AN EFFECTIVE MATRIX GEOMETRIC MEAN SATISFYING THE ANDO–LI–MATHIAS PROPERTIES

DARIO A. BINI, BEATRICE MEINI AND FEDERICO POLONI

ABSTRACT. We propose a new matrix geometric mean satisfying the ten properties given by Ando, Li and Mathias [Linear Alg. Appl. 2004]. This mean is the limit of a sequence which converges superlinearly with convergence of order 3 whereas the mean introduced by Ando, Li and Mathias is the limit of a sequence having order of convergence 1. This makes this new mean very easily computable. We provide a geometric interpretation and a generalization which includes as special cases our mean and the Ando-Li-Mathias mean.

1. INTRODUCTION

In several contexts, it is natural to generalize the geometric mean of two positive real numbers $a \# b := \sqrt{ab}$ to real symmetric positive definite $n \times n$ matrices as

$$(1.1) \qquad A \# B := A(A^{-1}B)^{1/2} = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}.$$

Several papers, e.g. [3, 4, 9], and a chapter of the book [2], are devoted to studying the geometry of the cone of positive definite matrices $\mathbb{P}^n$ endowed with the Riemannian metric defined by

$$ds = \left\| A^{-1/2}dAA^{-1/2} \right\|,$$

where $\|B\| = \sqrt{\sum_{i,j}|b_{i,j}|^2}$ denotes the Frobenius norm. The distance induced by this metric is

$$(1.2) \qquad d(A,B) = ||\log(A^{-1/2}BA^{-1/2})||$$

It turns out that on this manifold the geodesic joining X and Y has equation

$$\gamma(t) = X^{1/2}(X^{-1/2}YX^{-1/2})^t X^{1/2} = X(X^{-1}Y)^t =: X \#_t Y, \ t \in [0,1],$$

and thus $A \# B$ is the midpoint of the geodesic joining A and B. An analysis of numerical methods for computing the geometric mean of two matrices is carried out in [7].

It is less clear how to define the geometric mean of more than two matrices. In the seminal work [1], Ando, Li and Mathias list ten properties that a "good" matrix geometric mean should satisfy, and show that several natural approaches based on generalization of formulas working for the scalar case, or for the case of two matrices, do not work well. They propose a new definition of mean of k

matrices satisfying all the requested properties. We refer to this mean as to the Ando-Li-Mathias mean, or shortly ALM-mean.

The ALM-mean is the limit of a recursive iteration process where at each step of the iteration k geometric means of $k-1$ matrices must be computed. One of the main drawback of this iteration is its linear convergence. In fact, the large number of iterations needed to approximate each geometric mean at all the recursive steps makes it quite expensive to actually compute the ALM-mean with this algorithm. Moreover, no other algorithms endowed with a higher efficiency are known.

A class of geometric means satisfying the Ando, Li, Mathias requirements has been introduced in [8]. These means are defined in terms of the solution of certain matrix equations. This approach provides interesting theoretical properties concerning the means but no effective tools for their computation.

In this paper, we propose a new matrix geometric mean satisfying the ten properties of Ando, Li and Mathias. Likewise the ALM-mean, our mean is defined as the limit of an iteration process with the relevant difference that convergence is superlinear with order of convergence at least three. This property makes it much less expensive to compute this geometric mean since the number of iterations required to reach a high accuracy is dropped down to just a few ones.

The iteration on which our mean is based has a simple geometrical interpretation. In the case $k = 3$, given the positive definite matrices A_1, A_2, A_3, we generate three matrix sequences $A_1^{(m)}, A_2^{(m)}, A_3^{(m)}$ starting from $A_i^{(0)} = A_i$, $i = 1, 2, 3$. At the step $m+1$, the matrix $A_1^{(m+1)}$ is chosen along the geodesic which connects $A_1^{(m)}$ with the midpoint of the geodesic connecting $A_2^{(m)}$ to $A_3^{(m)}$ at distance $2/3$ from $A_1^{(m)}$. The matrices $A_2^{(m+1)}$ and $A_3^{(m+1)}$ are similarly defined. In the case of the Euclidean geometry, just one step of the iteration provides the value of the limit, i.e., the centroid of the triangle with vertices $A_1^{(m)}, A_2^{(m)}, A_3^{(m)}$. In fact, in a triangle the medians intersect each other at $2/3$ of their length. In the different geometry of the cone of positive definite matrices the geodesics which play the role of the medians, might even not intersect each other.

In the case of k matrices $A_1, A_2, \ldots, A_k$, the matrix $A_i^{(m+1)}$ is chosen along the geodesic which connects $A_i^{(m)}$ with the geometric mean of the remaining matrices, at distance $k/(k+1)$ from $A_i^{(m)}$. In the customary geometry, this point is the common intersection point of all the "medians" of the k-dimensional simplex formed by all the matrices $A_i^{(m)}$, $i = 1, \ldots, k$. We prove that the sequences $\{A_i^{(m)}\}_m$, $i = 1, \ldots, k$, converge to a common limit $\bar{A}$ with order of convergence at least 3. The limit $\bar{A}$ is our definition of geometric mean of $A_1, \ldots, A_k$.

It is interesting to point out that our mean and the ALM-mean of k matrices can be viewed as two specific instances of a class of more general means depending on $k-1$ parameters $s_i \in [0, 1]$, $i = 1, \ldots, k-1$. All the means of this class satisfy the requirements of Ando, Li and Mathias, moreover, the ALM-mean is obtained with $\boldsymbol{s} = (1/2, 1, 1, \ldots, 1)$, for $\boldsymbol{s} = (s_i)$, while our mean is obtained with $\boldsymbol{s} = ((k-1)/k, (k-2)/(k-1), \ldots, 1/2)$. The new mean is the only one in this class for which the matrix sequences generated at each recursive step converge superlinearly.

The article is structured as follows. After this introduction, in Section 2 we present the ten Ando–Li–Mathias properties and briefly describe the ALM-mean; then, in Section 3, we propose our new definition of a matrix geometric mean and

prove some of its properties by giving also a geometrical interpretation; in Section 4 we provide a generalization which includes the ALM-mean and our mean as two special cases. Finally, in Section 5 we present some numerical experiments of explicit computations involving this means concerning some problems from Physics. It turns out that, in the case of six matrices, the speed up reached by our approach with respect to the ALM-mean is by a factor greater than 200. We also experimentally demonstrate that the ALM-mean is different, even though very close, from our mean. Finally, for $k = 3$ we provide a pictorial description of the parametric family of geometric means.

2. Known results

Throughout this section we use the positive semidefinite ordering defined by $A \geq B$ if $A - B$ is positive semidefinite. We denote by A^* the transpose conjugate of A.

2.1. **Ando–Li–Mathias properties for a matrix geometric mean.** Ando, Li and Mathias [1] proposed the following list of properties that a "good" geometric mean $G(\cdot)$ of three matrices should satisfy.

