*IMA Journal of Numerical Analysis* (2013) **33**, 1436–1468
doi:10.1093/imanum/drs049
Advance Access publication on April 4, 2013

# Approximation schemes for functions of positive-definite matrix values

NIR SHARON*
*School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel*
*Corresponding author: nirsharo@post.tau.ac.il

AND

URI ITAI
*Department of Mathematics, Technion – Israel Institute of Technology, Haifa 32000, Israel*
uritai@cs.technion.ac.il

[Received on 4 February 2012; revised on 11 November 2012]

In recent years, there has been an enormous interest in developing methods for the approximation of manifold-valued functions. In this paper, we focus on the manifold of symmetric positive-definite (SPD) matrices. We investigate the use of SPD-matrix means to adapt linear positive approximation methods to SPD-matrix-valued functions. Specifically, we adapt corner-cutting subdivision schemes and Bernstein operators. We present the concept of admissible matrix means and study the adapted approximation schemes based on them. Two important cases of admissible matrix means are treated in detail: the exp–log and the geometric matrix means. We derive special properties of the approximation schemes based on these means. The geometric mean is found to be superior in the sense of preserving more properties of the data, such as monotonicity and convexity. Furthermore, we give error bounds for the approximation of univariate SPD-matrix-valued functions by the adapted operators.

*Keywords*: approximation of matrix-valued functions; symmetric positive-definite matrices; matrix means for symmetric positive-definite matrices; corner-cutting subdivision schemes; Bernstein operators.

## 1. Introduction

The adaptation of approximation operators to manifold-valued data and the development of analysis methods for such schemes have attracted a lot of attention in recent years (Xie & Yu, 2007, 2010; Grohs, 2009, 2010b; Shingel, 2009; Dyn *et al.*, 2010). These generalizations provide important tools for application and initiated a rich mathematical theory that combines aspects of analysis, geometry and algebra. We focus on adapting approximation methods for the manifold of symmetric positive-definite (SPD) matrices.

As a main prototype for our adaptation process, we use corner-cutting subdivision schemes (de Boor, 1987). These schemes are considered to be the most basic nontrivial subdivision schemes. Nevertheless, those schemes have several important geometric properties and are popular in practice. A specific corner-cutting scheme is the Chaikin (1974) scheme, which is also a special case of the Lane & Riesenfeld (1980) schemes.

A real matrix $A$ is an SPD matrix if $A = A^*$ and $x^*Ax > 0$ for any nonzero vector $x$ (here $X^*$ stands for the standard transpose operator $X^*_{i,j} = X_{j,i}$). The class of SPD matrices is ubiquitous in science and engineering. SPD matrices have a unique spectral structure of real positive eigenvalues, orthogonal eigenvectors and more. In spite of the fact that these matrices are well studied, approximating SPD-matrix-valued functions within the class is still a challenge.

Positive-definite matrices can be considered as a natural generalization of positive numbers. Schaefer *et al.* (2008) adapted the Lane–Riesenfeld algorithm for positive numbers, by replacing any arithmetic average $\frac{1}{2}(a+b)$ by an average from the family of $p$-averages $(\frac{1}{2}(a^p+b^p))^{1/p}$ for $p \neq 0$ and the geometric mean $\sqrt{ab}$ for $p=0$. A similar approach for constructing nonlinear subdivision schemes by using nonlinear averages has been proved to be efficient also in the case of data on general manifolds (see e.g. Wallner & Dyn, 2005).

In this paper, we use the means of SPD matrices in order to define approximation operators for SPD-matrix-valued functions. We show that many of the algebraic and geometric properties of such schemes are derived from the properties of the matrix means. Hence, we introduce a class of 'good' matrix means which will be referred to as admissible means.

Special attention is given to two particular admissible means: the exp–log mean and the geometric mean for matrices. The first one gained popularity in recent years (Rahman *et al.*, 2005; Grohs & Wallner, 2008) due to the use of the Lie group structure. The second one was used by the authors for interpolation in Itai & Sharon (2012). For the generated functions by the corner-cutting subdivision schemes based on the latter, we prove monotonicity, order preservation, Schur and pinching properties. In fact, it is easy to extend this result to every positive subdivision scheme. Furthermore, we show by counterexamples that these properties do not hold for the exp–log mean.

A secondary prototype of a positive, linear, sampled-based operator is the class of Bernstein operators. We use the De Casteljau (1959) algorithm which evaluates Bernstein polynomials by repeated binary means, and we replace the means of numbers by matrix means.

This paper is organized as follows. We present fundamental definitions and notation in Section 2. In Section 3, we define matrix means and admissible matrix means and use them to adapt subdivision schemes to SPD-matrix data. In Section 4, we investigate subdivision schemes based on the exp–log mean. Section 5 studies corner-cutting subdivision schemes based on the geometric matrix mean. In Section 6, we present special properties of these subdivision schemes. We conclude this paper in Section 7, where we adapt Bernstein operators to SPD-matrix-valued functions.

## 2. Preliminaries

The results of this paper involve some elementary notation and definitions. We denote an SPD matrix $A$ by $A \succ 0$; if $A$ is only positive semidefinite, we denote it by $A \succeq 0$. We use the Löwner partial order (Hauke & Markiewicz, 1994) for symmetric matrices:

$$A \succeq B \text{ if and only if } A - B \succeq 0. \tag{2.1}$$

Löwner order is only partial since there exist SPD matrices $A$ and $B$ such that $A \not\succeq B$ and $B \not\succeq A$. For example,

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

For any $X \succ 0$, we denote by $\lambda_{\max}(X)$ and $\lambda_{\min}(X)$ the maximal and minimal eigenvalue of $X$, respectively. For two matrices $X, Y \succ 0$, we use the notation $\lambda_{\max}(X,Y) = \max\{\lambda_{\max}(X), \lambda_{\max}(Y)\}$ and similarly $\lambda_{\min}(X,Y) = \min\{\lambda_{\min}(X), \lambda_{\min}(Y)\}$. We also denote the cone of $n \times n$ symmetric positive-definite matrices by $\mathsf{SPD}(n)$.

We conclude this section with the following two definitions.

Definition 2.1 Let $F : \mathsf{SPD}(n) \to \mathsf{SPD}(n)$ be a bijective map. If for every $A, B \in \mathsf{SPD}(n)$, we have $A \succeq B$ if and only if $F(A) \succeq F(B)$, then $F$ is called an *order-preserving function*. Similarly, if $A \succeq B$ and only if $F(A) \preceq F(B)$, then $F$ is called an *order-reversing function.*

Definition 2.2 Let $A : \mathbb{R} \to \mathsf{SPD}(n)$ be a matrix-valued function satisfying

$$\lim_{\triangle x \to 0} \|A(x + \triangle x) - A(x)\| = 0, \quad x \in I.$$

Then $A(x)$ is called a *continuous-matrix curve*. If in addition, the limit

$$B(x) = \lim_{\triangle x \to 0} \frac{A(x + \triangle x) - A(x)}{\triangle x}, \quad x \in I$$

exists and $B(x)$ is continuous for every $x$, then $A(x)$ is called a *smooth matrix curve* ($\mathbf{C}^1$).

Unless otherwise stated, we use the Frobenius matrix norm (see Golub & Van Loan, 1996, Chapter 2)

$$\|A\| = \sqrt{\sum A_{i,j}^2} = \sqrt{\operatorname{tr}(AA^*)},$$

where $\operatorname{tr}(X) = \sum_i X_{i,i}$ is the standard trace operator on a matrix $X$. Some authors refer to this norm as the Hilbert–Schmidt norm. The use of the Frobenius norm is arbitrary due to norm equivalence. Most of the results of this paper can be generalized for an arbitrary norm.

## 3. SPD-matrix-valued subdivision schemes

Our adaptation of positive operators to SPD-matrix-valued functions is based on binary averages of SPD matrices.

### 3.1 *Matrix means*

We start by defining the notion of an SPD-matrix mean.

Definition 3.1 A function $M_t : \mathsf{SPD}(n) \times \mathsf{SPD}(n) \to \mathsf{SPD}(n)$, where $t \in [0, 1]$, is called an SPD-matrix binary mean rule for $\mathsf{SPD}(n)$, or more simply a matrix mean, if for any $A, B \succ 0$ we have the following properties.

(1) $M_t(A, A) = A$.

(2) $M_t(A, B) = M_{1-t}(B, A)$.

(3) $M_0(A, B) = A$.

(4) (a) $\lambda_{\max}(M_t(A, B)) \leqslant \lambda_{\max}(A, B)$.

  (b) $\lambda_{\min}(M_t(A, B)) \geqslant \lambda_{\min}(A, B)$.

Note that by using Definition 3.1(2,3), we get $M_1(A, B) = B$.

We introduce a relatively simple construction to produce a matrix mean. This construction, which generalizes the $p$-averages (Schaefer *et al.*, 2008), is termed the *quasi-arithmetic mean* and used in the solution of the Matkowski–Sutô problem (Daróczy *et al.*, 2006).

THEOREM 3.2 Let $F$ be an order-preserving or -reversing function. Then

$$\mathscr{M}_t^F(A,B) = F((1-t)F^{-1}(A) + tF^{-1}(B))$$

is a matrix mean.

*Proof.* Definition 3.1(1–3) are straightforward. Thus, it is enough to show property (4). For any SPD matrices $A$ and $B$,

$$A, B \preceq \lambda_{\max}(A,B)I.$$

Since $F$ is an order-preserving function,

$$(1-t)F^{-1}(A) + tF^{-1}(B) \preceq (1-t)F^{-1}(\lambda_{\max}(A,B)I) + tF^{-1}(\lambda_{\max}(A,B)I).$$

Hence,

$$F((1-t)F^{-1}(A) + tF^{-1}(B)) \preceq \lambda_{\max}(A,B)I.$$

Since $X \preceq Y$ entails that $\lambda_{\max}(X) \leqslant \lambda_{\max}(Y)$, we get

$$\lambda_{\max}(\mathscr{M}_t^F(A,B)) \leqslant \lambda_{\max}(A,B).$$

Similarly,

$$\lambda_{\min}(\mathscr{M}_t^F(A,B)) \geqslant \lambda_{\min}(A,B).$$

The proof for an order-reversing function is analogous. □

Matrix means are a useful tool for constructing approximation operators for matrix-valued functions. However, if one wishes to retain matrix properties such as inverse or determinant, additional conditions for the matrix mean are required. Thus, we have the following definition.

DEFINITION 3.3 Let $M_t$ be a matrix mean for SPD matrices. We call $M_t$ an *admissible matrix mean* if we have the following properties.

P1 Commutativity with the inverse: $M_t(A^{-1}, B^{-1}) = (M_t(A,B))^{-1}$.

P2 Invariance to orthogonal coordinate changes:

$$Q^*M_t(A,B)Q = M_t(Q^*AQ, Q^*BQ), \quad QQ^* = I.$$

P3 Incompressibility: if $\det(A) = \det(B) = 1$, then

$$\det(M_t(A,B)) = 1.$$

P4 Homogeneity:

$$M_t(\alpha A, \beta B) = M_t(\alpha,\beta)M_t(A,B), \quad \alpha,\beta > 0,$$

where $M_t(\alpha,\beta)$ is defined for the $1 \times 1$ (SPD) matrices $\alpha, \beta$.

Admissible matrix means are at the centre of this paper. However, there exist many interesting matrix means which do not agree with Definition 3.3. Next, we show two examples of matrix means for matrices which are not admissible.

EXAMPLE 3.4 We introduce two common matrix versions for the well-known harmonic and the arithmetic means.

(1) (Harmonic mean) $H_t(A,B) = ((1-t)A^{-1} + tB^{-1})^{-1}$. This mean is generated by $F(X) = X^{-1}$, which is an order-reversing function. The harmonic mean satisfies Definition 3.1. However, one can verify that it is not homogeneous. For example, let

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix};$$

then

$$H_{1/2}(A^{-1},B^{-1}) = \begin{pmatrix} \frac{2}{3} & 0 \\ 0 & \frac{4}{3} \end{pmatrix} \neq (H_{1/2}(A,B))^{-1} = \begin{pmatrix} \frac{3}{4} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}.$$

Namely, P1 does not hold. Furthermore, we have $\det(A) = 1$ and $\det(B) = 1$ but $\det(H_t(A,B)) \neq 1$. Thus, P3 does not hold.

(2) (Arithmetic mean) $L_t(A,B) = (1-t)A + tB$. This mean is generated by the identity function in the sense of Theorem 3.2. The commutativity of matrix addition implies that $L_t$ satisfies Definition 3.1. Nevertheless, P1 and P3 of Definition 3.3 are not satisfied; $A$ and $B$ introduced in the harmonic case illustrate this.

Example 3.4 can be generalized to the matrix analogue of the $p$-averages,

$$(\tfrac{1}{2}(a^p + b^p))^{1/p}, \quad a,b > 0, \quad p \neq 0.$$

The first and useful conclusion about admissible mean is as follows.

