



---

Konrad-Zuse-Zentrum  
für Informationstechnik Berlin

Takustraße 7  
D-14195 Berlin-Dahlem  
Germany

GUILLAUME SAGNOL

**On the semidefinite representations of  
real functions applied to symmetric  
matrices**

Herausgegeben vom  
Konrad-Zuse-Zentrum für Informationstechnik Berlin  
Takustraße 7  
D-14195 Berlin-Dahlem

Telefon: 030-84185-0  
Telefax: 030-84185-125

e-mail: [bibliothek@zib.de](mailto:bibliothek@zib.de)  
URL: <http://www.zib.de>

ZIB-Report (Print) ISSN 1438-0064  
ZIB-Report (Internet) ISSN 2192-7782

# On the semidefinite representations of real functions applied to symmetric matrices

Guillaume Sagnol

Final version: August 2013

## Abstract

We present a new semidefinite representation for the trace of a real function  $f$  applied to symmetric matrices, when a semidefinite representation of the convex (or concave) function  $f$  is known. Our construction is intuitive, and yields a representation that is more compact than the previously known one. We also show with the help of matrix geometric means and a Riemannian metric over the set of positive definite matrices that for a rational exponent  $p$  in the interval  $(0, 1]$ , the matrix  $X$  raised to  $p$  is the largest element of a set represented by linear matrix inequalities. This result further generalizes to the case of the matrix  $A \sharp_p B$ , which is the point of coordinate  $p$  on the geodesic from  $A$  to  $B$ . We give numerical results for a problem inspired from the theory of experimental designs, which show that the new semidefinite programming formulation can yield an important speed-up factor.

**Keywords** semidefinite representability, optimal experimental designs, SDP, matrix geometric mean

**AMS Classification:** 90C22, 62K05

## 1 Introduction

In this article we discuss semidefinite representations of scalar functions applied to symmetric matrices. We recall that it is possible to extend the definition of a function  $f : I \mapsto \mathbb{R}, x \rightarrow f(x)$ , where  $I$  is a real interval, to the set  $\mathbb{S}_m^I$  of  $m \times m$ -symmetric matrices whose spectrum lies in  $I$  as follows: if  $X = U \text{Diag}(\lambda_1, \dots, \lambda_m) U^T$  is an eigenvalue decomposition of  $X$ , then we define  $f(X) := U \text{Diag}(f(\lambda_1), \dots, f(\lambda_m)) U^T$ . Throughout this article we denote by  $\mathbb{S}_m$  (resp.  $\mathbb{S}_m^+, \mathbb{S}_m^{++}$ ) the set of  $m \times m$  symmetric (resp. positive semidefinite, positive definite) matrices.

If the scalar function  $f$  is semidefinite representable, then a result of Ben-Tal and Nemirovski can be used to construct a semidefinite representation of  $X \rightarrow \text{trace } f(X)$ . Indeed,  $\text{trace } f(X)$  can be rewritten as  $\sum_i f(\lambda_i)$ , which is a symmetric and semidefinite representable function of the eigenvalues of  $X$ , so that Proposition 4.2.1. in [5] applies.

In this article, we show that the semidefinite representation of  $x \rightarrow f(x)$  can be lifted to the matrix case  $X \rightarrow \text{trace } f(X)$  by an intuitive transformation

which involves Kronecker products (Theorem 3.1). For a convex (resp. concave) function, the resulting semidefinite representation of the epigraph (resp. hypograph)

$$\begin{aligned} E &= \{(t, X) \in \mathbb{R} \times \mathbb{S}_m : \text{trace } f(X) \leq t\} \\ \left( \text{resp. } H &= \{(t, X) \in \mathbb{R} \times \mathbb{S}_m : \text{trace } f(X) \geq t\} \right) \end{aligned}$$

is more compact than the one obtained from the general construction of Ben-Tal and Nemirovski, in which the Ky-Fan  $k$ -norms of  $M$  must be bounded for  $k = 1, \dots, m$ . Our numerical results of Section 5 moreover show that the semidefinite programs (SDP) based on the present representation are solved in a shorter time than the former SDP formulations, and they are numerically more stable.

For the case where  $f(x) = x^p : \mathbb{R}^+ \mapsto \mathbb{R}^+$ , where  $p$  is a rational number in  $(0, 1]$ , we shall see that our construction yields a stronger result. Namely, we show in Theorem 4.4 that  $X^p$  has an extremal representation of the form

$$X^p = \max_{\preceq} \{T \in \mathbb{S}_m : T \in \mathcal{S}\},$$

where the set  $\mathcal{S}$  is semidefinite representable and  $\max_{\preceq}$  denotes the largest element with respect to the Löwner ordering, which is defined over  $\mathbb{S}_m$  as follows:

$$A \preceq B \iff (B - A) \in \mathbb{S}_m^+.$$

More generally, our result shows that the matrix  $B \sharp_p A := B^{\frac{1}{2}} (B^{-\frac{1}{2}} A B^{-\frac{1}{2}})^p B^{\frac{1}{2}}$  has an extremal representation. The proof of this result uses the notion of *matrix geometric mean*, and the Banach fixed point theorem in the space  $\mathbb{S}_m^{++}$  equipped with a Riemannian metric.

Our study is motivated by the *theory of optimal experimental designs*, where the general problem to solve takes the form

$$\begin{aligned} \max_{\mathbf{w} \in \mathbb{R}^s} \quad & \Phi_p \left( \sum_{i=1}^s w_i M_i \right), \\ \text{s. t.} \quad & \sum_{i=1}^s w_i = 1, \quad \mathbf{w} \geq \mathbf{0}, \end{aligned} \tag{1}$$

where  $M_1, \dots, M_s$  are given positive semidefinite matrices, and for  $p \in [-\infty, 1]$  the  $\Phi_p$ -criterion is defined over the set of positive definite matrices  $M \in \mathbb{S}_m^{++}$  as

$$\Phi_p(M) = \begin{cases} \lambda_{\min}(M) & \text{for } p = -\infty ; \\ \left( \frac{1}{m} \text{trace } M^p \right)^{\frac{1}{p}} & \text{for } p \in (-\infty, 1], p \neq 0 ; \\ (\det(M))^{\frac{1}{m}} & \text{for } p = 0. \end{cases} \tag{2}$$

The definition of  $\Phi_p$  is extended by continuity to singular matrices  $M \in \mathbb{S}_m^+$ , so that  $\Phi_p(M) = 0$  if  $M$  is singular and  $p \leq 0$ . This design problem arises when an experimenter must select a subset of trials to perform among a set of available experiments, and has many applications in various domains, such as dose-finding in clinical studies [11] or measurements in telecommunication networks [21]. We refer the reader to Pukelsheim [17] for more background on optimal experimental designs.

Note that any semidefinite representation of the function  $M \rightarrow \text{trace } MP$  yields a semidefinite programming (SDP) formulation of Problem (1). The cases  $p = -\infty$ ,  $p = -1$ , and  $p = 0$ , known as  $E$ -,  $A$ - and  $D$ -optimal design problems have been extensively studied in the literature, and SDP formulations are known for these problems [10]. We also point out that lighter Second Order Cone Programming (SOCP) formulations exist for  $p = -1$  and  $p = 0$  [19, 18]. The general case ( $p \in [-\infty, 1]$ ) deserved less attention. However, it was recently noticed by Papp [16] that the SDP formulation can be obtained by using Proposition 4.2.1. in [5]. Our numerical results (cf. Section 5) show that the new SDP formulation from this paper can improve the computation time by several orders of magnitude.

## 2 Preliminaries

In this section, we briefly recall some basic notion about semidefinite representability and matrix geometric means. A *Semidefinite Program* (SDP) is an optimization problem where a linear function  $\mathbf{c}^T \mathbf{x}$  must be maximized, among the vectors  $\mathbf{x}$  belonging to a set  $S$  defined by *linear matrix inequalities* (LMI):

$$S = \{\mathbf{x} \in \mathbb{R}^n : F_0 + \sum_i x_i F_i \succeq 0\}.$$

We now recall the definition of a semidefinite representable set, which was introduced by Ben-Tal and Nemirovski [5]:

**Definition 2.1** (Semidefinite representability). A convex set  $S \subset \mathbb{R}^n$  is said to be *semidefinite representable*, abbreviated *SDr*, if  $S$  is the projection of a set in a higher dimensional space which can be described by LMIs. In other words,  $S$  is *SDr* if and only if there exists symmetric matrices  $F_0, \dots, F_n, F'_1, \dots, F'_{n'}$  such that

$$\mathbf{x} \in S \iff \exists \mathbf{y} \in \mathbb{R}^{n'} : F_0 + \sum_{i=1}^n x_i F_i + \sum_{i=1}^{n'} y_i F'_i \succeq 0.$$

Such an LMI is called a *semidefinite representation* (SDR) of the set  $S$ .