P1: Consistency with scalars. If A, B, C commute then $G(A, B, C) = (ABC)^{1/3}$.
P2: Joint homogeneity. $G(\alpha A, \beta B, \gamma C) = (\alpha\beta\gamma)^{1/3} G(A, B, C)$.
P3: Permutation invariance. For any permutation $\pi(A, B, C)$ of A, B, C it holds $G(A, B, C) = G(\pi(A, B, C))$.
P4: Monotonicity. If $A \geq A'$, $B \geq B'$, $C \geq C'$, then $G(A, B, C) \geq G(A', B', C')$.
P5: Continuity from above. If A_n, B_n, C_n are monotonic decreasing sequences converging to A, B, C, respectively, then $G(A_n, B_n, C_n)$ converges to $G(A, B, C)$.
P6: Congruence invariance. For any nonsingular S, $G(S^*AS, S^*BS, S^*CS) = S^*G(A, B, C)S$.
P7: Joint concavity. If $A = \lambda A_1 + (1 - \lambda)A_2$, $B = \lambda B_1 + (1 - \lambda)B_2$, $C = \lambda C_1 + (1 - \lambda)C_2$, then $G(A, B, C) \geq \lambda G(A_1, B_1, C_1) + (1 - \lambda)G(A_2, B_2, C_2)$.
P8: Self-duality. $G(A, B, C)^{-1} = G(A^{-1}, B^{-1}, C^{-1})$.
P9: Determinant identity. $\det G(A, B, C) = (\det A \det B \det C)^{1/3}$.
P10: Arithmetic–geometric–harmonic mean inequality:

$$\frac{A + B + C}{3} \geq G(A, B, C) \geq \left(\frac{A^{-1} + B^{-1} + C^{-1}}{3}\right)^{-1}.$$

Moreover, it is proved in [1] that P5 and P10 are consequences of the others. Notice that all these properties can be easily generalized to the mean of any number of matrices. We will call a *geometric mean* of three or more matrices any map $G(\cdot)$ satisfying P1–P10 or their analogous for a number $k \neq 3$ of entries.

2.2. **The Ando–Li–Mathias mean.** Here and hereafter, we use the following notation. We denote by $G_2(A, B)$ the usual geometric mean $A \# B$ and, given the k-tuple $(A_1, \ldots, A_k)$, we define

$$\mathcal{Z}_i(A_1, \ldots, A_k) = (A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_k), \ i = 1, \ldots, k,$$

that is, the k-tuple where the i-th term has been dropped out.

In [1], Ando, Li and Mathias note that the previously proposed definitions of means of more than two matrices do not satisfy all the properties P1–P10, and propose a new definition that fulfills all of them. Their mean is defined inductively on the number of arguments k.

Given $A_1, \ldots, A_k$ positive definite, and given the definition of a geometric mean $G_{k-1}(\cdot)$ of $k-1$ matrices, they set $A_i^{(0)} = A_i$, $i = 1, \ldots, k$, and define for $r \geq 0$

$$(2.1) \qquad A_i^{(r+1)} := G_{k-1}(\mathcal{Z}_i(A_1^{(r)}, \ldots, A_k^{(r)})), \; i = 1, \ldots, k.$$

For $k = 3$, the iteration reads

$$\begin{bmatrix} A^{(r+1)} \\ B^{(r+1)} \\ C^{(r+1)} \end{bmatrix} = \begin{bmatrix} G_2(B^{(r)}, C^{(r)}) \\ G_2(A^{(r)}, C^{(r)}) \\ G_2(A^{(r)}, B^{(r)}) \end{bmatrix}.$$

Ando, Li and Mathias show that the k sequences $(A_i^{(r)})_{r=1}^{\infty}$ converge to the same matrix $\tilde{A}$, and finally define $G_k(A_1, \ldots, A_k) = \tilde{A}$. In the following, we shall denote by $G(\cdot)$ the Ando–Li–Mathias mean, dropping the subscript k when not essential.

An additional property of the Ando–Li–Mathias mean which will turn out to be important in the convergence proof is the following. Recall that $\rho(X)$ denotes the spectral radius of X, and let

$$R(A, B) := \max(\rho(A^{-1}B), \rho(B^{-1}A)).$$

This function is a *multiplicative metric*, that is, we have $R(A, B) \geq 1$ with equality iff $A = B$, and

$$R(A, C) \leq R(A, B)R(B, C).$$

The additional property is

P11: For each $k \geq 2$, and for each pair of sequences $(A_1, \ldots, A_k)$, $(B_1, \ldots, B_k)$ it holds

$$R\left(G(A_1, \ldots, A_k), G(B_1, \ldots, B_k)\right) \leq \left(\prod_{i=1}^{k} R(A_i, B_i)\right)^{1/k}.$$

3. A NEW MATRIX GEOMETRIC MEAN

3.1. **Definition.** We are going to define for each $k \geq 2$ a new mean $\bar{G}_k(\cdot)$ of k matrices satisfying P1–P11. Let $\bar{G}_2(A, B) = A \# B$, and suppose that the mean has already been defined for up to $k-1$ matrices. Let us denote shortly $T_i^{(r)} = \bar{G}_{k-1}(\mathcal{Z}_i(\bar{A}_1^{(r)}, \ldots, \bar{A}_k^{(r)}))$ and define $\bar{A}_i^{(r+1)}$ for $i = 1, \ldots, k$ as

$$(3.1) \qquad \bar{A}_i^{(r+1)} := \bar{G}_k(\bar{A}_i^{(r)}, \underbrace{T_i^{(r)}, T_i^{(r)}, \ldots, T_i^{(r)}}_{k-1 \text{ times}}),$$

with $\bar{A}_i^{(0)} = A_i$ for all i. Notice that apparently this needs the mean $\bar{G}_k(\cdot)$ to be already defined; in fact, in the special case in which $k-1$ of the k arguments are coincident, the properties P1 and P6 alone allow one to determine the common value of any geometric mean:

$$G(X, Y, Y, \ldots, Y) = X^{1/2}G(I, X^{-1/2}YX^{-1/2}, \ldots, X^{-1/2}YX^{-1/2})X^{1/2}$$

$$= X^{1/2}(X^{-1/2}YX^{-1/2})^{\frac{k-1}{k}}X^{1/2} = X \#_{\frac{k-1}{k}} Y.$$

Thus we can use this simpler expression directly in (3.1) and set

$$(3.2) \qquad \bar{A}_i^{(r+1)} = \bar{A}_i^{(r)} \#_{\frac{k-1}{k}} T_i^{(r)}.$$

In sections 3.3 and 3.4, we are going to prove that the k sequences $(\bar{A}_i^{(r)})_{r=1}^{\infty}$ converge to a common limit $\bar{A}$ with order of convergence at least three, and this will enable us to define $\bar{G}_k(A_1, \ldots, A_k) := \bar{A}$. In the following, we will drop the index k from $\bar{G}_k(\cdot)$ when it can be easily inferred from the context.