THEOREM 3.5 Let $M_t$ be an admissible matrix mean. Then, for any $A,B \succ 0$,

$$\det(M_t(A,B)) = M_t(\det(A),\det(B)).$$

*Proof.* Let $\alpha,\beta > 0$ and $n \in \mathbb{N}$. Reusing the homogeneity property,

$$M_t(\alpha^n,\beta^n) = M_t(\alpha,\beta)M_t(\alpha^{n-1},\beta^{n-1}) = \cdots = M_t(\alpha,\beta)^n. \tag{3.1}$$

For any nonsingular matrix $X$, $\det(\det(X)^{-1/n}X) = 1$. By the homogeneity property,

$$M_t(A,B) = M_t(\det(A)^{1/n},\det(B)^{1/n})M_t(A\det(A)^{-1/n},B\det(B)^{-1/n}).$$

Taking the determinant, using (3.1) and the incompressibility property conclude the proof. □

The homogeneity property of admissible means, by reducing the mean to scalars, yields (3.1). In many cases, this implies a common scalar version. In particular, we prove that any admissible matrix mean satisfies an additional condition: when reduced to the scalar case, it has the form of the standard geometric mean for numbers $\alpha^{1-t}\beta^t$. This condition is presented in the next lemma and is valid for the means introduced in Theorem 3.2.

LEMMA 3.6 Let $\mathscr{M}_t^F$ be a matrix mean, constructed as described in Theorem 3.2 with a corresponding order-preserving (or order-reversing) function $F$. Then

$$\mathscr{M}_{1/2}^F(\mathscr{M}_t^F(A,B),A) = \mathscr{M}_{t/2}^F(A,B), \quad A,B \succ 0, \ \ t \in [0,1]\,. \tag{3.2}$$

*Proof.* By the construction given in Theorem 3.2 and the correspondence function $F$,

$$\begin{aligned}
\mathscr{M}_{t/2}^F(A,B) &= F\left(\frac{t}{2}F^{-1}(B) + \left(1-\frac{t}{2}\right)F^{-1}(A)\right) \\
&= F\left(\frac{1}{2}F^{-1}(F((1-t)F^{-1}(A)+tF^{-1}(B))) + \frac{1}{2}F^{-1}(A)\right) \\
&= \mathscr{M}_{1/2}^F(F((1-t)F^{-1}(A)+tF^{-1}(B)),A) \\
&= \mathscr{M}_{1/2}^F(\mathscr{M}_t^F(A,B),A).
\end{aligned}$$

□

Note that (3.2) is a special case of the *quasi-linear* condition,

$$M_{1/2}(M_t(A,B),A) = M_{t/2}(A,B), \quad A,B \succ 0, \ \ t \in [0,1], \tag{3.3}$$

where $M_t$ is a matrix mean. This condition is true for a wider class of means than the one presented in Theorem 3.2. An example of such an admissible mean is provided in Section 5.

Matrix means that satisfy (3.3) are a generalization of the geometric scalar mean.

THEOREM 3.7 Let $M_t$ be a continuous, quasi-linear and admissible matrix mean for scalars, namely

$$M_{1/2}(M_t(\alpha,\beta),\alpha) = M_{t/2}(\alpha,\beta), \quad \alpha,\beta > 0, \ \ t \in [0,1].$$

Then

$$M_t(\alpha,\beta) = \alpha^{1-t}\beta^t, \quad \alpha,\beta > 0, \ \ t \in [0,1].$$

*Proof.* The homogeneity property P4 implies $M_t(\alpha,\beta) = M_t(\alpha,1)M_t(1,\beta)$. Thus, it is sufficient to show $M_t(1,\beta) = \beta^t$. We prove by induction that

$$M_{i/2^n}(1,\beta) = \beta^{i/2^n}, \quad i = 1,\ldots,2^n, \ \ n \in \mathbb{N}. \tag{3.4}$$

For $n = 1$, by Definition 3.1(1,2) and homogeneity,

$$\begin{aligned}
M_{1/2}(1,\beta) &= M_{1/2}(1,\sqrt{\beta})M_{1/2}(1,\sqrt{\beta}) \\
&= M_{1/2}(\sqrt{\beta},1)M_{1/2}(1,\sqrt{\beta}) \\
&= M_{1/2}(\sqrt{\beta},\sqrt{\beta}) = \sqrt{\beta}.
\end{aligned}$$

For the induction basis, let (3.4) be true for $n = m$; we prove it for $m+1$ in two steps. First, we consider the case $i = 1,\ldots,2^m$. By the quasi-linear property, the basis of the induction and the case $n = 1$, respectively,

$$\begin{aligned}
M_{i/2^{m+1}}(1,\beta) &= M_{1/2}(M_{i/2^m}(1,\beta),1) \\
&= M_{1/2}(\beta^{i/2^m},1) = \beta^{i/2^{m+1}}.
\end{aligned}$$

Second, we consider $i = 2^m + 1, \ldots, 2^{m+1}$. We use homogeneity, symmetry and the result of the first step:

$$\begin{aligned} M_{i/2^{m+1}}(1,\beta) &= M_{i/2^{m+1}}(\beta,\beta)M_{i/2^{m+1}}(\beta^{-1},1) \\ &= \beta M_{1-i/2^{m+1}}(1,\beta^{-1}) \\ &= \beta\beta^{i/2^{m+1}-1} = \beta^{i/2^{m+1}}. \end{aligned}$$

Note that in the second step, $2^{m+1} - i < 2^{m+1} - 2^m = 2^m$. Thus, (3.4) is valid, which implies that $M_t(1,\beta) = \beta^t$ holds for all the dyadic points of $[0,1]$. The continuity of $M_t$ ensures that $M_t(1,\beta) = \beta^t$ is true for all $t \in [0,1]$. □

By the above, we can rewrite Theorem 3.5 for the construction introduced in Theorem 3.2.

COROLLARY 3.8 Let $M_t$ be a continuous, quasi-linear admissible matrix mean. Then

$$\det(M_t(A,B)) = \det(A)^{1-t}\det(B)^t, \quad t \in [0,1].$$

A noteworthy fact of an admissible matrix mean $M_t$ is that for any given SPD matrix $A$, then $M_t(A, A^{-1}) = I$. The proof of this property is a direct result of P1 of Definition 3.3 and the fact that the only SPD matrix which is equal to its inverse is the identity.

### 3.2 *Adaptation of subdivision schemes to SPD-matrix-valued data*

In Section 3.1, we introduced matrix means. Here, we use them to construct approximating subdivision schemes for SPD matrices. Any such subdivision scheme consists of refinement operators. A refinement operator $\mathscr{S}$ of an approximation subdivision scheme for scalar initial data $\mathbf{P}_0 = \{p_{0,i}\}_{i\in\mathbb{Z}}$ is of the form

$$\begin{aligned} (\mathscr{S}(\mathbf{P}_0))_{2i} &= \sum_l a_{2l} p_{0,i-l}, \\ (\mathscr{S}(\mathbf{P}_0))_{2i+1} &= \sum_l a_{2l+1} p_{0,i-l}. \end{aligned} \tag{3.5}$$

We assume that the set $\{a_n \mid n \in \mathbb{Z}, a_n \neq 0\}$ is finite and consists of non-negative elements. A necessary condition for convergence (Dyn, 2006) is $\sum_i a_{2i} = \sum_i a_{2i+1} = 1$ and thus we can rewrite any refinement rule of the form (4.2) as (finite) repeated weighted averages (Wallner & Dyn, 2005, Theorem 1). Therefore, the adaptation for matrix-valued data using a matrix mean is plain sailing.

Our main prototype in this paper is the 'corner-cutting' subdivision schemes, which is the simplest nontrivial representative of the subdivision splines schemes. Other generalizations are the Lane–Riesenfeld schemes (Lane & Riesenfeld, 1980; Dyn & Goldman, 2011).

The corner-cutting schemes were first introduced by Chaikin (1974). Following de Boor (1987), the corner-cutting refinement rules are defined using a generalized arithmetic mean:

$$\begin{aligned} p_{j+1,2i} &= (1-\mu)p_{j,i} + \mu p_{j,i+1}; \\ p_{j+1,2i+1} &= \mu p_{j,i} + (1-\mu)p_{j,i+1}, \end{aligned} \tag{3.6}$$

with $i \in \mathbb{Z}$, $j \in \mathbb{Z}_+$ and $0 < \mu < \frac{1}{2}$. Note that (3.6) is a special case of (3.5) where all the nonzero coefficients are $a_{-2} = a_1 = \mu$ and $a_{-1} = a_0 = 1 - \mu$. For further information, see Dyn & Levin (2002).

Applying the matrix mean for (3.6), we get the matrix corner-cutting schemes,

$$\begin{aligned} A_{j+1,2i} &= M_{\mu}(A_{j,i}, A_{j,i+1}), \\ A_{j+1,2i+1} &= M_{1-\mu}(A_{j,i}, A_{j,i+1}). \end{aligned} \tag{3.7}$$

The subdivision process consists of repeating the refinement rule that generates a limit function $\mathscr{S}^{\infty}$ such that

$$\lim_{l\to\infty} \mathscr{S}^{\infty}(i2^{-j}) - A_{j+l,i2^l} = 0, \quad i \in \mathbb{Z}, \ j \in \mathbb{Z}_+.$$

Proving convergence of a scheme of the form (3.7) is done by finding a contraction factor, namely a scalar $0 < \gamma < 1$ such that

$$\sup_{i\in\mathbb{Z}} d(A_{j+1,i}, A_{j+1,i+1}) = \gamma \sup_{i\in\mathbb{Z}} d(A_{j,i}, A_{j,i+1}), \quad j \in \mathbb{Z}_+, \tag{3.8}$$

where $d(\cdot,\cdot)$ is a matrix metric. We construct such a proof in Section 5 for the admissible geometric mean of a matrix.

The corner-cutting scheme (3.7) is not interpolatory, yet we can find some of the limit values by the following lemma.

LEMMA 3.9 Let $\mathscr{S}^{\infty}$ be the limit matrix curve for the scheme (3.7), for the initial data $\mathbb{A}_0$. Let $\mathscr{M}_t^F$ be the matrix mean defined by the construction in Theorem 3.2 with a function $F$. Then

$$\mathscr{M}_{1/2}^F(A_{0,i}, A_{0,i+1}) = \mathscr{S}^{\infty}(i + \tfrac{1}{2}), \quad i \in \mathbb{Z}.$$

*Proof.* By the construction of $\mathscr{M}_t^F$ in Theorem 3.2, we have

$$\mathscr{M}_{1/2}^F(\mathscr{M}_t^F(A_{0,i}, A_{0,i+1}), \mathscr{M}_{1-t}^F(A_{0,i}, A_{0,i+1})) = \mathscr{M}_{1/2}^F(A_{0,i}, A_{0,i+1}).$$

We prove by induction that this property holds for each refinement step. Using continuity, the claim of the theorem follows. □

For a better understanding of the importance of admissibility we need the following definition.

DEFINITION 3.10 Let $\mathscr{S}$ be a subdivision refinement operator of the form (3.5), adapted for matrix data $\mathbf{A}_0 = \{A_{0,i}\}_{i\in\mathbb{Z}}$. We call $\mathscr{S}^{\infty}$ an *admissible* approximation scheme for a matrix if it satisfies the admissibility conditions, analogous to P1–P4, i.e. for all $x \in \mathbb{R}$, we have the following properties.

(1) Commutativity with the inverse:

$$\mathscr{S}^{\infty}(\mathbf{A}_0)(x) = (\mathscr{S}^{\infty}(\mathbf{A}_0^{-1})(x))^{-1},$$

where $\mathbf{A}_0^{-1} = \{A_{0,i}^{-1}\}_{i\in\mathbb{Z}}$.

(2) Invariance to orthogonal coordinate change:

$$Q^*\mathscr{S}^{\infty}(\mathbf{A}_0)(x)Q = \mathscr{S}^{\infty}(Q^*\mathbf{A}_0 Q)(x), \quad Q^*Q = I,$$

where $Q^*\mathbf{A}_0 Q = \{Q^* A_{0,i} Q\}_{i\in\mathbb{Z}}$.

(3) Incompressibility: if $\det(A_i) = 1$ for all $i$, then $\det(\mathscr{S}^{\infty}(\mathbf{A})(x)) = 1$.

(4) Homogeneity:

$$\mathscr{S}^{\infty}(\lambda \mathbf{A})(x) = \mathscr{S}^{\infty}(\lambda)(x)\mathscr{S}^{\infty}(\mathbf{A})(x),$$

where $\lambda = \{\lambda_i\}_{i\in\mathbb{Z}}$ is a sequence of positive scalars.[1]

In the case of a convergent corner-cutting scheme, satisfying P1–P4 in every refinement level ensures that the generated limit curve also satisfies this requirement. We have the following corollary.

COROLLARY 3.11 Subdivision schemes of the form (3.7), defined using an admissible matrix mean, are admissible schemes for matrices.

REMARK 3.12 We note three additional properties: according to the above.

(1) Theorem 3.5 implies that an admissible scheme with a refinement operator of the form (3.7) commutes with the determinant:

$$\det(\mathscr{S}^{\infty}(\{A_{0,i}\}_{i\in\mathbb{Z}})(x)) = \mathscr{S}^{\infty}(\{\det(A_{0,i})\}_{i\in\mathbb{Z}})(x), \quad x \in \mathbb{R}.$$

(2) Note that convergence does not guarantee that the limit is positive definite. Nevertheless, a convergent corner-cutting scheme (3.7) which is based on an admissible, quasi-linear matrix mean $M_t$ has a positive-definite limit. The latter is true due to Theorem 3.5 and the positivity of the scalar version of the corner cutting based on the geometric mean for numbers (for more details of such schemes see Schaefer *et al.*, 2008).

