A PANORAMA OF POSITIVITY

ALEXANDER BELTON, DOMINIQUE GUILLOT, APOORVA KHARE,
AND MIHAI PUTINAR

ABSTRACT. This survey contains a selection of topics unified by the concept of positive semi-definiteness (of matrices or kernels), reflecting natural constraints imposed on discrete data (graphs or networks) or continuous objects (probability or mass distributions). We put emphasis on entrywise operations which preserve positivity, in a variety of guises. Techniques from harmonic analysis, function theory, operator theory, statistics, combinatorics, and group representations are invoked. Some partially forgotten classical roots in metric geometry and distance transforms are presented with comments and full bibliographical references. Modern applications to high-dimensional covariance estimation and regularization are included.

CONTENTS

1. Introduction

Matrix positivity, or positive semidefiniteness, is one of the most wide-reaching concepts in mathematics, old and new. Positivity of a matrix is as natural as positivity of mass in statics or positivity of a probability distribution. It is a notion which has attracted the attention of many great minds. Yet, after at least two centuries of research, positive matrices still hide enigmas and raise challenges for the working mathematician.

The vitality of matrix positivity comes from its breadth, having many theoretical facets and also deep links to mathematical modelling. It is not our aim here to pay homage to matrix positivity in the large. Rather, the

present survey, split for technical reasons into two parts, has a limited but carefully chosen scope.

Our panorama focuses on entrywise transforms of matrices which preserve their positive character. In itself, this is a rather bold departure from the dogma that canonical transformations of matrices are not those that operate entry by entry. Still, this apparently esoteric topic reveals a fascinating history, abundant characteristic phenomena and numerous open problems. Each class of positive matrices or kernels (regarding the latter as continuous matrices) carries a specific toolbox of internal transforms. Positive Hankel forms or Toeplitz kernels, totally positive matrices, and group-invariant positive definite functions all possess specific *positivity preservers*. As we see below, these have been thoroughly studied for at least a century.

One conclusion of our survey is that the classification of positivity preservers is accessible in the dimension-free setting, that is, when the sizes of matrices are unconstrained. In stark contrast, precise descriptions of positivity preservers in fixed dimension are elusive, if not unattainable with the techniques of modern mathematics. Furthermore, the world of applications cares much more about matrices of fixed size than in the free case. The accessibility of the latter was by no means a sequence of isolated, simple observations. Rather, it grew organically out of distance geometry, and spread rapidly through harmonic analysis on groups, special functions, and probability theory. The more recent and highly challenging path through fixed dimensions requires novel methods of algebraic combinatorics and symmetric functions, group representations, and function theory.

As well as its beautiful theoretical aspects, our interest in these topics is also motivated by the statistics of big data. In this setting, functions are often applied entrywise to covariance matrices, in order to induce sparsity and improve the quality of statistical estimators (see [72, 73, 114]). Entrywise techniques have recently increased in popularity in this area, largely because of their low computational complexity, which makes them ideal to handle the ultra high-dimensional datasets arising in modern applications. In this context, the dimensions of the matrices are fixed, and correspond to the number of underlying random variables. Ensuring that positivity is preserved by these entrywise methods is critical, as covariance matrices must be positive semidefinite. Thus, there is a clear need to produce characterizations of entrywise preservers, so that these techniques are widely applicable and mathematically justified. We elaborate further on this in the second part of the survey.

We conclude by remarking that, while we have tried to be comprehensive in our coverage of the field of matrix positivity and the entrywise calculus, our panorama is far from being complete. We apologize for any omissions.

2. From metric geometry to matrix positivity

2.1. **Distance geometry.** During the first decade of the 20th century, the concept of a metric space emerged from the works of Fréchet and Hausdorff, each having different and well-anchored roots, in function spaces and in set theory and measure theory. We cannot think today of modern mathematics and physics without referring to metric spaces, which touch areas as diverse as economics, statistics, and computer science. Distance geometry is one of the early and ever-lasting by-products of metric-space theory. One of the key figures of the Vienna Circle, Karl Menger, started a systematic study in the 1920s of the geometric and topological features of spaces that are intrinsic solely to the distance they carry. Menger published his findings in a series of articles having the generic name "Untersuchungen über allgemeine Metrik," the first one being [99]; see also his synthesis [100]. His work was very influential in the decades to come [23], and by a surprising and fortunate stroke not often encountered in mathematics, Menger's distance geometry has been resurrected in recent times by practitioners of convex optimization and network analysis [39, 95].

Let (X, ρ) be a metric space. One of the naive, yet unavoidable, questions arising from the very beginning concerns the nature of operations $\phi(\rho)$ which may be performed on the metric and which enhance various properties of the topological space X. We all know that $\rho/(\rho + 1)$ and ρ^γ, if $\gamma \in (0, 1)$, also satisfy the axioms of a metric, with the former making it bounded. Less well known is an observation due to Blumenthal, that the new metric space (X, ρ^γ) has the four-point property if $\gamma \in (0, 1/2]$: every four-point subset of X can be embedded isometrically into Euclidean space [23, Section 49].

Metric spaces which can be embedded isometrically into Euclidean space, or into infinite-dimensional Hilbert space, are, of course, distinguished and desirable for many reasons. We owe to Menger a definitive characterization of this class of metric spaces. The core of Menger's theorem, stated in terms of certain matrices built from the distance function (known as Cayley–Menger matrices) was slightly reformulated by Fréchet and cast in the following simple form by Schoenberg.

Theorem 2.1 (Schoenberg [120])**.** *Let $d \geq 1$ be an integer and let (X, ρ) be a metric space. An $(n + 1)$-tuple of points x_0, x_1, ..., x_n in X can be isometrically embedded into Euclidean space $\mathbb{R}^d$, but not into $\mathbb{R}^{d-1}$, if and only if the matrix*

$$[\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2]_{j,k=1}^n,$$

is positive semidefinite with rank equal to d.

Proof. This is surprisingly simple. Necessity is immediate, since the Euclidean norm and scalar product in $\mathbb{R}^d$ give that

$$\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2$$
$$= \|x_0 - x_j\|^2 + \|x_0 - x_k\|^2 - \|(x_0 - x_j) - (x_0 - x_k)\|^2$$
$$= 2\langle x_0 - x_j, x_0 - x_k \rangle,$$

and the latter are the entries of a positive semidefinite Gram matrix of rank less than or equal to d.

For the other implication, we consider first a full-rank $d \times d$ matrix associated with a $(d+1)$-tuple. The corresponding quadratic form

$$Q(\lambda) = \frac{1}{2} \sum_{j,k=1}^{d} (\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2) \lambda_j \lambda_k$$

is positive definite. Hence there exists a linear change of variables

$$\lambda_k = \sum_{j=1}^{d} a_{jk} \mu_j \qquad (1 \le j \le d)$$

such that

$$Q(\lambda) = \mu_1^2 + \mu_2^2 + \cdots + \mu_d^2.$$

Interpreting $(\mu_1, \mu_2, \ldots, \mu_d)$ as coordinates in $\mathbb{R}^d$, the standard simplex with vertices

$$e_0 = (0, \ldots, 0), \quad e_1 = (1, 0, \ldots, 0), \quad \ldots, \quad e_d = (0, \ldots, 0, 1)$$

has the corresponding quadratic form (of distances) equal to $\mu_1^2 + \mu_2^2 + \cdots + \mu_d^2$. Now we perform the coordinate change $\mu_j \mapsto \lambda_j$. Specifically, set $P_0 = 0$ and let $P_j \in \mathbb{R}^d$ be the point with coordinates $\lambda_j = 1$ and $\lambda_k = 0$ if $k \ne j$. Then one identifies distances:

$$\|P_0 - P_j\| = \rho(x_0, x_j) \qquad (0 \le j \le d)$$
$$\text{and} \quad \|P_j - P_k\| = \rho(x_j, x_k) \qquad (1 \le j, k \le d).$$

The remaining case with $n > d$ can be analyzed in a similar way, after taking an appropriate projection. $\qquad \square$

In the conditions of the theorem, fixing a "frame" of d points and letting the $(d+1)$-th point float, one obtains an embedding of the full metric space (X, ρ) into $\mathbb{R}^d$. This idea goes back to Menger, and it led, with Schoenberg's touch, to the following definitive statement. Here and below, all Hilbert spaces are assumed to be separable.

Corollary 2.2 (Schoenberg [120], following Menger). *A separable metric space (X, ρ) can be isometrically embedded into Hilbert space if and only if, for every $(n+1)$-tuple of points $(x_0, x_1, \ldots, x_n)$ in X, where $n \ge 2$, the matrix*

$$[\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2]_{j,k=1}^{n}$$

is positive semidefinite.

The notable aspect of the two previous results is the interplay between purely geometric concepts and matrix positivity. This will be a recurrent theme of our survey.

2.2. **Spherical distance geometry.** One can specialize the embedding question discussed in the previous section to submanifolds of Euclidean space. A natural choice is the sphere.

For two points x and y on the unit sphere $S^{d-1} \subset \mathbb{R}^d$, the rotationally invariant distance between them is

$$\rho(x, y) = \sphericalangle(x, y) = \arccos\langle x, y\rangle,$$

where the angle between the two vectors is measured on a great circle and is always less than or equal to π.

A straightforward application of the simple, but central, Theorem 2.1] yields the following result.

Theorem 2.3 (Schoenberg [120]). *Let (X, ρ) be a metric space and let $(x_1, \ldots, x_n)$ be an n-tuple of points in X. For any integer $d \geq 2$, there exists an isometric embedding of $(x_1, \ldots, x_n)$ into S^{d-1} endowed with the geodesic distance but not S^{d-2} if and only if*

$$\rho(x_j, x_k) \leq \pi \qquad (1 \leq j, k \leq n)$$

and the matrix $\left[\cos\rho(x_j, x_k)\right]_{j,k=1}^n$ is positive semidefinite of rank d.

Indeed, the necessity is assured by choosing x_0 to be the origin in $\mathbb{R}^d$. In this case,

$$\begin{aligned}
\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2 &= \|x_j\|^2 + \|x_k\|^2 - \|x_j - x_k\|^2 \\
&= 2\langle x_j, x_k\rangle \\
&= 2\cos\rho(x_j, x_k).
\end{aligned}$$

The condition is also sufficient, by possibly adding an external point x_0 to the metric space, subject to the constraints that $\rho(x_0, x_j) = 1$ for all j. The details can be found in [120].[1]

2.3. **Distance transforms.** A notable step forward in the study of the existence of isometric embeddings of a metric space into Euclidean or Hilbert space was made by Schoenberg. In a series of articles [121, 123, 124, 136], he changed the set-theoretic lens of Menger, by initiating a harmonic-analysis interpretation of this embedding problem. This was a major turning point, with long-lasting, unifying, and unexpected consequences.

---

[1]An alternate proof of sufficiency is to note that $A := [\cos\rho(x_j, x_k)]_{j,k=1}^n$ is a Gram matrix of rank r, hence equal to $B^T B$ for some $r \times n$ matrix B with unit columns. Denoting these columns by $\mathbf{b}_1, \ldots, \mathbf{b}_n \in S^{r-1}$, the map $x_j \mapsto b_j$ is an isometry since $\rho(x_j, x_k)$ and $\sphericalangle(y_j, y_k) \in [0, \pi]$. Moreover, since A has rank r, the $\mathbf{b}_j$ cannot all lie in a smaller-dimensional sphere.

We return to a separable metric space (X, ρ) and seek distance-function transforms $\rho \mapsto \phi(\rho)$ which enhance the geometry of X, to the extent that the new metric space $(X, \phi(\rho))$ is isometrically equivalent to a subspace of Hilbert space. Schoenberg launched this whole new chapter from the observation that the Euclidean norm is such that the matrix

$$[\exp(-\|x_j - x_k\|^2)]_{j,k=1}^N$$

is positive semidefinite for any choice of points $x_1, \ldots, x_N$ in the ambient space. Once again, we see the presence of matrix positivity. While this claim may not be obvious at first sight, it is accessible once we recall a key property of Fourier transforms.

An even function $f : \mathbb{R}^d \to \mathbb{C}$ is said to be *positive definite* if the complex matrix $[f(x_j - x_k)]_{j,k=1}^N$ is positive semidefinite for any $N \geq 1$ and any choice of points $x_1, \ldots, x_N \in \mathbb{R}^d$. We will call $f(x-y)$ a *positive semidefinite kernel* on $\mathbb{R}^d \times \mathbb{R}^d$ in this case. (See [132] for a comprehensive survey of this class of maps.)

Bochner's theorem [25] characterizes positive definite functions on $\mathbb{R}^d$ as Fourier transforms of even positive measures of finite mass:

$$f(\xi) = \int e^{-ix\cdot\xi} \, d\mu(x).$$

Indeed,

$$f(\xi - \eta) = \int e^{-ix\cdot\xi} e^{ix\cdot\eta} \, d\mu(x)$$

is a positive semidefinite kernel because it is the average over μ of the positive kernel $(\xi, \eta) \mapsto e^{-ix\cdot\xi} e^{ix\cdot\eta}$. Since the Gaussian e^{-x^2} is the Fourier transform of itself (modulo constants), it turns out that it is a positive definite function on $\mathbb{R}$, whence $\exp(-\|x\|^2)$ has the same property as a function on $\mathbb{R}^d$. Taking one step further, the function $x \mapsto \exp(-\|x\|^2)$ is positive definite on any Hilbert space.