3.2. **Geometrical interpretation.** In [3], an interesting geometrical interpretation of the Ando–Li–Mathias mean is proposed for $k = 3$. We propose an interpretation of the new mean $\bar{G}(\cdot)$ in the same spirit. For $k = 3$, the iteration defining $\bar{G}(\cdot)$ reads

$$\begin{bmatrix} \bar{A}^{(r+1)} \\ \bar{B}^{(r+1)} \\ \bar{C}^{(r+1)} \end{bmatrix} = \begin{bmatrix} \bar{A}^{(r)} \#_{\frac{2}{3}} (\bar{B}^{(r)} \# \bar{C}^{(r)}) \\ \bar{B}^{(r)} \#_{\frac{2}{3}} (\bar{A}^{(r)} \# \bar{C}^{(r)}) \\ \bar{C}^{(r)} \#_{\frac{2}{3}} (\bar{A}^{(r)} \# \bar{B}^{(r)}) \end{bmatrix}.$$

We can interpret this iteration as a geometrical construction in the following way. To find e.g. $\bar{A}^{(r+1)}$, the algorithm is:

(1) Draw the geodesic joining $\bar{B}^{(r)}$ and $\bar{C}^{(r)}$, and take its midpoint $M^{(r)}$;
(2) Draw the geodesic joining $\bar{A}^{(r)}$ and $M^{(r)}$, and take the point lying at 2/3 of its length: this is $\bar{A}^{(r+1)}$.

If we execute the same algorithm on the Euclidean plane, replacing the word "geodesic" with "straight line segment", it turns out that $\bar{A}^{(1)}$, $\bar{B}^{(1)}$, and $\bar{C}^{(1)}$ coincide in the centroid of the triangle with vertices A, B, C. Thus, unlike the Euclidean counterpart of the Ando–Li–Mathias mean, this process converges in one step on the plane. Roughly speaking, when A, B and C are very close to each other, we can approximate (in some intuitive sense) the geometry on the Riemannian manifold $\mathbb{P}^n$ with the geometry on the Euclidean plane: since this construction to find the centroid of a plane triangle converges faster than the Ando–Li–Mathias one, we can expect that also the convergence speed of the resulting algorithm is faster. This is indeed what will result after a more accurate convergence analysis.

3.3. **Global convergence and properties P1–P11.** In order to prove that the iteration (3.2) is convergent (and thus that $\bar{G}(\cdot)$ is well defined), we are going to adapt a part of the proof of Theorem 3.2 of [1] (namely, Argument 1).

Theorem 3.1. *Let $A_1, \ldots A_k$ be positive definite.*

(1) *All the sequences $(\bar{A}_i^{(r)})_{r=1}^{\infty}$ converge for $r \to \infty$ to a common limit $\bar{A}$;*
(2) *the function $\bar{G}_k(A_1, \ldots, A_k)$ satisfies P1–P11.*

Proof. We work by induction on k. For $k = 2$, our mean coincides with the ALM-mean, so all the required work has been done in [1]. Let us now suppose that the thesis holds for all $k' \leq k - 1$. We have

$$\bar{A}_i^{(r+1)} \leq \frac{1}{k} \left(\bar{A}_i^{(r)} + (k-1)T_i^{(r)} \right) \leq \frac{1}{k} \sum_{i=1}^{k} \bar{A}_i^{(r)},$$

where the first inequality follows from P10 for the ALM-mean $G_k(\cdot)$ (remember that in the special case in which $k - 1$ of the arguments coincide, $G_k(\cdot) = \bar{G}_k(\cdot)$),

and the second from P10 for $\bar{G}_{k-1}(\cdot)$. Thus,

$$(3.3) \qquad \sum_{i=1}^{k} \bar{A}_i^{(r+1)} \leq \sum_{i=1}^{k} \bar{A}_i^{(r)} \leq \sum_{i=1}^{k} A_i.$$

Therefore, the sequence $(\bar{A}_1^{(r)}, \ldots, \bar{A}_k^{(r)})_{r=1}^{\infty}$ is bounded, and there must be a converging subsequence, say, converging to $(\bar{A}_1, \ldots, \bar{A}_k)$.

Moreover, for each $p, q \in \{1, \ldots, k\}$ we have

$$R(\bar{A}_p^{(r+1)}, \bar{A}_q^{(r+1)}) \leq R(\bar{A}_p^{(r)}, \bar{A}_q^{(r)})^{1/k} R(T_p^{(r)}, T_q^{(r)})^{\frac{k-1}{k}}$$
$$\leq R(\bar{A}_p^{(r)}, \bar{A}_q^{(r)})^{1/k} (R(\bar{A}_q^{(r)}, \bar{A}_p^{(r)})^{\frac{1}{k-1}})^{\frac{k-1}{k}} = R(\bar{A}_p^{(r)}, \bar{A}_q^{(r)})^{2/k},$$

where the first inequality follows from P11 in the special case, and the second from P11 in the inductive hypothesis. Passing at the limit of the converging subsequence, one can verify that

$$R(\bar{A}_p, \bar{A}_q) \leq R(\bar{A}_p, \bar{A}_q)^{2/k},$$

from which we get $R(\bar{A}_p, \bar{A}_q) \leq 1$, that is, $\bar{A}_p = \bar{A}_q$, because of the properties of R, i.e., the limit of the subsequence is in the form $(\bar{A}, \bar{A}, \ldots, \bar{A})$. Suppose there is another subsequence converging to $(\bar{B}, \bar{B}, \ldots, \bar{B})$; then, by (3.3), we have both $k\bar{A} \leq k\bar{B}$ and $k\bar{B} \leq k\bar{A}$, that is, $\bar{A} = \bar{B}$. Therefore, the sequence has only one limit point, thus it is convergent. This proves the first point of the theorem.

We now turn to show that P11 holds for our mean $\bar{G}_k(\cdot)$. Consider k-tuples $A_1, \ldots, A_k$ and $B_1, \ldots, B_k$, and let $\bar{B}_i^{(r)}$ be defined as $\bar{A}_i^{(r)}$ but starting the iteration from the k-tuple (B_i) instead of (A_i). We have for each i

$$R(\bar{A}_i^{(r+1)}, \bar{B}_i^{(r+1)}) \leq$$
$$\leq R(\bar{A}_i^{(r)}, \bar{B}_i^{(r)})^{1/k} R(\bar{G}(\mathcal{Z}_i(\bar{A}_1^{(r)}, \ldots, \bar{A}_k^{(r)})), \bar{G}(\mathcal{Z}_i(\bar{B}_1^{(r)}, \ldots, \bar{B}_k^{(r)})))^{\frac{k-1}{k}}$$
$$\leq R(\bar{A}_i^{(r)}, \bar{B}_i^{(r)})^{1/k} \left(\prod_{j \neq i} R(\bar{A}_j^{(r)}, \bar{B}_j^{(r)})^{\frac{1}{k-1}} \right)^{\frac{k-1}{k}}$$
$$= \prod_{j=1}^{k} R(\bar{A}_j^{(r)}, \bar{B}_j^{(r)})^{1/k}.$$

This yields

$$\prod_{i=1}^{k} R(\bar{A}_i^{(r+1)}, \bar{B}_i^{(r+1)}) \leq \prod_{i=1}^{k} R(\bar{A}_i^{(r)}, \bar{B}_i^{(r)});$$

chaining together these inequalities for successive values of r and passing to the limit, we get

$$R(G(A_1, \ldots, A_k), G(B_1, \ldots, B_k))^k \leq \prod_{i=1}^{k} R(A_i, B_i),$$

which is P11.