(3) P1 with the above observation implies commutativity with the adjoint operator:[2]

$$\mathrm{adj}(\mathscr{S}^{\infty}(\{A_{0,i}\}_{i\in\mathbb{Z}})(x)) = \mathscr{S}^{\infty}(\{\mathrm{adj}(A_{0,i})\}_{i\in\mathbb{Z}})(x), \quad x \in \mathbb{R}.$$

### 3.3 *Spectral properties of the adapted schemes*

A major aspect of matrix theory is spectral information. We derive spectral properties of a limit of a subdivision scheme, inherited from the data. The results of this section hold for a wider class of schemes than the admissible schemes, e.g. schemes that are based on the harmonic and arithmetic means.

For the rest of this section, we denote by $\mathscr{S}$ an operator of a convergent subdivision scheme, adapted by a matrix mean $M_t$ from a linear subdivision scheme of the form (3.5). We assume that $M_t$ preserves a common eigenvector, i.e.

$$M_t(A,B)v = M_t(\alpha,\beta)v, \quad t \in [0,1], \tag{3.9}$$

for any $A, B \succ 0$ such that $Av = \alpha v$ and $Bv = \beta v$, $v \neq 0$.

THEOREM 3.13 Let $M_t$, $t \in [0,1]$ be a matrix mean that satisfies P2 (invariant to orthogonal coordinate change) and condition (3.9). Then we have the following results.

(1) If $v$ is a common eigenvector of the data, $A_{0,i}v = \lambda_i v$, then

$$(\mathscr{S}^{\infty}(\{A_{0,i}\}_{i\in\mathbb{Z}})(x))v = \mathscr{S}^{\infty}(\{\lambda_i\}_{i\in\mathbb{Z}})(x)v, \quad x \in \mathbb{R}.$$

---

[1] The scheme is well defined for such scalars; see also P4 of Definition 3.3.

[2] For $X \succ 0$, $\mathrm{adj}(X)$ is defined as $\det(X)X^{-1}$.

(2) Let $V$ be an invariant subspace of the data, namely $v \in V$ implies $A_{0,i}v \in V$. Then $V$ is an invariant subspace of the matrices

$$\mathscr{S}^{\infty}(\{A_{0,i}\}_{i\in\mathbb{Z}})(x), \quad x \in \mathbb{R}.$$

(3) For a set of matrices $\mathscr{J}$, we define

$$\lambda_{\max}(\mathscr{J}) = \max_{X\in\mathscr{J}} \lambda_{\max}(X) \quad \text{and} \quad \lambda_{\min}(\mathscr{J}) = \min_{X\in\mathscr{J}} \lambda_{\min}(X).$$

Then

$$\lambda_{\max}(\mathscr{S}^{\infty}(\{A_{0,i}\}_{i\in\mathbb{Z}})(x)) \leqslant \lambda_{\max}(\{A_{0,i}\}_{i\in\mathbb{Z}})$$

and

$$\lambda_{\min}(\mathscr{S}^{\infty}(\{A_{0,i}\}_{i\in\mathbb{Z}})(x)) \geqslant \lambda_{\min}(\{A_{0,i}\}_{i\in\mathbb{Z}}).$$

*Proof.* The first part is a direct result of the convergence and definition of any subdivision of the form (3.7). For the second part, by the theorem's assumption, there exists an orthogonal matrix $Q$ such that $A_{0,i} = Q^{*}B_{0,i}Q$. $B_{0,i}$ is the block matrix

$$B_{0,i} = \left[\begin{array}{c|c} \bar{A}^{1}_{0,i} & 0 \\ \hline 0 & \bar{A}^{2}_{0,i} \end{array}\right], \quad i \in \mathbb{Z}, \tag{3.10}$$

with $\{\bar{A}^{1}_{0,i}\}_{i\in\mathbb{Z}}$ a set of SPD $k \times k$ matrices, $k = \dim(V)$. The blocks on the diagonal are the restrictions of the initial data for $V$ and its orthogonal complement, i.e.

$$A_{0,i}|_{V} = \bar{A}^{1}_{0,i}, \quad A_{0,i}|_{V^{\perp}} = \bar{A}^{2}_{0,i}.$$

By the P2 property,

$$\mathscr{S}(\{A_{0,i}\}_{i\in\mathbb{Z}}) = \left[\begin{array}{c|c} \mathscr{S}(\{\bar{A}^{1}_{0,i}\}_{i\in\mathbb{Z}}) & 0 \\ \hline 0 & \mathscr{S}(\{\bar{A}^{2}_{0,i}\}_{i\in\mathbb{Z}}) \end{array}\right].$$

This observation, when used repeatedly, leads to the claim of the second part of the theorem. The third claim follows from similar arguments combined with the Definition 3.1(4) and the scheme definition. □

REMARK 3.14 The finite set of nonzero coefficients in (3.5) implies locality of the subdivision process. Thus, we can update Theorem 3.13 as follows.

(1) Parts (1) and (2) of the theorem assumed a common eigenvector (or a common invariant subspace) for the whole data. We can relax the condition of Theorem 3.13 to get a local eigenvector property (or local invariant space).

(2) For the corner-cutting schemes, one can tighten the boundaries to

$$\lambda_{\max}(\mathscr{S}^{\infty}(\{A_{0,i}\}_{i\in\mathbb{Z}})(x)) \leqslant \max_{i=\lfloor x\rfloor-1,\lfloor x\rfloor,\lfloor x\rfloor+1,\lfloor x\rfloor+2} \lambda_{\max}(\{A_{0,i}\}_{i\in\mathbb{Z}}).$$

Analogously, we get the lower bound for $\lambda_{\min}$ and for the condition number.

### 3.4 *Approximation order of an adapted corner-cutting scheme*

The linear corner cutting for scalars reproduces any linear polynomial, sampled equidistantly. Thus, for equidistant samples of a $\mathbf{C}^1$ function with a distance $h$, the approximation order is $\mathscr{O}(h^2)$. We prove an analogue for the matrix setting, assuming only Lipschitz continuity for the sampled matrix function. For such functions, we have the approximation order $\mathscr{O}(h)$.

Let $M_t$ be a continuous matrix mean associated with a metric $d(\cdot,\cdot)$ of $\mathsf{SPD}(n)$. Such metrics are natural in many cases, as illustrated in Sections 4 and 5. Let $\Delta_j = \sup_{i\in\mathbb{Z}} d(A_{j,i}, A_{j,i+1})$; we assume $\Delta_0 < \infty$ and a contraction factor such as (3.8), in respect to the metric.

Consider the *metric property*,

$$d(M_t(A,B),A) = td(A,B), \quad A,B \in \mathsf{SPD}(n). \tag{3.11}$$

For a given corner-cutting scheme and initial data $\mathbf{A}_0 = \{A_{0,i}\}_{i\in\mathbb{Z}}$, we denote the sequence of continuous matrix functions using the matrix mean $M_t$ by

$$F_j(x) = M_{(x-i2^{-j})/((i+1)2^{-j}-i2^{-j})}(A_{j,i}, A_{j,i+1}), \quad x \in [i2^{-j}, (i+1)2^{-j}), \;\; j \in \mathbb{Z}_+. \tag{3.12}$$

These matrix curves are the analogue of piecewise linear curves in the Euclidean setting. The use of such curves is beneficial for proofs of convergence, e.g. the proof of Theorem 5.2. We use $\{F_j(\cdot)\}_{j\in\mathbb{Z}_+}$ for our proof of the approximation order as well.

THEOREM 3.15 Assume that $\mathscr{S}$ is a corner-cutting scheme, based on a matrix mean $M_t$ and which satisfies (3.8). Let $\mathbf{A}_0 = \{A_{0,i}\}_{i\in\mathbb{Z}}$ be initial data such that $A_{i,0} = F(ih)$, where $F : \mathbb{R} \to \mathsf{SPD}(n)$ is a Lipschitz continuous function, namely there exists a constant $\mathscr{L}_F$ such that

$$\|F(x) - F(y)\| \leqslant \mathscr{L}_F h.$$

Then

$$E(h,x) = d(\mathscr{S}^\infty(\mathbf{A_0})(x/h), F(x)) \leqslant C_F h,$$

where $C_F$ is a constant independent of $h$ and $d(\cdot,\cdot)$ is a metric satisfying the metric property introduced in (3.11).

*Proof.* By the triangle inequality, for any $x \in [ih, (i+1)h)$,

$$E(h,x) \leqslant d(\mathscr{S}^\infty(\mathbf{A_0})(x/h), F_0(x/h)) + d(F_0(x/h), A_{0,i}) + d(A_{0,i}, F(x)).$$

Using the contraction factor, one can deduce

$$d(F_j(x), F_{j+1}(x)) \leqslant 4\gamma\Delta_j, \quad j \in \mathbb{Z}_+.$$

Therefore,

$$d(\mathscr{S}^\infty(\mathbf{A_0})(x/h), F_0(x/h)) \leqslant 4\frac{1}{1-\gamma}\mathscr{L}_F h.$$

According to the Lipchitz condition and the metric property, we get

$$E(h,x) \leqslant \left(\frac{6-4\gamma}{1-\gamma}\right)\mathscr{L}_F h, \quad h \ll 1, \;\; x \in \mathbb{R}. \qquad \square$$

REMARK 3.16 Dyn & Farkhi (2001, Theorem 4.4) proved the approximation order of Lipschitz set-valued functions. We used similar arguments analogues results for SPD matrices. In addition, we can conclude as in Dyn & Farkhi (2001), that the matrix functions $\{F_j(\cdot)\}_{j\in\mathbb{Z}_+}$ are Lipschitz continuous with the same constant. The contraction factor implies uniform convergence. Hence, the limit of the corner-cutting scheme is Lipschitz continuous with the same constant as the approximant.

## 4. Subdivision schemes based on the exp–log mean

The space of SPD matrices under the operation $\exp(\log(A)+\log(B))$ forms a Lie group (Arsigny *et al.*, 2007). Following this approach, we introduce the next class of approximation schemes, adapted for SPD matrices, and based on the *exp–log mean*, i.e.

$$A \odot_t B = \exp((1-t)\log(A) + t\log(B)), \quad t \in [0,1]. \tag{4.1}$$

This mean is defined on the tangent plane of the space $\mathsf{SPD}(n)$, which is the space of symmetric matrices (Rahman *et al.*, 2005).

Based on the above, one can explain the popularity of adapting approximation operators for matrices using (4.1) (e.g. Rahman *et al.*, 2005; Navayazdani & Yu, 2010). This approach has been supported by several studies concerning analysis tools for such schemes, especially the high-proximity conditions (e.g. see Grohs & Wallner, 2008; Navayazdani & Yu, 2010). We refer to these schemes as *exp–log*-based schemes. We show that (4.1) is a special case of the construction presented in Theorem 3.2 and prove its admissibility and an additional property.

### 4.1 *Basic properties of the exp–log mean*

In the following, we introduce a fundamental characterization of the exp–log mean.

LEMMA 4.1 The operation defined in (4.1) is a matrix mean.

*Proof.* The operation (4.1) is generated by $F(X) = \exp(X)$ which is an order-preserving function. Thus, the claim follows by Theorem 3.2. □

Next we show the admissibility property of the exp–log.

THEOREM 4.2 The exp–log matrix mean is an admissible mean for matrices.

*Proof.* The matrix exponential is invertible and $(e^X)^{-1} = e^{-X}$ (Hall, 2003, Chapter 2). Thus, for P1, we have

$$\begin{aligned} A^{-1} \odot_t B^{-1} &= \exp((1-t)\log(A^{-1}) + t\log(B^{-1})) \\ &= \exp(-(1-t)\log(A) - t\log(B)) \\ &= (\exp((1-t)\log(A) + t\log(B)))^{-1} \\ &= (A \odot_t B)^{-1}. \end{aligned}$$

For P2, we have

$$Q^* \exp(X) Q = \exp(Q^* X Q), \quad Q^* \log(X) Q = \log(Q^* X Q).$$

For P3, we show a stronger result on the determinants. A well-known theorem (e.g. Hall, 2003, Chapter 2, Theorem 2.11) states that for any matrix $X$, we have $\det(\exp(X)) = \exp(\mathrm{tr}(X))$. Using the above and the log rules on $\mathsf{SPD}(n)$, we get

$$\begin{aligned}\det(A \odot_t B) &= \exp(\mathrm{tr}((1-t)\log(A) + t\log(B)))\\ &= \exp(\mathrm{tr}(\log(A^{1-t}) + \log(B^t)))\\ &= \exp(\mathrm{tr}(\log(A^{1-t})))\exp(\mathrm{tr}(\log(B^t)))\\ &= \det(A)^{1-t}\det(B)^t.\end{aligned}$$

We show P4 by using matrix–log rules (for more details see Hall, 2003, Chapter 2) and the fact that a scalar matrix commutes with any other matrix. Thus,

$$\begin{aligned}\alpha A \odot_t \beta B &= \exp\{\log(\alpha A)^{1-t} + \log(\beta B)^t\}\\ &= \exp\{\log \alpha^{1-t} I + \log A^{1-t} + \log \beta^t I + \log B^t\}\\ &= \exp\{(1-t)\log A + t\log B\}\exp\{\log \alpha^{1-t} I\}\exp\{\log \beta^t I\}\\ &= \alpha^{1-t}\beta^t A \odot_t B.\end{aligned}$$ □

In the above, we prove that the exp–log is admissible. In addition, in Lemma 4.1, we showed that the exp–log mean can be constructed in the fashion of Theorem 3.2. Thus the following corollary holds.