**Definition 2.2** (SDR of a function). A convex (resp. concave) function  $f : S \subset \mathbb{R}^n \mapsto \mathbb{R}$  is said *SDr* if and only if the epigraph of  $f$ ,  $\{(t, \mathbf{x}) : f(\mathbf{x}) \leq t\}$  (resp. the hypograph  $\{(t, \mathbf{x}) : t \leq f(\mathbf{x})\}$ ), is *SDr*.

It follows immediately from these two definitions that the problem of maximizing a concave *SDr* function (or minimizing a convex one) over a *SDr* set can be cast as an SDP.

We now give a short insight on the theory of matrix geometric means and the Riemannian metric of the set of positive definite matrices  $\mathbb{S}_m^{++}$ . We refer the reader to the book of Bhatia [6] and the references therein for more details on this subject. The Geometric mean of two positive definite matrices  $A, B \in \mathbb{S}_m^{++}$  was introduced by Ando [3]:

$$A \sharp B := A^{1/2} (A^{-1/2} B A^{-1/2})^{1/2} A^{1/2}.$$

In the latter paper, Ando shows that  $A \sharp B$  satisfies the following extremal property:

$$A \sharp B = \max_{\succeq} \left\{ X \in \mathbb{S}_m : \begin{pmatrix} A & X \\ X & B \end{pmatrix} \succeq 0 \right\}. \quad (3)$$

The space of positive definite matrices is equipped with the Riemannian metric

$$\delta_2(A, B) = \|\log A^{-1/2} B A^{-1/2}\|_F,$$

where  $\|M\|_F = \sqrt{\text{trace}(M^T M)}$  denotes the Frobenius norm of  $M$ . In this space, there exists a unique geodesic  $[A, B]$  between two matrices  $A$  and  $B$ , which can be parametrized as follows (using the common  $\sharp_t$ -notation):

$$\gamma_{A \rightarrow B}(t) = A \sharp_t B := A^{1/2} (A^{-1/2} B A^{-1/2})^t A^{1/2}, \quad 0 \leq t \leq 1.$$

Note that  $A \sharp B = A \sharp_{\frac{1}{2}} B$  is the midpoint of this geodesic. The geometric mean of two matrices has the following properties (cf. chapter 4 in [6]):

- *it is commutative:*

$$A \sharp B = B \sharp A; \quad (4)$$

- if  $A$ , and  $B$  commute, then

$$A \sharp B = A^{1/2} B^{1/2}; \quad (5)$$

- the map  $X \rightarrow A \sharp X$  is *operator monotone*, i.e.

$$Y \succeq X \implies A \sharp Y \succeq A \sharp X; \quad (6)$$

- the  $\sharp$ -operator is *invariant under congruent transformation*: for all invertible matrix  $U$  of size  $m$ ,

$$U^T (A \sharp B) U = (U^T A U) \sharp (U^T B U). \quad (7)$$

We also point out that the metric  $\delta_2$  enjoys an important convexity property (cf. §6.1.11 in [6]), which will be useful in the proof of Theorem 4.4:

$$\forall A, B, C, D \in \mathbb{S}_m^{++}, \quad \delta_2(A \sharp B, C \sharp D) \leq \frac{1}{2} \delta_2(A, C) + \frac{1}{2} \delta_2(B, D). \quad (8)$$

### 3 Lifting the SDR of a scalar function

In this section, we show that the *SDR* of a function  $f : I \mapsto \mathbb{R}$  can be transformed in a simple way to a *SDR* of trace  $f : \mathbb{S}_m^I \rightarrow \mathbb{R}$ :

**Theorem 3.1.** *Let  $f : I \mapsto \mathbb{R}$  be a scalar function, where  $I$  is a real interval. Assume that  $f$  admits the following SDR: for all  $x \in I$ ,*

$$f(x) \leq t \iff \exists \mathbf{y} \in \mathbb{R}^n : F_0 + xF_X + tF_T + \sum_{i=1}^n y_i F_i \succeq 0,$$

where the symmetric matrices  $F_0, \dots, F_n, F_X, F_T$  are given. Then, a SDR of the function  $g : \mathbb{S}_m^I \mapsto \mathbb{R}, X \rightarrow \text{trace } f(X)$  is given by: for all  $X \in \mathbb{S}_m^I$ ,

$$\text{trace } f(X) \leq t \iff \exists T, Y_1, \dots, Y_n \in \mathbb{S}_m :$$

$$(i) \quad F_0 \otimes I_m + F_X \otimes X + F_T \otimes T + \sum_{i=1}^n F_i \otimes Y_i \succeq 0;$$

$$(ii) \quad \text{trace } T \leq t,$$

where  $I_m$  denotes the  $m \times m$  identity matrix and  $\otimes$  is the Kronecker product. In other words, the SDR is lifted from scalar to matrices by replacing each scalar by a corresponding matrix block of size  $m \times m$ .

*Proof.* Let  $X$  be an arbitrary matrix in  $\mathbb{S}_m^I$ , and  $X = U \text{Diag}(\boldsymbol{\lambda}) U^T$  be an eigenvalue decomposition of  $X$ .

We first assume that  $\text{trace } f(X) \leq t$ . For  $k = 1, \dots, m$ , define  $t_k = f(\lambda_k)$ . By assumption there exists a vector  $\mathbf{y}^{(k)}$  such that

$$B_k := F_0 + \lambda_k F_X + t_k F_T + \sum_{i=1}^n y_i^{(k)} F_i \succeq 0.$$

Denote by  $\mathcal{B}$  the block diagonal matrix with blocks  $B_1, \dots, B_m$  on the diagonal, and by  $\mathbf{y}_i$  the vector of  $\mathbb{R}^m$  with components  $y_i^{(1)}, \dots, y_i^{(m)}$ . We may write

$$\mathcal{B} = I_m \otimes F_0 + \text{Diag}(\boldsymbol{\lambda}) \otimes F_X + \text{Diag}(\mathbf{t}) \otimes F_T + \sum_{i=1}^n \text{Diag}(\mathbf{y}_i) \otimes F_i \succeq 0.$$

In the previous expression, we may commute the Kronecker products, which is equivalent to pre- and post-multiplying by a permutation matrix:

$$F_0 \otimes I_m + F_X \otimes \text{Diag}(\boldsymbol{\lambda}) + F_T \otimes \text{Diag}(\mathbf{t}) + \sum_{i=1}^n F_i \otimes \text{Diag}(\mathbf{y}_i) \succeq 0.$$

Now, we multiply this expression to the left by the block diagonal matrix  $\text{Diag}(U, \dots, U) = I \otimes U$ , and to the right by its transpose. This gives:

$$F_0 \otimes I_m + F_X \otimes X + F_T \otimes T + \sum_{i=1}^n F_i \otimes Y_i \succeq 0,$$

where we have set  $T = U \text{Diag}(\mathbf{t}) U^T$  and  $Y_i = U \text{Diag}(\mathbf{y}_i) U^T$ . By construction, we have  $T = f(X)$ , so (ii) holds and we have proved the “ $\Rightarrow$ ” part of the theorem.

For the converse part, consider some matrices  $T', Y'_1, \dots, Y'_n \in \mathbb{S}_m$  such that the LMI (i) of the theorem is satisfied. Define  $H_T = T' - T$  and  $H_i = Y'_i - Y_i$ , where  $T = f(X)$  and  $Y_i = U \text{Diag}(\mathbf{y}_i) U^T$  are defined as in the first part of this proof. We will show that  $\text{trace } H_T \geq 0$ , which implies  $\text{trace } T' \geq \text{trace } f(X)$ , and the proof will be complete.

