E}\{X(\zeta)\} = Y. \tag{34}$$

Indeed,

$$\mathbf{E}\{X(\zeta)\} = \mathbf{E}\left\{\mathrm{Vec}(\zeta)\mathrm{Vec}^T(\zeta)\right\} = \mathbf{E}\left\{SU^T\xi\xi^TUS\right\} = SU^T\mathbf{E}\left\{\mathrm{Vec}(\xi)\mathrm{Vec}^T(\xi)\right\}US$$
$$= SUU^TS = Y.$$

2^0. We have

$$\mathcal{C}\zeta \equiv 0 \tag{35}$$

Indeed,

$$\mathbf{E}\left\{\left(\sum C_{\mu,ij}\zeta_{ij}\right)^2\right\} = \mathbf{E}\left\{\mathrm{Tr}(C^\mu\mathrm{Vec}(\zeta)\mathrm{Vec}^T(\zeta))\right\} = \mathrm{Tr}(C^\mu Y) = 0$$

(we have used (34) and the fact that Y is feasible for (SDP)).

Since the relations $\mathcal{C}x = 0$ imply that x is block-diagonal with $m_k \times n_k$ diagonal blocks, $k = 1, ..., K$, we conclude that all realizations of ζ are block-diagonal with $m_k \times n_k$ diagonal blocks ζ_k, $k = 1, ..., K$. Recalling (33) and the nature of ξ, we see that all combinations of the columns of the matrix SU^T with coefficients ± 1 are of the form $\mathrm{Vec}(z)$ with block-diagonal, of the block-diagonal structure $\mathbf{\Delta}$, $m \times n$ matrices z; this is possible if and only if every one of the columns in SU^T is of the form $\mathrm{Vec}(z)$ with block-diagonal, of the block-diagonal structure $\mathbf{\Delta}$, matrices z. Recalling (33), we arrive at

$$\zeta_k = \zeta_k(\xi) = \sum_{i,j} z_{k,ij}\xi_{ij},\ k = 1, ..., K, \tag{36}$$

with deterministic $m_k \times n_k$ matrices $z_{k,ij}$.

3^0. We have also

$$\begin{array}{ll} (a) & \langle\zeta, \mathcal{A}\zeta\rangle \equiv \mathrm{Opt(SDP)} \\ (b) & \mathbf{E}\{\langle\zeta, \mathcal{B}\zeta\rangle\} \le 1 \\ (b') & \mathbf{E}\{\langle\zeta, \mathcal{B}_\ell\zeta\rangle\} \le 1,\ \ell = 1, ..., L \end{array} \tag{37}$$

Indeed,

$$\begin{array}{rcl} \langle\zeta, \mathcal{A}\zeta\rangle & = & \mathrm{Tr}(A\mathrm{Vec}(\zeta)\mathrm{Vec}^T(\zeta)) = \mathrm{Tr}(ASU^T\xi\xi^TUS) \\ & = & \mathrm{Tr}(U(SAS)U^T\xi\xi^T) = \mathrm{Tr}(U(U^T\Lambda U)U^T\xi\xi^T) \\ & = & \mathrm{Tr}(\Lambda\xi\xi^T) = \mathrm{Tr}(\Lambda) = \mathrm{Tr}(U^T\Lambda U) = \mathrm{Tr}(SAS) = \mathrm{Tr}(AY) = \mathrm{Opt(SDP)}, \end{array}$$

as required in (37.a). Further,

$$\begin{array}{rcl} \mathbf{E}\{\langle\zeta, \mathcal{B}\zeta\rangle\} & = & \mathbf{E}\left\{\mathrm{Tr}(B\mathrm{Vec}(\zeta)\mathrm{Vec}^T(\zeta))\right\} = \mathbf{E}\left\{\mathrm{Tr}(BSU^T\xi\xi^TUS)\right\} \\ & = & \mathrm{Tr}(BSU^T\underbrace{\mathbf{E}\{\xi\xi^T\}}_{=I}US) = \mathrm{Tr}(BS^2) = \mathrm{Tr}(BY) \le 1 \end{array}$$

where the concluding $\le$ comes from the fact that Y is feasible for (SDP). We have arrived at (37.b); verification of (37.b') is completely similar.

4^0. Finally, we have

$$\mathbf{E}\left\{\zeta^T\zeta\right\} \preceq I_n,\ \mathbf{E}\left\{\zeta\zeta^T\right\} \preceq I_m. \tag{38}$$

Indeed, by the origin of $\mathcal{S}$ and $\mathcal{T}$, we have $\zeta\zeta^T = \mathcal{S}(X(\zeta))$, $\zeta^T\zeta = \mathcal{T}(X(\zeta))$, and (38) follows from (34).
Recalling that $\zeta = \mathrm{Diag}\{\zeta_1, ..., \zeta_K\}$, we have

$$\zeta\zeta^T = \mathrm{Diag}\{\zeta_1\zeta_1^T, ..., \zeta_K\zeta_K^T\},\ \zeta^T\zeta = \mathrm{Diag}\{\zeta_1^T\zeta_1, ..., \zeta_K^T\zeta_K\},$$

and (38) implies that

$$\mathbf{E}\left\{\zeta_k\zeta_k^T\right\} \preceq I_{m_k},\ \mathbf{E}\left\{\zeta_k^T\zeta_k\right\} \preceq I_{n_k},\ k = 1, ..., K. \tag{39}$$

Invoking (36), we have $\zeta_k = \sum\limits_{i,j} z_{k,ij}\xi_{ij}$ with deterministic $m_k \times n_k$ matrices $z_{k,ij}$, so that (39) implies

$$\sum_{i,j} z_{k,ij}z_{k,ij}^T \preceq I_{m_k},\ \sum_{i,j} z_{k,ij}^T z_{k,ij} \preceq I_{n_k},\ k = 1, ..., K. \tag{40}$$