COROLLARY 4.3 The exp–log matrix mean (4.1) satisfies the quasi-linear condition (3.3).

### 4.2 *Properties of the exp–log subdivision schemes*

The refinement operator $\mathscr{S}$ of an exp–log-based subdivision scheme is analogous to (3.5), adapted for SPD matrices initial data $\mathbf{A}_0 = \{A_{0,i}\}_{i\in\mathbb{Z}}$ is of the form

$$\begin{aligned}(\mathscr{S}(\mathbf{A}_0))_{2i} &= \exp\left(\sum_l a_{2l}\log(A_{0,i-l})\right),\\ (\mathscr{S}(\mathbf{A}_0))_{2i+1} &= \exp\left(\sum_l a_{2l+1}\log(A_{0,i-l})\right).\end{aligned} \tag{4.2}$$

The aspects of convergence and smoothness of the exp–log schemes are well studied (e.g. Grohs & Wallner, 2008; Navayazdani & Yu, 2010). Thus, we omit these issues.

We start with a corollary deduced in the same arguments as Corollary 3.11 for the exp–log subdivision using the result of Theorem 4.2. Furthermore, by Corollary 4.3, we also have Theorem 3.13 for the exp–log approximation schemes of the form (4.2).

COROLLARY 4.4 The matrix subdivision scheme (4.2) is an admissible scheme satisfying the spectral claims of Theorem 3.13.

Next, we discuss the corner-cutting subdivision scheme (3.7), using the exp–log matrix mean (4.1). As mentioned above, it is an admissible approximation scheme, for which Theorem 3.13 holds. Nevertheless, we have additional properties.

For the first property, the exp–log mean has an associated metric of the form

$$d(A,B) = \| \log(A) - \log(B) \|_2 = \lambda_{\max}(A-B, B-A),$$

where $\| \cdot \|_2$ is the induced norm (Golub & Van Loan, 1996, Chapter 2). For this metric, the metric property (3.11) holds. Thus, the exp–log-based corner-cutting scheme obtains the approximation order result of Theorem 3.15. The second property is stated in the following theorem.

THEOREM 4.5 Let $A, B \succ 0$. Then

$$\operatorname{tr}(A \odot_t B) \leqslant (\operatorname{tr}(A))^{1-t}(\operatorname{tr}(B))^t, \quad t \in [0,1]. \tag{4.3}$$

Moreover, for any initial data $\{A_{0,i}\}_{i\in\mathbb{Z}}$ and $\mathscr{S}^\infty$ a corner-cutting subdivision scheme (3.7) that is based on the exp–log matrix mean (4.1),

$$\operatorname{tr}(\mathscr{S}^\infty(\{A_{0,i}\}_{i\in\mathbb{Z}})(x)) \leqslant \mathscr{S}^\infty(\{\operatorname{tr}(A_{0,i})\}_{i\in\mathbb{Z}})(x), \quad x \in \mathbb{R}.$$

For the proof of the theorem, we use the following lemma.

LEMMA 4.6 Let $\{a_i\}_{i=1}^n, \{b_i\}_{i=1}^n, \{c_i\}_{i=1}^n$ be sequences of numbers such that

$$\sum_{i=1}^k b_i \geqslant \sum_{i=1}^k c_i, \quad 1 \leqslant k \leqslant n$$

and $a_1 \geqslant a_2 \geqslant \cdots \geqslant a_n \geqslant 0$. Then

$$\sum_{i=1}^n a_i b_i \geqslant \sum_{i=1}^n a_i c_i.$$

*Proof.* Let

$$d_k = \sum_{i=1}^k b_i - c_i, \quad 1 \leqslant k \leqslant n.$$

Then by the assumption of the lemma, $d_k \geqslant 0$ for all $k$. We use induction to prove that

$$\sum_{i=1}^k (b_i - c_i) a_i \geqslant a_k d_k.$$

The case $k = 1$ is trivially true. Now assume that the claim is true for $k \geqslant m$. Then

$$\begin{aligned}
\sum_{i=1}^{m+1} (b_i - c_i) a_i &= \sum_{i=1}^{m} (b_i - c_i) a_i + (b_{m+1} - c_{m+1}) a_{m+1} \\
&\geqslant a_m d_m + (b_{m+1} - c_{m+1}) a_{m+1} \\
&\geqslant a_{m+1} (d_m + b_{m+1} - c_{m+1}) \\
&= a_{m+1} d_{m+1}.
\end{aligned}$$

In the above, we used the basis of the induction, the monotonic decreasing of $a_i$ and that $d_k \geqslant 0$. Since $a_k, d_k \geqslant 0$, the claim follows. □

*Proof of Theorem* 4.5. By Hölder's inequality, for sequences of positive numbers $\{x_i\}_{i=1}^n, \{y_i\}_{i=1}^n$, we have

$$\sum_{i=1}^n x_i y_i \leqslant \left(\sum_{i=1}^n x_i^q\right)^{1/q} \left(\sum_{i=1}^n y_i^p\right)^{1/p}, \quad \frac{1}{p}+\frac{1}{q}=1.$$

On these grounds, for $\beta_i = y_i^p$, $\alpha_i = x_i^q$,

$$\sum_{i=1}^n \alpha_i^{1/q} \beta_i^{1/p} \leqslant \left(\sum_{i=1}^n \alpha_i\right)^{1/q} \left(\sum_{i=1}^n \beta_i\right)^{1/p}. \tag{4.4}$$

For any $X \succ 0$ we denote the full set of its eigenvalues by

$$\lambda_{\max}(X) = \lambda_1(X) \geqslant \lambda_2(X) \geqslant \cdots \lambda_n(X) = \lambda_{\min}(X) > 0.$$

Let $A, B \succ 0$. Thus, for $t = 1/p$, we get from (4.4),

$$\sum_{i=1}^n \lambda_i(A)^{1-t} \lambda_i(B)^t \leqslant (\mathrm{tr}(A))^{1-t} (\mathrm{tr}(B))^t, \quad t \in [0,1]. \tag{4.5}$$

Recall P2 of admissibility for the exp–log mean and the fact the trace is invariant for orthogonal similarity, namely that for any matrix $M$,

$$\mathrm{tr}(Q^* M Q) = \mathrm{tr}\, M, \quad QQ^* = I.$$

Thus, when considering $\mathrm{tr}(A \odot_t B)$, we can assume without loss of generality that $A$ is a diagonal matrix. The Golden–Thompson inequality (e.g. Petz, 1994) yields

$$\begin{aligned} \mathrm{tr}(\exp((1-t)\log A + t \log B)) &\leqslant \mathrm{tr}\exp((1-t)\log A)\exp(t \log B) \\ &= \mathrm{tr}(A^{1-t} B^t) = \sum_{i=1}^n A_{ii}^{1-t} (B^t)_{ii} \\ &= \sum_{i=1}^n \lambda_i^{1-t}(A) (B^t)_{ii}. \end{aligned}$$

We use the following claim by Schur (Horn & Johnson, 1990, Chapter 4, Section 3, p. 193). The claim states that for any $X \in \mathsf{SPD}(n)$,

$$\sum_{i=1}^k X_{ii} \leqslant \sum_{i=1}^k \lambda_i(X), \quad k \leqslant n.$$

Recall that $B^t \in \mathsf{SPD}(n)$, hence, all its diagonal elements are strictly positive. By Lemma 4.6, we have

$$\sum_{i=1}^n \lambda_i^{1-t}(A)(B^t)_{ii} \leqslant \sum_{i=1}^n \lambda_i^{1-t}(A) \lambda_i^t(B), \quad t \in [0,1].$$

The above with (4.5) implies

$$\mathrm{tr}(A \odot_t B) \leqslant (\mathrm{tr}(A))^{1-t}(\mathrm{tr}(B))^t.$$

Repeating the previous argument yields the second claim for the subdivision scheme. □

Using Theorem 4.5, we give a bound for the exp–log-based corner-cutting scheme in the Frobenius norm.

THEOREM 4.7 Let $M_t$ be a matrix mean that agrees with

$$\mathrm{tr}(M_t(A,B)) \leqslant M_t(\mathrm{tr}(A), \mathrm{tr}(B)), \quad A, B \succ 0.$$

Then for any initial data $\mathbf{A}_0$ and for the corner cutting based on $M_t$,

$$\|\mathscr{S}^\infty(\mathbf{A}_0)(x)\| \leqslant \mathscr{S}^\infty(\{\|A_{0,i}^{1/2}\|^2\}_{i\in\mathbb{Z}})(x), \quad x \in \mathbb{R}.$$

*Proof.* For any $X \succ 0$,

$$\mathrm{tr}(XX^*) = \sum_{i=1}^n \lambda_i(X)^2, \quad (\mathrm{tr}(X))^2 = \left(\sum_{i=1}^n \lambda_i(X)\right)^2.$$

Since $\lambda_i(X) > 0$, we have

$$\mathrm{tr}(XX^*) \leqslant (\mathrm{tr}\, X)^2.$$

Thus,

$$\begin{aligned}\|M_t(A,B)\| &= \sqrt{\mathrm{tr}(M_t(A,B)M_t(A,B)^*)}\\ &\leqslant \mathrm{tr}(M_t(A,B))\\ &\leqslant M_t(\mathrm{tr}(A), \mathrm{tr}(B))\\ &= M_t(\|A^{1/2}\|^2, \|B^{1/2}\|^2).\end{aligned}$$

By the definition of the corner-cutting scheme, the claim follows. □

It is worth mentioning that Theorem 4.7 holds for any scheme that is based on a mean which agrees with (4.3).

## 5. Subdivision schemes based on the geometric mean

For the manifold of the SPD matrices, there exists a Riemann metric,

$$d(A,B) = \|\log(A^{-1/2}BA^{-1/2})\|. \tag{5.1}$$

The geodesic line between $A$ and $B$ with respect to the Riemann metric (5.1) is

$$G_t(A,B) = A(A^{-1}B)^t = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}, \quad 0 \leqslant t \leqslant 1. \tag{5.2}$$

In this section, we investigate an admissible mean reduced by the geodesic line (5.2).

### 5.1 *The geometric mean for SPD matrices*

The theory of means for positive operators was developed by Kubo & Ando (1980) and led to the study of the special case of the geometric mean for SPD matrices (Ando *et al.*, 2004; Bini *et al.*, 2010).

An interesting special case is $t=\frac{1}{2}$, which is the midpoint of the geodesic line. We present three alternative definitions for the midpoint, which emphasize the various interpretations of the geometric mean (for more details, see Ando *et al.*, 2004 and Bini *et al.*, 2010).

The first is $G_{1/2}(A,B)=A^{1/2}QB^{1/2}$, where $Q$ is any orthogonal matrix that ensures that the product is symmetric positive definite. Although the choice of $Q$ is arbitrary, the value of $A^{1/2}QB^{1/2}$ is unique.

The second is the value of the integral

$$G_{1/2}(A,B)=\frac{1}{\Gamma(1/2)^2}\int_0^1 (xB^{-1}+(1-x)A^{-1})^{-1}(x(1-x))^{-1/2}\,\mathrm{d}x,$$

where $\Gamma(x)=\int_0^\infty w^{x-1}e^{-w}\,\mathrm{d}w$ is the well-known gamma function.

The third is the solution of the following optimization problem:

$$G_{1/2}(A,B)=\max\left\{X\succ 0\colon \begin{pmatrix} A & X\\ X & B\end{pmatrix}\succeq 0\right\}. \tag{5.3}$$

Equation (5.3) is equivalent to the existence and uniqueness of a solution to the matrix Riccati equation $XA^{-1}X=B$ (Lawson & Lim, 2001, Lemma 2.4).

A list of fundamental properties of the geometric mean (5.2) can be found in Itai & Sharon (2012). We note that there is no trivial choice of function to construct the geometric mean as suggested in Theorem 3.2. Nevertheless, the quasi-linear condition (3.3) holds and $G_t(x,y)=x^{1-t}y^t$ for all $x,y>0$.

A fundamental result is as follows.

THEOREM 5.1 The geometric mean for matrix $G_t$ is an admissible matrix mean.

*Proof.* First, we show that $G_t$ is a matrix mean. The symmetry property can be found in Itai & Sharon (2012, Lemma 3.1(2)). Ando *et al.* (2004) show that[3]

$$G_t(A,B)\preceq(1-t)A+tB\preceq\lambda_{\max}(A,B)I,\qquad t\in[0,1].$$

Definition 3.3 follows by the same arguments as Itai & Sharon (2012, Theorems 5.3, 5.10, 5.2), respectively. The homogeneity property follows by (5.2). □

### 5.2 *The geometric corner-cutting schemes*

Constructing a corner-cutting scheme (3.7) using the geometric mean for matrix (5.2) is well defined. We name such a scheme a *geometric corner-cutting* subdivision scheme.