The other properties P1–P4 and P6–P9 (remember that P5 and P10 are consequences of these) are not difficult to prove. All the proofs are quite similar, and can be established by induction, using also the fact that since they hold for the ALM-mean, they can be applied to the mean $\bar{G}(\cdot)$ appearing in (3.2) (since we just proved that all possible geometric means take the same value if applied with $k-1$

equal arguments). For the sake of brevity, we provide only the proof for three of these properties.

P1: We need to prove that if the A_i commute then $\bar{G}(A_1, \ldots, A_k) = (A_1 \cdots A_k)^{1/k}$.

Using the inductive hypothesis, we have $T_i^{(1)} = \prod_{j \neq i} \bar{A}_i^{\frac{1}{k-1}}$. Using the fact that P1 holds for the ALM-mean, we have

$$\bar{A}_i^{(1)} = A_i^{1/k} \left(\prod_{j \neq i} A_j^{\frac{1}{k-1}} \right)^{\frac{k-1}{k}} = \prod_{i=1}^{k} A_i^{1/k},$$

as needed. So, from the second iteration on, we have $\bar{A}_1^{(r)} = \bar{A}_2^{(r)} = \cdots = \bar{A}_k^{(r)} = \prod_{i=1}^{k} A_i^{1/k}$.

P4: Let $T_i'^{(r)}$ and $\bar{A}_i'^{(r)}$ be defined as $T_i^{(r)}$ and $\bar{A}_i^{(r)}$ but starting from $A_i' \leq A_i$. Using monotonicity in the inductive case and in the ALM-mean, we have for each $s \leq 1$ and for each i

$$T_i^{(r+1)} \leq T_i'^{(r+1)}$$

and thus

$$\bar{A}_i^{(r+1)} \leq \bar{A}_i'^{(r+1)}.$$

Passing to the limit for $r \to \infty$, we obtain P4.

P7: Suppose $A_i = \lambda A_i' + (1 - \lambda) A_i''$, and let $T_i'^{(r)}$, (resp. $T_i''^{(r)}$) and $\bar{A}_i'^{(r)}$, (resp. $\bar{A}_i''^{(r)}$) be defined as $T_i^{(r)}$ and $\bar{A}_i^{(r)}$ but starting from A_i' (resp. A_i''). Suppose that for some r it holds $\bar{A}_i^{(r)} \geq \lambda \bar{A}_i'^{(r)} + (1 - \lambda) \bar{A}_i''^{(r)}$ for all i. Then by joint concavity and monotonicity in the inductive case we have

$$\begin{aligned}
T_i^{(r+1)} =& \bar{G}(\mathcal{Z}_i(\bar{A}_1^{(r)}, \ldots, \bar{A}_k^{(r)})) \\
\geq& \bar{G}(\mathcal{Z}_i(\lambda \bar{A}_1'^{(r)} + (1 - \lambda) \bar{A}_1''^{(r)}, \ldots, \lambda \bar{A}_k'^{(r)} + (1 - \lambda) \bar{A}_k''^{(r)})) \\
\geq& \lambda T_i'^{(r)} + (1 - \lambda) T_i''^{(r)},
\end{aligned}$$

and by joint concavity and monotonicity of the Ando–Li–Mathias mean we have

$$\begin{aligned}
\bar{A}_i^{(r+1)} =& \bar{A}_i^{(r)} \#_{\frac{k-1}{k}} T_i^{(r)} \\
\geq& \left(\lambda \bar{A}_i'^{(r)} + (1 - \lambda) \bar{A}_i''^{(r)} \right) \#_{\frac{k-1}{k}} \left(\lambda T_i'^{(r)} + (1 - \lambda) T_i''^{(r)} \right) \\
\geq& \lambda \bar{A}_i'^{(r+1)} + (1 - \lambda) \bar{A}_i''^{(r+1)}.
\end{aligned}$$

Passing to the limit for $r \to \infty$, we obtain P7.

$\square$

3.4. Cubic convergence.

In this section, we will use the big-O notation in the norm sense, that is, we will write $X = Y + O(\varepsilon^h)$ to denote that there are universal positive constants $\varepsilon_0 < 1$ and θ such that for each $0 < \varepsilon < \varepsilon_0$ it holds $\|X - Y\| \leq \theta \varepsilon^h$. The usual arithmetic rules involving this notation hold. In the following, these constants may depend on k, but not on the specific choice of the matrices involved in the formulas.

Theorem 3.2. *Let $0 < \varepsilon < 1$, M and $\bar{A}_i^{(0)} = A_i$, $i = 1, \ldots, k$, be positive definite $n \times n$, and $E_i := M^{-1}A_i - I$. Suppose that $\|E_i\| \le \varepsilon$ for all $i = 1, \ldots, k$. Then, for the matrices $\bar{A}_i^{(1)}$ defined in (3.2) the following hold.*

C1: *We have*

$$(3.4) \qquad M^{-1}\bar{A}_i^{(1)} - I = T_k + O(\varepsilon^3)$$

where

$$T_k := \frac{1}{k}\sum_{j=1}^{k} E_j - \frac{1}{4k^2}\sum_{i,j=1}^{k}(E_i - E_j)^2.$$

C2: *There are positive constants θ, σ and $\bar{\varepsilon} < 1$ (all of which may depend on k) such that for all $\varepsilon \le \bar{\varepsilon}$ it holds*

$$\left\|M_1^{-1}\bar{A}_i^{(1)} - I\right\| \le \theta\varepsilon^3$$

for a suitable matrix M_1 satisfying $\left\|M^{-1}M_1 - I\right\| \le \sigma\varepsilon$.

C3: *The iteration (3.2) converges at least cubically.*

C4: *We have*

$$(3.5) \qquad M_1^{-1}\bar{G}(A_1, \ldots, A_k) - I = O(\varepsilon^3).$$

Proof. Let us first find a local expansion of a generic point on the geodesic $A \#_t B$: let $M^{-1}A = I + F_1$ and $M^{-1}B = I + F_2$ with $\|F_1\| \le \delta$, $\|F_2\| \le \delta$, $0 < \delta < 1$. Then we have

$$
\begin{aligned}
M^{-1}(A \#_t B) &= M^{-1}A(A^{-1}B)^t = (I + F_1)\left((I + F_1)^{-1}(I + F_2)\right)^t \\
&= (I + F_1)\left((I - F_1 + F_1^2 + O(\delta^3))(I + F_2)\right)^t \\
&= (I + F_1)\left(I + F_2 - F_1 - F_1 F_2 + F_1^2 + O(\delta^3)\right)^t \\
(3.6) \qquad &= (I + F_1)\Big(I + t(F_2 - F_1 - F_1 F_2 + F_1^2) \\
&\qquad + \frac{t(t-1)}{2}(F_2 - F_1)^2 + O(\delta^3)\Big) \\
&= I + (1 - t)F_1 + tF_2 + \frac{t(t-1)}{2}(F_2 - F_1)^2 + O(\delta^3),
\end{aligned}
$$

where we made use of the matrix series expansion $(I + X)^t = I + tX + \frac{t(t-1)}{2}X^2 + O(X^3)$.