Applying Propositions 2, 3, we extract from (40) that

$$t > \mu_k \Rightarrow \mathrm{Prob}\left\{\xi : \|\zeta_k(\xi)\| \geq t\right\} \leq 22\exp\left\{-\frac{t^2}{32}\right\},\ k = 1, ..., K. \tag{41}$$

6^0. We are basically done; the only additional element we need to complete the proof of (i) is the following simple fact:

Lemma 3 *One has*

$$\begin{array}{ll}(a) & \mathrm{Prob}\underbrace{\left\{\xi : \langle\zeta(\xi), \mathcal{B}\zeta(\xi)\rangle \leq 1\right\}}_{E_a} \geq \frac{1}{3} \\ (b) & t \geq 8 \Rightarrow \mathrm{Prob}\left\{\xi : \langle\zeta(\xi), \mathcal{B}_\ell\zeta(\xi)\rangle > t^2\right\} \leq \frac{4}{3}\exp\{-\frac{t^2}{32}\},\ \ell = 1, ..., L.\end{array} \tag{42}$$

Proof. Recall that

$$\langle\zeta(\xi), \mathcal{B}\zeta(\xi)\rangle = \mathrm{Tr}(B\mathrm{Vec}(\zeta(\xi))\mathrm{Vec}^T(\zeta(\xi))) = \mathrm{Tr}(BSU^T\xi\xi^TUS) = \mathrm{Tr}((USBSU^T)\xi\xi^T). \tag{43}$$

Since B is positive semidefinite dyadic matrix, so is the matrix $USBSU^T$, that is, $USBSU^T = dd^T$ for a mn-dimensional deterministic vector d with entries d_{ij}, $1 \leq i \leq m$, $1 \leq j \leq n$. It follows that

$$\langle\zeta(\xi), \mathcal{B}\zeta(\xi)\rangle = (d^T\xi)^2 = \left(\sum_{i,j} d_{ij}\xi_{ij}\right)^2 \tag{44}$$

Applying (37.b), we derive from (44) that $\sum\limits_{i,j} d_{ij}^2 = \mathbf{E}\left\{\langle\zeta(\xi), \mathcal{B}\zeta(\xi)\rangle\right\} \leq 1$. Invoking Lemma A.1 in [1], we conclude that $\mathrm{Prob}\left\{|\sum\limits_{i,j} d_{ij}\xi_{ij}| \leq 1\right\} \geq \frac{1}{3}$, and (42.$a$) follows from (44). Similarly to (43), we have

$$\langle\zeta(\xi), \mathcal{B}_\ell\zeta(\xi)\rangle = \xi^T\underbrace{USB_\ell SU^T}_{D_\ell}\xi.$$

The matrix D_ℓ is symmetric positive semidefinite along with B_ℓ; setting $F_\ell = D_\ell^{1/2}$, we arrive at the identity

$$\langle\zeta(\xi), \mathcal{B}_\ell\zeta(\xi)\rangle = \|F_\ell\xi\|_2^2. \tag{45}$$

Invoking (37.b'), we derive from the latter inequality that

$$\|F_\ell\|_2^2 = \mathbf{E}\left\{\langle \zeta(\xi), \mathcal{B}_\ell \zeta(\xi)\rangle\right\} \le 1. \tag{46}$$

We are about to use the following simple fact:

Lemma 4 *Let $b_p \in \mathbf{R}^\pi$, $p = 1, ..., P$, be deterministic vectors such that $\sum_p \|b_p\|_2^2 \le 1$, and δ_p, $p = 1, ..., P$, be independent random scalars taking values ± 1 with probabilities $1/2$. Then*

$$t \ge 8 \Rightarrow \mathrm{Prob}\left\{\|\sum_{p=1}^P \delta_p b_p\|_2 > t\right\} \le \frac{4}{3}\exp\{-\frac{t^2}{32}\}. \tag{47}$$

Proof. Let $Q = \{\gamma \in \mathbf{R}^P : \|\sum_p \gamma_p b_p\|_2 \le 1\}$, and let μ be the distribution of the random vector $\gamma = (\delta_1/2, ..., \delta_P/2)$. Observe that $\mathbf{E}\left\{\|\sum_p \gamma_p b_p\|_2^2\right\} = \frac{1}{4}\sum_p \|b_p\|_2^2 \le \frac{1}{4}$, whence $\mu\{\gamma \notin Q\} = \mu\{\gamma : \|\sum_p \gamma_p b_p\|_2 > 1\} \le \frac{1}{4}$, so that $\mu(Q) \ge \frac{3}{4}$. Further, Q clearly is a closed convex set. We claim that this set contains the unit Euclidean ball in $\mathbf{R}^P$. Indeed, if $u \in \mathbf{R}^P$ and $\|u\|_2 \le 1$, then

$$\|\sum_p u_p b_p\|_2 \le \sum_p |u_p| \|b_p\|_2 \le \sqrt{\sum_p u_p^2}\sqrt{\sum_p \|b_p\|_2^2} \le 1,$$

so that $u \in Q$. Now, if $s > 1$ and $u \in \mathbf{R}^P$ is such that $\|\sum_p u_p b_p\|_2 > 2s$, then the vector $u/2$ does not belong to sQ, so that $\mathrm{dist}_{\|\cdot\|_2}(u, Q) > s - 1$ (since Q contains the unit Euclidean ball in $\mathbf{R}^P$). Applying the Talagrand Inequality to the distribution μ, we get