With this preparation we are ready for a second characterization of metric subspaces of Hilbert space.

Theorem 2.4 (Schoenberg [123]). *A separable metric space (X, ρ) can be embedded isometrically into Hilbert space if and only if the kernel*

$$X \times X \to (0, \infty); \ (x, y) \mapsto \exp(-\lambda^2 \rho(x, y)^2)$$

is positive semidefinite for all $\lambda \in \mathbb{R}$.

Proof. Necessity follows from the positive definiteness of the Gaussian discussed above. (We also provide an elementary proof below; see Lemma 5.7 and the subsequent discussion). To prove sufficiency, we recall the Menger–Schoenberg characterization of isometric subspaces of Hilbert space. We have to derive, from the positivity assumption, the positivity of the matrix

$$[\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2]_{j,k=1}^n.$$

Elementary algebra transforms this constraint into the requirement that

$$\sum_{j,k=0}^{n} \rho(x_j, x_k)^2 c_j c_k \leq 0 \qquad \text{whenever } \sum_{j=0}^{n} c_j = 0.$$

By expanding $\exp(-\lambda^2 \rho(x_j, x_k)^2)$ as a power series in λ^2, and invoking the positivity of the exponential kernel, we see that

$$0 \leq -\lambda^2 \sum_{j,k=0}^{n} \rho(x_j, x_k)^2 c_j c_k + \frac{\lambda^4}{2} \sum_{j,k=0}^{n} \rho(x_j, x_k)^4 c_j c_k - \cdots$$

for all $\lambda > 0$. Hence the coefficient of $-\lambda^2$ is non-positive. $\qquad \square$

The flexibility of the Fourier-transform approach is illustrated by the following application, also due to Schoenberg [123].

Corollary 2.5. *Let H be a Hilbert space with norm $\|\cdot\|$. For every $\delta \in (0,1)$, the metric space $(H, \|\cdot\|^{\delta})$ is isometric to a subspace of a Hilbert space.*

Proof. Note first the identity

$$\xi^{\alpha} = c_{\alpha} \int_0^{\infty} (1 - e^{-s^2 \xi^2}) s^{-1-\alpha} \, ds \qquad (\xi > 0, \ 0 < \alpha < 2),$$

where c_{α} is a normalization constant. Consequently,

$$\|x - y\|^{\alpha} = c_{\alpha} \int_0^{\infty} (1 - e^{-s^2 \|x-y\|^2}) s^{-1-\alpha} \, ds.$$

Let $\delta = \alpha/2$. For points $x_0, x_1, \ldots, x_n$ in H and weights $c_0, c_1, \ldots, c_n$ satisfying

$$c_0 + c_1 + \cdots + c_n = 0,$$

it holds that

$$\sum_{j,k=0}^{n} \|x_j - x_k\|^{2\delta} c_j c_k = -c_{\alpha} \int_0^{\infty} \sum_{j,k=0}^{n} c_j c_k e^{-s^2 \|x_j - x_k\|^2} s^{-1-\alpha} \, ds \leq 0,$$

and the proof is complete. $\qquad \square$

Several similar consequences of the Fourier-transform approach are within reach. For instance, Schoenberg observed in the same article that if the L^p norm is raised to the power γ, where $0 < \gamma \leq p/2$ and $1 \leq p \leq 2$, then $L^p(0,1)$ is isometrically embeddable into Hilbert space.

2.4. **Altering Euclidean distance.** By specializing the theme of the previous section to Euclidean space, Schoenberg and von Neumann discovered an arsenal of powerful tools from harmonic analysis that were able to settle the question of whether Euclidean space equipped with the altered distance $\phi(\|x - y\|)$ may be isometrically embedded into Hilbert space [122, 136]. The key ingredients are characterizations of Laplace and Fourier transforms of positive measures, that is, Bernstein's completely monotone functions [17] and Bochner's positive definite functions [25].

Here we present some highlights of the Schoenberg–von Neumann framework. First, we focus on an auxiliary class of distance transforms. A real continuous function ϕ is called *positive definite in Euclidean space* $\mathbb{R}^d$ if the kernel

$$(x, y) \mapsto \phi\big(\|x - y\|\big)$$

is positive semidefinite. Bochner's theorem and the rotation-invariance of this kernel prove that such a function ϕ is characterized by the representation

$$\phi(t) = \int_0^\infty \Omega_d(tu)\,\mathrm{d}\mu(u),$$

where μ is a positive measure and

$$\Omega_d\big(\|x\|\big) = \int_{\|\xi\|=1} e^{\mathrm{i}x\cdot\xi}\,\mathrm{d}\sigma(\xi),$$

with σ the normalized area measure on the unit sphere in $\mathbb{R}^d$; see [122, Theorem 1]. By letting d tend to infinity, one finds that positive definite functions on infinite-dimensional Hilbert space are precisely of the form

$$\phi(t) = \int_0^\infty e^{-t^2 u^2}\,\mathrm{d}\mu(u),$$

with μ a positive measure on the semi-axis. Notice that positive definite functions in $\mathbb{R}^d$ are not necessarily differentiable more than $(d-1)/2$ times, while those which are positive definite in Hilbert space are smooth and even complex analytic in the sector $|\arg t| < \pi/4$.

The class of functions f which are continuous on $\mathbb{R}_+ := [0, \infty)$, smooth on the open semi-axis $(0, \infty)$, and such that

$$(-1)^n f^{(n)}(t) \geq 0 \qquad \text{for all } t > 0$$

was studied by S. Bernstein, who proved that they coincide with Laplace transforms of positive measures on $\mathbb{R}_+$:

$$f(t) = \int_0^\infty e^{-tu}\,\mathrm{d}\mu(u). \tag{2.1}$$

Such functions are called *completely monotonic* and have proved highly relevant for probability theory and approximation theory; see [17] for the foundational reference. Thus we have obtained a valuable equivalence.

Theorem 2.6 (Schoenberg)**.** *A function f is completely monotone if and only if $t \mapsto f(t^2)$ is positive definite on Hilbert space.*

The direct consequences of this apparently innocent observation are quite deep. For example, the isometric-embedding question for altered Euclidean distances is completely answered via this route. The following results are from [122] and [136].

Theorem 2.7 (Schoenberg–von Neumann)**.** *Let H be a separable Hilbert space with norm $\|\cdot\|$.*

(1) *For any integers $n \geq d > 1$, the metric space $(\mathbb{R}^d, \phi(\|\cdot\|))$ may be isometrically embedded into $(\mathbb{R}^n, \|\cdot\|)$ if and only if $\phi(t) = ct$ for some $c > 0$.*

(2) *The metric space $(\mathbb{R}^d, \phi(\|\cdot\|))$ may be isometrically embedded into H if and only if*

$$\phi(t)^2 = \int_0^\infty \frac{1 - \Omega_d(tu)}{u^2} \, \mathrm{d}\mu(u),$$

where μ is a positive measure on the semi-axis such that

$$\int_1^\infty \frac{1}{u^2} \, \mathrm{d}\mu(u) < \infty.$$

(3) *The metric space $(H, \phi(\|\cdot\|))$ may be isometrically embedded into H if and only if*

$$\phi(t)^2 = \int_0^\infty \frac{1 - e^{-t^2 u}}{u} \, \mathrm{d}\mu(u),$$

where μ is a positive measure on the semi-axis such that

$$\int_1^\infty \frac{1}{u} \, \mathrm{d}\mu(u) < \infty.$$

In von Neumann and Schoenberg's article [136], special attention is paid to the case of embedding a modified distance on the line into Hilbert space. This amounts to characterizing all *screw lines* in a Hilbert space H: the continuous functions

$$f : \mathbb{R} \to H; \ t \mapsto f_t$$

with the translation-invariance property

$$\|f_s - f_t\| = \|f_{s+r} - f_{t+r}\| \qquad \text{for all } s, r, t \in \mathbb{R}.$$

In this case, the gauge function ϕ is such that $\phi(t-s) = \|f_s - f_t\|$ and $t \mapsto f_t$ provides the isometric embedding of $(\mathbb{R}, \phi(|\cdot|))$ into H. Von Neumann seized the opportunity to use Stone's theorem on one-parameter unitary groups, together with the spectral decomposition of their unbounded self-adjoint generators, to produce a purely operator-theoretic proof of the following result.

Corollary 2.8. *The metric space $(\mathbb{R}, \phi(|\cdot|))$ isometrically embeds into Hilbert space if and only if*

$$\phi(t)^2 = \int_0^\infty \frac{\sin^2(tu)}{u^2} \, \mathrm{d}\mu(u) \qquad (t \in \mathbb{R}),$$

where μ is a positive measure on $\mathbb{R}_+$ satisfying

$$\int_1^\infty \frac{1}{u^2} \, \mathrm{d}\mu(u) < \infty.$$

Moreover, in the conditions of the corollary, the space $(\mathbb{R}, \phi(|\cdot|))$ embeds isometrically into $\mathbb{R}^d$ if and only if the measure μ consists of finitely many point masses, whose number is roughly $d/2$; see [136, Theorem 2] for the precise statement. To give a simple example, consider the function

$$\phi : \mathbb{R} \to \mathbb{R}_+; \ t \mapsto \sqrt{t^2 + \sin^2 t}.$$

This is indeed a screw function, because

$$\phi(t-s)^2 = (t-s)^2 + \sin^2(t-s)$$
$$= (t-s)^2 + \frac{1}{4}\big(\cos(2t) - \cos(2s)\big)^2 + \frac{1}{4}\big(\sin(2t) - \cos(2s)\big)^2.$$

Note that a screw line is periodic if and only if it is not injective. Furthermore, one may identify screw lines with period $\tau > 0$ by the geometry of the support of the representing measure: this support must be contained in the lattice $(\pi/\tau)\mathbb{Z}_+$, where $\mathbb{Z}_+ := \mathbb{Z} \cap \mathbb{R}_+ = \{0, 1, 2, \dots\}$. Consequently, all periodic screw lines in Hilbert space have a gauge function ϕ such that

$$\phi(t)^2 = \sum_{k=1}^{\infty} c_k \sin^2(k\pi t/\tau), \tag{2.2}$$

where $c_k \geq 0$ and $\sum_{k=1}^{\infty} c_k < \infty$; see [136, Theorem 5].

2.5. **Positive definite functions on homogeneous spaces.** Having resolved the question of isometrically embedding Euclidean space into Hilbert space, a natural desire was to extend the analysis to other special manifolds with symmetry. This was done almost simultaneously by Schoenberg on spheres [125] and by Bochner on compact homogeneous spaces [26].

Let X be a compact space endowed with a transitive action of a group G and an invariant measure. We seek G-invariant distance functions, and particularly those which identify X with a subspace of a Hilbert space. To simplify terminology, we call the latter *Hilbert distances*.

The first observation of Bochner is that a G-invariant symmetric kernel $f : X \times X \to \mathbb{R}$ satisfies the Hilbert-space embeddability condition,

$$\sum_{k=0}^{n} c_k = 0 \quad \Longrightarrow \quad \sum_{j,k=0}^{n} f(x_j, x_k) c_j c_k \geq 0,$$

for all choices of weights c_j and points $x_j \in X$, if and only if f is of the form

$$f(x, y) = h(x, y) - h(x_0, x_0) \qquad (x, y \in X),$$

where h is a G-invariant positive definite kernel and x_0 is a point of X. One implication is clear. For the other, we start with a G-invariant function f subject to the above constraint and prove, using G-invariance and integration over X, the existence of a constant c such that $h(x, y) = f(x, y) + c$ is a positive semidefinite kernel. This gives the following result.

Theorem 2.9 (Bochner [26]). *Let X be a compact homogeneous space. A continuous invariant function ρ on $X \times X$ is a Hilbert distance if and only if there exists a continuous, real-valued, invariant, positive definite kernel h on X and a point $x_0 \in X$, such that*

$$\rho(x, y) = \sqrt{h(x_0, x_0) - h(x, y)} \qquad (x, y \in X).$$

Privileged orthonormal bases of G-invariant functions, in the L^2 space associated with the invariant measure, provide a canonical decompositions of positive definite kernels. These generalized spherical harmonics were already studied by E. Cartan, H. Weyl and J. von Neumann; see, for instance [138]. We elaborate on two important particular cases.

Let $X = \mathbb{T} = \{e^{i\theta} : \theta \in \mathbb{R}\}$ be the unit torus, endowed with the invariant arc-length measure. A continuous positive definite function $h : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$ admits a Fourier decomposition

$$h(e^{ix}, e^{iy}) = \sum_{j,k \in \mathbb{Z}} a_{jk} e^{ijx} e^{-iky}.$$

If h is further required to be rotation invariant, we find that

$$h(e^{ix}, e^{iy}) = \sum_{k \in \mathbb{Z}} a_k e^{ik(x-y)},$$

where $a_k \geq 0$ for all $k \in \mathbb{Z}$ and $a_k = a_{-k}$ because h takes real values. Moreover, the series is Abel summable: $\sum_{k=0}^{\infty} a_k = h(1, 1) < \infty$. Therefore, a rotation-invariant Hilbert distance ρ on the torus has the expression (after taking its square):

$$\rho(e^{ix}, e^{iy})^2 = h(1, 1) - h(e^{ix}, e^{iy}) = \sum_{k=1}^{\infty} a_k(2 - e^{ik(x-y)} - e^{-ik(x-y)})$$

$$= 2 \sum_{k=1}^{\infty} a_k(1 - \cos k(x - y))$$

$$= 4 \sum_{k=1}^{\infty} a_k \sin^2(k(x - y)/2).$$

These are the periodic screw lines (2.2) already investigated by von Neumann and Schoenberg.