Now, we are going to prove the theorem by induction on k in the following way. Let Ci_k denote the assertion Ci of the theorem (for $i = 1, \ldots 4$) for a given value of k. We show that

 (1) C1_2 holds;
 (2) $\text{C1}_k \implies \text{C2}_k$;
 (3) $\text{C2}_k \implies \text{C3}_k, \text{C4}_k$;
 (4) $\text{C4}_k \implies \text{C1}_{k+1}$.

It is clear that these claims imply that the results C1—C4 hold for all $k \ge 2$ by induction; we will now turn to prove them one by one.

 (1) This is simply equation (3.6) for $t = \frac{1}{2}$.
 (2) It is obvious that $T_k = O(\varepsilon)$; thus, choosing $M_1 := M(I + T_k)$ one has

$$(3.7) \quad \bar{A}_i^{(1)} = M(I + T_k + O(\varepsilon^3)) = M_1(I + (I + T_k)^{-1}O(\varepsilon^3)) = M_1(I + O(\varepsilon^3)).$$

Using explicit constants in the big-O estimates, we get

$$\left\| M_1^{-1}\bar{A}_i^{(1)} - I \right\| \le \theta\varepsilon^3, \quad \left\| M^{-1}M_1 - I \right\| \le \sigma\varepsilon$$

for suitable constants θ and σ.

(3) Suppose ε is small enough to have $\theta\varepsilon^3 \le \varepsilon$. We shall apply C2 with initial matrices $\bar{A}_i^{(1)}$, with $\varepsilon_1 = \theta\varepsilon^3$ in lieu of ε and M_1 in lieu of M, getting

$$\left\| M_2^{-1}\bar{A}_i^{(2)} - I \right\| \le \theta\varepsilon_1^3, \quad \left\| M_1^{-1}M_2 - I \right\| \le \sigma\varepsilon_1.$$

Repeating again for all the steps of our iterative process, we get for all $s = 1, 2, \ldots$

(3.8) $$\left\| M_s^{-1}\bar{A}_i^{(s)} - I \right\| \le \theta\varepsilon_{s-1}^3 = \varepsilon_s, \quad \left\| M_s^{-1}M_{s+1} - I \right\| \le \sigma\varepsilon_s$$

with $\varepsilon_{s+1} := \theta\varepsilon_s^3$ and $M_0 := M$.

For simplicity's sake, we introduce the notation

$$d(X, Y) := \left\| X^{-1}Y - I \right\|$$

for any two $n \times n$ symmetric positive definite matrices X and Y. It will be useful to notice that $\|X - Y\| \le \|X\| \left\| X^{-1}Y - I \right\| \le \|X\| \, d(X, Y)$ and

(3.9) $$\begin{aligned} d(X, Z) &= \left\| (X^{-1}Y - I)(Y^{-1}Z - I) + X^{-1}Y - I + Y^{-1}Z - I \right\| \\ &\le d(X, Y)d(Y, Z) + d(X, Y) + d(Y, Z). \end{aligned}$$

With this notation, we can restate (3.8) as

$$d(M_s, \bar{A}_i^{(s)}) \le \varepsilon_s, \quad d(M_s, M_{s+1}) \le \sigma\varepsilon_s.$$

We will now prove by induction that, for ε smaller than a fixed constant, it holds

(3.10) $$d(M_s, M_{s+t}) \le \left(2 - \frac{1}{2^t} \right)\sigma\varepsilon_s.$$

First of all, it holds for all $t \ge 1$

$$\varepsilon_{s+t} = \theta^{\frac{3^t-1}{2}}\varepsilon^{3^t},$$

which, for ε smaller than $\min(1/8, \theta^{-1})$, implies $\frac{\varepsilon_{s+t}}{\varepsilon_s} \le \varepsilon_s^{\frac{3^t-1}{2}} \le \varepsilon_s^t \le \frac{1}{2^{t+2}}$.

Now, using (3.9), and supposing additionally $\varepsilon \le \sigma^{-1}$, we have

$$\begin{aligned} d(M_s, M_{s+t+1}) &\le d(M_s, M_{s+t})d(M_{s+t}, M_{s+t+1}) \\ &\qquad + d(M_s, M_{s+t}) + d(M_{s+t}, M_{s+t+1}) \\ &\le \left(2 - \frac{1}{2^t} \right)\sigma\varepsilon_s + \sigma\varepsilon_s \left(\sigma\varepsilon_{s+t} + \frac{\varepsilon_{s+t}}{\varepsilon_s} \right) \\ &\le \left(2 - \frac{1}{2^t} \right)\sigma\varepsilon_s + \sigma\varepsilon_s \left(2\frac{\varepsilon_{s+t}}{\varepsilon_s} \right) \\ &\le \left(2 - \frac{1}{2^t} \right)\sigma\varepsilon_s + \sigma\varepsilon_s\frac{1}{2^{t+1}} = \left(2 - \frac{1}{2^{t+1}} \right)\sigma\varepsilon_s \end{aligned}$$

Thus, we have for each t

$$\| M_t - M \| \le \|M\| \left\| M^{-1}M_t - I \right\| \le 2\sigma \|M\| \varepsilon,$$

which implies $\|M_t\| \le 2\|M\|$ for all t. By a similar argument,

$$(3.11) \qquad \|M_{s+t} - M_s\| \le \|M_s\| \, d(M_{s+t}, M_s) \le 2\sigma \|M\| \, \varepsilon_s,$$

Due to the bounds already imposed on ε, the sequence ε_s tends monotonically to zero with cubic convergence rate; thus (M_t) is a Cauchy sequence and therefore converges. In the following, let M^* be its limit. The convergence rate is cubic, since passing to the limit (3.11) we get

$$\|M^* - M_s\| \le 2\sigma \|M\| \, \varepsilon_s.$$

Now, using the other relation in (3.8), we get

$$\left\| \bar{A}_i^{(s)} - M^* \right\| \le \left\| \bar{A}_i^{(s)} - M^s \right\| + \|M^* - M_s\|$$

$$\le 2\|M\| \, d(M_s, \bar{A}_i^{(s)}) + 2\sigma \|M\| \, \varepsilon_s$$

$$\le (2\sigma + 2)\|M\| \, \varepsilon_s,$$

that is, $\bar{A}_i^{(s)}$ converges with cubic convergence rate to M^*. Thus C3 is proved. By (3.9), (3.10), and (3.8), we have

$$d(M_1, \bar{A}_i^{(t)}) \le d(M_1, M_t) d(M_t, \bar{A}_i^{(t)}) + d(M_1, M_t) + d(M_t, \bar{A}_i^{(t)})$$

$$\le 2\sigma \varepsilon_1 \varepsilon_t + 2\sigma \varepsilon_1 + \varepsilon_t \le (4\sigma + 1)\varepsilon_1 = O(\varepsilon^3),$$

which is C4.