$$\begin{array}{rcl} s > 1 \Rightarrow \mathrm{Prob}\left\{\delta : \|\sum_p \delta_p b_p\|_2 > 2s\right\} & \le & \exp\{-\frac{(s-1)^2}{4}\}\int \exp\{\frac{\mathrm{dist}^2_{\|\cdot\|_2}(\gamma, Q)}{4}\}\mu(d\gamma) \le \frac{\exp\{-\frac{(s-1)^2}{4}\}}{\mu(Q)} \\ & \le & \frac{4\exp\{-\frac{(s-1)^2}{4}\}}{3}, \end{array}$$

and (47) follows. □

Specifying in Lemma 4 P as mn, b_p as the columns of F_ℓ and δ_p as the random scalars ξ_{ij}, we derive from (46) that

$$t \ge 8 \Rightarrow \mathrm{Prob}\left\{\xi : \|F_\ell \xi\|_2 > t\right\} \le \frac{4}{3}\exp\{-\frac{t^2}{32}\},$$

which combines with (45) to imply (42.b). Lemma 3 is proved. □

7^0. We are ready to complete the proof of (i). Let Ω be given by (31). By (42), with this Ω for every $\ell \le L$ we have $\mathrm{Prob}\left\{\xi : \langle \zeta(\xi), \mathcal{B}_\ell \zeta(\xi)\rangle > \Omega^2\right\} \le \frac{1}{8(L+1)}$, whence

$$\mathrm{Prob}\underbrace{\left\{\xi : \langle \zeta(\xi), \mathcal{B}_\ell \zeta(\xi)\rangle \le \Omega^2,\ 1 \le \ell \le L\right\}}_{E_b} \ge \frac{7}{8}. \tag{48}$$

By (41), with our Ω for every $k \le K$ we have also $\mathrm{Prob}\left\{\xi : \|\zeta_k(\xi)\| > \Omega\right\} \le \frac{1}{6K}$, whence

$$\mathrm{Prob}\underbrace{\left\{\xi : \|\zeta(\xi)\| \equiv \max_{k \le K} \|\zeta_k(\xi)\| \le \Omega\right\}}_{E_d} \ge \frac{5}{6}. \tag{49}$$

Combining (42.a), (48) and (49), we see that the events E_a, E_b and E_d have a point ξ_* in common. Setting $\bar{x} = \zeta(\xi_*)$, we see that $\bar{x}$ satisfies all the requirements in (30) ($(*)$ – by (37.a), $(a), (b), (d)$ – due to $\xi_* \in E_a \cap E_b \cap E_d$, and (c) by (35)). ■

4.2.1 Comments

A. With norm constraint (d) in (P) eliminated, (P) becomes the purely quadratic program

$$\max_{x\in\mathbf{M}^{m,n}}\left\{\langle x,\mathcal{A}x\rangle:\ \begin{array}{ll}\langle x,\mathcal{B}x\rangle\le 1 & (a)\\ \langle x,\mathcal{B}_\ell x\rangle\le 1,\ \ell=1,...,L & (b)\\ \mathcal{C}x=0 & (c)\end{array}\right\};\qquad (P')$$

its semidefinite relaxation (SDP$'$) is obtained from (SDP) by eliminating constraints (SDP.d). From the proof of Proposition 6 it follows that

$$\mathrm{Opt}(P')\le\mathrm{Opt}(\mathrm{SDP}')\le\Omega^2\mathrm{Opt}(P'),\ \Omega=\sqrt{32\ln(12(L+1))};$$

the resulting statement is a slightly improved version of "Approximate $\mathcal{S}$-Lemma" from [1] (in the original statement, L in the formula for Ω is replaced with the total rank of mappings $\mathcal{B}_\ell$, $\ell=1,...,L$).

B. The proof of Proposition 6 goes along the lines of the proof of Approximate $\mathcal{S}$-Lemma [1]; the crucial new component (bound (41) based on Theorem 2.2) allows to treat the norm constraint $(P.d)$.

C. Any further progress towards the proof of Conjecture 1.1 would result in improving the result of Proposition 6. For example, if Conjecture were true, we would be able to replace the terms $7(m_k+n_k)^{\frac{1}{6}}$ in (31) with a much nicer, from the theoretical viewpoint, terms $O(1)\sqrt{\ln(m_k+n_k)}$.

D. As it is usually the case with semidefinite relaxations of difficult optimization problems, (SDP) not only provides us with an efficiently computable upper bound on the optimal value of (P), but offers as well a randomized algorithm for building suboptimal feasible solutions to (P). Such an algorithm is suggested by the proof of Proposition 6; specifically, we generate a sample of, say, $M=1000$ realizations $\zeta^1,...,\zeta^M$ of the random matrix $\zeta(\xi)$ (see (33)), choose the largest possible scale factors λ_p such that the scaled matrices $\widehat{\zeta}^p=\lambda_p\zeta^p$ are feasible for (P), thus getting a sample of feasible solutions to (P), and then choose among these feasible solutions the one with the best – the largest – value of the objective. Note that the required "feasible scalings" indeed are possible, since the only potentially dangerous in this respect constraint in (P) – the system of homogeneous linear equations $(P.c)$ – is satisfied by every realization of $\zeta(\xi)$.

E. Under favourable circumstances, the outlined randomized algorithm can be further improved by a kind of "purification". Specifically, assume that

- (P) has no quadratic constraints $(P.a$-$b)$ (that is, $\mathcal{B}=0$, $\mathcal{B}_\ell=0$, $\ell=1,...,L$);
- The linear homogeneous constraints $(P.c)$ say that a feasible solution x to (P) possesses certain block-diagonal structure $\mathbf{\Delta}=\{(m_k,n_k)\}_{k=1}^K$ and impose no further restrictions on x;
- The objective $f(x)=f(x_1,...,x_K)$ we are maximizing in (P) is convex w.r.t. every one of the diagonal blocks x_k in a feasible (and thus block-diagonal) candidate solution x, the remaining components being fixed.