The next theorem ensures convergence and smoothness of the scheme.

THEOREM 5.2 The geometric corner-cutting scheme (3.7) converges to a smooth $\mathbf{C}^1$ matrix curve, for any initial data of SPD matrices with eigenvalues that are bounded away from zero.

[3] Ando proved the matrix means inequality $H_t(A,B)\preceq G_t(A,B)\preceq L_t(A,B)$.

For the proof of the theorem, we use the following notation for given initial data $\{A_{0,i}\}_{i\in\mathbb{Z}}$, as well as for any refined set $\{A_{j,i}\}_{i\in\mathbb{Z}}$,

$$\varDelta_j = \sup_{i\in\mathbb{Z}} d(A_{j,i}, A_{j,i+1}), \quad j \in \mathbb{Z}_+.$$

We assume that our initial data satisfy $\varDelta_0 < \infty$ and use two lemmas.

LEMMA 5.3 The geometric corner-cutting scheme (3.7) satisfies

$$\varDelta_{j+1} < \gamma \varDelta_j, \quad j \in \mathbb{Z}_+, \ 0 < \gamma < 1. \tag{5.4}$$

*Proof.* The geometric corner cutting schemes (3.7) generate refined matrices on the geodesic line. Therefore, for any initial data $\mathbf{A}^0 = \{A_i\}_{i\in\mathbb{Z}}$, we have

$$d(A_{j+1,2i}, A_{j+1,2i+1}) = (1-2\mu) d(A_{j,i}, A_{j,i+1}), \quad j \in \mathbb{Z}_+, \ i \in \mathbb{Z}.$$

Following the same argument and the triangle inequality,

$$d(A_{j+1,2i}, A_{j+1,2i-1}) \leqslant 2\mu d(A_{j,i}, A_{j,i+1}).$$

By $\mu \in (0, \frac{1}{2})$, we get $\gamma = \max\{1-2\mu, 2\mu\} < 1$. This completes the proof. □

The distance between the arithmetic matrix mean $L_t$ (see Example 3.4) and the geometric mean for matrices, for sufficiently close matrices is given in the next lemma. A similar result is partially proved by the authors in Itai & Sharon (2012).

LEMMA 5.4 Let $A, B \in \mathsf{SPD}(n)$ be bounded matrices $m < \lambda_{\min}(A), \lambda_{\min}(B)$ with $0 < m$, and such that

$$\|A^{-1}B - I\| < 1.$$

Then

$$\|G_t(A,B) - L_t(A,B)\| < C\|A-B\|^2, \tag{5.5}$$

where $C$ depends only on $n$ and $m$.

*Proof.* Using the Taylor expansion,

$$(x+1)^t = 1 + tx + \tfrac{1}{2}t(t-1)x^2 + \cdots, \quad |x-1| < 1,$$

implies (Mathias, 1993, Corollary 2) that for $\|A^{-1}B - I\| < 1$, the truncated Taylor matrix approximation with a remainder $R$ has the form

$$\begin{aligned}
\|R\| &= \|(A^{-1}B)^t - (I + t(A^{-1}B - I))\| \\
&= \|(A^{-1}B)^t - ((1-t)I + t(A^{-1}B))\| \\
&= \tfrac{1}{2} \max_{s\in[0,1]} \|(A^{-1}B - I)^2 t(t-1)((1-s)I + s(A^{-1}B))\| \\
&= \tfrac{1}{2}|t(t-1)| \max_{s\in[0,1]} \|(A^{-1}B - I)^2\| \, \|L_s(I, A^{-1}B)\|.
\end{aligned}$$

The sub-multiplication of the Frobenius norm entails

$$\|G_t(Ae,B) - L_t(A,B)\| \leqslant \frac{1}{2}|t(t-1)|\|A^{-1}\|^2 \left(\max_{s\in[0,1]} \|L_s(I,A^{-1}B)\|\right) \|A-B\|^2.$$

Consider $t \in [0,1]$, and with $\|A^{-1}B - I\| < 1$ implies $\|L_s(I,A^{-1}B)\| \leqslant 1+\sqrt{n}$, one can set $C = n(1+\sqrt{n})/8m^2$. Thus, the claim follows. □

*Proof of Theorem* 5.2. Let $x_{j,i} = 2^{-j}i$, $j \in \mathbb{Z}_+$, $i \in \mathbb{Z}$ and $F_j(x)$, $j \in \mathbb{Z}$ be the continuous matrix functions defined in (3.12). Similar arguments as in Lemma 5.3 imply

$$\sup_{x\in\mathbb{R}} d(F_{j+1}(x_{j+1,2i(x)-1}), F_{j+1}(x)) \leqslant 2\mu\Delta_j, \quad j \in \mathbb{Z}_+,$$

where $i(x) = \arg\min_k \{d(F_j(x), F_j(x_{j,k}))\}$. Moreover, by the definition of $F_j$, one gets $F_j(x_{j+1,2i(x)-1}) = F_{j+1}(x_{j+1,2i(x)-1})$. Then

$$\begin{aligned} d(F_j(x), F_{j+1}(x)) &\leqslant d(F_j(x), F_j(x_{j,i(x)})) + d(F_j(x_{j,i(x)}), F_{j+1}(x_{j+1,2i(x)-1})) \\ &\quad + d(F_{j+1}(x_{j+1,2i(x)-1}), F_{j+1}(x)) \\ &\leqslant 1/2\Delta_j + \gamma\Delta_j + 2\gamma\Delta_j \leqslant 4\Delta_j. \end{aligned}$$

Combining the latter with Lemma 5.3 and the triangle inequality, we have for any $j_1, j_2 \in \mathbb{Z}_+$ (without loss of generality $j_2 > j_1$),

$$d(F_{j+j_1}(x), F_{j+j_2}(x)) \leqslant \sum_{l=0}^{j_2-j_1-1} d(F_{j+j_1+k}(x), F_{j+j_1+k+1}(x)) \leqslant 4\frac{1}{1-\gamma}\Delta_j.$$

Since $\Delta_j$ tends to zero and is independent of $x$, we get that $\{F_j(x)\}_{j\in\mathbb{N}}$ is a Cauchy series under the uniform norm. The completeness of $\mathbb{R}^{n\times n}$ implies that the sequence $\{F_j(x)\}_{j\in\mathbb{N}}$ converges uniformly to a well-defined function $F(x)$.

For the $\mathbf{C}^1$ smoothness, we use the notion of proximity, introduced by Wallner & Dyn (2005). Lemma 5.4 guarantees the conditions for proximity. This is true due to locality and convergence of the corner cutting. □

The geometric mean preserves common eigenvectors. Therefore, the claims of Theorem 3.13 hold for the geometric corner cutting.

During the work on the paper, the authors learned about the work of Ebner (2012), which generalized the proof of convergence given here. However, for the sake of making the paper as self-contained as possible, we decided to include the full proof.

### 5.3 *Approximation order of the geometric corner cutting*

In Section 3.4, we showed that for a scheme based on an admissible mean, the approximation order is $\mathscr{O}(h)$. This result, however, was not tight for many adapted corner-cutting schemes. For example, we show in this section that for the geometric corner cutting the approximation order is $\mathscr{O}(h^2)$.

The manifold of $\mathsf{SPD}(n)$ matrices is locally Euclidean (Spivak, 2005, Chapter 9). Therefore, for any compact set $\mathscr{P} \subset \mathsf{SPD}(n)$, there exist two positive constants $c$ and $C$ depending only on $\mathscr{P}$

such that

$$c\|A-B\| \leqslant d(A,B) \leqslant C\|A-B\|, \quad A,B \in \mathscr{P}. \tag{5.6}$$

Here, $d(\cdot,\cdot)$ is the Riemannian metric (5.1). We denote by $\mathscr{S}_L$ the corner-cutting refinement operator based on the arithmetic matrix mean $L_t$ (see Example 3.4). The next three lemmas study the connection between the linear and geometric schemes.

LEMMA 5.5 Let $\mathbf{A}_0$ and $\mathbf{B}_0$ be initial data sequences such that $\|A_{0,i}-B_{0,i}\| < \epsilon$. Then

$$\|\mathscr{S}_L^{\infty}(\mathbf{A}_0)(x) - \mathscr{S}_L^{\infty}(\mathbf{B}_0)(x)\| < \epsilon, \quad x \in \mathbb{R}.$$

*Proof.* By the linearity of the scheme,

$$\mathscr{S}_L(\mathbf{A}_0)_j - \mathscr{S}_L(\mathbf{B}_0)_j = \mathscr{S}_L(\mathbf{A}_0 - \mathbf{B}_0)_j, \quad j \in \mathbb{Z}.$$

Thus, for the first refinement step, we have

$$\|\mathscr{S}_L(\mathbf{A}_0)_j - \mathscr{S}_L(\mathbf{B}_0)_j\| < \epsilon, \quad j \in \mathbb{Z}.$$

With repetition of the last observation, the lemma follows. □

By using Lemma 5.4, we deduce the following Lemma 5.6.

LEMMA 5.6 Let $\mathbf{A}_0$ be initial bounded data satisfying $\|A_{0,i}-A_{0,i+1}\| \leqslant h$ and $\|A_{0,i}^{-1}A_{0,i+1} - I\| < 1$. Then

$$\|\mathscr{S}_L(\mathbf{A}_p)_j - \mathscr{S}(\mathbf{A}_p)_j\| < C(\gamma^p h)^2, \quad j \in \mathbb{Z}, \ \ p \in \mathbb{Z}_+,$$

where $\mathbf{A}_p = \mathscr{S}^p(\mathbf{A}_0)$ and $C$ is the constant of (5.5).

*Proof.* By the contraction factor (5.4) and Lemma 5.5, the claim follows. □

Note that Lemma 5.6 combined with Lemma 5.5 ensures that

$$\|\mathscr{S}_L^{m+1}(\mathbf{A}_p)_j - \mathscr{S}_L^{m}(\mathbf{A}_p)_j\| < C(\gamma^p h)^2, \quad j \in \mathbb{Z}, \ \ \forall\, m \in \mathbb{N}.$$

LEMMA 5.7 In the notation of Lemma 5.6,

$$\|\mathscr{S}^{\infty}(\mathbf{A}_0)(x) - \mathscr{S}_L^{\infty}(\mathbf{A}_0)(x)\| < \mathscr{L}h^2, \quad x \in \mathbb{R},$$

where $\mathscr{L} = C(1/(1-\gamma^2))$.

*Proof.* Denote by

$$I_{k,j} = \|\mathscr{S}_M^k(\mathbf{A}_0)_j - \mathscr{S}_L^k(\mathbf{A}_0)_j\|, \quad j \in \mathbb{R}.$$

By Lemma 5.6, we have

$$\begin{aligned} I_{k,j} &\leqslant \|\mathscr{S}_M^k(\mathbf{A}_0)_j - \mathscr{S}_L(\mathbf{A}_{k-1})_j\| + \|\mathscr{S}_L(\mathbf{A}_{k-1})_j - \mathscr{S}_L^2(\mathbf{A}_{k-2})_j\| \\ &\quad + \cdots + \|\mathscr{S}_L^{k-1}(\mathbf{A}_1)_j - \mathscr{S}_L^k(\mathbf{A}_0)_j\| \\ &\leqslant Ch^2\left(\sum_{p=0}^{k-1}\gamma^{2p}\right) \leqslant \mathscr{L}h^2. \end{aligned}$$

□

Next is the main theorem of this section.

THEOREM 5.8 Let $\mathbf{A}_0 = \{A_{0,i}\}_{i\in\mathbb{Z}}$ be initial data sampled uniformly from a smooth function $F:\mathbb{R}\to \mathsf{SPD}(n)$, in the following fashion:

$$A_{i,0} = F(ih), \quad i \in \mathbb{Z}.$$

Then for sufficiently small $h$,

$$E(h,x) = \|\mathscr{S}^{\infty}(\mathbf{A_0})(x/h) - F(x)\| \leqslant C_F h^2, \quad x \in \mathbb{R}.$$

*Proof.* By the triangle inequality,

$$E(h,x) \leqslant \|\mathscr{S}^{\infty}(\mathbf{A_0})(x/h) - \mathscr{S}_L^{\infty}(\mathbf{A_0})(x/h)\| + \|\mathscr{S}_L^{\infty}(\mathbf{A_0})(x/h) - F(x)\|.$$

For the linear scheme, the differences are elementwise. Thus,

$$\|\mathscr{S}_L^{\infty}(\mathbf{A_0})(x) - F(x)\| < C_1 h^2,$$

where $C_1$ depends on the maximal elementwise differentiation of $F$ and the order of the matrices. Next, we use that $\|A_{0,i} - A_{0,i+1}\| \leqslant \mathscr{L}_F h$, where $\mathscr{L}_F$ is a constant depending only on $F$. For a sufficiently small $h$, we can apply Lemma 5.7 and the theorem follows. □

REMARK 5.9 The properties of the geometric mean that we have used in this section are: a proximity condition of the form given in Lemma 5.4, a contraction factor (see (3.8)), and a metric equivalence of the form (5.6). Hence, any corner cutting scheme, based on a matrix mean that agrees with such properties, has an approximation order of $\mathscr{O}(h^2)$. In addition, one can weaken the assumption of a contraction factor to a proper uniform convergence condition. For further discussion on the approximation order of nonlinear subdivision schemes using proximity conditions, see Dyn *et al.* (2010) and Grohs (2010a).