As a second example, we follow Bochner in examining a separable, compact group G. A real-valued, continuous, positive definite and G-invariant kernel h admits the decomposition

$$h(x, y) = \sum_{k \in \mathbb{Z}} c_k \chi_k(yx^{-1}),$$

where $c_k \geq 0$ for all $k \in \mathbb{Z}$, $\sum_{k \in Z} c_k < \infty$ and χ_k denote the characters of irreducible representations of G. In conclusion, an invariant Hilbert distance ρ

on G is characterized by the formula

$$\rho(x,y)^2 = \sum_{k\in\mathbb{Z}} c_k\big(1 - \frac{\chi_k(yx^{-1}) + \chi_k(xy^{-1})}{2\chi_k(1)}\big),$$

where $c_k \geq 0$ and $\sum_{k\in\mathbb{Z}} c_k < \infty$.

For details and an analysis of similar decompositions on more general homogeneous spaces, we refer the reader to [26].

The above analysis of positive definite functions on homogeneous spaces was carried out separately by Schoenberg in [125]. First, he remarks that a continuous, real-valued, rotationally invariant and positive definite kernel f on the sphere S^{d-1} has a distinguished Fourier-series decomposition with non-negative coefficients. Specifically,

$$f(\cos\theta) = \sum_{k=0}^{\infty} c_k P_k^{(\lambda)}(\cos\theta) \tag{2.3}$$

where $\lambda = (d-2)/2$, $P_k^{(\lambda)}$ are the ultraspherical orthogonal polynomials, $c_k \geq 0$ for all $k \geq 0$ and $\sum_{k=0}^{\infty} c_k < \infty$. This decomposition is in accord with Bochner's general framework, with the difference lying in Schoenberg's elementary proof, based on induction on dimension. As with all our formulas concerning the sphere, θ represents the geodesic distance (arc length along a great circle) between two points.

To convince the reader that expressions in the cosine of the geodesic distance are positive definite, let us consider points $x_1, \ldots, x_n \in S^{d-1}$. The Gram matrix with entries

$$\langle x_j, x_k \rangle = \cos\theta(x_j, x_k)$$

is obviously positive semidefinite, with constant diagonal elements equal to 1. According to the Schur product theorem [129], all functions of the form $\cos^k\theta$, where k is a non-negative integer, are therefore positive definite on the sphere.

At this stage, Schoenberg makes a leap forward and studies invariant positive definite kernels on S^∞, that is, functions $f(\cos\theta)$ which admit representations as above for all $d \geq 2$. His conclusion is remarkable in its simplicity.

Theorem 2.10 (Schoenberg [125])**.** *A real-valued function $f(\cos\theta)$ is positive definite on all spheres, independent of their dimension, if and only if*

$$f(\cos\theta) = \sum_{k=0}^{\infty} c_k \cos^k\theta, \tag{2.4}$$

where $c_k \geq 0$ for all $k \geq 0$ and $\sum_{k=0}^{\infty} c_k < \infty$.

This provides a return to the dominant theme, of isometric embedding into Hilbert space.

Corollary 2.11. *The function $\rho(\theta)$ is a Hilbert distance on S^∞ if and only if*

$$\rho(\theta)^2 = \sum_{k=0}^\infty c_k(1 - \cos^k \theta),$$

where $c_k \geq 0$ for all $k \geq 0$ and $\sum_{k=0}^\infty c_k < \infty$.

However, there is much more to derive from Schoenberg's theorem, once it is freed from the spherical context.

Theorem 2.12 (Schoenberg [125]). *Let $f : [-1, 1] \to \mathbb{R}$ be a continuous function. If the matrix $[f(a_{jk})]_{j,k=1}^n$ is positive semidefinite for all $n \geq 1$ and all positive semidefinite matrices $[a_{jk}]_{j,k=1}^n$ with entries in $[-1, 1]$, then, and only then,*

$$f(x) = \sum_{k=0}^\infty c_k x^k \qquad (x \in [-1, 1]),$$

where $c_k \geq 0$ for all $k \geq 0$ and $\sum_{k=0}^\infty c_k < \infty$.

Proof. One implication follows from the Schur product theorem [129], which says that if the $n \times n$ matrices A and B are positive semidefinite, then so is their entrywise product $A \circ B := [a_{jk}b_{jk}]_{j,k=1}^n$. Indeed, inductively setting $B = A^{\circ k} = A \circ \cdots \circ A$, the k-fold entrywise power, shows that every monomial x^k preserves positivity when applied entrywise. That the same property holds for functions $f(x) = \sum_{k \geq 0} c_k x^k$, with all $c_k \geq 0$, now follows from the fact that the set of positive semidefinite $n \times n$ matrices forms a closed convex cone, for all $n \geq 1$.

For the non-trivial, reverse implication we restrict the test matrices to those with leading diagonal terms all equal to 1. By interpreting such a matrix A as a Gram matrix, we identify n points on the sphere x_1, ..., $x_n \in S^{n-1}$ satisfying

$$a_{jk} = \langle x_j, x_k \rangle = \cos \theta(x_j, x_k) \quad (1 \leq j, k \leq n).$$

Then we infer from Schoenberg's theorem that f admits a uniformly convergent Taylor series with non-negative coefficients. $\qquad \square$

We conclude this section by mentioning some recent avenues of research that start from Bochner's theorem (and its generalization in 1940, by Weil, Povzner, and Raikov, to all locally compact abelian groups) and Schoenberg's classification of positive definite functions on spheres. On the theoretical side, there has been a profusion of recent mathematical activity on classifying positive definite functions (and strictly positive definite functions) in numerous settings, mostly related to spheres [9, 10, 32, 141, 142, 144], two-point homogeneous spaces[2][7, 8, 28], locally compact abelian groups and homogeneous spaces [45, 64], and products of these [15, 16, 63, 65, 67, 66].

---

[2]Recall [137] that a metric space (X, ρ) is *n-point homogeneous* if, given finite sets X_1, $X_2 \subset X$ of equal size no more than n, every isometry from X_1 to X_2 extends to a self-isometry of X. This property was first considered by Birkhoff [21], and of course differs

Moreover, this line of work directly impacts applied fields. For instance, in climate science and geospatial statistics, one uses positive definite kernels and Schoenberg's results (and their sequels) to study trends in climate behavior on the Earth, since it can be modelled by a sphere, and positive definite functions on $S^2 \times \mathbb{R}$ characterize space-time covariance functions on it. See [62, 101, 108, 109] for more details on these applications. There is a natural connection to probability theory, through the work of Lévy; see e.g. [56]. Other applied fields include genomics and finance, through high-dimensional covariance estimation. We elaborate on this in Chapter 7 below.

There are several other applications of Schoenberg's work on positive definite functions on spheres (his paper [125] has more than 160 citations) and we mention here just a few of them. Schoenberg's results were used by Musin [102] to compute the kissing number in four dimensions, by an extension of Delsarte's linear-programming method. Moreover, the results also apply to obtain new bounds on spherical codes [103], with further applications to sphere packing [35, 36, 37, 38]. There are also applications to approximating functions and interpolating data on spheres, pseudodifferential equations with radial basis functions, and Gaussian random fields.

Remark 2.13. Another modern-day use of Schoenberg's results in [125] is in Machine Learning; see [131, 133], for example. Given a real inner-product space H and a function $f : \mathbb{R} \to \mathbb{R}$, an alternative notion of f being *positive definite* is as follows: for any finite set of vectors $x_1, \ldots, x_n \in H$, the matrix

$$[f(\langle x_j, x_k \rangle)]_{j,k=1}^n$$

is positive semidefinite. This is in contrast to the notion promoted by Bochner, Weil, Schoenberg, Pólya, and others, which concerns positivity of the matrix with entries $f(\langle x_j - x_k, x_j - x_k \rangle^{1/2})$. It turns out that every positive definite kernel on H, given by

$$(x, y) \mapsto f(\langle x, y \rangle)$$

for a function f which is positive definite in this alternate sense, gives rise to a reproducing-kernel Hilbert space, which is a central concept in Machine Learning. We restrict ourselves here to mentioning that, in this setting, it is desirable for the kernel to be strictly positive definite; see [105] for further clarification and theoretical results along these lines.

2.6. **Connections to harmonic analysis.** Positivity and sharp continuity bounds for linear transformations between specific normed function spaces go hand in hand, especially when focusing on the kernels of integral transforms. The end of 1950s marked a fortunate condensation of observations, leading to a quasi-complete classification of preservers of positive or bounded convolution transforms acting on spaces of functions on locally compact

---

from the more common usage of the terminology of a homogeneous space G/H, whose study by Bochner was mentioned above.

abelian groups. In particular, these results can be interpreted as Schoenberg-type theorems for Toeplitz matrices or Toeplitz kernels. We briefly recount the main developments.

A groundbreaking theorem of the 1930s attributed to Wiener and Levy asserts that the pointwise inverse of a non-vanishing Fourier series with coefficients in L^1 exhibits the same summability behavior of the coefficient sequence. To be more precise, if ϕ is never zero and has the representation

$$\phi(\theta) = \sum_{n=-\infty}^{\infty} c_n e^{in\theta}, \quad \text{where} \quad \sum_{n=-\infty}^{\infty} |c_n| < \infty,$$

then its reciprocal has a representation of the same form:

$$(1/\phi)(\theta) = \sum_{n=-\infty}^{\infty} d_n e^{in\theta}, \quad \text{where} \quad \sum_{n=-\infty}^{\infty} |d_n| < \infty.$$

It was Gelfand [61] who in 1941 cast this permanence phenomenon in the general framework of commutative Banach algebras. Gelfand's theory applied to the Wiener algebra $W := \widehat{L^1(\mathbb{Z})}$ of Fourier transforms of L^1 functions on the dual of the unit torus proves the following theorem.

Theorem 2.14 (Gelfand [61]). *Let $\phi \in W$ and let $f(z)$ be an analytic function defined in a neighborhood of $\phi(\mathbb{T})$. Then $f(\phi) \in W$.*

The natural inverse question of deriving smoothness properties of inner transformations of Lebesgue spaces of Fourier transforms was tackled almost simultaneously by several analysts. For example, Rudin proved in 1956 [115] that a coefficient-wise transformation $c_n \mapsto f(c_n)$ mapping the space $\widehat{L^1(\mathbb{T})}$ into itself implies the analyticity of f in a neighborhood of zero. In a similar vein, Rudin and Kahane proved in 1958 [84] that a coefficient-wise transformation $c_n \mapsto f(c_n)$ which preserves the space of Fourier transforms $\widehat{M(\mathbb{T})}$ of finite measures on the torus implies that f is an entire function. In the same year, Kahane [83] showed that no quasi-analytic function (in the sense of Denjoy–Carleman) preserves the space $\widehat{L^1(\mathbb{Z})}$ and Katznelson [87] refined an inverse to Gelfand's theorem above, by showing the semi-local analyticity of transformers of elements of $\widehat{L^1(\mathbb{Z})}$ subject to some support conditions.

Soon after, the complete picture emerged in full clarity. It was unveiled by Helson, Kahane, Katznelson and Rudin in an Acta Mathematica article [74]. Given a function f defined on a subset E of the complex plane, we say that f operates on the function algebra A, if $f(\phi) \in A$ for every $\phi \in A$ with range contained in E. The following metatheorem is proved in the cited article.

Theorem 2.15 (Helson–Kahane–Katznelson–Rudin [74]). *Let G be a locally compact abelian group and let Γ denote its dual, and suppose both are*

endowed with their respective Haar measures. Let $f : [-1,1] \to \mathbb{C}$ be a function satisfying $f(0) = 0$.

(1) *If Γ is discrete and f operates on $\widehat{L^1(G)}$, then f is analytic in some neighborhood of the origin.*

(2) *If Γ is not discrete and f operates on $\widehat{L^1(G)}$, then f is analytic in $[-1,1]$.*

(3) *If Γ is not compact and f operates on $\widehat{M(G)}$, then f can be extended to an entire function.*

Rudin refined the above results to apply in the case of various L^p norms [117, 118], by stressing the lack of continuity assumption for the transformer f in all results (similar in nature to the statements in the above theorem). From Rudin's work we extract a highly relevant observation, *à la* Schoenberg's theorem, aligned to the spirit of the present survey.