Note that outlined assumptions are satisfied in all problems of quadratic optimization under orthogonality constraints mentioned in the beginning of section 4. Now, given a feasible solution x with components x_k, $k=1,...,K$, purification converts x to a feasible solution $\widehat{x}$ with the same or better value of the objective in such a way that $\widehat{x}$ is an "extreme point" feasible solution to

(P). The latter means that every component $\widehat{x}_k$ of $\widehat{x}$ satisfies the orthogonality relation, namely, $\widehat{x}_k\widehat{x}_k^T = I_{m_k}$ when $m_k \le n_k$ and $\widehat{x}_k^T\widehat{x}_k = I_{n_k}$ when $m_k \ge n_k$.

The conversion $x \mapsto \widehat{x}$ takes K steps. At the first step, we represent x_1 as a convex combination of a moderate number Q of matrices x_1^q, $q = 1, ..., Q$, satisfying the orthogonality constraint (this is possible and can be done efficiently, see below); note that every one of the Q candidate solutions $x^q = (x_1^q, x_2, ..., x_K)$ is feasible for (P). We compute the value of f at these solutions and find the one, let it be x^{q_*}, with the largest value of f. Since $f(\cdot, x_2, ..., x_K)$ is convex and x_1 is a convex combination of x_1^q, $q = 1, ..., Q$, we have $f(x^{q_*}) \ge f(x)$. Thus, we have found a feasible solution $x^{(1)} = x^{q_*}$ to (P) with the same or better value of the objective than the one at x and with the first block satisfying the orthogonality constraint. Now we repeat this procedure with $x^{(1)}$ in the role of x and x_2 in the role of x_1, thus getting a feasible solution $x^{(2)}$ which is at least as good as $x^{(1)}$ in terms of the objective and has two blocks satisfying the orthogonality constraints. Proceeding in this fashion, we in K steps end up with an extreme point feasible solution $\widehat{x} = x^{(K)}$ which is at least as good as x.

> For the sake of completeness, we present here the standard algorithm for representing a given $\mu \times \nu$ matrix z, $\|z\| \le 1$, as a convex combination of matrices satisfying the orthogonality constraint. W.l.o.g., let us assume that $\mu \ge \nu$, so that the orthogonality constraint required form a $\mu \times \nu$ matrix w to satisfy the relation $w^T w = I_\nu$. We first find the singular value decomposition $z = U\mathrm{Diag}\{\sigma\}V^T$, where V is orthogonal $\nu \times \nu$ matrix, $\sigma \ge 0$ is the ν-dimensional vector of singular values of z and the $\mu \times \nu$ matrix U satisfies the relation $U^T U = I_\nu$. Since $\|z\| \le 1$, we have $0 \le \sigma_i \le 1$. Now observe that whenever $\gamma \in \mathbf{R}^\nu$ has entries ± 1, the matrix $U\mathrm{Diag}\{\gamma\}V^T$ satisfies the orthogonality constraint. Thus, all we need is to represent σ as a convex combination $\sum_q \lambda_q \gamma^q$ of a moderate number of vectors γ^q with entries ± 1, thus inducing the desired representation $z = \sum_q \lambda_q U\mathrm{Diag}\{\gamma^q\}V^T$. A required representation of σ is immediate. W.l.o.g. we may assume that $\sigma_1 \le \sigma_2 \le ... \le \sigma_\nu$. Let us define $\nu + 1$ ν-dimensional vectors δ^i as
>
> $$\delta^1 = (1, ..., 1)^T, \delta^2 = (0, 1, ..., 1)^T, \delta^3 = (0, 0, 1, ..., 1)^T, ..., \delta^{\nu+1} = (0, ..., 0)^T;$$
>
> observe that δ^i is the half-sum of two vectors $\delta_\pm^i$ with coordinates ± 1. The required representation of σ is merely
>
> $$\sigma = \sum_{i=1}^{\nu+1} (\sigma_i - \sigma_{i-1})\delta^i = \sum_{i=1}^{\nu+1} \frac{\sigma_i - \sigma_{i-1}}{2}[\delta_+^i + \delta_-^i]$$
>
> (we have set $\sigma_0 = 0$, $\sigma_{\nu+1} = 1$). This representation involves $Q = 2\nu + 1$ vectors with coordinates ± 1 (note that $\delta_+^1 = \delta_-^1$).

F. Finally, when f is affine in every one of x_k (as it is the case in the Procrustes problem), the purification can be simplified and improved – here we can at every step easily maximize f in the block to be updated. To simplify notation, consider the first step. Given $x_2, ..., x_K$, we can represent the affine function $\phi(y) = f(y, x_2, ..., x_K)$ of $m_1 \times n_1$ matrix y as $\phi(y) = \mathrm{Tr}(ya^T) + c$ with a, c readily given by $x_2, ..., x_K$. Assuming w.l.o.g. that $m_1 \ge n_1$, let us compute the singular value decomposition $a = U\mathrm{Diag}\{\sigma\}V^T$ of a, so that $\sigma \ge 0$. It is immediately seen that the maximum of $\phi(y)$ over y's satisfying the constraint $\|y\| \le 1$ is equal to $\sum_i \sigma_i$ and is attained at the matrix $y_* = UV^T$ satisfying the orthogonality constraint; y_* is clearly the best possible extreme point updating of x_1.