(4) Using C4$_k$ and (3.6) with $F_1 = E_{k+1}$, $F_2 = M^{-1}\bar{G}(A_1, \ldots, A_k) = T_k + O(\varepsilon^3)$, $\delta = 2k\varepsilon$, we have

$$M^{-1}\bar{A}_{k+1}^{(1)} = M^{-1}\left(A_{k+1} \#_{\frac{k}{k+1}} \bar{G}(A_1, \ldots, A_k) \right)$$

$$(3.12) \qquad = I + \frac{1}{k+1}E_{k+1} + \frac{k}{k+1}T_k$$

$$- \frac{k}{2(k+1)^2}\left(E_{k+1} - \frac{1}{k}\sum_{i=1}^{k} E_i \right)^2 + O(\varepsilon^3).$$

Observe that

$$T_k = \frac{1}{k}S_k + \frac{P_k - (k-1)Q_k}{2k^2}$$

where $S_k = \sum_{i=1}^{k} E_i$, $Q_k = \sum_{i=1}^{k} E_i^2$, $P_k = \sum_{i,j=1,\ i\ne j}^{k} E_i E_j$. Since $S_k^2 = P_k + Q_k$ and $S_{k+1} = S_k + E_{k+1}$, $Q_{k+1} = Q_k + E_{k+1}^2$, $P_{k+1} = P_k + E_{k+1}S_k + S_k E_{k+1}$, from (3.12) one finds that

$$M^{-1}\bar{A}_{k+1}^{(1)} = I + \frac{1}{k+1}S_{k+1} - \frac{k}{2(k+1)^2}Q_{k+1} + \frac{1}{2(k+1)^2}P_{k+1} + O(\varepsilon^3)$$

$$= I + T_{k+1} + O(\varepsilon^3).$$

Since the expression we found is symmetric with respect to the E_i, it follows that $\bar{A}_j^{(1)}$ has the same expansion for any j.

$\square$

Observe that Theorems 3.1 and 3.2 imply that the iteration (3.2) is globally convergent with order of convergence at least 3.

It is worth to point out that, in the case where the matrices A_i, $i = 1, \ldots, A_k$, commute each other, the iteration (3.2) converges in just one step, i.e., $\bar{A}_i^{(1)} = \bar{A}$

for any i. In the noncommutative general case, one has $\det(\bar{A}_i^{(s)}) = \det(\bar{A})$ for any i and for any $s \geq 1$, i.e., the determinant converges in one single step to the determinant of the matrix mean.

Our mean is different from the ALM-mean, as we will show with some numerical experiments in Section 5. In the next section 4 we prove that our mean and the ALM-mean belong to a general class of matrix geometric means, which depends on a set of $k - 1$ parameters.

4. A NEW CLASS OF MATRIX GEOMETRIC MEANS

In this section we introduce a new class of matrix means depending on a set of parameters $s_1, \ldots, s_{k-1}$ and show that the ALM-mean and our mean are two specific instances of this class.

For the sake of simplicity, we describe this generalization in the case of $k = 3$ matrices A, B, C. The case $k > 3$ is outlined. Here, the distance between two matrices is defined in (1.2).

For $k = 3$, the algorithm presented in Section 3 replaces the triple A, B, C with A', B', C' where A' is chosen in the geodesic connecting A with the midpoint of the geodesic connecting B and C, at distance $2/3$ from A, and similarly is made for B' and C'. In our generalization we use two parameters $s, t \in [0, 1]$. We consider the point $P_t = B\#_t C$ in the geodesic connecting B to C at distance t from B. Then we consider the geodesic connecting A to P_t and define A' the matrix on this geodesic at distance s from A. That is, we set $A' = A\#_s(B\#_t C)$. Similarly we do with B and C. This transformation is recursively repeated so that the matrix sequences $A^{(r)}$, $B^{(r)}$, $C^{(r)}$ are generated by means of

$$
\begin{aligned}
(4.1) \qquad & A^{(r+1)} = A^{(r)}\#_s(B^{(r)}\#_t C^{(r)}), \\
& B^{(r+1)} = B^{(r)}\#_s(C^{(r)}\#_t A^{(r)}), \quad r = 0, 1, \ldots \\
& C^{(r+1)} = C^{(r)}\#_s(A^{(r)}\#_t B^{(r)}),
\end{aligned}
$$

starting with $A^{(0)} = A$, $B^{(0)} = B$, $C^{(0)} = C$.

By following the same arguments of Section 3 it can be shown that the three sequences have a common limit $G_{s,t}$ for any $s, t \in [0, 1]$. Moreover, for $s = 1$, $t = 1/2$ one obtains the ALM-mean, i.e., $G = G_{1,\frac{1}{2}}$, while for $s = 2/3$, $t = 1/2$ the limit coincides with our mean, i.e., $\bar{G} = G_{\frac{2}{3},\frac{1}{2}}$. Moreover, it is possible to prove that for any $s, t \in [0, 1]$ the limit satisfies the conditions P1–P11 so that it can be considered a good geometric mean.

Concerning the convergence speed of the sequence generated by (4.1) we may perform a more accurate analysis. Assume that $A = M(I + E_1)$, $B = M(I + E_2)$, $C = M(I + E_3)$, where $\|E_i\| \leq \varepsilon < 1$, $i = 1, 2, 3$. Then, applying (3.6) in (4.1) yields

$$
A' \doteq M(I + (1-s)E_1 + s(1-t)E_2 + stE_3 + \tfrac{st(t-1)}{2}H_2^2 + \tfrac{s(s-1)}{2}(H_1 + tH_2)^2)
$$

$$
B' \doteq M(I + (1-s)E_2 + s(1-t)E_3 + stE_1 + \tfrac{st(t-1)}{2}H_3^2 + \tfrac{s(s-1)}{2}(H_2 + tH_3)^2)
$$

$$
C' \doteq M(I + (1-s)E_3 + s(1-t)E_1 + stE_2 + \tfrac{st(t-1)}{2}H_1^2 + \tfrac{s(s-1)}{2}(H_3 + tH_1)^2)
$$

where $\doteq$ denotes equality up to $O(\varepsilon^3)$ terms, with $H_1 = E_1 - E_2$, $H_2 = E_2 - E_3$, $H_3 = E_3 - E_1$. Whence we have $A' = M(I + E_1')$, $B' = M(I + E_2')$, $C' = $

$M(I + E_3')$, with

$$
\begin{bmatrix} E_1' \\ E_2' \\ E_3' \end{bmatrix} \doteq C(s,t) \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix} + \frac{st(t-1)}{2} \begin{bmatrix} H_2^2 \\ H_3^2 \\ H_1^2 \end{bmatrix} + \frac{s(s-1)}{2} \begin{bmatrix} (H_1 - tH_2)^2 \\ (H_2 - tH_3)^2 \\ (H_3 - tH_1)^2 \end{bmatrix}
$$

where

$$
C(s,t) = \begin{bmatrix} (1-s)I & s(1-t)I & stI \\ stI & (1-s)I & s(1-t)I \\ s(1-t)I & stI & (1-s)I \end{bmatrix}.
$$

Observe that the block circulant matrix $C(s,t)$ has eigenvalues $\lambda_1 = 1$, $\lambda_2 = (1 - \frac{3}{2}s) + \boldsymbol{i}\frac{\sqrt{3}}{2}s(2t-1)$, and $\lambda_3 = \bar{\lambda}_2$, with multiplicity n, where $\boldsymbol{i}^2 = -1$. Moreover, the pair $(s,t) = (2/3, 1/2)$ is the only one which yields $\lambda_2 = \lambda_3 = 0$. In fact $(2/3, 1/2)$ is the only pair which provides superlinear convergence. For the ALM-mean, where $t = 1/2$ and $s = 1$ it holds $|\lambda_2| = |\lambda_3| = 1/2$ which is the rate of convergence of the ALM iteration [1].