Theorem 2.16 (Rudin [116]). *Suppose $f : (-1,1) \to \mathbb{R}$ maps every positive semidefinite Toeplitz kernel with elements in $(-1,1)$ into a positive semidefinite kernel:*

$$[a_{j-k}]_{j,k=-\infty}^{\infty} \geq 0 \qquad \Longrightarrow \qquad [f(a_{j-k})]_{j,k=-\infty}^{\infty} \geq 0.$$

Then f is absolutely monotonic, that is analytic on $(-1,1)$ with a Taylor series having non-negative coefficients:

$$f(x) = \sum_{n=0}^{\infty} c_k x^k, \qquad \text{where } c_k \geq 0 \text{ for all } k \geq 0.$$

In the above Theorem and henceforth, the notations $A \leq B$ and $B \geq A$ mean that the (possibly semi-infinite or bi-infinite) matrix $B - A$ is positive semidefinite, i.e. has all principal minors non-negative.

The converse is obviously true by the Schur product theorem. The elementary proof, quite independent of the derivation of the metatheorem stated above, is contained in [116]. Notice again the lack of a continuity assumption in the hypotheses.

In fact, Rudin proves more, by restricting the test domain of positive semidefinite Toeplitz kernels to the two-parameter family

$$a_n = \alpha + \beta \cos(n\theta) \qquad (n \in \mathbb{Z}) \tag{2.5}$$

with θ fixed so that θ/π is irrational and $\alpha, \beta \geq 0$ such that $\alpha + \beta < 1$. Rudin's proof commences with a mollifier argument to deduce the continuity of the transformer, then uses a development in spherical harmonics very similar to the original argument of Schoenberg. We will resume this topic in Section 3.3, setting it in a wider context.

With the advances in abstract duality theory for locally convex spaces, it is not surprising that proofs of Schoenberg-type theorems should be accessible with the aid of such versatile tools. We will confine ourselves here

to mentioning one pertinent convexity-theoretic proof of Schoenberg's theorem, due to Christensen and Ressel [33]. (See also [34] for a complex sphere variant.)

Skipping freely over the details, the main observation of these two authors is that the multiplicatively closed convex cone of positivity preservers of positive semidefinite matrices of any size, with entries in $[-1, 1]$, is closed in the product topology of $\mathbb{R}^{[-1,1]}$, with a compact base K defined by the normalization $f(1) = 1$. The set of extreme points of K is readily seen to be closed, and an elementary argument identifies it as the set of all monomials x^n, where $n \geq 0$, plus the characteristic functions $\chi_1 \pm \chi_{-1}$. An application of Choquet's representation theorem now provides a proof of a generalization of Schoenberg's theorem, by removing the continuity assumption in the statement.

3. Entrywise functions preserving positivity in all dimensions

3.1. **History.** With the above history to place the present survey in context, we move to its dominant theme: entrywise positivity preservers. In analysis and in applications in the broader mathematical sciences, one is familiar with applying functions to the spectrum of diagonalizable matrices: $A = UDU^*$ then $f(A) = Uf(D)U^*$. More formally, one uses the Riesz–Dunford holomorphic functional calculus to define $f(A)$ for classes of matrices A and functions f.

Our focus in this survey will be on the parallel philosophy of *entrywise calculus*. To differentiate this from the functional calculus, we use the notation $f[A]$.

Definition 3.1. Fix a domain $I \subset \mathbb{C}$ and integers m, $n \geq 1$. Let $\mathcal{P}_n(I)$ denote the set of $n \times n$ Hermitian positive semidefinite matrices with all entries in I.

A function $f : I \to \mathbb{C}$ acts *entrywise* on a matrix

$$A = [a_{jk}]_{1 \leq j \leq m,\ 1 \leq k \leq n} \in I^{m \times n}$$

by setting

$$f[A] := [f(a_{jk})]_{1 \leq j \leq m,\ 1 \leq k \leq n} \in \mathbb{C}^{m \times n}.$$

Below, we allow the dimensions m and n to vary, while keeping the uniform notation $f[-]$.

We also let $\mathbf{1}_{m \times n}$ denote the $m \times n$ matrix with each entry equal to one. Note that $\mathbf{1}_{n \times n} \in \mathcal{P}_n(\mathbb{R})$.

In this survey, we explore the following overarching question in several different settings.

Which functions preserve positive semidefiniteness when applied entrywise to a class of positive matrices?

This question was first asked by Pólya and Szegö in their well-known book [107]. The authors observed that Schur's product theorem, together with

the fact that the positive matrices form a closed convex cone, has the following consequence: if $f(x)$ is any power series with non-negative Maclaurin coefficients that converges on a domain $I \subset \mathbb{R}$, then f preserves positivity (that is, preserves positive semidefiniteness) when applied entrywise to positive semidefinite matrices with entries in I. Pólya and Szegö then asked if there are any other functions that possess this property. As discussed above, Schoenberg's theorem 2.12 provides a definitive answer to their question (together with the improvements by Rudin or Christensen–Ressel to remove the continuity hypothesis). Thanks to Pólya and Szegö's observation, Schoenberg's result may be considered as a rather challenging converse to the Schur product theorem.

In a similar vein, Rudin [116] observed that if one moves to the complex setting, then the conjugation map also preserves positivity when applied entrywise to positive semidefinite complex matrices. Therefore the maps

$$z \mapsto z^j \overline{z}^k \qquad (j, k \geq 0)$$

preserve positivity when applied entrywise to complex matrices of all dimensions, again by the Schur product theorem. The same property is now satisfied by non-negative linear combinations of these functions. In [116], Rudin made this observation and conjectured, à la Pólya–Szegö, that these are all of the preservers. This was proved by Herz in 1963.

Theorem 3.2 (Herz [77]). *Let $D(0,1)$ denote the open unit disc in $\mathbb{C}$, and suppose $f : D(0,1) \to \mathbb{C}$. The entrywise map $f[-]$ preserves positivity on $\mathcal{P}_n\big(D(0,1)\big)$ for all $n \geq 1$, if and only if*

$$f(z) = \sum_{j,k \geq 0} c_{jk} z^j \overline{z}^k \qquad \text{for all } z \in D(0,1),$$

where $c_{jk} \geq 0$ for all $j, k \geq 0$.

Akin to the above results by Schoenberg, Rudin, Christensen and Ressel, and Herz, we mention one more Schoenberg-type theorem, for matrices with positive entries. The following result again demonstrates the rigid principle that analyticity and absolute monotonicity follow from the preservation of positivity in all dimensions.

Theorem 3.3 (Vasudeva [134]). *Let $f : (0, \infty) \to \mathbb{R}$. Then $f[-]$ preserves positivity on $\mathcal{P}_n\big((0, \infty)\big)$ for all $n \geq 1$, if and only if $f(x) = \sum_{k=0}^{\infty} c_k x^k$ on $(0, \infty)$, where $c_k \geq 0$ for all $k \geq 0$.*

3.2. The Horn–Loewner necessary condition in fixed dimension.
The previous section contains several variants of a "dimension-free" result: namely, the classification of entrywise maps that preserve positivity on test sets of matrices of all sizes. In the next section, we discuss a dimension-free result that parallels Rudin's work in [116], by approaching the problem via preservers of moment sequences for positive measures on the real line. In other words, we will work with Hankel instead of Toeplitz matrices.

In the later part of this survey, we focus on entrywise functions that pre-serve positivity when the test set consists of matrices of a fixed size. For both of these settings, the starting point is an important result first published by R. Horn (who in [80] attributes it to his PhD advisor C. Loewner).

Theorem 3.4 ([80]). *Let $f : (0, \infty) \to \mathbb{R}$ be continuous. Fix a positive integer n and suppose $f[-]$ preserves positivity on $\mathcal{P}_n\big((0, \infty)\big)$. Then $f \in C^{n-3}((0, \infty))$,*

$$f^{(k)}(x) \geq 0 \qquad \text{whenever } x \in (0, \infty) \text{ and } 0 \leq k \leq n - 3,$$

and $f^{(n-3)}$ is a convex non-decreasing function on $(0, \infty)$. Furthermore, if $f \in C^{n-1}\big((0, \infty)\big)$, then $f^{(k)}(x) \geq 0$ whenever $x \in (0, \infty)$ and $0 \leq k \leq n-1$.

This result and its variations are the focus of the present section.
Theorem 3.4 is remarkable for several reasons.

(1) Modulo variations, it remains to this day the only known criterion for a general entrywise function to preserve positivity in a fixed di-mension. Later on, we will see more precise conclusions drawn when f is a polynomial or a power function, but for a general function there are essentially no other known results.

(2) While Theorem 3.4 is a fixed-dimension result, it can be used to prove some of the aforementioned dimension-free characterizations. For instance, if $f[-]$ preserves positivity on $\mathcal{P}_n\big((0, \infty)\big)$ for all $n \geq 1$, then, by Theorem 3.4, the function f is absolutely monotonic on $(0, \infty)$. A classical result of Bernstein on absolutely monotonic func-tions now implies that f is necessarily given by a power series with non-negative coefficients, which is precisely Vasudeva's Theorem 3.3.

In the next section, we will outline an approach to prove a stronger version of Schoenberg's theorem 2.12 (in the spirit of Theorem 2.16 by Rudin), starting from Theorem 3.3.

(3) Theorem 3.4 is also significant because there is a sense in which it is sharp. We elaborate on this when studying polynomial and power-function preservers; see Chapters 4 and 6.

Remark 3.5. There are other, rather unexpected consequences of Theo-rem 3.4 as well. It was recently shown that the key determinant compu-tation underlying Theorem 3.4 can be generalized to yield a new class of symmetric function identities for any formal power series. The only such identities previously known were for the case $f(x) = \frac{1-cx}{1-x}$. This is discussed in Section 4.6.

We next explain the steps behind the proof of the Horn–Loewner theo-rem 3.4. These also help in proving certain strengthenings of Theorem 3.4, which are mentioned below. In turn, these strengthenings additionally serve to clarify the nature of the Horn–Loewner necessary condition.

Proof of Theorem 3.4. The proof by Loewner is in two steps. First he as-sumes f to be smooth and shows the result by induction on n. The base

case of $n = 1$ is immediate, and for the induction step one proceeds as follows. Fix $a > 0$, choose any vector $\mathbf{u} = (u_1, \ldots, u_n)^T \in \mathbb{R}^n$ with distinct coordinates, and define

$$\Delta(t) := \det[f(a + tu_j u_k)]_{j,k=1}^n = \det f[a\mathbf{1}_{n \times n} + \mathbf{u}\mathbf{u}^T] \qquad (0 < t \ll 1).$$

Then Loewner shows that

$$\Delta(0) = \Delta'(0) = \cdots = \Delta^{\binom{n}{2}-1}(0) = 0,$$
$$\Delta^{\binom{n}{2}}(0) = cf(a)f'(a) \cdots f^{(n-1)}(a) \qquad \text{for some } c > 0. \tag{3.1}$$

(See Remark 3.5 above.)

Returning to the proof of Theorem 3.4 for smooth functions: apply the above treatment not to f but to $g_\tau(x) := f(x) + \tau x^n$, where $\tau > 0$. By the Schur product theorem, g_τ satisfies the hypotheses, whence $\Delta(t)/t^{\binom{n}{2}} \geq 0$ for $t > 0$. Taking $t \to 0^+$, by L'Hôpital's rule we obtain

$$g_\tau(a)g_\tau'(a) \cdots g_\tau^{(n-1)}(a) \geq 0, \qquad \text{for all } \tau > 0.$$

Finally, the induction hypothesis implies that $f, f', \ldots, f^{(n-2)}$ are non-negative at a, whence $g_\tau(a), \ldots, g_\tau^{(n-2)}(a) > 0$. It follows that $g_\tau^{(n-1)}(a) \geq 0$ for all $\tau > 0$, and hence, $f^{(n-1)}(a) \geq 0$, as desired.

Remark 3.6. The above argument is amenable to proving more refined results. For example, it can be used to prove the positivity of the first n non-zero derivatives of a smooth preserver f; see Theorem 3.10.

The second step of Loewner's proof begins by using mollifiers. Suppose f is continuous; approximate it by a mollified family $f_\delta \to f$ as $\delta \to 0^+$. Thus f_δ is smooth and its first n derivatives are non-negative on $(0, \infty)$. By the mean-value theorem for divided differences, this implies that the divided differences of each f_δ, of orders up to $n - 1$ are non-negative. Since f is continuous, the same holds for f.

Now one invokes a rather remarkable result by Boas and Widder [24], which can be viewed as a converse to the mean-value theorem for divided differences. It asserts that given an integer $k \geq 2$ and an open interval $I \subset \mathbb{R}$, if all kth order "equi-spaced" forward differences (whence divided differences) of a continuous function $f : I \to \mathbb{R}$ are non-negative on I, then f is $k - 2$ times differentiable on I; moreover, $f^{(k-2)}$ is continuous and convex on I, with non-decreasing left- and right-hand derivatives. Applying this result for each $2 \leq k \leq n - 1$ concludes the proof of Theorem 3.4. $\qquad \square$

Note that this proof only uses matrices of the form $a\mathbf{1}_{n \times n} + t\mathbf{u}\mathbf{u}^T$, and the arguments are all local. Thus it is unsurprising that strengthened versions of the Horn–Loewner theorem can be found in the literature; see [12, 71], for example. We present here the stronger of these variants.