4.3 Numerical illustration: Procrustes problem

To illustrate the outlined considerations, we are about to present numerical results for the Procrustes problem

$$\mathrm{Opt(Pr)} = \min_{x[1],...,x[K]} \left\{ \sum_{1\le k<k'\le K} \|a[k]x[k] - a[k']x[k']\|_2^2 : x[k]x^T[k] = I_n, k = 1, ..., K \right\} \quad \text{(Pr)}$$

where $a[\cdot]$ are given $N \times n$ matrices. The problem is equivalent to the quadratic problem with orthogonality constraints

$$\max_{x[1],...,x[K]} \left\{ 2 \sum_{1\le k<k'\le K} \mathrm{Tr}(a[k]x[k]x^T[k']a^T[k']) : x[k]x^T[k] = I_n, k = 1, ..., K \right\}; \quad (50)$$

as we have already explained, relaxing the orthogonality constraints in the latter problem to $\|x[k]\| \le 1$, we preserve the optimal value, so that (Pr) is equivalent to the problem

$$\mathrm{Opt(Pr_+)} = \max_{x[1],...,x[K]} \left\{ 2 \sum_{1\le k<k'\le K} \mathrm{Tr}(a[k]x[k]x^T[k']a^T[k']) : \|x[k]\| \le 1,\ k = 1, ..., K \right\} \quad \text{(Pr}_+\text{)}$$

of the form of (P); the optimal values in (Pr) and (Pr$_+$) are linked by the relation

$$\mathrm{Opt(Pr)} = \underbrace{(K-1)\sum_{k=1}^{K} \mathrm{Tr}(a[k]a^T[k])}_{C} - \mathrm{Opt(Pr_+)}. \quad (51)$$

In our experiments, we generated random instances of (Pr), solved the semidefinite relaxations (SDP) of the resulting instances of (Pr$_+$), thus obtaining upper bounds on the optimal values of the latter instances (which, in turn, induce via (51) lower bounds on the optimal values in (Pr)), and used the randomized algorithm outlined in item D, section 4.2.1, to get suboptimal solutions to (Pr). The details are as follows.

Generating instances. Given "sizes" K, N, n of (Pr), we generated the data $a[1], ..., a[K]$ of (Pr) as follows: entries of $a[1]$ were picked at random from the standard Gaussian distribution $\mathcal{N}(0, 1)$, while the remaining matrices were generated as $a[k] = a[1]U_k + \epsilon Q_k$, $2 \le k \le K$, with randomly chosen orthogonal matrix U_k and random matrix Q_k generated in the same fashion as $a[1]$. The "closeness parameter" ϵ was chosen at random according to $\epsilon = \exp\{\xi\}$ with ξ uniformly distributed in $[-3, 3]$. The sizes K, n of (Pr) were limited by the necessity to end up with semidefinite relaxation not too difficult for the SDP solver `mincx` (LMI Toolbox for MATLAB) we used, which means at most 1000 – 1200 free entries in X. This restriction allows to handle the sizes (K, n) with $Kn \le 50$ (see below). The column size N of $a[\cdot]$ was always set to 20.

The relaxation of (Pr$_+$) as given by the above construction is the semidefinite problem

$$\mathrm{Opt(SDP)} = \max_X \left\{ F(X) \equiv 2\mathrm{Tr}(AX) : \begin{array}{rl} X \equiv [X_{kij,k'i'j'}]_{\substack{k,k'\le K \\ i,i',j,j'\le n}} & \succeq 0 \\ \mathcal{S}_k(X) \equiv \left[\sum_{p=1}^{n} X_{kpi,kpj}\right]_{i,j\le n} & \preceq I_n,\ k \le K \\ \mathcal{T}_k(X) \equiv \left[\sum_{p=1}^{n} X_{kip,kjp}\right]_{i,j\le n} & \preceq I_n,\ k \le K \end{array} \right\} \quad (52)$$

(see (SDP) and Remark 4.1), where A is the symmetric $Kn^2 \times Kn^2$ matrix with the entries

$$A_{kij,k'i'j'} = \begin{cases} \frac{1}{2}\sum_{\ell=1}^{N} a[k]_{\ell i} a[k']_{\ell i'}, & j = j' \text{ and } k \neq k' \\ 0, & j \neq j' \text{ or } k = k' \end{cases}. \tag{53}$$

In fact (52) can be significantly simplified. Specifically, let us treat $Kn \times Kn$ symmetric matrices Y as $K \times K$ block matrices with $n \times n$ blocks $Y^{k,k'} = [Y_{ij}^{k,k'}]_{i,j=1}^{n}$, $1 \leq k, k' \leq K$, and consider the semidefinite program

$$\max_{Y} \Big\{ G(Y) = \mathrm{Tr}(BY) : Y = \{Y^{k,k'}\} \in \mathbf{S}^{Kn}, Y \succeq 0, Y^{k,k} \preceq I_n,\ k = 1, ..., K \Big\} \tag{54}$$

where $B \in \mathbf{S}^{Kn}$ is the block matrix with $n \times n$ blocks $B^{k,k'} = \begin{cases} a^T[k]a[k'], & k \neq k' \\ 0, & k = k' \end{cases}$, $1 \leq k, k' \leq K$. Note that the design dimension of (54) is less than the one of (52) by factor $\approx n^2$.