## 6. Special properties of the geometric corner-cutting schemes

Ando *et al.* (2004) introduced several important properties for the geometric mean of a matrix. In this section, we show that those properties hold for the geometric corner cutting. We also prove additional properties.

### 6.1 *Properties related to the Löwner partial ordering*

The Löwner partial order (2.1) is strongly related to the geometric mean for matrices (5.2). In the next theorem, we introduce a list of properties relating the geometric corner-cutting scheme to the Löwner partial order.

THEOREM 6.1 Let $\mathbf{A}_0 = \{A_{0,i}\}_{i\in\mathbb{Z}}$ be initial data for the geometric corner-cutting scheme and denote by $\mathscr{S}$ its refinement rule.

(1) Hierarchy preserving: let $\mathbf{B}_0 = \{B_{0,i}\}_{i\in\mathbb{Z}}$ be sequences of SPD matrices such that $A_{0,i} \succeq B_{0,i}$ for all $i \in \mathbb{Z}$. Then

$$(\mathscr{S}^{\infty}\mathbf{B}_0)(x) \preceq (\mathscr{S}^{\infty}\mathbf{A}_0)(x), \quad x \in \mathbb{R}.$$

(2) Monotonicity: if $A_{0,i} \preceq A_{0,i+1}$ for all $i \in \mathbb{Z}$, then

$$(\mathscr{S}^\infty \mathbf{A}_0)(x_1) \preceq (\mathscr{S}^\infty \mathbf{A}_0)(x_2), \quad x_1 \leqslant x_2.$$

(3) Schur complement: let $X = \begin{pmatrix} X_{11} & X_{12} \\ X_{12}^* & X_{22} \end{pmatrix}$. Then[4]

$$\mathrm{Sc}(\mathscr{S}^\infty(\{(A_{0,i})\}_{i\in\mathbb{Z}})) \preceq (\mathscr{S}^\infty(\{\mathrm{Sc}(A_{0,i})\}_{i\in\mathbb{Z}})).$$

(4) Pinching: let $X = \begin{pmatrix} X_{11} & X_{12} \\ X_{12}^* & X_{22} \end{pmatrix}$ be an SPD matrix. Denote the pinching operator by $\Phi(X) = \begin{pmatrix} X_{11} & 0 \\ 0 & X_{22} \end{pmatrix}$. Then

$$\Phi((\mathscr{S}^\infty \mathbf{A}_0)(x)) \preceq \mathscr{S}^\infty(\{\Phi(\mathbf{A}_0)\}_{i\in\mathbb{Z}})(x), \quad x \in \mathbb{R}.$$

(5) Trace:

$$\mathrm{tr}(\mathscr{S}^\infty(\{A_{0,i}\}_{i\in\mathbb{Z}})(x)) \leqslant \mathscr{S}^\infty(\{\mathrm{tr}(A_{0,i})\}_{i\in\mathbb{Z}})(x), \quad x \in \mathbb{R}.$$

*Proof.* By Ando *et al.* (2004), we get for any SPD matrices $X_2 \preceq X_1$, $Y_2 \preceq Y_1$,

$$G_t(X_2, Y_2) \preceq G_t(X_1, Y_1), \quad t \in [0,1]. \tag{6.1}$$

Following (3.7), we have $B_{1,i} \preceq A_{1,i}$.

For monotonicity, we show $A_{j,2i-1} \preceq A_{j,2i} \preceq A_{j,2i+1}$, based on the monotonicity of the $(j-1)$th level. According to (6.1), if $A \preceq B$, then

$$A = G_t(A,A) \preceq G_t(A,B) \preceq G_t(B,B) = B, \quad t \in [0,1],$$

which yields

$$A \preceq G_{1/4}(A,B) = G_{1/2}(G_{1/2}(A,B),B) \preceq G_{1/2}(A,B).$$

Using induction over the dyadic values and the dyadic density yields

$$G_{t_1}(A,B) \preceq G_{t_2}(A,B), \quad 0 \leqslant t_1 < t_2 \leqslant 1.$$

By the above, for the initial data, we have

$$\begin{aligned} A_{0,i} &= G_t(A_{0,i}, A_{0,i}) \\ &\preceq G_t(A_{0,i}, A_{0,i+1}) \\ &\preceq G_t(A_{0,i+1}, A_{0,i+1}) = A_{0,i+1}. \end{aligned}$$

Thus, for any $j \in \mathbb{Z}_+$,

$$\begin{aligned} A_{j+12i-1} &= G_{1-\mu}(A_{j,i-1}, A_{j,i}) \\ &\preceq G_{1/2}(A_{j,i-1}, A_{j,i}) \\ &\preceq G_\mu(A_{j,i-1}, A_{j,i}) = A_{j+1,2i}. \end{aligned}$$

The claim $A_{j,2i} \preceq A_{j,2i+1}$ is similar.

---

[4] Let $\mathrm{Sc}(X) = X_{22} - X_{12}^* X_{11}^{-1} X_{12}$ be the Schur complement of $X$. For more details on the Schur complement, see Boyd & Vandenberghe (2004).

For the Schur property, the Schur complement can be written as (Li & Mathias, 2000)

$$\mathrm{Sc}(X) = \max \left\{ C \mid X \succeq \begin{pmatrix} 0 & 0 \\ 0 & C \end{pmatrix} \right\}.$$

Hence, $X \succeq \left(\begin{smallmatrix} 0 & 0 \\ 0 & \mathrm{Sc}(X) \end{smallmatrix}\right)$. Thus, according to claim (1),

$$\mathscr{S}^\infty(\{A_{0,i}\}_{i\in\mathbb{Z}}) \succeq \mathscr{S}^\infty \left( \left\{ \begin{pmatrix} 0 & 0 \\ 0 & \{\mathrm{Sc}(A_{0,i})\}_{i\in\mathbb{Z}} \end{pmatrix} \right\}_{i\in\mathbb{Z}} \right).$$

Using the Schur complement on both sides implies (3).

For the pinching property, we recall that

$$\Phi(X) = (X + S^*XS)/2, \quad S = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}.$$

Therefore, for $X \succeq 0$,

$$\Phi(X) = \tfrac{1}{2}(X + (X^{1/2}S)^*(X^{1/2}S)) \succeq 0.$$

The linearity of $\Phi$ entails that $\Phi$ is a monotonicity-preserving operator. By Ando *et al.* (2004),

$$\Phi(G_t(X, Y) \preceq G_t(\Phi(X), \Phi(Y)).$$

Now by the monotonicity and the definition of the scheme (3.7), we have

$$\Phi(A_{j+1,i})) \preceq \mathscr{S}(\Phi(A_{j,i})), \quad j \in \mathbb{Z}_+,$$

which concludes the case for the pinching property.

For the trace property, we apply the pinching operator $\lfloor \log(n) \rfloor + 1$ times for any $n \times n$ matrices, and use the pinching property. □

To emphasize some of the corner-cutting-scheme properties, we present a version of Lemma 3.9 for geometric corner cutting.

LEMMA 6.2 Let $\mathscr{S}^\infty$ be the limit matrix curve for the geometric corner-cutting schemes with the initial data $\mathbf{A}_0$. Then

$$G_{1/2}(A_{0,i}, A_{0,i+1}) = \mathscr{S}^\infty(\mathbf{A}_0)(i + \tfrac{1}{2}), \quad i \in \mathbb{Z}.$$

*Proof.* The proof is a direct consequence of the fact that the geometric mean is the geodesic line in the Riemann metric (5.1) for $t \in [0, 1]$. □

In Section 4, we described the *exp–log*-based subdivision scheme. This setting provides many fruitful properties, e.g. Theorem 4.5, which is the analogue of the trace property of Theorem 6.1. Nevertheless, for this class of subdivision schemes, Properties (1–4) of Theorem 6.1 do not hold. We now supply counterexamples.

EXAMPLE 6.3 By Lemma 3.9, we can construct a counterexample based on the fact that

$$A_{0,k} \odot_{1/2} A_{0,k+1} = \mathscr{S}^{\infty}(\{A_{0,i}\}_{i\in\mathbb{Z}})(k+\tfrac{1}{2}), \quad k \in \mathbb{Z},$$

where $\mathscr{S}$ is the refinement operator of the corner-cutting scheme for matrices (3.7) under the exp–log mean (4.1). For the first two examples, we use the matrices

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

Clearly $A \succeq B$.

(1) Consider the initial data

$$A_{0,i} = A, \quad B_{0,i} = \begin{cases} A & \text{if } i \leqslant k, \\ B & \text{otherwise.} \end{cases}$$

Then

$$A = (A_{0,k} \odot_{1/2} A_{0,k+1}) \not\succeq (B_{0,k} \odot_{1/2} B_{0,k+1}) = (A \odot_{1/2} B) \approx \begin{pmatrix} 2.40 & 0.39 \\ 0.39 & 1.38 \end{pmatrix}.$$

This is a counterexample for the hierarchy-preserving property.

(2) Let the initial data be

$$C_{0,i} = \begin{cases} A & \text{if } i \leqslant k, \\ B & \text{otherwise.} \end{cases}$$

Clearly, $C_{0,i} \succeq C_{0,i+1}$. However, $C_{0,k} \not\succeq (A \odot_{1/2} B)$ which violates the monotonicity property.

(3) By calculation, for a counterexample of the Schur property for the exp–log corner cutting, let

$$A = \begin{pmatrix} 1.0744 & 1.3925 & 0.5498 \\ 1.3925 & 1.8048 & 0.7127 \\ 0.5498 & 0.7127 & 0.3507 \end{pmatrix}, \quad B = \begin{pmatrix} 1.0342 & 0.7715 & 0.5365 \\ 0.7715 & 0.5911 & 0.4960 \\ 0.5365 & 0.4960 & 1.1395 \end{pmatrix}.$$

Thus,

$$\mathrm{Sc}(A \odot_{1/2} B) \not\succeq (\mathrm{Sc}(A) \odot_{1/2} \mathrm{Sc}(B)).$$

(4) For the pinching property, let

$$A = \begin{pmatrix} 0.5729 & 0.3775 & 0.7131 \\ 0.3775 & 0.3748 & 0.3884 \\ 0.7131 & 0.3884 & 1.0861 \end{pmatrix}, \quad B = \begin{pmatrix} 1.5520 & 1.0121 & 2.0427 \\ 1.0121 & 0.8086 & 1.2249 \\ 2.0427 & 1.2249 & 2.9252 \end{pmatrix}.$$

Similarly to the above, we have $\Phi(A \odot_{1/2} B) \not\succeq (\Phi(A) \odot_{1/2} \Phi(B))$.

## 6.2 *The convex hull property*

Convex properties are fundamental in the field of geometric analysis. A generalization to manifolds is made by replacing the straight line by geodesic lines. As in the classical case of corner cutting (de Boor, 1987), there exists a strong relation between the convex hull of the data and the convex hull of the generated curve.

DEFINITION 6.4 Let $K$ be a nonempty closed set of SPD matrices. If $A, B \in K$ implies $G_t(A,B) \in K$, $t \in [0,1]$, then $K$ is a *convex set*. For any nonempty set $P$ of SPD matrices, we denote by $\mathrm{conv}(P)$ the minimal convex set containing $P$ and call it a *convex hull*.

THEOREM 6.5 Let $\mathbf{A}_0 = \{A_{0,i}\}_{i\in\mathbb{Z}}$ be initial data for the geometric corner-cutting scheme (3.7) and denote by $\mathscr{S}$ its refinement rule. Then the following conditions hold.

(1) $\mathrm{conv}(\mathscr{S}^\infty(\mathbf{A}_0)) \subseteq \mathrm{conv}(\mathbf{A}_0)$.

(2) $\mathrm{conv}(\mathscr{S}^\infty(\mathbf{A}_0)) = \bigcap_{j=0}^{\infty} \mathrm{conv}(\mathbf{A}_j)$, where $\mathbf{A}_j = \mathscr{S}^j(\mathbf{A}_0),\ j \in \mathbb{N}$.

*Proof.* By the scheme definition (3.7), it is clear that

$$A_{1,i} \in \mathrm{conv}(\mathbf{A}^0), \quad i \in \mathbb{Z}.$$

Thus,

$$\mathrm{conv}(\mathscr{S}^j(\mathbf{A}_0)) \subseteq \mathrm{conv}(\mathbf{A}_0), \quad i \in \mathbb{Z},\ j \in \mathbb{Z}_+.$$

The convex hull is a closed set and therefore the first part follows.