Theorem 3.7 (See [12, Section 3]). *Suppose $0 < \rho \leq \infty$, $I = (0, \rho)$, and $f : I \to \mathbb{R}$. Fix $u_0 \in (0, 1)$ and an integer $n \geq 1$, and define $\mathbf{u} :=$*

$(1, u_0, \ldots, u_0^{n-1})^T$. *Suppose $f[A] \in \mathcal{P}_2(\mathbb{R})$ for all $A \in \mathcal{P}_2(I)$, and also that $f[A] \in \mathcal{P}_n(\mathbb{R})$ for all Hankel matrices $A = a\mathbf{1}_{n \times n} + t\mathbf{u}\mathbf{u}^T$, with a, $t \geq 0$ such that $a + t \in I$. Then the conclusions of Theorem 3.4 hold.*

Beyond the above strengthenings, the notable feature here is that the continuity hypothesis has been removed, akin to the Rudin and Christensen–Ressel results. We reproduce here an elegant argument to show continuity; this can be found in Vasudeva's paper [134], and uses only the test set $\mathcal{P}_2(I)$. By considering $f[A]$ for $A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$ with $0 < b < a < \rho$, it follows that f is non-negative and non-decreasing on I. One also shows that f is either identically zero or never zero on I. In the latter case, considering $f[A]$ for $A = \begin{bmatrix} a & \sqrt{ab} \\ \sqrt{ab} & b \end{bmatrix} \in \mathcal{P}_2(I)$ shows that f is *multiplicatively mid-convex*: the function

$$g(y) := \log f(e^y) \qquad (y < \log \rho)$$

is midpoint convex and locally bounded on the interval $\log I$. Now the following classical result [113, Theorem 71.C] shows that g is continuous on $\log I$, so f is continuous on I.

Proposition 3.8. *Let U be a convex open set in a real normed linear space. If $g : U \to \mathbb{R}$ is midpoint convex on U and bounded above in an open neighborhood of a single point in U, then g is continuous, so convex, on U.*

We now move to variants of the Horn–Loewner result. Notice that Theorems 3.4 and 3.7 are results for arbitrary positivity preservers $f(x)$. When more is known about f, such as smoothness or even real analyticity, stronger conclusions can be drawn from smaller test sets of matrices. A recent variant is the following lemma, shown by evaluating $f[-]$ at matrices $(tu_ju_k)_{j,k=1}^n$ and using the invertibility of "generic" generalized Vandermonde matrices.

Lemma 3.9 (Belton–Guillot–Khare–Putinar [11] and Khare–Tao [89]). *Let $n \geq 1$ and $0 < \rho \leq \infty$. Suppose $f(x) = \sum_{k \geq 0} c_k x^k$ is a convergent power series on $I = [0, \rho)$ that is positivity preserving entrywise on rank-one matrices in $\mathcal{P}_n(I)$. Further assume that $c_{m'} < 0$ for some m'.*

(1) *If $\rho < \infty$, then we have $c_m > 0$ for at least n values of $m < m'$. (In particular, the first n non-zero Maclaurin coefficients of f, if they exist, must be positive.)*

(2) *If instead $\rho = \infty$, then we have $c_m > 0$ for at least n values of $m < m'$ and at least n values of $m > m'$. (In particular, if f is a polynomial, then the first n non-zero coefficients and the last n non-zero coefficients of f, if they exist, are all positive.)*

Notice that this lemma (a) talks about the derivatives of f at 0 and not in $(0, \rho)$; and moreover, (b) considers not the first few derivatives, but the first few *non-zero* derivatives. Thus, it is morally different from the preceding

two theorems, and one naturally seeks a common unification of these three results. This was recently achieved:

Theorem 3.10 (Khare [88])**.** *Let* $0 \le a < \infty, \epsilon \in (0, \infty), I = [a, a + \epsilon),$ *and let* $f : I \to \mathbb{R}$ *be smooth. Fix integers* $n \ge 1$ *and* $0 \le p \le q \le n$, *with* $p = 0$ *if* $a = 0$, *and such that* $f(x)$ *has* $q - p$ *non-zero derivatives at* $x = a$ *of order at least* p. *Now let*

$$m_0 := 0, \quad \ldots \quad m_{p-1} := p - 1;$$

suppose further that

$$p \le m_p < m_{p+1} < \cdots < m_{q-1}$$

are the lowest orders (above p) *of the first* $q - p$ *non-zero derivatives of* $f(x)$ *at* $x = a$.

Also fix distinct scalars $u_1, \ldots, u_n \in (0, 1)$, *and let* $\mathbf{u} := (u_1, \ldots, u_n)^T$. *If* $f[a\mathbf{1}_{n \times n} + t\mathbf{u}\mathbf{u}^T] \in \mathcal{P}_n(\mathbb{R})$ *for all* $t \in [0, \epsilon)$, *then the derivative* $f^{(k)}(a)$ *is non-negative whenever* $0 \le k \le m_{q-1}$.

Notice that varying p allows one to control the number of initial derivatives versus the number of subsequent non-zero derivatives of smallest order. In particular, if $p = q = n$, then the result implies the "stronger" Horn–Loewner theorem 3.7 (and so Theorem 3.4) pointwise at every $a > 0$. At the other extreme is the special case of $p = 0$ (at any $a \ge 0$), which strengthens the conclusions of Theorems 3.4 and 3.7 for smooth functions.

Corollary 3.11. *Suppose* a, ϵ, I, f, n *and* $\mathbf{u}$ *are as in Theorem 3.10. If* $f[a\mathbf{1}_{n \times n} + t\mathbf{u}\mathbf{u}^T] \in \mathcal{P}_n(\mathbb{R})$ *for all* $t \in [0, \epsilon)$, *then the first* n *non-zero derivatives of* $f(x)$ *at* $x = a$ *are positive.*

Remark 3.12. Theorem 3.10 further clarifies the nature of the Horn–Loewner result and its proof. The reduction from arbitrary functions, to continuous functions, to smooth functions, requires an open domain $(0, \rho)$, in order to use mollifiers, for example. However, the result for smooth functions actually holds pointwise, as shown by Theorem 3.10.

The proof of Theorem 3.10 combines novel arguments together with the previously mentioned techniques of Loewner. The refinement of the determinant computations (3.1) is of particular note; see Theorem 4.20 and its consequence, Theorem 4.22.

3.3. **Schoenberg redux: moment sequences and Hankel matrices.**
In this section, we outline another approach to proving Schoenberg's theorem 2.12, which yields a stronger version parallel to the strengthening by Rudin of Theorem 2.16. The present section reveals connections between positivity preservers, totally non-negative Hankel matrices, moment sequences of positive measures on the real line, and also a connection to semi-algebraic geometry.

We begin with Rudin's Theorem 2.16 and the family (2.5). Notice that the positive definite sequences in (2.5) give rise to the Toeplitz matrices

$A(n, \alpha, \beta, \theta)$ with (j, k) entry equal to $\alpha + \beta \cos\big((j - k)\theta\big)$. From the elementary identity

$$\cos(p - q) = \cos p \cos q + \sin p \sin q \qquad (p, q \in \mathbb{R}),$$

it follows that these Toeplitz matrices have rank at most three:

$$A(n, \alpha, \beta, \theta) = \alpha \mathbf{1}_{n \times n} + \beta \mathbf{u}\mathbf{u}^T + \beta \mathbf{v}\mathbf{v}^T, \tag{3.2}$$

where

$$\mathbf{u} := \big(\cos\theta, \cos(2\theta), \ldots, \cos(n\theta)\big)^T \quad \text{and} \quad \mathbf{v} := \big(\sin\theta, \sin(2\theta), \ldots, \sin(n\theta)\big)^T.$$

In particular, Rudin's work (see Theorem 2.16 and the subsequent discussion) implies the following result.

Proposition 3.13. *Let $\theta \in \mathbb{R}$ such that θ/π is irrational. An entrywise map $f : \mathbb{R} \to \mathbb{R}$ preserves positivity on the set of Toeplitz matrices*

$$\{A(n, \alpha, \beta, \theta) : n \geq 1, \ \alpha, \beta > 0\}$$

if and only if $f(x) = \sum_{k=0}^{\infty} c_k x^k$ is a convergent power series on $\mathbb{R}$, with $c_k \geq 0$ for all $k \geq 0$.

Thus, one can significantly reduce the set of test matrices.

Proof. Given $0 < \rho < \infty$, let the restriction $f_\rho := f|_{(-\rho, \rho)}$. Observe from the discussion following Theorem 2.16 that Rudin's work explicitly shows the result for f_1, whence for any f_ρ by a change of variables. Thus,

$$f_\rho(x) = \sum_{k=0}^{\infty} c_{k,\rho} x^k, \qquad c_{k,\rho} \geq 0 \text{ for all } k \geq 0 \text{ and } \rho > 0.$$

Given $0 < \rho < \rho' < \infty$, it follows by the identity theorem that $c_{k,\rho} = c_{k,\rho'}$ for all k. Hence $f(x) = \sum_{k \geq 0} c_{k,1} x^k$ (which was Rudin's $f_1(x)$), now on all of $\mathbb{R}$. $\qquad \square$

In a parallel vein to Rudin's results and Proposition 3.13, the following strengthening of Schoenberg's result can be shown, using a different (and perhaps more elementary) approach than those of Schoenberg and Rudin.

Theorem 3.14 (Belton–Guillot–Khare–Putinar [12])**.** *Suppose $0 < \rho \leq \infty$ and $I = (-\rho, \rho)$. Then the following are equivalent for a function $f : I \to \mathbb{R}$.*

(1) *The entrywise map $f[-]$ preserves positivity on $\mathcal{P}_n(I)$, for all $n \geq 1$.*

(2) *The entrywise map $f[-]$ preserves positivity on the Hankel matrices in $\mathcal{P}_n(I)$ of rank at most 3, for all $n \geq 1$.*

(3) *The function f is real analytic on I and absolutely monotonic on $(0, \rho)$. In other words, $f(x) = \sum_{k \geq 0} c_k x^k$ on I, with $c_k \geq 0 \ \forall k$.*

Remark 3.15. Recall the alternate notion of positive definite functions discussed in Remark 2.13. In [105] and related works, Pinkus and other authors study this alternate notion of positive definite functions on H. Notice that such matrices form precisely the set of positive semidefinite symmetric

matrices of rank at most $\dim H$. In particular, Theorem 3.14 and the far earlier 1959 paper [116] of Rudin both provide a characterization of these functions, on every Hilbert space of dimension 3 or more.

Parallel to the discussions of the proofs of Schoenberg's and Rudin's results (see the previous chapter), we now explain how to prove Theorem 3.14. Clearly, (3) $\implies$ (1) $\implies$ (2) in the theorem. We first outline how to weaken the condition (2) even further and still imply (3). The key idea is to consider *moment sequences* of certain non-negative measures on the real line. This parallels Rudin's considerations of Fourier–Stieltjes coefficients of non-negative measures on the circle.

Definition 3.16. A measure μ with support in $\mathbb{R}$ is said to be *admissible* if $\mu \geq 0$ on $\mathbb{R}$, and all moments of μ exist and are finite:

$$s_k(\mu) := \int_{\mathbb{R}} x^k \, \mathrm{d}\mu(x) < \infty \qquad (k \geq 0).$$

The sequence $\mathbf{s}(\mu) := \big(s_k(\mu)\big)_{k=0}^{\infty}$ is termed the *moment sequence* of μ. Corresponding to μ and this moment sequence is the *moment matrix* of μ:

$$H_\mu := \begin{bmatrix} s_0(\mu) & s_1(\mu) & s_2(\mu) & \cdots \\ s_1(\mu) & s_2(\mu) & s_3(\mu) & \cdots \\ s_2(\mu) & s_3(\mu) & s_4(\mu) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix};$$

note that $H_\mu = [s_{i+j}(\mu)]_{i,j \geq 0}$ is a semi-infinite Hankel matrix. Finally, a function $f : \mathbb{R} \to \mathbb{R}$ acts entrywise on moment sequences, to yield real sequences:

$$f[\mathbf{s}(\mu)] := (f\big(s_0(\mu)\big), \ldots, f\big(s_k(\mu)\big), \ldots).$$

We are interested in understanding which entrywise functions preserve the space of moment sequences of admissible measures. The connection to positive semidefinite matrices is made through Hamburger's theorem, which says that a real sequence $(s_0, s_1, \ldots)$ is the moment sequence of an admissible measure on $\mathbb{R}$ if and only if every (finite) principal minor of the moment matrix H_μ is positive semidefinite. For simplicity, this last will be reformulated below to saying that H_μ is positive semidefinite.

The weakening of Theorem 3.14(2) is now explained: it suffices to consider the reduced test set of those Hankel matrices, which arise as the moment matrices of admissible measures supported at three points. Henceforth, let δ_x denote the Dirac probability measure supported at $x \in \mathbb{R}$. It is not hard to verify that the m-point measure $\mu = \sum_{j=1}^{m} c_j \delta_{x_j}$ has Hankel

matrix H_μ with rank no more than m:

$$
\begin{aligned}
s_k(\mu) &= \sum_{j=1}^m c_j x_j^k \qquad (k \geq 0) \\
\implies \quad H_\mu &= \sum_{j=1}^m c_j \mathbf{u}_j \mathbf{u}_j^T, \text{ where } \mathbf{u}_j := (1, x_j, x_j^2, \ldots)^T.
\end{aligned}
\tag{3.3}
$$

Thus, a further strengthening of Schoenberg's result is as follows.