Lemma 5 *Problems* (52), (54) *are equivalent to each other. Specifically, if a matrix* $X = [X_{kij,k'i'j'}]$ *is a feasible solution to* (52)*, then the matrix* $Y = \mathcal{Y}[X] \equiv \{Y^{k,k'}\}_{k,k'=1}^{K} \in \mathbf{S}^{Kn}$ *given by*

$$Y_{ii'}^{k,k'} = \sum_{p=1}^{n} X_{kip,k'i'p},\ 1 \leq i, i' \leq n, 1 \leq k, k' \leq K \tag{55}$$

is a feasible solution to (54)*, and* $F(X) = G(Y)$*. Moreover, every feasible solution* Y *to* (54) *is of the form* $\mathcal{Y}[X]$ *for an appropriate feasible solution* X *of* (52)*.*

Proof. Let X be a feasible solution to (52), $Y = \mathcal{Y}[X]$. Then $Y \succeq 0$. Indeed, since $X \succeq 0$, we have $X_{kij,k'i'j'} = \sum_{\ell=1}^{L} v_{kij}^{\ell} v_{k'i'j'}^{\ell}$ for appropriately chosen L and v_{kij}^{ℓ}. It follows that

$$Y \equiv \mathcal{Y}[X] = \sum_{\ell=1}^{L} \underbrace{\{ \left[\sum_{j=1}^{n} v_{kij}^{\ell} v_{k'i'j}^{\ell} \right]_{i,i'} \}_{k,k'=1}^{K}}_{Y^{\ell}};$$

it remains to note that the matrices Y^{ℓ} are sums of dyadic matrices and thus are symmetric positive semidefinite. Further, we have $\mathcal{T}_k(X) = Y^{k,k}$ (see (52)), so that Y is feasible for (54). The relation $F(X) = G(Y)$ is readily given by (55) and (53).

Now let $Y = \{Y^{k,k'}\}_{k,k'=1}^{K}$ be feasible for (54), and let us set

$$X_{kij,k'i'j'} = \frac{1}{n} \delta_{j'}^{j} Y_{ii'}^{k,k'} \tag{56}$$

(δ_q^p is the Kronecker symbol), so that $Y = \mathcal{Y}[X]$ by (55). It remains to prove that X is feasible for (52). Indeed, X is the Kronecker product of positive semidefinite matrix Y and $n^{-1} I_n$ and thus is positive semidefinite. Further, by (52) we clearly have $\mathcal{T}_k(X) = Y^{k,k} \preceq I_n$, and

$$(\mathcal{S}_k(X))_{jj'} = \sum_{p=1}^{n} X_{kpj,kpj'} = \sum_{p=1}^{n} \delta_{j'}^{j} \frac{1}{n} Y_{pp}^{k,k} \Rightarrow \mathcal{S}_k(X) = \frac{\mathrm{Tr}(Y^{k,k})}{n} I_n \preceq I_n,$$

where the concluding $\preceq$ is given by $Y^{k,k} \preceq I_n$ (recall that Y is feasible for (54)). ∎

Remark 4.2 The origin of (54) is as follows. The objective in (Pr_+) is a linear function of the matrix products $x[k]x^T[k']$, $k,k'=1,...,K$, which are nothing but the blocks $Y^{k,k'}$ in the positive semidefinite block matrix $Y=Y[x]=\left[\begin{array}{c}x[1]\\ \vdots\\ x[K]\end{array}\right]\left[\begin{array}{c}x[1]\\ \vdots\\ x[K]\end{array}\right]^T\in\mathbf{S}^{Kn}$, while the norm bounds in (Pr_+) translate into the constraints $\|Y^{k,k}\|\le 1$. Thus, (54) is obtained from (Pr_+) by passing to Y-variable and subsequent eliminating the nonconvex constraint "Y should be $Y[x]$ for some x".

Note that the outlined "recipe" for simplifying the semidefinite relaxation works in the case of general problem (P), provided that the constraints $(P.c)$ say *exactly* that the structure of (P) is $\{(m_k=\mu,n_k=\nu)\}_{k=1}^K$ and that the objective and the left hand sides in the constraints $(P.a-b)$ of (P) are linear functions of the matrix products $x[k]x^T[k']$, $k,k'=1,...,K$. Note also that under the latter assumptions the reasoning completely similar to the one in Lemma 5 demonstrates that the outlined simplification of (52) is in fact equivalent to (52).

Recovering suboptimal solutions to (Pr_+), (Pr) was implemented according to the randomized algorithm with purification outlined in section 4.2.1, items D, F. Specifically, after (high-accuracy approximation to) optimal solution Y_* of (54) was found, we "lifted" it, according to (56), to an optimal solution X_* of (52). Then we used X_* to generate a sample of $M=1000$ feasible solutions $x^\ell=\{x_k^\ell\}_{k=1}^K$ to (Pr_+) as explained in item 4.2.1.D and purified these solutions as explained in item 4.2.1.F, thus obtaining feasible solutions $\widehat{x}^\ell$ to (Pr_+) which satisfy the orthogonality constraints and thus are feasible for (50) and (Pr). The resulting suboptimal solution $\widehat{x}$ to (Pr) was the best (with the smallest value of the objective) of the feasible solutions $\widehat{x}^\ell$, $\ell=1,...,M$. The value of the objective of (Pr) at $\widehat{x}$ is an upper bound $\mathrm{Opt}^{\mathrm{up}}(\mathrm{Pr})$ on $\mathrm{Opt}(\mathrm{Pr})$, while the value of the objective of (Pr_+) at $\widehat{x}$ is a lower bound $\mathrm{Opt}_{\mathrm{lw}}(\mathrm{Pr}_+)$ on the optimal value of (Pr_+). Thus, we end up with brackets

$$\begin{array}{rcl}[L(\mathrm{Pr}_+),U(\mathrm{Pr}_+)] & \equiv & [\mathrm{Opt}_{\mathrm{lw}}(\mathrm{Pr}_+),\mathrm{Opt}(\mathrm{SDP})],\\ [L(\mathrm{Pr}),U(\mathrm{Pr})] & \equiv & [C-\mathrm{Opt}(\mathrm{SDP}),\mathrm{Opt}^{\mathrm{up}}(\mathrm{Pr})]\end{array}$$

on the optimal values of (Pr_+), (Pr), respectively, along with a feasible suboptimal solution $\widehat{x}$ to (Pr_+), (Pr); the values of the objectives of (Pr_+), (Pr) at $\widehat{x}$ are appropriate endpoints of the corresponding brackets.