So from (i) we have:

$$F_0 \otimes I_m + F_X \otimes X + F_T \otimes (T + H_T) + \sum_{i=1}^n F_i \otimes (Y_i + H_i) \succeq 0.$$

Again, we multiply this expression to the left by  $I \otimes U^T$  and to the right by  $I \otimes U$ , and then we commute the Kronecker products. This gives:

$$\text{Diag}(B_1, \dots, B_m) + U^T H_T U \otimes F_T + \sum_{i=1}^n U^T H_i U \otimes F_i \succeq 0.$$

For all  $k = 1, \dots, m$ , this implies that the  $k$ th diagonal block is positive semidefinite:

$$B_k + (U^T H_T U)_{k,k} F_T + \sum_{i=1}^n (U^T H_i U)_{k,k} F_i \succeq 0.$$

According to the *SDR* of the scalar function  $f$ , it means that

$$f(\lambda_k) \leq t_k + (U^T H_T U)_{k,k},$$

and since  $f(\lambda_k) = t_k$  we obtain  $(U^T H_T U)_{k,k} \geq 0$ . From there, it is easy to conclude:

$$\text{trace } H_T = \text{trace } H_T U U^T = \text{trace } U^T H_T U = \sum_{k=1}^m (U^T H_T U)_{k,k} \geq 0.$$

□

**Example 3.2.** A *SDR* of the function  $x \rightarrow x^p$ , where  $p \in \mathbb{Q}$  is briefly sketched in [5] (§3.3.1., examples 12 to 15) and given with more details in [1] (§2.3.h). (Note that this function is concave for  $p \in [0, 1]$  and convex for other values of  $p$ .) For example, the epigraph of the convex function  $x \rightarrow x^{-4/3}$  mapping  $(0, \infty)$  onto itself, may be represented as follows: for all  $t \geq 0, x > 0$ :

$$\begin{aligned} x^{-4/3} \leq t &\iff 1 \leq x^4 t^3 \\ &\iff \exists u \geq 0, v \geq 0 : 1 \leq xu, u^2 \leq tv, v^2 \leq t \\ &\iff \exists u \in \mathbb{R}, v \in \mathbb{R} : \begin{pmatrix} x & 1 \\ 1 & u \end{pmatrix} \succeq 0, \begin{pmatrix} t & u \\ u & v \end{pmatrix} \succeq 0, \begin{pmatrix} t & v \\ v & 1 \end{pmatrix} \succeq 0 \end{aligned}$$

For the second equivalence, the “ $\Leftarrow$ ” part is clear and the “ $\Rightarrow$ ” part is obtained by setting  $u = t^{3/4}$  and  $v = t^{1/2}$ . The third equivalence is a standard use of the Schur complement lemma. By using Theorem 3.1, we obtain a *SDR* of the function  $X \rightarrow \text{trace } X^{-4/3}$ :

$$\text{trace } X^{-4/3} \leq t \iff \exists U, V, T \in \mathbb{S}_m : \left\{ \begin{array}{l} \begin{pmatrix} X & I_m \\ I_m & U \end{pmatrix} \succeq 0 \\ \begin{pmatrix} T & U \\ U & V \end{pmatrix} \succeq 0, \\ \begin{pmatrix} T & V \\ V & I_m \end{pmatrix} \succeq 0, \\ \text{trace } T \leq t \end{array} \right.$$

□

Note however that LMI (i) of Theorem 3.1 does not imply the stronger property  $f(X) \preceq T$ . As a counter-example, consider the function  $f(x) = x^4$ , which admits the *SDR*

$$x^4 \leq t \iff \exists u \in \mathbb{R} : \begin{pmatrix} u & x \\ x & 1 \end{pmatrix} \succeq 0, \begin{pmatrix} t & u \\ u & 1 \end{pmatrix} \succeq 0.$$



If we set  $X = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ ,  $U = \begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix}$  and  $T = \begin{pmatrix} 18 & 24 \\ 24 & 34 \end{pmatrix}$ , the reader can check that the LMI (i) of Theorem 3.1 holds:

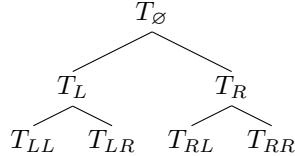
$$\begin{pmatrix} U & X \\ X & I_2 \end{pmatrix} \succeq 0, \quad \begin{pmatrix} T & U \\ U & I_2 \end{pmatrix} \succeq 0,$$

but  $X^4 \not\preceq T$ . In the next section, we show that this stronger property holds for  $f : x \rightarrow x^p$  when  $p \in \mathbb{Q} \cap (0, 1]$ .

## 4 Semidefinite representation of $A \#_p B$

Throughout this section,  $p$  denotes a rational number in  $(0, 1]$ , and we choose two integers  $\alpha$  and  $\beta$  such that  $p = \frac{\alpha}{\beta}$  and  $0 < \alpha \leq \beta$ . We are going to show that the lifted *SDR* of the function  $f_p$  mapping  $\mathbb{R}_+$  onto itself and defined by  $f(x) = x^p$ , also provides an extremal representation of  $X^p$ . In other words, there is a *SDr* set  $\mathcal{S} \subset \mathbb{S}_m^+$  for which  $X^p$  is the largest element with respect to Löwner ordering. In fact, we will prove a stronger statement: for any  $A \in \mathbb{S}_m^+$  and  $B \in \mathbb{S}_m^{++}$ , the matrix  $B \#_p A = B^{\frac{1}{2}} (B^{-\frac{1}{2}} A B^{-\frac{1}{2}})^p B^{\frac{1}{2}}$ , which is the point of coordinate  $p$  on the geodesic from  $B$  to  $A$ , is the largest element of a *SDr* set. The matrix power  $A^p$  arises as the particular case  $B = I_m$ .

To do this, we first present the construction of the *SDR* of  $\tilde{f}_p : (a, b) \rightarrow a^p b^{1-p}$ . As explained in [1], this *SDR* is based on binary trees whose nodes contain variables. Note that in a perfect binary tree, every node of depth  $k$  can be indexed by an element of  $\Gamma_k := \{L, R\}^k$ , which indicates the sequence of *left* or *right* turns needed to reach this node from the root of the tree. For example, a perfect binary tree  $T$  of depth 2 is indexed as follows:



We denote by  $\mathcal{T}_n = \Gamma_0 \cup \dots \cup \Gamma_n$  the set of node indices in a perfect binary trees of depth  $n$ , so that  $(\mathbb{S}_m)^{\mathcal{T}_n}$  represents the set of binary trees of depth  $n$ , whose nodes are matrices of  $\mathbb{S}_m$ . The concatenation of tree indices is denoted by  $\sqcup$ , so that for example,  $LR \sqcup L = LRL \in \Gamma_3$ . We define  $n$  as the integer such that  $2^{n-1} < \beta \leq 2^n$ , and  $\chi_{\alpha, \beta}$  as the following formal sequence of length  $2^n$ , which is formed with characters in the alphabet  $\{\mathbf{A}, \mathbf{B}, \mathbf{X}\}$ :

$$\chi_{\alpha, \beta} := (\underbrace{\mathbf{A}, \dots, \mathbf{A}}_{\alpha \text{ times}}, \underbrace{\mathbf{B}, \dots, \mathbf{B}}_{(\beta - \alpha) \text{ times}}, \underbrace{\mathbf{X}, \dots, \mathbf{X}}_{(2^n - \beta) \text{ times}}).$$

Now, let  $\sigma$  be any one of the permutations of  $\chi_{\alpha, \beta}$ , and let  $A, B$  be two arbitrary matrices (respectively in  $\mathbb{S}_m^+$  and  $\mathbb{S}_m^{++}$ ). The elements of  $\sigma$  are indexed by  $\gamma \in \Gamma_n$ , in the order corresponding to the leaves of a tree of depth  $n$  from left to right. For example, if  $\sigma = (\mathbf{A}, \mathbf{B}, \mathbf{X}, \mathbf{B})$ , we have  $\sigma_{LL} = \mathbf{A}$ ,  $\sigma_{RL} = \mathbf{X}$ , and  $\sigma_{LR} = \sigma_{RR} = \mathbf{B}$ . Let us now define for all  $\gamma \in \Gamma_n$  the function  $\hat{\sigma}_\gamma$ , mapping  $\mathbb{S}_m^+$  to either  $A, B$ , or

$X$  as follows:

$$\forall X \in \mathbb{S}_m^+, \hat{\sigma}_\gamma(X) = \begin{cases} A & \text{if } \sigma_\gamma = \mathbf{A} \\ B & \text{if } \sigma_\gamma = \mathbf{B} \\ X & \text{if } \sigma_\gamma = \mathbf{X} \end{cases} \quad (9)$$

With a slight abuse of notation, we will write  $\sigma_\gamma(X)$  instead of  $\hat{\sigma}_\gamma(X)$  in the remaining of this paper. (So  $\sigma_\gamma$  denotes a character in  $\{\mathbf{A}, \mathbf{B}, \mathbf{X}\}$ , while  $\sigma_\gamma(X)$  denotes the corresponding matrix in  $\mathbb{S}_m^+$ .) Similarly, we will simply write  $\sigma(X)$  to denote the sequence  $(\hat{\sigma}_\gamma(X))_{\gamma \in \Gamma_n}$ .