Theorem 3.17 (Belton–Guillot–Khare–Putinar [12]). *In the setting of Theorem 3.14, the three assertions contained therein are also equivalent to*

(4) *For each measure*

$$
\mu = a\delta_1 + b\delta_{u_0} + c\delta_{-1}, \quad \text{with } u_0 \in (0,1), \ a, b, c \geq 0, \ a+b+c \in (0, \rho), \tag{3.4}
$$

there exists an admissible measure σ_μ on $\mathbb{R}$ such that $f(s_k(\mu)) = s_k(\sigma_\mu)$ for all $k \geq 0$.

In fact, we will see in Section 3.4 below that this assertion (4) can be simplified to just assert that $f[H_\mu]$ is positive semidefinite, and so completely avoid the use of Hamburger's theorem.

We now discuss the proof of these results, working with $\rho = \infty$ for ease of exposition. The first observation is that the strengthening of the Horn–Loewner theorem 3.7, together with the use of Bernstein's theorem (see remark (2) following Theorem 3.4), implies the following "stronger" form of Vasudeva's theorem 3.3:

Theorem 3.18 (see [12]). *Suppose $I = (0, \infty)$ and $f : I \to \mathbb{R}$. Also fix $u_0 \in (0,1)$. The following are equivalent:*

(1) *The entrywise map $f[-]$ preserves positivity on $\mathcal{P}_n(I)$ for all $n \geq 1$.*
(2) *The entrywise map $f[-]$ preserves positivity on all moment matrices H_μ for $\mu = a\delta_1 + b\delta_{u_0}$, $a, b > 0$.*
(3) *The function f equals a convergent power series $\sum_{k=0}^\infty c_k x^k$ for all $x \in I$, with the Maclaurin coefficients $c_k \geq 0$ for all $k \geq 0$.*

Notice that the test matrices in assertion (2) are all Hankel, and of rank at most two. This severely weakens Vasudeva's original hypotheses.

Now suppose the assertion in Theorem 3.17(4) holds. By the preceding result, $f(x)$ is given on $(0, \infty)$ by an absolutely monotonic function $\sum_{k \geq 0} c_k x^k$. The next step is to show that f is continuous. For this, we will crucially use the following "integration trick". Suppose for each admissible measure μ as in (3.4), there is a non-negative measure σ_μ supported on $[-1, 1]$ such that $f(s_k(\mu)) = s_k(\sigma_\mu)$ for all $k \geq 0$. (Note here that it is not immediate that the support is contained in $[-1, 1]$.)

Now let $p(t) = \sum_{k \geq 0} b_k t^k$ be a polynomial that takes non-negative values on $[-1, 1]$. Then,

$$0 \leq \int_{-1}^{1} p(t)\, d\sigma_\mu(t) = \sum_{k=0}^{\infty} \int_{-1}^{1} b_k t^k \, d\sigma_\mu(t) = \sum_{k=0}^{\infty} b_k s_k(\sigma_\mu) = \sum_{k=0}^{\infty} b_k f\big(s_k(\mu)\big).$$
(3.5)

Remark 3.19. For example, suppose $p(t) = 1 - t^d$ for some $d \geq 1$. If $\mu = a\delta_1 + b\delta_{u_0} + c\delta_{-1}$, where $u_0 \in (0, 1)$ and a, b, $c > 0$, then the inequality (3.5) gives that

$$0 \leq f\big(s_0(\mu)\big) - f\big(s_d(\mu)\big) = f(a + b + c) - f(a + bu_0^d + c(-1)^d).$$

It is not clear *a priori* how to deduce this inequality using the fact that $f[-]$ preserves matrix positivity and the Hankel moment matrix of μ. The explanation, which we provide in Section 3.4 below, connects moment problems, matrix positivity, and real algebraic geometry.

We now outline how (3.5) can be used to prove of the continuity of f. First note that $|s_k(\mu)| \leq s_0(\mu)$ for μ as above and all $k \geq 0$. This fact and the easy observation that f is bounded on compact subsets of $\mathbb{R}$ together imply that all moments of σ_μ are uniformly bounded. From this we deduce that σ_μ is necessarily supported on $[-1, 1]$.

The inequality (3.5) now gives the left-continuity of f at $-\beta$, for every $\beta \geq 0$. Fix $u_0 \in (0, 1)$, and let

$$\mu_b := (\beta + bu_0)\delta_{-1} + b\delta_{u_0} \qquad (b > 0).$$

Applying (3.5) to the polynomials $p_{\pm,1}(t) := (1 \pm t)(1 - t^2)$, we deduce that

$$f\big(\beta + b(1 + u_0)\big) - f\big(\beta + b(u_0 + u_0^2)\big) \geq |f(-\beta) - f\big(-\beta - bu_0(1 - u_0^2)\big)|.$$

Letting $b \to 0^+$, the left continuity of f at $-\beta$ follows. Similarly, to show that f is right continuous at $-\beta$, we apply the integral trick to $p_{\pm,1}(t)$ and to $\mu_b' := (\beta + bu_0^3)\delta_{-1} + b\delta_{u_0}$ instead of μ_b.

Having shown continuity, to prove the stronger Schoenberg theorem, we next assume that f is smooth on $\mathbb{R}$. For all $a \in \mathbb{R}$, define the function

$$H_a : \mathbb{R} \to \mathbb{R}; \ x \mapsto f(a + e^x).$$

The function H_a satisfies the estimates

$$|H_a^{(n)}(x)| \leq H_{|a|}^{(n)}(x) \qquad (a, x \in \mathbb{R}, \ n \in \mathbb{Z}_+).$$
(3.6)

This is shown by another use of the integration trick (3.5), this time for the polynomials $p_{\pm,n}(t) := (1 \pm t)(1 - t^2)^n$ for all $n \geq 0$. In turn, the estimates (3.6) lead to showing that H_a is real analytic on $\mathbb{R}$, for all $a \in \mathbb{R}$. Now composing H_{-a} for $a > |x|$ with the function $L_a(y) := \log(a + y)$ shows that $f(x)$ is real analytic on $\mathbb{R}$ and agrees with $\sum_{k \geq 0} a_k x^k$ on $(0, \infty)$. This concludes the proof for smooth functions.

Finally, to pass from smooth functions to continuous functions, we again use a mollified family $f_\delta \to f$ as $\delta \to 0^+$. Each f_δ is the restriction of an

entire function, say $\widetilde{f}_\delta$, and the family $\{\widetilde{f}_{1/n} : n \geq 1\}$ forms a normal family on each open disc $D(0, r)$. It follows from results by Montel and Morera that $\widetilde{f}_{1/n}(z)$ converges uniformly to a function g_r on each closed disc $\overline{D(0, r)}$, and g_r is analytic. Since g_r restricts to f on $(-r, r)$, it follows that f is necessarily also real analytic on $\mathbb{R}$, and we are done.

3.4. **The integration trick, and positivity certificates.** Observe that the inequality (3.5) can be written more generally as follows.

   Given a polynomial $p(t) = \sum_{k \geq 0} b_k t^k$ which takes non-negative values on $[-1, 1]$, as well as a positive semidefinite Hankel matrix $H = (s_{i+j})_{i,j \geq 0}$, we have that

$$\sum_{k \geq 0} b_k s_k \geq 0. \tag{3.7}$$

As shown in (3.5), this assertion is clear via an application of Hamburger's theorem. We now demonstrate how the assertion can instead be derived from first principles, with interesting connections to positivity certificates.

First note that the inequality (3.7) holds if $p(t)$ is the square of a polynomial. For instance, if $p(t) = (1 - 3t)^2 = 1 - 6t + 9t^2$ on $[-1, 1]$, then

$$s_0 - 6s_1 + 9s_2 = (e_0 - 3e_1)^T H (e_0 - 3e_1), \tag{3.8}$$

where $e_0 = (1, 0, 0, \ldots)$ and $e_1 = (0, 1, 0, 0, \ldots)$. The non-negativity of (3.8) now follows immediately from the positivity of the matrix H. The same reasoning applies if $p(t)$ is a sum of squares of polynomials, or even the limit of a sequence of sums of squares. Thus, one approach to showing the inequality (3.7) for an arbitrary polynomial $p(t)$ which is non-negative on $[-1, 1]$ is to seek a *limiting sum-of-squares representation*, which is also known as a *positivity certificate*, for p.

If a d-variate real polynomial is a sum of squares of real polynomials, then it is clearly non-negative on $\mathbb{R}^d$, but the converse is not true for $d > 1$.[3] Even when $d = 1$, while a sum-of-squares representation is an equivalent characterization for one-variable polynomials that are non-negative on $\mathbb{R}$, here we are working on the compact semi-algebraic set $[-1, 1]$. We now give three proofs of the existence of such a positivity certificate in the setting used above.

Proof 1. A result of Berg, Christensen, and Ressel (see the end of [14]) shows more generally that, for every dimension $d \geq 1$, any non-negative polynomial on $[-1, 1]^d$ has a limiting sum-of-squares representation. $\qquad \square$

---

[3]This is connected to semi-algebraic geometry and to Hilbert's seventeenth problem: recall the famous result of Motzkin that there are non-negative polynomials on $\mathbb{R}^d$ that are not sums of squares, such as $x^4 y^2 + x^2 y^4 - 3x^2 y^2 + 1$. Such phenomena have been studied in several settings, including polytopes (by Farkas, Handelman, and Pólya) and more general semi-algebraic sets (by Putinar, Schmüdgen, Stengel, Vasilescu, and others).

Proof 2. The only polynomials used in proving the stronger form of Schoenberg's theorem, Theorems 3.14 and 3.17, appear following (3.6):

$$p_{\pm,n}(t) := (1 \pm t)(1 - t^2)^n \qquad (n \geq 0).$$

Each of these polynomials is composed of factors of the form $p_{\pm,0}(t) = 1 \pm t$, so it suffices to produce a limiting sum-of-squares representation for these two polynomials on $[-1, 1]$. Note that

$$\frac{1}{2}(1 \pm t)^2 = \frac{1}{2} \pm t + \frac{t^2}{2},$$

$$\frac{1}{4}(1 - t^2)^2 = \frac{1}{4} - \frac{t^2}{2} + \frac{t^4}{4},$$

$$\frac{1}{8}(1 - t^4)^2 = \frac{1}{8} - \frac{t^4}{4} + \frac{t^8}{8},$$

and so on. Adding the first n equations shows that $(1 \pm t) + 2^{-n}(t^{2^n} - 1)$ is a sum-of-squares polynomial for all n. Taking $n \to \infty$ finishes the proof. $\qquad\square$

Proof 3. In fact, for any $d \geq 1$ and any compact set $K \subset \mathbb{R}^d$, if f is a non-negative continuous function on K, then f has a positivity certificate. The Stone–Weierstrass theorem gives a sequence of polynomials which converges to $\sqrt{f}$, and the squares of these polynomials then provide the desired limiting representation for f. This is a simpler proof than Proof 1 from [14], but the convergence here is uniform, whereas the convergence in [14] is stronger. $\qquad\square$

Remark 3.20. In (3.5), we used $H = H_{\sigma_\mu}$, which was positive semidefinite by assumption. The previous discussion shows that Theorem 3.17(4) can be further weakened, by requiring only that $f[H_\mu]$ is positive semidefinite, as opposed to being equal to H_σ for some admissible measure σ. Hence we do not require Hamburger's theorem in order to prove the strengthening of Schoenberg's theorem that uses the test set of low-rank Hankel matrices.

3.5. **Variants of moment-sequence transforms.** We now present a trio of results on functions which preserve moment sequences.

For $K \subset \mathbb{R}$, let $\mathcal{M}(K)$ denote the set of moment sequences corresponding to admissible measures with support in K. We say that F maps $\mathcal{M}(K)$ into $\mathcal{M}(L)$, where $K, L \subset \mathbb{R}$, if for every admissible measure μ with support in K there exists an admissible measure σ with support in L such that

$$F(s_k(\mu)) = s_k(\sigma) \quad \text{for all } k \in \mathbb{Z}_+,$$

where $s_k(\mu)$ is the kth-power moment of μ, as in Definition 3.16.

Theorem 3.21. *A function $F : \mathbb{R} \to \mathbb{R}$ maps $\mathcal{M}([-1, 1])$ into itself if and only if F is the restriction to $\mathbb{R}$ of an absolutely monotonic entire function.*

Theorem 3.22. *A function $F : \mathbb{R}_+ \to \mathbb{R}$ maps $\mathcal{M}([0, 1])$ into itself if and only if F is absolutely monotonic on $(0, \infty)$ and $0 \leq F(0) \leq \lim_{\epsilon \to 0^+} F(\epsilon)$.*

Theorem 3.23. *A function $F : \mathbb{R} \to \mathbb{R}$ maps $\mathcal{M}([-1,0])$ into $\mathcal{M}((-\infty, 0])$ if and only if there exists an absolutely monotonic entire function $\widetilde{F} : \mathbb{C} \to \mathbb{C}$ such that*

$$F(x) = \begin{cases} \widetilde{F}(x) & \text{if } x \in (0, \infty), \\ 0 & \text{if } x = 0, \\ -\widetilde{F}(-x) & \text{if } x \in (-\infty, 0). \end{cases}$$

It is striking to observe the possibility of a discontinuity at the origin which may occur in the latter two of these three theorems.