For the second part, we note that

$$\mathrm{conv}(\mathbf{A}_j) \supseteq \mathrm{conv}(\mathbf{A}_{j+1}), \quad j \in \mathbb{Z}_+.$$

Hence, $\{\mathrm{conv}(\mathbf{A}_j)\}_{j\in\mathbb{Z}_+}$ is a nested sequence of closed sets. Cantor's lemma yields that

$$\bigcap_{j\in\mathbb{Z}_+} \mathrm{conv}(\mathbf{A}^j) \neq \emptyset.$$

Therefore, the right-hand side of the second claim is well defined. Next, we show the equality. On the one hand,

$$\mathrm{conv}(\mathscr{S}^\infty(\mathbf{A}_0)) \subseteq \mathrm{conv}(\mathbf{A}_k) = \bigcap_{j=0}^{k} \mathrm{conv}(\mathbf{A}_j), \quad k \in \mathbb{Z}_+;$$

by taking the limit, we get

$$\mathrm{conv}(\mathscr{S}^\infty(\mathbf{A}_0)) \subseteq \bigcap_{j=0}^{\infty} \mathrm{conv}(\mathbf{A}_j).$$

On the other hand,

$$\mathrm{conv}(\mathbf{A}_k) \supseteq \bigcap_{j=0}^{\infty} \mathrm{conv}(\mathbf{A}_j), \quad k \in \mathbb{Z}_+.$$

Following the above,

$$\mathrm{conv}(\mathscr{S}^\infty(\mathbf{A}_0)) \supseteq \bigcap_{j=0}^{\infty} \mathrm{conv}(\mathbf{A}_j). \tag{6.2}$$

The convergence of the scheme implies the existence of the right-hand side of (6.2). $\square$

### 6.3 *Extension to positive-semidefinite matrix data*

The closure of the set of SPD matrices is the closed set of symmetric positive-semidefinite (SPSD) matrices. Let us consider initial data consisting of SPSD matrices. These matrices are singular and thus the definition (5.2) of the geometric mean is not valid.

We suggest using a modified version of the geometric mean,

$$G_s^{\dagger}(A,B) = A^{1/2}((A^{\dagger})^{1/2}B(A^{\dagger})^{1/2})^s A^{1/2}, \quad s \in (0,1]. \tag{6.3}$$

Here, $A^{\dagger}$ is the generalized inverse or 'pseudo' inverse matrix, also known as the Moore–Penrose inverse. For further information, see Ben-Israel & Greville (2003) and Itai & Sharon (2012) and references within. Hence, a generalized geometric mean is defined for any $A, B \succeq 0$.

For any symmetric matrix $X = Q^* DQ$, with

$$D = \begin{pmatrix} d_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & d_n \end{pmatrix},$$

we have (Penrose, 1955) $X^{\dagger} = Q^* D^{\dagger} Q$, where

$$D^{\dagger} = \begin{pmatrix} d_1^{\dagger} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & d_n^{\dagger} \end{pmatrix}, \quad d_i^{\dagger} = \begin{cases} d_i^{-1} & \text{if } d_i \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, $X^{\dagger} \succeq 0$, which yields that Definition 3.1(1,4) hold for (6.3). However, the symmetry condition does not hold. A counterexample for that is

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then

$$\begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} = G_{1/2}(A,B) \neq G_{1/2}(B,A) = B.$$

Moreover, one can verify that $G_1^{\dagger}(A,B) \neq B$. Thus, Definition 3.1(3) does not hold either. The above example illustrates that for an arbitrary set of SPSD matrices, the use of (6.3) as a matrix mean is not well defined. Thus, an additional condition is required.

A sufficient condition for (6.3) to become a matrix mean is to define it over the set of SPSD matrices with a common kernel. Henceforth, we consider initial data of SPSD matrices $\mathbf{A_0}$ such that

$$\ker(A_{0,i}) = \ker(A_{0,k}), \quad i,k \in \mathbb{Z}. \tag{6.4}$$

We note that a detailed consideration of computational issues concerning the construction of such a matrix mean is available in Bonnabel & Sepulchre (2009).

Extending the admissible mean concept to SPSD is done using the pseudo-inverse and the pseudo-determinant, that is the product of all nonzero eigenvalues (Minka, 2000). Thus, we have the following theorem, i.e.

THEOREM 6.6 The geometric scheme (3.7) using (6.3) is well defined and admissible for initial data of SPSD matrices satisfying (6.4).

*Proof.* As in the proof of Theorem 3.13 and using the common kernel, the properties follow from the geometric mean properties. □

We note that similarly to the geometric corner cutting defined by (5.2), the scheme defined by (6.3) converges to a smooth matrix curve $A(x)$ such that $A(x)$ is an SPSD matrix for any real $x$. Furthermore, this implies that Theorem 3.13 holds. The next theorem points out the connection between the two geometric corner-cutting schemes.

THEOREM 6.7 Let $\mathbf{A_0}$ be initial data consisting of SPSD matrices satisfying (6.4). Then the corner-cutting scheme defined by (6.3) and denoted by $\mathscr{S}$ satisfies

$$\lim_{\epsilon\to 0} \mathscr{S}^{\infty}(\{A_{0,i}+\epsilon M_i\}) = \mathscr{S}^{\infty}(\{A_{0,i}\}),$$

where $\{M_i\}_{i\in\mathbb{Z}}$ is a sequence of SPD matrices with $\lambda_{\min}(M_i) \geqslant m > 0$.

To prove Theorem 6.7, we use the following two lemmas. First is the Löwner partial-order version for the squeeze theorem.

LEMMA 6.8 Let $\{A_n\}_{n=0}^{\infty}$, $\{B_n\}_{n=0}^{\infty}$ and $\{C_n\}_{n=0}^{\infty}$ be an SPD-matrix sequence such that

$$A_n \preceq B_n \preceq C_n, \quad n \in \mathbb{Z}_+.$$

Assume that

$$\lim_{n\to\infty} A_n = \lim_{n\to\infty} C_n = L.$$

Then

$$\lim_{n\to\infty} B_n = L.$$

*Proof.* On the one hand, we have

$$\lim_{n\to\infty} C_n - B_n \preceq \lim_{n\to\infty} C_n - A_n = 0.$$

On the other hand,

$$\lim_{n\to\infty} B_n - A_n \preceq \lim_{n\to\infty} C_n - A_n = 0.$$

Thus,

$$0 \preceq \lim_{n\to\infty} B_n - L \preceq 0.$$ □

The second lemma is as follows.

LEMMA 6.9 Let $A$ and $B$ be two positive-semidefinite matrices such that

$$\ker(A) = \ker(B).$$

In addition, $M_A$ and $M_B$ are two SPD matrices. Then

$$G_t^{\dagger}(A,B) = \lim_{\epsilon\to 0} G_t(A+\epsilon M_A, B+\epsilon M_B), \quad t\in(0,1].$$

*Proof.* First, we note that $A+\epsilon M_A, B+\epsilon M_B$ are clearly SPD. We have

$$A+\epsilon\lambda_{\min}(M_A,M_B)I \leqslant A+\epsilon M_A \leqslant A+\epsilon\lambda_{\max}(M_A,M_B)I,$$
$$B+\epsilon\lambda_{\min}(M_A,M_B)I \leqslant B+\epsilon M_B \leqslant B+\epsilon\lambda_{\max}(M_A,M_B)I.$$

By (6.1), we get that

$$\begin{aligned} &G_t(A+\epsilon\lambda_{\min}(M_A,M_B)I, B+\epsilon\lambda_{\min}(M_A,M_B)I) \leqslant G_t(A+\epsilon M_A, B+\epsilon M_B)\\ &\quad\leqslant G_t(A+\epsilon\lambda_{\max}(M_A,M_B)I, B+\epsilon\lambda_{\max}(M_A,M_B)I). \end{aligned}$$

Moreover,

$$\begin{aligned} G_t^{\dagger}(A,B) &= \lim_{\epsilon\to 0} G_t(A+\epsilon\lambda_{\min}(M_A,M_B)I, B+\epsilon\lambda_{\min}(M_A,M_B)I)\\ &= \lim_{\epsilon\to 0} G_t(A+\epsilon\lambda_{\max}(M_A,M_B)I, B+\epsilon\lambda_{\max}(M_A,M_B)I). \end{aligned}$$

Lemma 6.8 (the Löwner squeeze lemma) completes the proof. □

REMARK 6.10 For a sequence of SPSD matrices with a common kernel, it is possible to extend the exp–log operation (4.1) to

$$A\odot B = \lim_{\epsilon\to 0} \exp((1-t)\log(A+\epsilon I) + t\log(B+\epsilon I)), \quad A,B\succeq 0.$$

One can show that the limit exists by using the common kernel property and the identities

$$\exp\begin{pmatrix} X & 0\\ 0 & Y\end{pmatrix} = \begin{pmatrix} \exp(X) & 0\\ 0 & \exp(Y)\end{pmatrix},$$

and $\exp(\log(\epsilon I)) = \epsilon I$. However, this approach has a major numerical drawback. Due to floating-point representation, for sufficiently small $\epsilon$, the output of $\exp(\log(\epsilon I))$ is not $\epsilon I$.

## 7. Bernstein operators for SPD-matrix-valued functions

Up to now, we have used matrix means to construct subdivision schemes for SPD matrices. This approach can be generalized to additional classes of positive linear approximation operators. Consider sample-based approximation operators for real-valued functions,

$$\mathscr{A}(f)(x) = \sum_{i=0}^{k} a_i(x) f(x_i), \quad a_i(x) \geqslant 0, \ \sum_{i=0}^{k} a_i(x) = 1. \tag{7.1}$$

Such operators can be rewritten as repeated weighted (binary) averages and thus can be adapted to SPD matrices using the matrix mean. As a representative, we use Bernstein operator (Lorentz, 1953). These operators, unlike the subdivision operators, are not local and depend on the entire data.

### 7.1 *Definition of Bernstein operators*

The classic Bernstein operator to approximate $f : [0, 1] \to \mathbb{R}$ is

$$B_N(f; x) = \sum_{i=0}^{N} b_{N,i}(x) f\left(\frac{i}{N}\right), \quad x \in [0, 1],$$

where

$$b_{N,i}(x) = \binom{N}{i} x^i (1-x)^{N-i}.$$

For an easy evaluation of $B_N(f; x)$, the De De Casteljau (1959) algorithm is needed:

$$\beta_i^0 = f\left(\frac{i}{N}\right), \quad i = 0, \dots, N,$$

and

$$\beta_i^{j+1} = (1-x)\beta_i^j + x\beta_{i+1}^j, \quad i = 0, \dots, N-j, \ j = 0, \dots, N-1.$$

The output is $\beta_0^N = B_N(f; x)$. The proof of the correctness of the algorithms is based on the recursive relation

$$b_{k,i}(x) = (1-x) b_{k-1,i}(x) + x b_{k-1,i-1}(x), \quad x \in [0, 1].$$

Since $\beta_i^{j+1} = (1-x)\beta_i^j + x\beta_{i+1}^j$ is the $x$-weighted average of $\beta_i^j$ and $\beta_{i+1}^j$, one can replace it with any desired $x$-weighted mean.

For the matrix case, we assume initial data $\{A_{0,i}\}_{i=0}^N$ sampled uniformly from

$$F : [0, 1] \to \mathsf{SPD}(n), \quad F\left(\frac{i}{N}\right) = A_{0,i}, \quad i = 0, \dots, N.$$

By using an admissible matrix mean, the Bernstein operator for an SPD matrix at $x \in [0, 1]$ can be evaluated with the De Casteljau recurrence relation,

$$A_{j,i} = M_x(A_{j-1,i}, A_{j-1,i+1}), \quad j = 1, \dots, N, \ i = 0, \dots, N-j, \tag{7.2}$$

and the required result is obtained in $B_N(F; x) = A_{N,0}$, i.e. after $N$ steps. We call the result a Bernstein operator for matrices.

The De Casteljau algorithm requires only a finite number of calculations. Thus, by using similar arguments as in the previous sections, one can deduce the following two corollaries.

Corollary 7.1 Bernstein operators for matrices based on an admissible matrix mean satisfy the admissible conditions of Definition 3.10, adjusted to operators of the form (7.1).

In addition, for a matrix mean that satisfies the conditions of Theorem 3.13, the spectral results of the theorem are valid for the Bernstein operator as well.

In Section 6, we proved several special properties for the operator of geometric corner cutting for a matrix. We call the *geometric Bernstein operator* the Bernstein operator for a matrix, based on a geometric mean for matrix (5.2). Then the following corollary holds.

Corollary 7.2 The results of Theorems 6.1 and 6.5 hold for the geometric Bernstein operator.

Remark 7.3 We show the results over the interval [0, 1]. However, they are valid for a general interval by replacing the De Casteljau algorithm with the de Boor algorithm (Cohen *et al.*, 2001, Chapter 4).

### 7.2 *Error bounds for the approximation by Bernstein operators*

The classical Bernstein operator converges to the sampled function as the uniform samples get denser. Next, we present a similar result for SPD-matrix Bernstein operators under two mild assumptions on the matrix mean.

The first assumption is true for the geometric mean and proved in Lemma 5.4. It is also true for any twice-differentiated average for numbers (Dyn & Goldman, 2011, Proposition 2.6). We generalize this approach for a general matrix mean.