We can now construct the *SDR* of  $\tilde{f}_p$  (already lifted to  $\mathbb{S}_m^+$  by considering matrix blocks instead of scalar variables). It involves a tree whose root is  $X$ , leaves are defined by  $\sigma$ , and a LMI related to the matrix geometric mean must be satisfied at each node:

$$\begin{aligned} \mathcal{S}(\sigma) = \{X \in \mathbb{S}_m^+ : \exists T \in (\mathbb{S}_m)^{\mathcal{T}_n} : \\ (i) \ T_\emptyset = X; \\ (ii) \ \forall \gamma \in \Gamma_n, \ T_\gamma = \sigma_\gamma(X); \\ (iii) \ \forall k = 0, \dots, n-1, \ \forall \gamma \in \Gamma_k, \ \begin{pmatrix} T_{\gamma \sqcup L} & T_\gamma \\ T_\gamma & T_{\gamma \sqcup R} \end{pmatrix} \succeq 0\} \end{aligned}$$

**Example 4.1.** If  $p = 1/3$ , we have  $\alpha = 1$ ,  $\beta = 3$ ,  $n = 2$ , and  $\sigma$  must contain respectively  $\alpha = 1$ ,  $(\beta - \alpha) = 2$ , and  $(2^n - \beta) = 1$  copies of  $\mathbf{A}, \mathbf{B}$ , and  $\mathbf{X}$ . If  $\sigma = (\mathbf{A}, \mathbf{X}, \mathbf{B}, \mathbf{B})$ , the set  $\mathcal{S}(\sigma)$  is defined through a tree of the form

$$\begin{array}{c} X \\ \swarrow \quad \searrow \\ T_L \quad T_R \\ \swarrow \quad \searrow \quad \swarrow \quad \searrow \\ A \quad X \quad B \quad B \end{array} \quad (10)$$

The property (iii) in the definition of  $\mathcal{S}(\sigma)$  implies that  $T_R$  satisfies

$$\begin{pmatrix} B & T_R \\ T_R & B \end{pmatrix} \succeq 0.$$

So by Equation (3) we have  $T_R \preceq B$ , and the definition of  $\mathcal{S}(\sigma)$  simplifies to:

$$X \in \mathcal{S}(\sigma) \iff \exists T_L \in \mathbb{S}_m : \begin{pmatrix} T_L & X \\ X & B \end{pmatrix} \succeq 0, \begin{pmatrix} A & T_L \\ T_L & X \end{pmatrix} \succeq 0.$$

□

Generally speaking, we point out that the order of the elements in the permutation  $\sigma$  can be chosen such that the definition of  $\mathcal{S}(\sigma)$  involves no more than  $2(n-1) = O(\log \beta)$  LMIs of size  $2m \times 2m$ :

**Proposition 4.2.** *The permutation  $\sigma$  of  $\chi_{\alpha, \beta}$  can always be chosen such that  $\mathcal{S}(\sigma)$  admits an *SDR* involving no more than  $2(n-1) = O(\log \beta)$  LMIs of size  $2m \times 2m$ , and no more than  $2n-3 = O(\log \beta)$  additional variables  $T_\gamma \in \mathbb{S}_m$ .*

A constructive proof of this result is sketched in appendix.

Now, as a consequence of Equation (3), observe that property (iii) in the definition of  $\mathcal{S}(\sigma)$  implies  $T_\gamma \preceq T_{\gamma \sqcup L} \# T_{\gamma \sqcup R}$  (if the geometric mean is well defined, i.e.  $T_{\gamma \sqcup L}, T_{\gamma \sqcup R} \in \mathbb{S}_m^{++}$ ). By operator monotonicity of the matrix geometric mean (see (6)), we see that if the matrices  $A, B, X, T_L$  and  $T_R$  of Tree (10) are positive definite, then:

$$X \preceq T_L \# T_R \preceq (A \# X) \# (B \# B).$$

In the general case, we can bound  $X$  from above (with respect to Löwner ordering) by an expression with nested “ $\#$ -operations” in the binary tree whose leaves are defined through  $\sigma$ . Formally, for any two collections  $\varsigma, \varsigma' \in (\mathbb{S}_m^+)^{\Gamma_n}$  of  $2^n$  positive definite matrices indexed by tree indices, i.e.

$$\forall \gamma \in \Gamma_n, \quad \varsigma_\gamma \in \mathbb{S}_m^+, \varsigma'_\gamma \in \mathbb{S}_m^+,$$

we define the juxtaposition  $\varsigma \sqcap \varsigma' \in (\mathbb{S}_m^+)^{\Gamma_{n+1}}$  as follows:

$$\forall \gamma \in \Gamma_n, \quad (\varsigma \sqcap \varsigma')_{L \sqcup \gamma} = \varsigma_\gamma, \quad (\varsigma \sqcap \varsigma')_{R \sqcup \gamma} = \varsigma'_\gamma.$$

Note that if  $n \geq 1$ , any sequence  $\varsigma \in (\mathbb{S}_m^+)^{\Gamma_n}$  can be uniquely decomposed as  $\varsigma = \varsigma^{[L]} \sqcap \varsigma^{[R]}$ , where the subsequences  $\varsigma^{[L]}$  and  $\varsigma^{[R]}$  are elements of  $(\mathbb{S}_m^+)^{\Gamma_{n-1}}$ . This allows us to define inductively the *nested sharp operator*  $\#$ , which acts from  $(\mathbb{S}_m^{++})^{\Gamma_n}$  onto  $\mathbb{S}_m^{++}$  as follows:

$$\begin{cases} \forall \varsigma \in (\mathbb{S}_m^{++})^{\Gamma_0}, & \#(\varsigma) := \varsigma_\emptyset; \\ \forall \varsigma = \varsigma^{[L]} \sqcap \varsigma^{[R]} \in (\mathbb{S}_m^{++})^{\Gamma_n}, \quad n \geq 1, & \#(\varsigma) := (\#(\varsigma^{[L]})) \# (\#(\varsigma^{[R]})). \end{cases}$$

With a straightforward induction, we can now obtain the following lemma:

**Lemma 4.3.** *If  $A$  and  $B$  are positive definite, then*

$$X \in \mathcal{S}(\sigma) \implies X \preceq \#(\sigma(X)).$$

We can finally give the main result of this section:

**Theorem 4.4** (Extremal representation of  $B \#_p A$ ). *Let  $p = \frac{\alpha}{\beta}$ ,  $0 < \alpha \leq \beta$ ,  $\alpha \in \mathbb{N}, \beta \in \mathbb{N}$ ,  $A \in \mathbb{S}_m^+$ ,  $B \in \mathbb{S}_m^{++}$ , and let  $\sigma$  be an arbitrary permutation of  $\chi_{\alpha, \beta}$ . Then,  $B \#_p A$  satisfies the following extremal property*

$$B \#_p A = \max_{\preceq} \{X \in \mathbb{S}_m^+ : X \in \mathcal{S}(\sigma)\}.$$