We will content ourselves here with sketching the proof of the second result. For the others, see [12], noting that the first of the results follows from Theorems 3.14 and 3.17 for $\rho = \infty$.

Proof of Theorem 3.22. Note that the moment matrix corresponding to an element of $\mathcal{M}([0,1])$ has a zero entry if and only if $\mu = a\delta_0$ for some $a \geq 0$. This and the Schur product theorem give one implication.

For the converse, suppose F preserves $\mathcal{M}([0,1])$. Fix finitely many scalars $c_j, t_j > 0$ and an integer $n \geq 0$, and set

$$p(t) = (1 - t)^n \quad \text{and} \quad \mu = \sum_j e^{-t_j \alpha} c_j \delta_{e^{-t_j h}}, \tag{3.9}$$

where $\alpha > 0$ and $h > 0$. If $g(x) := \sum_j c_j e^{-t_j x}$ then the integration trick (3.5), but working on $[0,1]$, shows that the forward finite differences of $F \circ g$ alternate in sign:

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} F\left(\sum_j c_j e^{-t_j(\alpha + kh)} \right) \geq 0,$$

so $(-1)^n \Delta_h^n (F \circ g)(\alpha) \geq 0$. As this holds for all $\alpha, h > 0$ and all $n \geq 0$, it follows that $F \circ g : (0, \infty) \to (0, \infty)$ is completely monotonic. The weak density of measures of the form μ, together with Bernstein's theorem (2.1), gives that $F \circ g$ is completely monotonic on $(0, \infty)$ for every completely monotonic function $g : (0, \infty) \to (0, \infty)$. Finally, a theorem of Lorch and Newman [96, Theorem 5] now gives that $F : (0, \infty) \to (0, \infty)$ is absolutely monotonic. $\square$

3.6. **Multivariable positivity preservers and moment families.** We now turn to the multivariable case, and begin with two results of FitzGerald, Micchelli, and Pinkus [52]. We first introduce some notation and a piece of terminology.

Fix $I \subset \mathbb{C}$ and an integer $m \geq 1$, and let

$$A^k = (a_{ij}^k)_{i,j=1}^N \in I^{N \times N} \quad \text{for } k = 1, \dots, m.$$

For any function $f : I^m \to \mathbb{C}$, we have the $N \times N$ matrix

$$f(A^1, \dots, A^m) := \left(f(a_{ij}^1, \dots, a_{ij}^m) \right)_{i,j=1}^N \in \mathbb{C}^{N \times N}.$$

We say that $f : \mathbb{R}^m \to \mathbb{R}$ is *real positivity preserving* if

$$f(A^1, \ldots, A^m) \in \mathcal{P}_N(\mathbb{R}) \text{ for all } A^1, \ldots, A^m \in \mathcal{P}_N(\mathbb{R}) \text{ and all } N \geq 1,$$

where, as above $\mathcal{P}_N(\mathbb{R})$ is the collection of $N \times N$ positive semidefinite matrices with real entries. Similarly, we say that $f : \mathbb{C}^m \to \mathbb{C}$ is *positivity preserving* if

$$f(A^1, \ldots, A^m) \in \mathcal{P}_N \qquad \text{for all } A^1, \ldots, A^m \in \mathcal{P}_N \text{ and all } N \geq 1,$$

where $\mathcal{P}_N$ is the collection of $N \times N$ positive semidefinite matrices with complex entries. Finally, recall that a function $f : \mathbb{R}^m \to \mathbb{R}$ is said to be *real entire* if there exists an entire function $F : \mathbb{C}^m \to \mathbb{C}$ such that $F|_{\mathbb{R}^m} = f$. We will also use the multi-index notation

$$\mathbf{x}^\alpha := x_1^{\alpha_1} \cdots x_m^{\alpha_m} \quad \text{if } \mathbf{x} = (x_1, \ldots, x_m) \text{ and } \alpha = (\alpha_1, \ldots, \alpha_m).$$

The following theorems are natural extensions of Schoenberg's theorem and Herz's theorem, respectively.

Theorem 3.24 ([52, Theorem 2.1]). *Let $f : \mathbb{R}^m \to \mathbb{R}$, where $m \geq 1$. Then f is real positivity preserving if and only f is real entire of the form*

$$f(\mathbf{x}) = \sum_{\alpha \in \mathbb{Z}_+^m} c_\alpha \mathbf{x}^\alpha \qquad (\mathbf{x} \in \mathbb{R}^m),$$

where $c_\alpha \geq 0$ for all $\alpha \in \mathbb{Z}_+^m$.

Theorem 3.25 ([52, Theorem 3.1]). *Let $f : \mathbb{C}^m \to \mathbb{C}$, where $m \geq 1$. Then f is positivity preserving if and only f is of the form*

$$f(\mathbf{z}) = \sum_{\alpha, \beta \in \mathbb{Z}_+^m} c_{\alpha\beta} \mathbf{z}^\alpha \overline{\mathbf{z}}^\beta \qquad (\mathbf{z} \in \mathbb{C}^m),$$

where $c_{\alpha\beta} \geq 0$ for all $\alpha, \beta \in \mathbb{Z}_+^m$ and the power series converges absolutely for all $\mathbf{z} \in \mathbb{C}$.

We now consider the notion of moment family for measures on $\mathbb{R}^d$. As above, a measure on $\mathbb{R}^d$ is said to be *admissible* if it is non-negative and has moments of all orders. Given such a measure μ, we define the *moment family*

$$s_\alpha(\mu) := \int \mathbf{x}^\alpha \, d\mu(\mathbf{x}) \qquad \text{for all } \alpha \in \mathbb{Z}_+^m.$$

In line with the above, we let $\mathcal{M}(K)$ denote the set of all moment families of admissible measures supported on $K \subset \mathbb{R}^d$.

Note that a measure μ is supported in $[-1, 1]^d$ if and only if its moment family is uniformly bounded:

$$\sup\{|s_\alpha(\mu)| : \alpha \in \mathbb{Z}_+^m\} < \infty.$$

Theorem 3.26 ([12, Theorem 8.1]). *A function $F : \mathbb{R} \to \mathbb{R}$ maps $\mathcal{M}([-1, 1]^d)$ to itself if and only if F is absolutely monotonic and entire.*

Proof. Since $[-1, 1]$ can be identified with $[-1, 1] \times \{0\}^{d-1} \subset [-1, 1]^d$, the forward implication follows from the one-dimensional result, Theorem 3.21.

For the converse, we use the fact [112] that a collection of real numbers $(s_\alpha)_{\alpha \in \mathbb{Z}_+^d}$ is an element of $\mathcal{M}([-1, 1]^d)$ if and only if the weighted Hankel-type kernels on $\mathbb{Z}_+^d \times \mathbb{Z}_+^d$

$$(\alpha, \beta) \mapsto s_{\alpha+\beta} \quad \text{and} \quad (\alpha, \beta) \mapsto s_{\alpha+\beta} - s_{\alpha+\beta+2\mathbf{1}_j} \quad (1 \leq j \leq d)$$

are positive semidefinite, where

$$\mathbf{1}_j := (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{Z}_+^d$$

with 1 in the jth position. Now suppose F is absolutely monotonic and entire; given a family $(s_\alpha)_{\alpha \in \mathbb{Z}_+^d}$ subject to these positivity constraints, we have to verify that the family $(F(s_\alpha))_{\alpha \in \mathbb{Z}_+^d}$ satisfies them as well.

Theorem 3.14 gives that $(\alpha, \beta) \mapsto F(s_{\alpha+\beta})$ and $(\alpha, \beta) \mapsto F(s_{\alpha+\beta+2\mathbf{1}_j})$ are positive semidefinite, so we must show that

$$(\alpha, \beta) \mapsto F(s_{\alpha+\beta}) - F(s_{\alpha+\beta+2\mathbf{1}_j})$$

is positive semidefinite for $j = 1, \ldots, d$. As F is absolutely monotonic and entire, it suffices to show that

$$(\alpha, \beta) \mapsto (s_{\alpha+\beta})^{\circ n} - (s_{\alpha+\beta+2\mathbf{1}_j})^{\circ n}$$

is positive semidefinite for any $n \geq 0$, but this follows from the Schur product theorem: if $A \geq B \geq 0$, then

$$A^{\circ n} \geq A^{\circ(n-1)} \circ B \geq A^{\circ(n-2)} \circ B^{\circ 2} \geq \cdots \geq B^{\circ n}. \qquad \square$$

We next consider characterizations of real-valued multivariable functions which map tuples of moment sequences to moment sequences.

Let $K_1, \ldots, K_m \subset \mathbb{R}$. A function $F : \mathbb{R}^m \to \mathbb{R}$ acts on tuples of moment sequences of (admissible) measures $\mathcal{M}(K_1) \times \cdots \times \mathcal{M}(K_m)$ as follows:

$$F[\mathbf{s}(\mu_1), \ldots, \mathbf{s}(\mu_m)]_k := F\big(s_k(\mu_1), \ldots, s_k(\mu_m)\big) \quad \text{for all } k \geq 0. \qquad (3.10)$$

Given $I \subset \mathbb{R}^m$, a function $F : I \to \mathbb{R}$ is *absolutely monotonic* if F is continuous on I, and for all interior points $\mathbf{x} \in I$ and $\alpha \in \mathbb{Z}_+^m$, the mixed partial derivative $D^\alpha F(\mathbf{x})$ exists and is non-negative, where

$$D^\alpha F(\mathbf{x}) := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_m^{\alpha_m}} F(x_1, \ldots, x_m) \quad \text{and} \quad |\alpha| := \alpha_1 + \cdots + \alpha_m.$$

With this definition, the multivariable analogue of Bernstein's theorem is as one would expect; see [27, Theorem 4.2.2].

To proceed further, it is necessary to introduce the notion of a *facewise absolutely monotonic function* on $\mathbb{R}_+^m$. Observe that the orthant $\mathbb{R}_+^m$ is a convex polyhedron, and is therefore the disjoint union of the relative interiors of its faces. These faces are in one-to-one correspondence with subsets of $[m] := \{1, \ldots, m\}$:

$$J \mapsto \mathbb{R}_+^J := \{(x_1, \ldots, x_m) \in \mathbb{R}_+^m : x_i = 0 \text{ if } i \notin J\}; \qquad (3.11)$$

note that this face has relative interior $\mathbb{R}^J_{>0} := (0, \infty)^J \times \{0\}^{[m]\setminus J}$.

Definition 3.27. A function $F : \mathbb{R}^m_+ \to \mathbb{R}$ is *facewise absolutely monotonic* if, for every $J \subset [m]$, there exists an absolutely monotonic function g_J on $\mathbb{R}^J_+$ which agrees with F on $\mathbb{R}^J_{>0}$.

Thus a facewise absolutely monotonic function is piecewise absolutely monotonic, with the pieces being the relative interiors of the faces of the orthant $\mathbb{R}^m_+$. See [12, Example 8.4] for further discussion. In the special case $m = 1$, this broader class of functions (than absolutely monotonic functions on $\mathbb{R}_+$) coincides precisely with the maps which are absolutely monotonic on $(0, \infty)$ and have a possible discontinuity at the origin, as in Theorem 3.22 above.

This definition allows us to characterize the preservers of m-tuples of elements of $\mathcal{M}\big([0, 1]\big)$; the preceding observation shows that Theorem 3.22 is precisely the $m = 1$ case.

Theorem 3.28 ([12, Theorem 8.5]). *Let $F : \mathbb{R}^m_+ \to \mathbb{R}$, where the integer $m \geq 1$. The following are equivalent.*

(1) *F maps $\mathcal{M}([0, 1])^m$ into $\mathcal{M}([0, 1])$.*
(2) *F is facewise absolutely monotonic, and the functions $\{g_J : J \subset [m]\}$ are such that $0 \leq g_J \leq g_K$ on $\mathbb{R}^J_+$ whenever $J \subset K \subset [m]$.*
(3) *F is such that*

$$F\big(\sqrt{x_1 y_1}, \ldots, \sqrt{x_m y_m}\big)^2 \leq F(x_1, \ldots, x_m) F(y_1, \ldots, y_m)$$

for all $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^m_+$ and there exists some $\mathbf{z} \in (0, 1)^m$ such that the products $\mathbf{z}^\alpha := z_1^{\alpha_1} \cdots z_m^{\alpha_m}$ are distinct for all $\alpha \in \mathbb{Z}^m_+$ and F maps $\mathcal{M}\big(\{1, z_1\}\big) \times \cdots \times \mathcal{M}\big(\{1, z_m\}\big) \cup \mathcal{M}(\{0, 1\})^m$ to $\mathcal{M}(\mathbb{R})$.

The heart of Theorem 3.28 can be deduced from the following result on positivity preservation on tuples of low-rank Hankel matrices. In a sense, it is the multi-dimensional generalization of the 'stronger Vasudeva theorem' 3.18.