Sizes of instances. The design dimension of (54) is $\frac{Kn(Kn+1)}{2}$; in order for it to be at most about 1200 (the limitation imposed by the SDP solver we used), Kn should be at most 50. In our experiments, we used Pareto-maximal pairs (K,n) with $Kn\le 50$, specifically, the 10 pairs $(K=\lfloor\frac{50}{n}\rfloor,n)$ given by $n=2,3,4,5,6,7,8,10,12,15$, and for every pair solved 20 instances of the corresponding sizes (which amounts to the total of 200 instances).

The results of our numerical experiments were surprisingly good, much better than one could expect looking at the bound (32). Indeed, the latter bound guarantees that Opt(SDP) is at most by factor Ω^2 greater than $\mathrm{Opt}(\mathrm{Pr}_+)$, with Ω^2 slowly growing with K,n and thus being a "moderate" constant, provided that K,n are not too large[2)]. As about (Pr), Proposition 6

[2)] For the values of K and $m=n$ we used in our experiments, the bound (32) results in Ω^2 varying from ≈ 243 ($K=3,n=15$) to ≈ 301 ($K=25,n=2$); more accurate (and more messy) numerics in the proof of Proposition 6 reduces the range of Ω^2 for our K,n to $[\approx 84,\approx 102]$.

yields no bounds on the ratio of the true optimal value in (Pr) and its efficiently computable lower bound $C - \mathrm{Opt}(\mathrm{SDP})$. The outlined theoretical guarantees, if any, are not too optimistic, which is in sharp contrast with the actual numerical results we got. In 200 experiments we have run, the largest relative error $\frac{U(\mathrm{Pr}_+)-L(\mathrm{Pr}_+)}{\max[1,U(\mathrm{Pr}_+)]}$ in solving (Pr_+) was as small as 9.0%, while the largest relative error $\frac{U(\mathrm{Pr})-L(\mathrm{Pr})}{\max[1,U(\mathrm{Pr})]}$ in solving (Pr) was as small as 2.4%. These data correspond to the best of the purified solutions $\widehat{x}^\ell$. As far as problem (Pr_+) is concerned, already the unpurified solutions x^ℓ were not so bad: the relative error of the best, in terms of the objective, of these solutions $\widetilde{x}$ was at worst 62.2%. Thus, in our experiments we did not observe ratios $\mathrm{Opt}(\mathrm{SDP})/\mathrm{Opt}(\mathrm{Pr}_+)$ exceeding 1.09 (cf. with the theoretical upper bound ≈ 100 on this ratio). The histograms of the relative errors are presented on Fig. 1. Fig. 2 presents an illustrative 3D plot.

References

[1] Ben-Tal, A., Nemirovski, A., Roos, C., "Robust solutions of uncertain quadratic and conic-quadratic problems", *Mathematics of Operations Research* **v. 28** (2003), 497–523.

[2] Browne, M.W., "On oblique Procrustes rotation", *Psychometrika* 32 (1967), 125-132.

[3] Edelman, A., Arias, T. Smith, S.T., "The geometry of algorithms with orthogonality constraints", *SIAM J. Matrix Anal. Appl.*, 20 (1999), 303–353.

[4] W.B. Johnson, G. Schechtman, "Remarks on Talagrand's deviation inequality for Rademacher functions", Banach Archive 2/16/90, *Springer Lecture Notes* 1470 (1991), pp. 72-77.

[5] Nemirovski, A., "On tractable approximations of randomly perturbed convex constraints" – *Proceedings of the 42nd IEEE Conference on Decision and Control Maui, Hawaii USA, December 2003*, 2419-2422.

[6] Nemirovski, A., "Regular Banach spaces and large deviations of random sums", Working paper
`http://iew3.technion.ac.il/Labs/Opt/index.php?4`