*Proof.* Let  $A \in \mathbb{S}_m^{++}$  be an arbitrary positive definite matrix. We are first going to show that  $B \#_p A = \max_{\preceq} \{X \in \mathbb{S}_m^{++} : X \in \mathcal{S}(\sigma)\}$ . The general statement where  $A \in \mathbb{S}_m^+$  may be singular will be obtained at the end of this proof by continuity. Our proof relies on the following lemma:

**Lemma 4.5.** *If  $A$  and  $B$  are positive definite, then  $X = B \#_p A$  is a solution of the equation*

$$X = \#(\sigma(X)).$$

*Proof of the lemma.*

Define a sequence  $\sigma'(X)$ , such that for all  $\gamma \in \Gamma_n$ ,

$$\sigma'_\gamma(X) = B^{-\frac{1}{2}} \sigma_\gamma(X) B^{-\frac{1}{2}}.$$

Set  $Y = B^{-\frac{1}{2}}AB^{-\frac{1}{2}}$ , so that the elements of  $\sigma'(B \#_p A)$  are all in  $\{Y, I_m, Y^p\}$ , and hence they commute. A simple induction shows that  $\#(\sigma'(B \#_p A)) = Y^p$  (the geometric means are easy to compute when the matrices commute, see (5)). By congruence invariance (see (7)),

$$\begin{aligned}\#(\sigma(B \#_p A)) &= B^{\frac{1}{2}} \#(\sigma'(B \#_p A)) B^{\frac{1}{2}} \\ &= B^{\frac{1}{2}} Y^p B^{\frac{1}{2}} \\ &= B \#_p A.\end{aligned}$$

□

*Proof of Theorem 4.4 (continued).* Let  $X \in \mathbb{S}_m^{++}$  such that  $X \in \mathcal{S}(\sigma)$ , and let  $T \in (\mathbb{S}_m)^{\mathcal{T}_n}$  be a tree satisfying properties (i) – (iii) of the definition of  $\mathcal{S}(\sigma)$ . Define a new tree  $T'$  as follows:

$$\forall \gamma \in \Gamma_{n-1}, T'_\gamma := T_{\gamma \sqcup L} \# T_{\gamma \sqcup R} \succeq T_\gamma,$$

and

$$\forall \gamma \in \Gamma_k, T'_\gamma := T'_{\gamma \sqcup L} \# T'_{\gamma \sqcup R} \succeq T_\gamma$$

for  $k = (n-2), \dots, 0$ . By construction, the root of  $T'$  is  $X' := T'_\emptyset = \#(\sigma(X)) \succeq X$ . It remains to define the leaves of  $T'$ , which we do according to  $\sigma(X')$ :

$$\forall \gamma \in \Gamma_n, T'_\gamma := \sigma_\gamma(X') \succeq T_\gamma.$$

By construction, it is clear that  $T'$  satisfies the property (iii) of the definition of  $\mathcal{S}(\sigma)$  for the depth levels  $k = 0, \dots, n-2$ . For a  $\gamma \in \Gamma_{n-1}$ , (iii) also holds, because

$$\begin{pmatrix} T'_{\gamma \sqcup L} & T'_\gamma \\ T'_\gamma & T'_{\gamma \sqcup R} \end{pmatrix} \succeq \begin{pmatrix} T_{\gamma \sqcup L} & T'_\gamma \\ T'_\gamma & T_{\gamma \sqcup R} \end{pmatrix} \succeq 0,$$

where the first inequality follows from  $T'_{\gamma \sqcup L} \succeq T_{\gamma \sqcup L}$ ,  $T'_{\gamma \sqcup R} \succeq T_{\gamma \sqcup R}$ , and the second inequality is a consequence of  $T'_\gamma = T_{\gamma \sqcup L} \# T_{\gamma \sqcup R}$ . This shows that  $X'$  belongs to  $\mathcal{S}(\sigma)$ , too.

Define  $h : \mathbb{S}_m^{++} \mapsto \mathbb{S}_m^{++}$ ,  $X \mapsto \#(\sigma(X))$ . So far, we have shown that  $X \in \mathcal{S}(\sigma) \implies [h(X) \succeq X \text{ and } h(X) \in \mathcal{S}(\sigma)]$ . By using the convexity property of the Riemannian metric (Equation (8)), a simple induction shows that  $h$  is a contraction mapping with a contraction factor equal to the fraction of the number of elements of  $\sigma$  that take the value  $X$ :

$$\forall X, X' \in \mathbb{S}_m^{++}, \delta_2(h(X), h(X')) \leq \frac{2^n - \beta}{2^n} \delta_2(X, X') < \delta_2(X, X').$$

Hence, the mapping  $X \mapsto h(X)$  is contractive in the space  $\mathbb{S}_m^{++}$  equipped with the Riemannian metric  $\delta_2$ . It is known that this space is complete (see e.g. [15]), and hence we can apply the Banach fixed point theorem: the fixed point equation  $X = h(X)$  has a unique solution  $X^* \in \mathbb{S}_m^{++}$ . Moreover for all  $X \in \mathbb{S}_m^{++}$  the sequence defined by  $X_0 = X$ ,  $X_{i+1} = h(X_i)$  converges to  $X^*$ . In particular, if  $X \in \mathcal{S}(\sigma)$ , our previous discussion shows that  $X \preceq X^*$ . Moreover, we can construct a tree with leaves defined by  $\sigma(X^*)$ , and such that all non-leaf node is the geometric mean of its two children, so that the root will be  $h(X^*) = X^*$ .

So,  $X^* \in \mathcal{S}(\sigma)$  is the largest element of  $\mathcal{S}(\sigma)$ . Finally, we know from Lemma 4.5 that  $X^* = B \sharp_p A$ .

It remains to show that the statement of the theorem remains valid when the matrix  $A \in \mathbb{S}_m^+$  is singular. We will first show that  $B \sharp_p A \in \mathcal{S}(\sigma)$ . Consider a tree  $T$  with leaves  $\sigma(B \sharp_p A)$ . Since the matrix  $B \sharp_p A = B^{\frac{1}{2}} (B^{-\frac{1}{2}} A B^{-\frac{1}{2}})^p B^{\frac{1}{2}}$  is singular, we may not define the non-leaf nodes of this tree using geometric means. However, observe that every leaf is of the form  $B \sharp_k A$ , with  $k \in \{0, p, 1\}$ . Hence, we define the non-leaf nodes of  $T$  by the following relation: if  $T_{\gamma \sqcup L} = B \sharp_{k_1} A$  and  $T_{\gamma \sqcup R} = B \sharp_{k_2} A$ , then  $T_\gamma := B \sharp_{\frac{k_1+k_2}{2}} A$ . It is easy to check that the property (iii) of the definition of  $\mathcal{S}(\sigma)$  is satisfied for the resulting tree  $T$ . A simple induction shows that the root of this tree is

$$T_\emptyset = B \sharp_{\frac{n_A(\sigma) + p n_X(\sigma)}{2^n}} A,$$

where  $n_A(\sigma)$  and  $n_X(\sigma)$  represent the number of times that  $\mathbf{A}$  and  $\mathbf{X}$  appear in the sequence  $\sigma$ . Replacing  $n_A(\sigma)$  by  $\alpha$  and  $n_X(\sigma)$  by  $2^n - \beta$ , we find  $T_\emptyset = B \sharp_p A$ . This shows that  $B \sharp_p A \in \mathcal{S}(\sigma)$ .

Until now the matrices  $A$  and  $B$  were fixed, but now we need to let  $A$  vary so we introduce the notation  $\bar{\sigma}_{A'}(X)$  to represent the sequence of length  $2^n$  which is identical to  $\sigma(X)$ , except for the elements  $A$  that are replaced by  $A'$ :

$$(\bar{\sigma}_{A'}(X))_\gamma := \begin{cases} A' & \text{if } \sigma_\gamma = \mathbf{A}; \\ B & \text{if } \sigma_\gamma = \mathbf{B}; \\ X & \text{if } \sigma_\gamma = \mathbf{X}. \end{cases}$$

In particular,  $\bar{\sigma}_A(X) = \sigma(X)$  for all  $X \in \mathbb{S}_m^+$ . We also define the set

$$\bar{\mathcal{S}}(\sigma) := \{(A', X) \in \mathbb{S}_m^+ \times \mathbb{S}_m^+ : X \in \mathcal{S}(\bar{\sigma}_{A'})\}.$$

Note that  $\bar{\mathcal{S}}(\sigma)$  is *SDr* by construction, and hence convex.