In the case of $k > 3$ matrices, given the $(k-1)$-tuple $(s_1, s_2, \ldots, s_{k-1})$ we may recursively define $G_{s_1,\ldots,s_{k-1}}(A_1, \ldots, A_k)$ as the common limit of the sequences generated by

$$
A_i^{(r+1)} = A_i^{(r)} \#_{s_1} G_{s_2,\ldots,s_{k-1}}(\mathcal{Z}_i(A_1^{(r)}, \ldots, A_k^{(r)})), \; i = 1, \ldots, k.
$$

Observe that with $(s_1, \ldots, s_{k-1}) = (1/2, 1, 1, \ldots, 1)$ one obtains the ALM-mean, while with $(s_1, \ldots, s_{k-1}) = ((k-1)/k, (k-2)/(k-1), \ldots, 1/2)$ one obtains the new mean introduced in Section 3.

5. Numerical experiments

We have implemented the two iterations converging to the ALM mean and to the newly defined geometric mean in Matlab, and run some numerical experiments on a quad-Xeon 2.8Ghz computer. To compute matrix square roots we used Matlab's built-in `sqrtm` function, while for p-th roots with $p > 2$ we used the `rootm` function in Nicholas Higham's Matrix Computation Toolbox [6]. To counter the loss of symmetry due to the accumulation of computational errors, we chose to discard the imaginary part of the computed roots.

The experiments have been performed on the same data set as the paper [9]. It consists of five sets each composed of four to six 6×6 positive definite matrices, corresponding to physical data from elasticity experiments conducted by Hearmon [5]. The matrices are composed of smaller diagonal blocks of sizes 1×1 up to 4×4, depending on the symmetries of the involved materials. Two to three significative digits are reported for each experiments.

We have computed both the ALM mean and the newly defined mean of these sets; as a stopping criterion for each computed mean, we chose

$$
\max_i \left| A_i^{(r+1)} - A_i^{(r)} \right| < \varepsilon,
$$

where $|X| := \max_{i,j} |X_{ij}|$, with $\varepsilon = 10^{-10}$. The CPU times, in seconds, are reported in Table 1. For four matrices, the speed gain is a factor of 20, and it increases even more for more than four matrices.

We then focused on Hearmon's second data set (ammonium dihydrogen phosphate), composed of four matrices. In Table 2, we reported the number of outer ($k = 4$) iterations needed and the average number of iterations needed to reach

Data set (number of matrices)	ALM mean	New mean
$NaClO_3$ (5)	230.0	1.30
Ammonium dihydrogen phosphate (4)	9.9	0.39
Potassium dihydrogen phosphate (4)	9.7	0.38
Quartz (6)	6700.0	30.00
Rochelle salt (4)	10.0	0.53

TABLE 1. CPU times in seconds for the Hearmon elasticity data

	ALM mean	New mean
Outer iterations	23	3
Avg. inner iterations	18.3	2
Matrix square roots (`sqrtm`)	5052	72
Matrix p-th roots (`rootm`)	0	84

TABLE 2. Number of inner and outer iterations needed, and number of matrix roots needed

convergence in the inner ($k = 3$) iterations (remember that the computation of a mean of four matrices requires the computation of three means of three matrices at each of its steps). Moreover, we measured the number of square and p-th roots needed by the two algorithms, since they are the most expensive operation in the algorithm. From the results, it is evident that the speed gain in the new mean is due not only to the reduction of the number of outer iterations, but also of the number of inner iterations needed to get convergence at each step of the inner mean calculations. When the number of involved matrices becomes larger, these speedups add up at each level.

Hearmon's elasticity data are not suitable to measure the accuracy of the algorithm, since the results to be obtained are not known. To measure the accuracy of the computed results, we computed instead $\left|G(A^4, I, I, I) - A\right|$, which should yield zero in exact arithmetic (due to P1), and its analogue with the new mean. We chose A to be the first matrix in Hearmon's second data set. Moreover, in order to obtain results closer to machine precision, in this experiment we changed the stopping criterion choosing $\varepsilon = 10^{-13}$

Operation	Result
$\left|G(A^4, I, I, I) - A\right|$	3.6E-13
$\left|\bar{G}(A^4, I, I, I) - A\right|$	1.8E-14

The results are well within the errors permitted by the stopping criterion, and show that both algorithms can reach a satisfying precision.

The following examples provide an experimental proof that our mean is different from the ALM-mean.

Consider the following matrices

$$A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}, \ B = \begin{bmatrix} a & -b \\ -b & a \end{bmatrix}, \ C = \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}.$$

Observe that the triple (A, B, C) is transformed into (B, A, C) under the map $X \to S^{-1}XS$, for $S = \mathrm{diag}(1, -1)$. In this way, any matrix mean $G(A, B, C)$

satisfying condition P3 is such that $G = S^{-1}GS$, that is, the off-diagonal entries of G are zero. Whence, G must be diagonal. With $a = 2, b = 1, c = 24$, for the ALM-mean G and our mean $\bar{G}$ one finds that

$$\bar{G} = \begin{bmatrix} 1.487443626 & 0 \\ 0 & 4.033766318 \end{bmatrix}, \; G = \begin{bmatrix} 1.485347837 & 0 \\ 0 & 4.039457861 \end{bmatrix},$$

where we reported the first 10 digits. Observe that the determinant of both the matrices is 6, that is, the geometric mean of $\det A, \det B, \det C$, moreover, $\rho(\bar{G}) < \rho(G)$.

For the matrices

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 3 & -2 \\ 0 & -2 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 2 \\ 0 & 2 & 2 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 10 & 0 \\ 1 & 0 & 50 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 10 & 0 \\ -1 & 0 & 50 \end{bmatrix},$$

one has

$$\bar{G} = \begin{bmatrix} 1.3481 & 0 & -0.3016 \\ 0 & 3.8452 & 0 \\ -0.3016 & 0 & 6.1068 \end{bmatrix}, \; G = \begin{bmatrix} 1.3472 & 0 & -0.3106 \\ 0 & 3.8796 & 0 \\ -0.3106 & 0 & 6.0611 \end{bmatrix}.$$

Their eigenvalues are $(6.1258, 3.8452, 1.3290)$, and $(6.0815, 3.8796, 1.3268)$, respectively. Observe that, unlike in the previous example, it holds $\rho(\bar{G}) > \rho(G)$.