Assumption 7.4 Let $A, B \in \mathsf{SPD}(n)$, $\Delta = \|A - B\|$ and $M_t$ be an admissible matrix mean. Then there exists a constant $C_M$, independent of $\Delta$, such that

$$\|M_t(A, B) - L_t(A, B)\| < C_M \Delta^2, \quad t \in [0, 1].$$

The second assumption is the metric property (3.11) in respect of the Euclidean distance, $d(X, Y) = \|X - Y\|$. Note that the exp–log mean (4.1) and the geometric mean for matrix (5.2) both satisfy this assumption.

Assumption 7.5 For any $A, B \in \mathsf{SPD}(n)$,

$$\|M_t(A, B) - A\| = t\|A - B\|, \quad t \in [0, 1].$$

The error bound is as follows.

Theorem 7.6 Let $F : [0, 1] \to \mathsf{SPD}(n)$ be a Lipschitz (elementwise) continuous matrix function, namely there exists a constant $C_F$ such that

$$\|F(x) - F(y)\| < C_F|x - y|, \quad x, y \in \mathbb{R}.$$

Then

$$\|F(x) - B_N(F; x)\| \leqslant \mathcal{O}\left(\frac{1}{\sqrt{N}}\right),$$

where $B_N$ is a Bernstein operator, defined using any admissible matrix mean and obtained by the De Casteljau algorithm.

To prove Theorem 7.6, we use the following two lemmas.

LEMMA 7.7 Let $\{A_i\}_{i=0}^N, \{B_i\}_{i=0}^N \subset \mathsf{SPD}(n)$ and $\Delta = \max_{i\in\{0,\ldots,N\}} \|A_i - B_i\|$. Then

$$\|L_t(A_i, A_{i+1}) - L_t(B_i, B_{i+1})\| < \Delta, \quad t \in [0,1].$$

*Proof.* By the definition of the arithmetic mean,

$$\begin{aligned}\|L_t(A_i, A_{i+1}) - L_t(B_i, B_{i+1})\| &= \|(1-t)A_i + tA_{i+1} - ((1-t)B_i + tB_{i+1})\| \\ &\leqslant (1-t)\|A_i - B_i\| + t\|A_{i+1} - B_{i+1}\| < \Delta.\end{aligned}$$ □

LEMMA 7.8 Let $A, B, C \in \mathsf{SPD}(n)$ such that

$$\|A - B\|, \|A - C\| \leqslant \delta, \quad \delta > 0.$$

Assume that $M_t$ satisfies Assumptions 7.4 and 7.5. Then

$$\|M_t(A, B) - M_t(B, C)\| \leqslant \delta.$$

*Proof.* By the triangle inequality and Assumption 7.5,

$$\|M_t(A, B) - B + B - M_t(B, C)\| \leqslant (\xi_{1-t} + \xi_t)\delta \leqslant \delta.$$ □

A result of the previous lemma is that if we start the De Casteljau algorithm with data sampled uniformly, each of the levels withholds the same uniform distance.

In Theorem 7.6, we show that Assumption 7.4 is in fact a proximity condition to the case of linear arithmetic means of matrices. In the latter, the convergence is entrywise. Thus, one can apply the theorems of the Bernstein polynomials theory.

*Proof of Theorem* 7.6. We use the notation of (7.2) with a minor addition: we denote by $A^L_{j,i}$ the refined $j$th level, using the arithmetic matrix mean $L_t$. The key argument is

$$\|A^L_{N,0} - A_{N,0}\| < C_M N \Delta^2, \quad N \in \mathbb{N}. \tag{7.3}$$

Proof by induction. For the initial step, Assumption 7.4 ensures that $\|A^L_{1,0} - A_{1,0}\| < C_M \Delta^2$. For $N = 2$ we use the following:

$$\|L_t(A^L_{1,0}, A^L_{1,1}) - L_t(A_{1,0}, A_{1,1})\| < C_M \Delta^2.$$

Thus, the quantity $\|L_t(A^L_{1,0}, A^L_{1,1}) - M_t(A_{1,0}, A_{1,1})\|$ is bounded by

$$\|L_t(A^L_{1,0}, A^L_{1,1}) - L_t(A_{1,0}, A_{1,1}) + L_t(A_{1,0}, A_{1,1}) - M_t(A_{1,0}, A_{1,1})\| < 2C_M \Delta^2.$$

Consider the claim to be true for $N = m$:

$$\|A^L_{m,0} - A_{m,0}\| < mC_M \Delta^2.$$

Then for $N = m+1$, $\|A^L_{m+1,0} - A_{m+1,0}\|$, we have

$$\|L_t(A^L_{m,0}, A^L_{m,1}) - L_t(A_{m,0}, A_{m,1}) + L_t(A_{m,0}, A_{m,1}) - M_t(A_{m,0}, A_{m,1})\| < mC_M \Delta^2 + C_M \Delta^2,$$

which proves (7.3). By using the argument of Lemma 7.8, we get

$$\Delta = \|A_{k,i} - A_{k,i+1}\|, \quad k = 0, \ldots, N.$$

The Lipschitz condition of $F$ combined with norm equivalence yields an entrywise Lipschitz condition. Therefore, we can use the classical Bernstein convergence (Mathé, 1999) to get

$$\|F(x) - A_{N,0}\| \leqslant \mathscr{O}\left(\frac{1}{\sqrt{N}}\right).$$

The data are sampled uniformly from a univariate parametric curve of matrices. Thus, we deduce that $\Delta = \mathscr{O}(1/N)$. Hence, for sufficiently large $N$,

$$\|F(x) - A^L_{N,0} + A^L_{N,0} - A_{N,0}\| \leqslant \mathscr{O}\left(\frac{1}{\sqrt{N}}\right) + \mathscr{O}\left(\frac{1}{N}\right) = \mathscr{O}\left(\frac{1}{\sqrt{N}}\right). \qquad \square$$

**Acknowledgements**

We wish to express our gratitude to Professor Nira Dyn, who supervised this work along its various stages, for her substantial advice, involvement and encouragement.

## References

Ando, T., Li, C.-K. & Mathias, R. (2004) Geometric means. *Linear Algebra Appl.*, **385**, 305–334.

Arsigny, V., Fillard, P., Pennec, X. & Ayache, N. (2007) Geometric means in a novel vector space structure on symmetric positive definite matrices. *SIAM J. Matrix Anal. Appl.*, **29**, 328–347.

Ben-Israel, A. & Greville, T. N. E. (2003) *Generalized Inverses* (2nd edn). CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Vol. 15. New York: Springer, pp. xvi+420. Theory and Application.

Bini, D. A., Meini, B. & Poloni, F. (2010) An effective matrix geometric mean satisfying the Ando–Li–Mathias properties. *Math. Comp.*, **79**, 437–452.

Bonnabel, S. & Sepulchre, R. (2009) Riemannian metric and geometric mean for positive semidefinite matrices of fixed rank. *SIAM J. Matrix Anal. Appl.*, **31**, 1055–1070.

Boyd, S. & Vandenberghe, L. (2004) *Convex Optimization*. Cambridge: Cambridge University Press, pp. xiv+716.

Chaikin, G. (1974) An algorithm for high-speed curve generation. *Comput. Graph. Image Process*, **3**, 346–349.

Cohen, E., Riesenfeld, R. F. & Elber, G. (2001) *Geometric Modeling with Splines*. Natick, MA: A K Peters Ltd., pp. xxii+616. An introduction, with a foreword by Tom Lyche.

Daróczy, Z., Maksa, G. & Páles, Z. (2006) Functional equations involving means and their Gauss composition. *Proc. Amer. Math. Soc.*, **134**, 521–530.

de Boor, C. (1987) Cutting corners always works. *Comput. Aided Geom. Design*, **4**, 125–131. Special issue on topics in computer aided geometric design (Wolfenbüttel, 1986).

de Casteljau, P. (1959) Outillages méthodes calcul. *Technical Report*. Paris: Andre Citroen Automobiles SA.

Dyn, N. (2006) Linear subdivision schemes for the refinement of geometric objects. *International Congress of Mathematicians*, vol. III. Zürich: Eur. Math. Soc., pp. 1201–1226.

Dyn, N. & Farkhi, E. (2001) Spline subdivision schemes for compact sets with metric averages. *Trends in Approx. Theory*, 95–104.

Dyn, N. & Goldman, R. (2011) Convergence and smoothness of nonlinear Lane–Riesenfeld algorithms in the functional setting. *Found. Comput. Math.*, **11**, 79–94.

Dyn, N., Grohs, P. & Wallner, J. (2010) Approximation order of interpolatory nonlinear subdivision schemes. *J. Comput. Appl. Math.*, **233**, 1697–1703.

DYN, N. & LEVIN, D. (2002) Subdivision schemes in geometric modelling. *Acta Numer.*, **11**, 73–144.

EBNER, O. (2012) Convergence of refinement schemes on metric spaces. *Proc. Amer. Math. Soc.* To appear.

GOLUB, G. & VAN LOAN, C. (1996) *Matrix Computations*, vol. 3. Baltimore: Johns Hopkins University Press.

GROHS, P. (2009) Smoothness of interpolatory multivariate subdivision in Lie groups. *IMA J. Numer. Anal.*, **29**, 760–772.

GROHS, P. (2010a) Approximation order from stability for nonlinear subdivision schemes. *J. Approx. Theory*, **162**, 1085–1094.

GROHS, P. (2010b) A general proximity analysis of nonlinear subdivision schemes. *SIAM J. Math. Anal.*, **42**, 729–750.

GROHS, P. & WALLNER, J. (2008) Log-exponential analogues of univariate subdivision schemes in Lie groups and their smoothness properties. *Approximation Theory XII: San Antonio 2007*. Mod. Methods Math. Brentwood, TN: Nashboro Press, pp. 181–190.

HALL, B. C. (2003) Lie groups, Lie algebras, and representations. *Graduate Texts in Mathematics*, vol. 222. New York: Springer, pp. xiv+351. An elementary introduction.

HAUKE, J. & MARKIEWICZ, A. (1994) On orderings induced by the Löwner partial ordering. *Appl. Math. (Warsaw)*, **22**, 145–154.

HORN, R. & JOHNSON, C. (1990) *Matrix Analysis*. Cambridge: Cambridge University Press.

ITAI, U. & SHARON, N. (2012) Subdivision schemes for positive definite matrices. *Found. Comput. Math.*, 1–23.

KUBO, F. & ANDO, T. (1980) Means of positive linear operators. *Math. Ann.*, **246**, 205–224.

LANE, J. M. & RIESENFELD, R. F. (1980) A theoretical development for the computer generation and display of piecewise polynomial surfaces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35–46.

LAWSON, J. & LIM, Y. (2001) The geometric mean, matrices, metrics, and more. *Amer. Math. Monthly*, **108**, 797–812.

LI, C.-K. & MATHIAS, R. (2000) Extremal characterizations of the Schur complement and resulting inequalities. *SIAM Rev.*, **42**, 233–246 (electronic).

LORENTZ, G. G. (1953) Bernstein polynomials. *Mathematical Expositions*, vol. 8. Toronto: University of Toronto Press, pp. x+130.

MATHÉ, P. (1999) Approximation of Hölder continuous functions by Bernstein polynomials. *Amer. Math. Monthly*, **106**, 568–574.

MATHIAS, R. (1993) Approximation of matrix-valued functions. *SIAM J. Matrix Anal. Appl.*, **14**, 1061–1063.

MINKA, T. (2000) Inferring a Gaussian distribution. *Technical Report*. MIT.

NAVAYAZDANI, E. & YU, T. P.-Y. (2010) On Donoho's log-exp subdivision scheme: choice of retraction and time-symmetry. *Multiscale Model. Simul.*, **9**, 1801–1828.

PENROSE, R. (1955) A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.*, vol. 51. Cambridge: Cambridge University Press, p. C655.

PETZ, D. (1994) A survey of certain trace inequalities. *Functional Analysis and Operator Theory (Warsaw, 1992)*, vol. 30. Banach Center Publ. Warsaw: Polish Acad. Sci., pp. 287–298.

RAHMAN, I. U., DRORI, I., STODDEN, V. C., DONOHO, D. L. & SCHRÖDER, P. (2005) Multiscale representations for manifold-valued data. *Multiscale Model. Simul.*, **4**, 1201–1232.

SCHAEFER, S., VOUGA, E. & GOLDMAN, R. (2008) Nonlinear subdivision through nonlinear averaging. *Comput. Aided Geom. Design*, **25**, 162–180.

SHINGEL, T. (2009) Interpolation in special orthogonal groups. *IMA J. Numer. Anal.*, **29**, 731–745.

SPIVAK, M. (2005) *A Comprehensive Introduction to Differential Geometry*, vol. 1, 3rd edn. Berkeley, CA: Publish or Perish.

WALLNER, J. & DYN, N. (2005) Convergence and $C^1$ analysis of subdivision schemes on manifolds by proximity. *Comput. Aided Geom. Design*, **22**, 593–622.

XIE, G. & YU, T. P.-Y. (2007) Smoothness equivalence properties of manifold-valued data subdivision schemes based on the projection approach. *SIAM J. Numer. Anal.*, **45**, 1200–1225 (electronic).

XIE, G. & YU, T. P.-Y. (2010) Smoothness equivalence properties of interpolatory Lie group subdivision schemes. *IMA J. Numer. Anal.*, **30**, 731–750.