Now, choose a sequence  $A_i \in \mathbb{S}_m^+$  such that  $A_i \rightarrow A$  as  $i \rightarrow \infty$ , as well as a sequence  $\epsilon_i > 0$  such that  $\epsilon_i \rightarrow 0$ . We know that  $(A_i, B \sharp_p A_i) \in \bar{\mathcal{S}}(\sigma)$  for all  $i$ . Let  $X \in \mathcal{S}(\sigma)$ , so that  $(A, X) \in \bar{\mathcal{S}}(\sigma)$  and define  $A'_i := (1 - \epsilon_i)A + \epsilon_i A_i$ ,  $X'_i := (1 - \epsilon_i)X + \epsilon_i B \sharp_p A_i$ . By convexity of  $\bar{\mathcal{S}}(\sigma)$ , we have  $(A'_i, X'_i) \in \bar{\mathcal{S}}(\sigma)$ . Moreover, since the matrices  $A'_i$  and  $X'_i$  are positive definite, we know that  $X'_i \preceq B \sharp_p A'_i$ . By taking the limit, we obtain  $X \preceq B \sharp_p A$ . This completes the proof.  $\square$

**Corollary 4.6.** *Let  $p = \frac{\alpha}{\beta}$ ,  $0 < \alpha \leq \beta$ , and for all  $U, X \in \mathbb{S}_m^+$  let  $\sigma_X(U)$  be a permutation of  $\chi_{\alpha, \beta}(X, I_m, U)$  (the order of the elements in  $\sigma_X(U)$  does not depend on  $X$  and  $U$ ). If  $K$  is a  $m \times r$ -matrix, then the concave function  $X \rightarrow \text{trace } K^T X^p K$ , which maps  $\mathbb{S}_m^+$  to  $\mathbb{R}^+$ , has the following semidefinite representation: for all  $X \in \mathbb{S}_m^+$ ,*

$$t \leq \text{trace } K^T X^p K \iff \exists U \in \mathcal{S}(\sigma_X), t \leq \text{trace } K^T U K.$$

*Proof.* If  $t \leq \text{trace } K^T X^p K$ , we set  $U = X^p$ , so that  $t \leq \text{trace } K^T U K$  and by Theorem 4.4  $U \in \mathcal{S}(\sigma_X)$ . Conversely, assume that  $U \in \mathcal{S}(\sigma_X)$ . We know from previous theorem that  $U \preceq X^p$ . Hence, we have  $\text{trace } M U \leq \text{trace } M X^p$  for all positive semidefinite matrix  $M$ . In particular,

$$\text{trace } K^T U K = \text{trace } K K^T U \leq \text{trace } K K^T X^p = \text{trace } K^T X^p K,$$

from which the conclusion follows.  $\square$

**Remark 4.7.** In the recent years, many authors have proposed to generalize the definition of matrix geometric means for three or more matrices [4, 7, 9, 8]. In a seminal work, Ando, Li and Mathias [4] have given a list of 10 properties that a “good” geometric mean should satisfy. A natural question arising from this paper is the following: *is there a matrix geometric mean satisfying the ten properties of [4], that coincides with the largest element of a SDR set?* Interestingly, our result can be generalized in a straightforward way to give an extremal representation of the unique fixed point  $F(A_1, \dots, A_N)$  of the map

$$X \rightarrow \# \left( (A_1, A_2, \dots, A_N, \underbrace{X, \dots, X}_{2^n - N}) \right),$$

where the operand of  $\#$  is a sequence of  $(\mathbb{S}_m^{++})^{\Gamma_n}$ , with  $n \in \mathbb{N}$  such that  $2^{n-1} < N \leq 2^n$ . We claim that this fixed point satisfies most of the Ando-Li-Mathias properties, but the permutation invariance is not fulfilled: for a permutation  $\sigma$ ,  $F(A_1, \dots, A_N) \neq F(A_{\sigma(1)}, \dots, A_{\sigma(N)})$  in general.

## 5 Numerical Results

In this section, we compare the CPU time required to solve problems of the form

$$\min_{\substack{\mathbf{w} \geq \mathbf{0} \\ \sum_i w_i = 1}} \text{trace } f\left(\sum_{k=1}^s w_k M_k\right), \quad (P_f)$$

by using the semidefinite representation of Theorem 3.1, and the one of Ben-Tal and Nemirovski [5]. This problem is inspired from the application to optimal experimental design that is presented in the introduction. For the sake of variety, we do not limit ourselves to power functions  $x \rightarrow x^p$  with  $p < 1$ . More precisely, assume that  $f : I \rightarrow \mathbb{R}$  is a convex real valued function defined on the interval  $I$ , an SDR of  $f$  is known:

$$\forall x \in I, f(x) \leq t \iff \exists \mathbf{y} \in \mathbb{R}^n : F_0 + xF_X + tF_T + \sum_{i=1}^n y_i F_i \succeq 0,$$

and the matrices  $M_1, \dots, M_s \in \mathbb{S}_m^I$  are given. We compare the efficiency of the following two SDP formulations of Problem  $(P_f)$ : the one with block matrices resulting from Theorem 3.1,

$$\begin{aligned} & \min_{X, T, \{Y_i\}, \mathbf{w}} \text{trace } T && (SDP_f - 1) \\ & \text{s. t.} && F_0 \otimes I_m + F_X \otimes X + F_T \otimes T + \sum_{i=1}^n F_i \otimes Y_i \succeq 0; \\ & && X = \sum_{k=1}^s w_k M_k, \quad \mathbf{w} \geq \mathbf{0}, \quad \sum_{k=1}^s w_k = 1, \end{aligned}$$

$f(x)$	$I$	$m$	CPU time (s) [SeDuMi]		CPU time (s) [MOSEK]	
			$(SDP_f - 1)$	$(SDP_f - 2)$	$(SDP_f - 1)$	$(SDP_f - 2)$
$-x^{\frac{1}{3}}$	$[0, \infty)$	10	0.40	0.80	0.05	0.18
		25	5.16	40.85	0.72	6.92
		40	59.19	<i>706.43<sup>†</sup></i>	3.47	<i>71.25<sup>b</sup></i>
$-x^{\frac{2}{5}}$	$[0, \infty)$	10	0.58	1.28	0.08	0.22
		25	20.38	39.57	1.68	5.88
		40	298.90	<i>799.77<sup>†</sup></i>	<i>19.33<sup>b</sup></i>	<i>66.69<sup>b</sup></i>
$x^{\frac{-8}{7}}$	$(0, \infty)$	10	0.49	0.90	0.10	0.19
		25	22.38	<i>40.07<sup>†</sup></i>	1.65	4.96
		40	357.22	<i>691.75<sup>†</sup></i>	14.65	61.87
$x^{\frac{7}{4}}$	$[0, \infty)$	10	0.41	1.23	0.07	0.17
		25	8.71	39.95	<i>1.19<sup>b</sup></i>	<i>5.90<sup>b</sup></i>
		40	120.16	<i>741.15<sup>†</sup></i>	<i>6.38<sup>b</sup></i>	<i>68.63<sup>b</sup></i>
$\frac{1}{x(1-x)}$	$(0, 1)$	10	0.30	0.76	0.05	0.15
		25	4.31	37.21	0.37	4.97
		40	51.79	<i>607.57<sup>†</sup></i>	2.27	<i>61.44<sup>b</sup></i>
$g_{\text{conv}}(x)$	$\mathbb{R}$	10	0.75	1.50	0.12	0.24
		25	63.62	<i>43.08<sup>†</sup></i>	2.94	4.37
		40	1019.70	<i>903.55<sup>†</sup></i>	31.06	<i>58.46<sup>b</sup></i>

Table 1: CPU time of two SDP formulations for Problem  $(P_f)$  with SeDuMi and MOSEK. The second column indicates the interval  $I$  where the function  $f$  is defined, and the third column specifies the size of the matrices  $M_i \in \mathbb{S}_m^I$ . The function  $g_{\text{conv}}$  in the last row is defined in (11). <sup>†</sup>The numbers displayed in italics for SeDuMi indicate that the solver stopped before reaching the optimality tolerance, because of numerical problems. <sup>b</sup>The numbers displayed in italics for MOSEK indicate that the solver returned the near-optimal status.