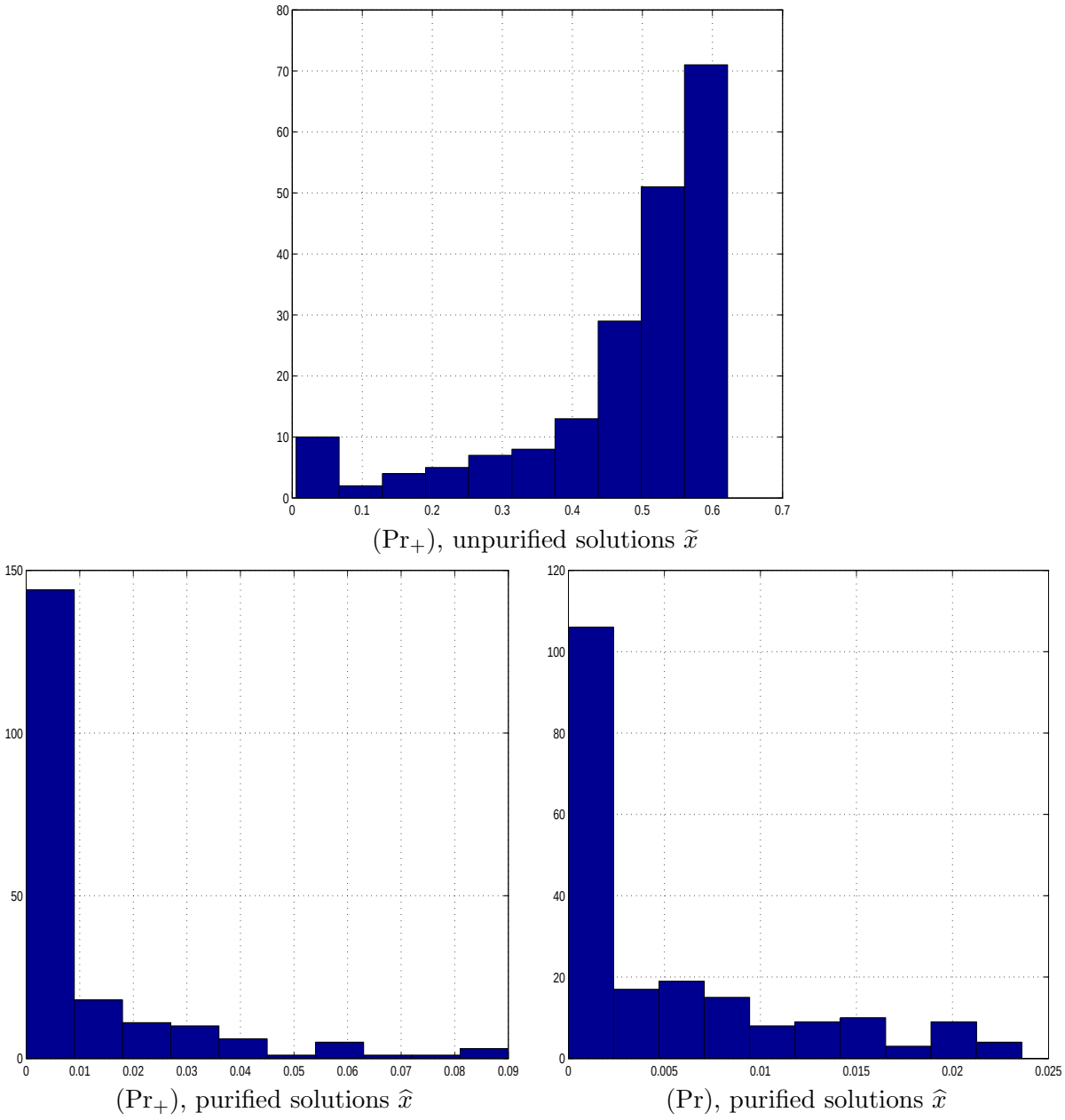
[7] Shapiro, A., "Extremal Problems on the Set of Nonnegative Definite Matrices", *Linear Algebra and Applications*, 67 (1985), 7–18.

[8] Shapiro, A., Botha, J.D., "Dual Algorithms for Orthogonal Procrustes Rotations", *SIAM Journal on Matrix Analysis and Applications* 9 (1988), 378–383.

[9] Ten Berge, J.M.F., Nevels, K., "A general solution to Mosiers oblique Procrustes problem", *Psychometrika* 42 (1977), 593–600.

[10] Wolkowicz, H., "Semidefinite Programming Approaches to the Quadratic Assignment Problem", in: R. Saigal, H. Wolkowitcz, L. Vandenberghe, Eds. *Handbook on Semidefinite Programming*, Kluwer Academic Publishers, 2000.

[11] Wolkowicz, H., Zhao, Q., "Semidefinite programming relaxations for graph partitioning problems", *Discrete Appl. Math.* 96/97 (1999), 461–479.



(Pr_+), unpurified solutions $\widetilde{x}$

(Pr_+), purified solutions $\widehat{x}$

(Pr), purified solutions $\widehat{x}$

Figure 1: Histograms of relative errors in 200 experiments

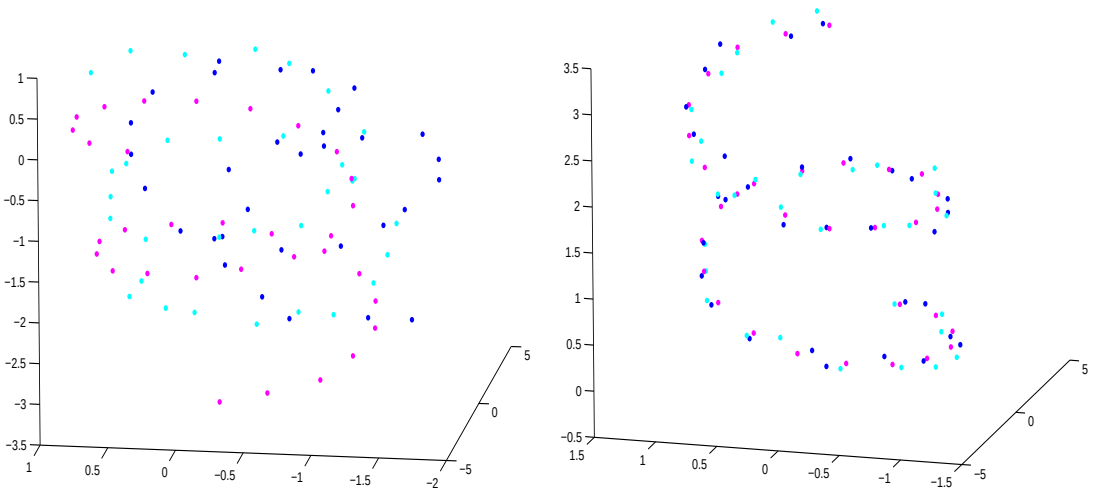


Figure 2: Procrustes problem with three 32×3 matrices with rows treated as coordinates of 3D points. Rotations are given by purified solutions to the Procrustes problem.

[12] Zhao, Q., Karisch, S. E., Rendl, F., Wolkowicz, H., "Semidefinite programming relaxations for the quadratic assignment problem", *J. Comb. Optim.* 2 (1998), 71–109.