In order to illustrate the properties of the set

$$\{G_{s,t} : \quad (s,t) \in (0,1] \times (0,1)\},$$

where $G_{s,t}$ is the mean of three matrices defined in Section 4, we considered the intervals $[1/15, 1], [1/15, 14/15]$ and discretized them into two sets $\mathcal{S}, \mathcal{T}$ of 15 equidistant points $\{1/15 = s_1 < s_2 < \cdots < s_{15} = 1\}$, $\{1/15 = t_1 < t_2 < \cdots < t_{15} = 14/15\}$, respectively. For each pair $(s_i, t_j) \in \mathcal{S} \times \mathcal{T}$, $i, j = 1, \ldots, 15$, we computed G_{s_i, t_j} and the orthogonal projection $(x(i,j), y(i,j), z(i,j))$ of the matrix $G_{s_i, t_j} - G_{\frac{2}{3}, \frac{1}{2}}$, over a three dimensional fixed randomly generated subspace. The set

$$\mathcal{V} = \{(x(i,j), y(i,j), z(i,j)) \in \mathbb{R}^3, \; i, j = 1, \ldots, 15\}$$

has been plotted with the Matlab command `mesh(x,y,z)` which connects each point of coordinates $(x(i,j), y(i,j), z(i,j))$ to its four neighborhoods of coordinates $(x(i + \delta, j + \gamma), y(i + \delta, j + \gamma), z(i + \delta, j + \gamma)))$ for $\delta, \gamma \in \{1, -1\}$.

Figure 1 displays the set $\mathcal{V}$ from six different points of view, where the matrices A, B and C of size 3, have been randomly generated. The set appears to be a flat surface with part of the edge tightly folded on itself. The geometric mean $G_{\frac{2}{3}, \frac{1}{2}}$ corresponds to the point of coordinates $(0, 0, 0)$ which is denoted by a small circle and seems to be located in the central part of the figure. These properties, reported for only one triple (A, B, C), are maintained with very light differences in all the plots that we have performed.

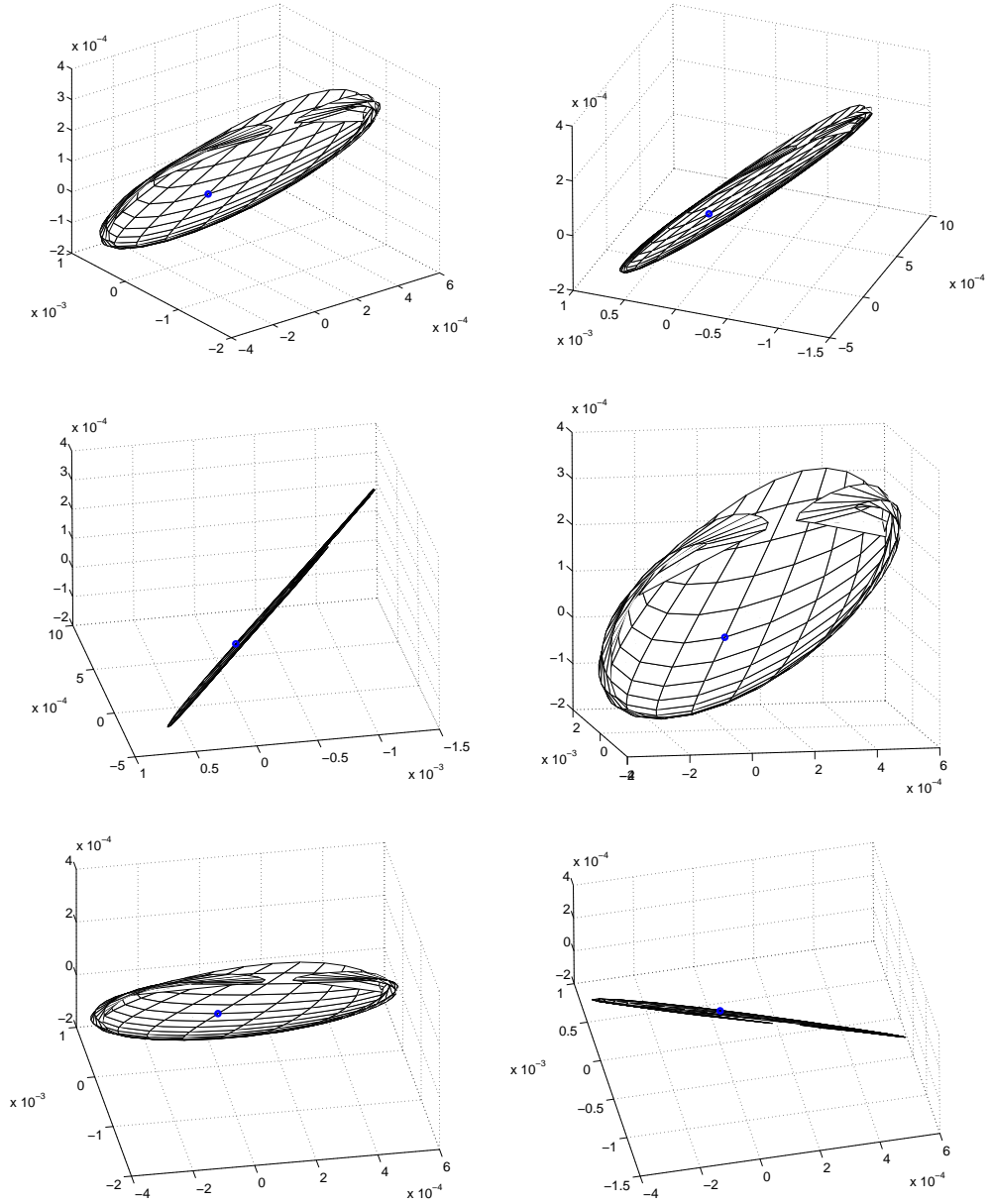The software concerning our experiments can be delivered upon request.

FIGURE 1. Plot of the set $\mathcal{V}$, the small circle corresponds to $G_{2/3,1/2}$.

References

1. T. Ando, Chi-Kwong Li, and Roy Mathias, *Geometric means*, Linear Algebra Appl. **385** (2004), 305–334. MR MR2063358 (2005f:47049)
2. Rajendra Bhatia, *Positive definite matrices*, Princeton Series in Applied Mathematics, Princeton University Press, Princeton, NJ, 2007. MR MR2284176 (2007k:15005)
3. Rajendra Bhatia and John Holbrook, *Noncommutative geometric means*, Math. Intelligencer **28** (2006), no. 1, 32–39. MR MR2202893 (2007g:47023)
4. _____ , *Riemannian geometry and matrix geometric means*, Linear Algebra Appl. **413** (2006), no. 2-3, 594–618. MR MR2198952 (2007c:15030)
5. R. F. S. Hearmon, *The elastic constants of piesoelectric crystals*, J. Appl. Phys. **3** (1952), 120–123.
6. Nicholas J. Higham, *The Matrix Computation Toolbox*, http://www.ma.man.ac.uk/~higham/mctoolbox.
7. Bruno Iannazzo and Beatrice Meini, *The matrix geometric mean and other matrix functions: a unifying framework*, Tech. report, Dipartimento di Matematica, Università di Pisa, 2009.
8. Yongdo Lim, *On Ando-Li-Mathias geometric mean equations*, Linear Algebra Appl. **428** (2008), no. 8-9, 1767–1777. MR MR2398117
9. Maher Moakher, *On the averaging of symmetric positive-definite tensors*, J. Elasticity **82** (2006), no. 3, 273–296. MR MR2231065 (2007a:74007)

Dipartimento di Matematica, Università di Pisa, Largo B. Pontecorvo 5, 56127 Pisa, Italy

E-mail address: `bini, meini@dm.unipi.it`

Scuola Normale Superiore, Piazza dei Cavalieri 6, 56126 Pisa, Italy

E-mail address: `poloni@sns.it`