and the SDP from [5] that bounds each Ky-Fan Norm of  $X$ :

$$\begin{aligned}
& \min_{X, t, x, y, \sigma, \{Z_j\}} \sum_{j=1}^m t_j && (SDP_f - 2) \\
& \text{s. t.} && F_0 + x_j F_X + t_j F_T + \sum_{i=1}^n y_i^{(j)} F_i \succeq 0, \quad (j = 1, \dots, m); \\
& && x_1 \geq x_2 \geq \dots \geq x_m; \\
& && \sum_{k=1}^j x_k - j\sigma_j - \text{trace}(Z_j) \geq 0, \quad (j = 1, \dots, m-1); \\
& && Z_j \succeq 0, \quad (j = 1, \dots, m-1); \\
& && Z_j - X + \sigma_j I_m \succeq 0, \quad (j = 1, \dots, m-1); \\
& && \text{trace } X = \sum_{j=1}^m x_j; \\
& && X = \sum_{k=1}^s w_k M_k, \quad \mathbf{w} \geq \mathbf{0}, \quad \sum_{k=1}^s w_k = 1.
\end{aligned}$$

Our computational results are summarized in Table 1. Besides rational power functions, we have also consider the function  $f : (0, 1) \mapsto \mathbb{R}$ ,  $x \mapsto \frac{1}{x(x-1)}$ ,

which has the *SDR*

$$\begin{aligned} \forall x \in (0, 1), \quad f(x) \leq t &\iff \exists u \in \mathbb{R} : 1 \leq u(1-x), \quad 1 \leq (t-u)x \\ &\iff \exists u \in \mathbb{R} : \begin{pmatrix} u & 1 \\ 1 & 1-x \end{pmatrix} \succeq 0, \quad \begin{pmatrix} t-u & 1 \\ 1 & x \end{pmatrix} \succeq 0, \end{aligned}$$

as well as the convex envelope of a polynomial of degree 6. The fact that convex envelopes of univariate rational functions are *SDr* was proved by Laraki and Lasserre [13]. For the function

$$g_{\text{conv}} : \mathbb{R} \mapsto \mathbb{R}, x \rightarrow \text{convex-envelope}\left(\frac{x^6}{6} - 3\frac{x^4}{2} + 4x^2 + x\right), \quad (11)$$

the *SDR* given in [13] is:

$$\begin{aligned} g_{\text{conv}}(x) \leq t &\iff \exists y_2, \dots, y_6 \in \mathbb{R} : \begin{pmatrix} 1 & x & y_2 & y_3 \\ x & y_2 & y_3 & y_4 \\ y_2 & y_3 & y_4 & y_5 \\ y_3 & y_4 & y_5 & y_6 \end{pmatrix} \succeq 0, \\ &t \geq \frac{y_6}{6} - 3\frac{y_4}{2} + 4y_2 + x. \end{aligned}$$

For all our instances, we have generated  $s = 25$  random matrices  $M_i \in \mathbb{S}_m^I$ . We solved the SDPs by using SeDuMi [22] interfaced by YALMIP [14], and MOSEK 7.0 [2] interfaced by PICOS [20], with the default settings on a PC with 8 processors at 2.2GHz. Our experiments show that the block matrix formulation (*SDP<sub>f</sub> - 1*) improves the CPU time by a factor that vary between 2 and 12 for SeDuMi, and between 1.5 and 27 with MOSEK. An exception is the case  $f = g_{\text{conv}}$  with SeDuMi, but in this case the solver encountered numerical problems with (*SDP<sub>f</sub> - 2*) and stopped the computation before reaching the optimality tolerance.

Also note that SeDuMi was always able to compute an optimal solution with (*SDP<sub>f</sub> - 1*) but not with (*SDP<sub>f</sub> - 2*), and that MOSEK returned the *near-optimal* status twice more often with the formulation (*SDP<sub>f</sub> - 2*). This suggests that the formulation from this paper is numerically more stable.

To confirm this fact, we have computed the behavioural measures studied in [12], for the two SDP formulations of several of the above instances. The authors of this article evidence a positive correlation between the number of interior point method iterations and some measures of the SDP instances, such as the Renegar condition number  $C$  and some geometry measures  $D_p^\epsilon, g_p, D_d^\epsilon, g_d$  related to the primal and dual feasible regions and the norm of the solutions. We used a matlab code written by an author of [12] to compute these behavioural measures and compare them. For all the instances we considered, the Renegar condition number  $C$  was infinite, as well as the geometric measures  $D_p^\epsilon$  and  $g_d$ . However,  $g_p$  and  $D_d^\epsilon$  had finite values for most instances, see Table 2. These quantities are always orders of magnitude smaller for the formulation (*SDP<sub>f</sub> - 1*), which suggests that the formulation of this paper has better numerical properties indeed.



$f(x)$	$D_d^\epsilon$		$g_p$	
	$(SDP_f - 1)$	$(SDP_f - 2)$	$(SDP_f - 1)$	$(SDP_f - 2)$
$-x^{1/3}$	319.2	2438.4	639.0	38292
$-x^{2/5}$	463.5	2115.7	1143.0	38420
$x^{-8/7}$	$+\infty$	$+\infty$	$+\infty$	$+\infty$
$x^{7/4}$	674.4	2087.8	297.4	8117.7
$\frac{x}{1-x}$	167.7	215.2	128.9	4867.8
$g_{\text{conv}}(x)$	1506.8	2582.4	5522.8	12472

Table 2: Behavioural measures  $D_d^\epsilon$  and  $g_p$  of [12], computed for the same instances as in Table 1, for  $m = 25$ .

## 6 Acknowledgement

The author warmly thanks Stéphane Gaubert for his suggestion to generalize the second result of this paper, which was originally reduced to the case  $B = I_m$ . His thanks also go to Fernando Ordóñez, who has provided the matlab code to compute the behavioural measures of [12], and to an anonymous referee, whose contribution helped to improve the clarity of this paper.

## References

- [1] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Mathematical programming*, 95(1):3–51, 2003.
- [2] E.D. Andersen, B. Jensen, J. Jensen, R. Sandvik, and U. Worsøe. Mosek version 6. Technical report, Technical Report TR–2009–3, MOSEK, 2009.
- [3] T. Ando. Concavity of certain maps on positive definite matrices and applications to hadamard products. *Linear Algebra and its Applications*, 26:203–241, 1979.
- [4] T. Ando, C. Li, and R. Mathias. Geometric means. *Linear algebra and its applications*, 385:305–334, 2004.
- [5] A. Ben-Tal and A. Nemirovski. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, volume 2. Society For Industrial Mathematics, 1987.
- [6] R. Bhatia. *Positive definite matrices*. Princeton University Press, 2008.
- [7] R. Bhatia and J. Holbrook. Riemannian geometry and matrix geometric means. *Linear algebra and its applications*, 413(2):594–618, 2006.
- [8] D. Bini and B. Iannazzo. Computing the karcher mean of symmetric positive definite matrices. *Linear Algebra and its Applications*, 438(4):1700–1710, 2013.
- [9] D. Bini, B. Meini, and F. Poloni. An effective matrix geometric mean satisfying the ando-li-mathias properties. *Mathematics of Computation*, 79(269):437–452, 2010.

- [10] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [11] V. Fedorov, Y. Wu, and R. Zhang. Optimal dose-finding designs with correlated continuous and discrete responses. *Statistics in Medicine*, 2012.
- [12] R.M. Freund, F. Ordóñez, and K. Toh. Behavioral measures and their correlation with ipm iteration counts on semi-definite programming problems. *Mathematical programming*, 109(2-3):445–475, 2007.
- [13] R. Laraki and J.B. Lasserre. Computing uniform convex approximations for convex envelopes and convex hulls. *Journal of Convex Analysis*, 15(3):635–654, 2008.
- [14] J. Löfberg. Yalmip: A toolbox for modeling and optimization in matlab. In *Computer Aided Control Systems Design, 2004 IEEE International Symposium on*, pages 284–289. IEEE, 2004.
- [15] M. Moakher and M. Zéraï. The riemannian geometry of the space of positive-definite matrices and its application to the regularization of positive-definite matrix-valued data. *Journal of Mathematical Imaging and Vision*, 40(2):171–187, 2011.
- [16] D. Papp. Optimal designs for rational function regression. *Journal of the American Statistical Association*, 107(497):400–411, 2012.
- [17] F. Pukelsheim. *Optimal Design of Experiments*. Wiley, 1993.
- [18] G. Sagnol. A class of semidefinite programs with a rank-one solution. *Linear Algebra and its Applications*, 435(6):1446–1463, 2011.
- [19] G. Sagnol. Computing optimal designs of multiresponse experiments reduces to second-order cone programming. *Journal of Statistical Planning and Inference*, 141(5):1684 – 1708, 2011.
- [20] G. Sagnol. Picos, a python interface to conic optimization solvers. Technical Report 12-48, ZIB, 2012. <http://picos.zib.de>.
- [21] G. Sagnol, S. Gaubert, and M. Bouhtou. Optimal monitoring on large networks by successive c-optimal designs. In *22nd international teletraffic congress (ITC22), Amsterdam, The Netherlands*, September 2010.
- [22] J.F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12:625–653, 1999.

## Appendix

*Proof of Proposition 4.2.* To prove this result, the key point is that whenever both children of a variable node are the same matrix  $Z \in \{A, B, X\}$ , we can assume without loss of generality that the variable takes the same value  $Z$ . To see this, let  $T_\gamma$  be a variable node, and assume that  $T_{\gamma \sqcup L} = T_{\gamma \sqcup R} = Z$ . Denote respectively by  $\gamma_0$  and  $\gamma_1$  the parent and sibling nodes of  $\gamma$ . The variable  $T_\gamma$  is involved in the 2 LMIs

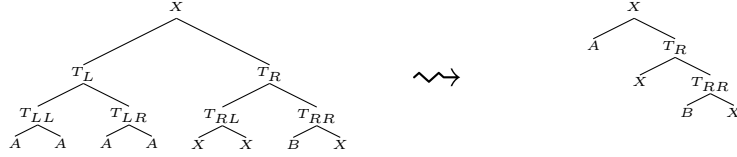
$$\begin{pmatrix} Z & T_\gamma \\ T_\gamma & Z \end{pmatrix} \succeq 0 \text{ and } \begin{pmatrix} T_\gamma & T_{\gamma_0} \\ T_{\gamma_0} & T_{\gamma_1} \end{pmatrix} \succeq 0. \quad (12)$$

We know that  $T_\gamma = Z \# Z = Z$  is the largest matrix (in the Löwner sense) satisfying the first LMI (see (3)), and if  $T_\gamma$  satisfies the second LMI, then  $Z \succeq T_\gamma$  implies

$$\begin{pmatrix} Z & T_{\gamma_0} \\ T_{\gamma_0} & T_{\gamma_1} \end{pmatrix} \succeq 0.$$

Hence we can assume without loss of generality that  $T_\gamma$  takes the value  $Z$ .

This technique can be used to reduce the number of matrix variables and LMIs to represent the set  $\mathcal{S}(\sigma)$ . For example, the tree involved in the representation of  $\mathcal{S}((A, A, A, A, X, X, B, X))$  can be reduced as follows



so that this set can be represented using only 3 LMIs and 2 additional matrix variables  $T_R \in \mathbb{S}_m$  and  $T_{RR} \in \mathbb{S}_m$ .

We next show how to construct a permutation  $\sigma$  of  $\chi_{\alpha, \beta}$ , such that  $\mathcal{S}(\sigma)$  has a representation involving no more than  $2(n-1)$  LMIs of size  $2m \times 2m$  and  $2n-3$  additional variables  $T_\gamma \in \mathbb{S}_m$ .

Let  $n_A := \alpha$ ,  $n_B := \beta - \alpha$  and  $n_X := 2^n - \beta$  represent the number of occurrences of A, B and X in  $\chi_{\alpha, \beta}$  (Recall that  $n$  is the largest integer such that  $2^{n-1} < \beta \leq 2^n$ ). Let the binary representation of these numbers be

$$n_A = \sum_{k \in K_A} 2^k, n_B = \sum_{k \in K_B} 2^k, \text{ and } n_X = \sum_{k \in K_X} 2^k,$$

where  $K_A, K_B$  and  $K_X$  are (finite) subsets of  $\mathbb{N}$ . Denote by  $\varsigma_k^Z$  the sequence containing  $2^k$  times the character Z. Our permutation  $\sigma$  is obtained by concatenating the sequences  $(\varsigma_k^A)_{k \in K_A}, (\varsigma_k^B)_{k \in K_B}, (\varsigma_k^X)_{k \in K_X}$ , ordered by decreasing values of  $k$  (the order of the subsequences  $\varsigma_k^{Z_1}$  and  $\varsigma_k^{Z_2}$  that have the same index  $k$  but different characters  $Z_1 \neq Z_2$  has no importance). For example, if  $n_A = 4 = 2^2$ ,  $n_B = 1 = 2^0$ , and  $n_X = 3 = 2^1 + 2^0$ , the permutation  $\sigma$  is obtained by arranging the subsequences  $\{(A, A, A, A), (B), (X, X), (X)\}$  by decreasing lengths, which gives e.g.  $\sigma = (A, A, A, A, X, X, B, X)$ .

Observe that we need at most one LMI to define the variables  $T_\gamma$  for the depth level  $\gamma \in \Gamma_{n-1}$ . Indeed, the variables are arranged by groups of  $2^k$

variables, so if  $k > 0$  this number is even and according to our claim at the beginning of the appendix we do not need an LMI, nor an additional matrix variable to define the parents of the variables in this group. There are at most three groups  $\zeta_0^Z = (Z)$  of cardinality one (one for each  $Z \in \{A, B, X\}$ ), but  $0 \in K_A \cap K_B \cap K_X$  would imply that  $n_A, n_B$  and  $n_X$  are odd, and so  $n = n_A + n_B + n_X$  would be odd, a contradiction. So there are at most two singleton groups of the form  $(Z)$  in  $\sigma$ , and we need a single LMI to define their parent.

We next show by induction on  $j$  that the number of LMIs needed at each depth level  $n - j$  of the binary tree is bounded by 2. This was readily shown for  $j = 1$ . The leftmost variables in the tree at the depth  $n - j$  are arranged by groups of  $2^{k-j}$  identical variables  $Z$  ordered by decreasing length, for all  $k \in K_Z \geq j$ ,  $Z \in \{A, B, X\}$ . In addition, by our induction hypothesis we have 0, 1 or 2 additional variables  $T_{\gamma_i}$ , which are at the rightmost locations of  $\Gamma_{n-j}$ . So the number of *singleton variables* at depth  $n - j$ , i.e. the variables who can't be paired with an identical sibling to define their parent, is bounded by 5 (up to 2 additional variables and a singleton  $(Z)$  for each  $Z$  such that  $j \in K_Z$ ), but as above it can't be 5 otherwise there would be an odd number of nodes at level  $n - j$ . This shows that the *singleton variables* of depth  $n - j$  are at most 4, they occupy the rightmost positions of  $\Gamma_{n-j}$ , and we need no more than 2 LMIs and 2 additional variables indexed in  $\Gamma_{n-(j+1)}$  to handle them.

To sum up, we need at most one LMI at level  $n - 1$ , and at most two LMIs at levels  $1, 2, \dots, n - 2$ , plus one LMI at the root of the tree, so the total number of LMIs required to define  $\mathcal{S}(\sigma)$  is bounded by  $2(n - 1)$ . The number of required additional variable is the same number minus one, because  $T_\emptyset = X$  and we do not need an extra variable for the LMI at the root.

Before concluding the proof, we point out that this bound can be attained. For example, take  $\alpha = 3, \beta = 5$ , so that  $n = 3$ ,  $n_A = n_X = 3$  and  $n_B = 2$ . This yields the permutation  $\sigma = (A, A, B, B, X, X, A, X)$  of  $\chi_{\alpha, \beta}$ , and the reader can verify that  $2(n - 1) = 4$  LMIs of size  $2m \times 2m$  are needed to define  $\mathcal{S}(\sigma)$ .  $\square$