Fix $\rho \in (0, \infty]$, an integer $m \geq 1$ and a point $\mathbf{z} \in (0, 1)^m$ with distinct products, as in Theorem 3.28(3). For all $N \geq 1$, let

$$\mathcal{H}_N := \{a\mathbf{1}_{N\times N} + b\mathbf{u}_{l,N}\mathbf{u}_{l,N}^T : a \in (0, \rho),\ b \in [0, \rho - a),\ 1 \leq l \leq m\},$$

where $\mathbf{u}_{l,N} := (1, z_l, \ldots, z_l^{N-1})^T$.

Theorem 3.29 ([12, Theorem 8.6]). *If $F : (0, \rho)^m \to \mathbb{R}$ preserves positivity on $\mathcal{P}_2\big((0, \rho)\big)^m$ and $\mathcal{H}_N^m$ for all $N \geq 1$, then F is absolutely monotonic and is the restriction of an analytic function on the polydisc $D(0, \rho)^m$.*

The notion of facewise absolute monotonicity emerges from the study of positivity preservers of tuples of moment sequences. If one focuses instead on maps preserving positivity of tuples of all positive semidefinite matrices, or even all Hankel matrices, then this richer class of maps does not appear.

Proposition 3.30. *Suppose $\rho \in (0, \infty]$ and $F : [0, \rho)^m \to \mathbb{R}$. The following are equivalent.*

(1) *$F[-]$ preserves positivity on the space of m-tuples of Hankel matrices with entries in $[0, \rho)$.*

(2) *F is absolutely monotonic on $[0, \rho)^m$.*

(3) *$F[-]$ preserves positivity on the space of m-tuples of all matrices with entries in $[0, \rho)$.*

Proof. Clearly (2) $\implies$ (3) $\implies$ (1), so suppose (1) holds. It follows from Theorem 3.29 that F is absolutely monotonic on the domain $(0, \rho)^m$ and agrees there with an analytic function $g : D(0, \rho)^m \to \mathbb{C}$. To see that $F \equiv g$ on $[0, \rho)^m$, we use induction on m, with the $m = 1$ case being left as an exercise (see [12, Proof of Proposition 7.3]).

Now suppose $m > 1$, let $\mathbf{c} = (c_1, \ldots, c_m) \in [0, \rho)^m \setminus (0, \rho)^m$ and define

$$
H := \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad \text{and} \quad A_i := \begin{cases} \mathbf{1}_{3 \times 3} & \text{if } c_i > 0, \\ H & \text{if } c_i = 0. \end{cases}
$$

Choosing $\mathbf{u}_n = (u_{1,n}, \ldots, u_{m,n}) \in (0, \rho)^m$ such that $\mathbf{u}_n \to \mathbf{c}$, it follows that

$$
\lim_{n \to \infty} F[u_{1,n} A_1, \ldots, u_{m,n} A_m] = \begin{bmatrix} g(\mathbf{c}) & F(\mathbf{c}) & g(\mathbf{c}) \\ F(\mathbf{c}) & g(\mathbf{c}) & g(\mathbf{c}) \\ g(\mathbf{c}) & g(\mathbf{c}) & g(\mathbf{c}) \end{bmatrix} \in \mathcal{P}_3,
$$

where the $(1, 2)$ and $(2, 1)$ entries are as claimed by the induction hypothesis. The determinants of the first and last principal minors now give that

$$
g(\mathbf{c}) \geq 0 \quad \text{and} \quad -g(\mathbf{c})\big(g(\mathbf{c}) - F(\mathbf{c})\big)^2 \geq 0,
$$

whence $F(\mathbf{c}) = g(\mathbf{c})$. $\qquad\qquad\square$

Having considered functions defined on the positive orthant, we now look at the situation for functions defined over the whole of $\mathbb{R}^m$.

Theorem 3.31 ([12, Theorem 8.9]). *Suppose $F : \mathbb{R}^m \to \mathbb{R}$ for some integer $m \geq 1$. The following are equivalent.*

(1) *F maps $\mathcal{M}\big([-1, 1]\big)^m$ into $\mathcal{M}(\mathbb{R})$.*

(2) *The function F is real positivity preserving.*

(3) *The function F is absolutely monotonic on $\mathbb{R}_+^m$ and agrees with an entire function on $\mathbb{R}^m$.*

As before, the proof reveals that verifying positivity preservation for tuples of low-rank Hankel matrices suffices. The following notation and corollary make this precise.

For all $u \in (0, \infty)$, let $\mathcal{M}_u := \mathcal{M}\big(\{-1, u, 1\}\big)$ and

$$
\mathcal{M}_{[u]} := \bigcup \{\mathcal{M}\big(\{s_1, s_2\}\big) : s_1 \in \{-1, 0, 1\}, \ s_2 \in \{-u, 0, u\}\}.
$$

Corollary 3.32 ([12, Theorem 8.10]). *The hypotheses in Theorem 3.31 are also equivalent to the following.*

(4) *There exist $u_0 \in (0,1)$ and $\epsilon > 0$ such that F maps*

$$\mathcal{M}^m_{[u_0]} \cup \bigcup \{\mathcal{M}_{v_1} \times \cdots \times \mathcal{M}_{v_m} : v_1, \ldots, v_m \in (0, 1 + \epsilon)\}$$

into $\mathcal{M}(\mathbb{R})$.

4. Entrywise polynomials preserving positivity in fixed dimension

Having discussed at length the dimension-free setting, we now turn our attention to functions that preserve positivity in a fixed dimension $N \geq 2$. This is a natural question from the standpoint of both theory as well as applications. This latter connection to applied fields and to high-dimensional covariance estimation will be explained below in Chapter 7.

Mathematically, understanding the functions f such that $f[-] : \mathcal{P}_N \to \mathcal{P}_N$ for fixed $N \geq 2$, is a non-trivial and challenging refinement of Schoenberg's 1942 theorem. A complete characterization was found for $N = 2$ by Vasudeva [134]:

Theorem 4.1 (Vasudeva [134]). *Given a function $f : (0, \infty) \to \mathbb{R}$, the entrywise map $f[-]$ preserves positivity on $\mathcal{P}_2\big((0,\infty)\big)$ if and only f is non-negative, non-decreasing, and multiplicatively mid-convex:*

$$f(x)f(y) \geq f\big(\sqrt{xy}\big)^2 \qquad \text{for all } x, y > 0.$$

In particular, f is either identically zero or never zero on $(0, \infty)$, and f is also continuous.

On the other hand, if $N \geq 3$, then such a characterization remains open to date. As mentioned above, perhaps the only known result for general entrywise preservers is the Horn–Loewner theorem 3.4 (or its more general variants such as Theorem 3.10).

In light of this challenging scarcity of results in fixed dimension, a strategy adopted in the literature has been to further refine the problem, in one of several ways:

(1) Restrict the class of functions, while operating entrywise on all of $\mathcal{P}_N$ (over some given domain I, say $(0, \rho)$ or $(-\rho, \rho)$ for $0 < \rho \leq \infty$). For example, in this survey we consider possibly non-integer power functions, polynomials and power series, and even linear combinations of real powers.

(2) Restrict the class of matrices and study entrywise functions over this class in a fixed dimension. For instance, popular sub-classes of matrices include positive matrices with rank bounded above, or with a given sparsity pattern (zero entries), or classes such as Hankel or Toeplitz matrices; or intersections of these classes. For instance, in discussing the Horn–Loewner and Schoenberg–Rudin results, we encountered Toeplitz and Hankel matrices of low rank.

(3) Study the problem under both of the above restrictions.

In this chapter we begin with the first of these restrictions. Specifically, we will study polynomial maps that preserve positivity, when applied entrywise to $\mathcal{P}_N$. Recall from the Schur product theorem that if the polynomial f has only non-negative coefficients then $f[-]$ preserves positivity on $\mathcal{P}_N$ for every dimension $N \geq 1$. It is natural to expect that if one reduces the test set, from all dimensions to a fixed dimension, then the class of polynomial preservers should be larger. Remarkably, until 2016 not a single example was known of a polynomial positivity preserver with a negative coefficient. Then, in quick succession, the two papers [11, 89] provided a complete understanding of the sign patterns of entrywise polynomial preservers of $\mathcal{P}_N$. The goal of this chapter is to discuss some of the results in these works.

4.1. **Characterizations of sign patterns.** Until further notice, we work with entrywise polynomial or power-series maps of the form

$$f(x) = c_0 x^{n_0} + c_1 x^{n_1} + \cdots, \quad \text{with } 0 \leq n_0 < n_1 < \cdots, \qquad (4.1)$$

and $c_j \in \mathbb{R}$ typically non-zero, which preserve $\mathcal{P}_N(I)$ for various I. Our goal is to try and understand their sign patterns, that is, which c_j can be negative. The first observation is that as soon as I contains the interval $(0, \rho)$ for any $\rho > 0$, by the Horn–Loewner type necessary conditions in Lemma 3.9, the lowest N non-zero coefficients of $f(x)$ must be positive.

The next observation is that if $I \not\subset \mathbb{R}_+$, then, in general, there is no structured classification of the sign patterns of the power series preservers on $\mathcal{P}_N(I)$. For example, let k be a non-negative integer; the polynomials

$$f_{k,t}(x) := t(1 + x^2 + \cdots + x^{2k}) - x^{2k+1} \qquad (t > 0)$$

do not preserve positivity entrywise on $\mathcal{P}_N\big((-\rho, \rho)\big)$ for any $N \geq 2$. This may be seen by taking $\mathbf{u} := (1, -1, 0, \ldots, 0)^T$ and $A := \eta \mathbf{u}\mathbf{u}^T$ for some $0 < \eta < \rho$, and noting that

$$\mathbf{u}^T f_{k,t}[A]\mathbf{u} = -4\eta^{2k+1} < 0.$$

Similarly, if one allows complex entries and uses higher-order roots of unity, such negative results (vis-à-vis Lemma 3.9) are obtained for complex matrices.

Given this, in the rest of the chapter we will focus on $I = (0, \rho)$ for $0 < \rho \leq \infty$.[4] As mentioned above, if f as in (4.1) entrywise preserves positivity even on rank-one matrices in $\mathcal{P}_N\big((0, \rho)\big)$ then its first N non-zero Maclaurin coefficients are positive. Our goal is to understand if any other coefficient can be negative (and if so, which of them). This has at least two ramifications:

(1) It would yield the first example of a polynomial entrywise map (for a fixed dimension) with at least one negative Maclaurin coefficient.

---

[4]That said, we also briefly discuss the one situation in which our results do apply more generally, even to $I = D(0, \rho) \subset \mathbb{C}$ (an open complex disc).

Recall the contrast to Schoenberg's theorem in the dimension-free setting.

(2) This also yields the first example of a polynomial (or power series) that entrywise preserves positivity on $\mathcal{P}_N(I)$ but not $\mathcal{P}_{N+1}(I)$. In particular it would imply that the Horn–Loewner type necessary condition in Lemma 3.9(1) is "sharp".

These goals are indeed achieved in the particular case $n_0 = 0, \ldots, n_{N-1} = N - 1$ in [11], and subsequently, for arbitrary $n_0 < \cdots < n_{N-1}$ in [89]. (In fact, in the latter work the n_j need not even be integers; this is discussed below.) Here is a 'first' result along these lines. Henceforth we assume that $\rho < \infty$; we will relax this assumption midway through Section 4.5 below.

Theorem 4.2 (Belton–Guillot–Khare–Putinar [11] and Khare–Tao [89]). *Suppose $N \geq 2$ and $n_0 < \cdots < n_{N-1}$ are non-negative integers, and ρ, c_0, $\ldots$, c_{N-1} are positive scalars. Given $\epsilon_M \in \{0, \pm 1\}$ for all $M > n_{N-1}$, there exists a power series*

$$f(x) = c_0 x^{n_0} + \cdots + c_{N-1} x^{n_{N-1}} + \sum_{M > n_{N-1}} d_M x^M$$

such that f is convergent on $(0, \rho)$, the entrywise map $f[-]$ preserves positivity on $\mathcal{P}_N\big((0, \rho)\big)$ and d_M has the same sign (positive, negative or zero) as ϵ_M for all $M > n_{N-1}$.

Outline of proof. The claim is such that it suffices to show the result for exactly one $\epsilon_M = -1$. Indeed, given the claim, for each $M > n_{N-1}$ there exists $\delta_M \in (0, 1/M!)$ such that $\sum_{j=0}^{N-1} c_j x^{n_j} + d x^M$ preserves positivity entrywise on $\mathcal{P}_N\big((0, \rho)\big)$ whenever $|d| \leq \delta_M$. Now let $d_M := \epsilon_M \delta_M$ for all $M > n_{N-1}$, and define

$$f_M(x) := \sum_{j=0}^{N_1} c_j x^{n_j} + d_M x^M \quad \text{and} \quad f(x) := \sum_{M > n_{N-1}} 2^{n_{N-1} - M} f_M(x).$$

Then it may be verified that $|f(x)| \leq \sum_{j=0}^{N-1} c_j x^{n_j} + 2^{n_{N-1}} e^{x/2}$, and hence f has the desired properties. □

Thus it suffices to show the existence of a polynomial positivity preserver on $\mathcal{P}_N\big((0, \rho)\big)$ with precisely one negative Maclaurin coefficient, the leading term. In the next few sections we explain how to achieve this goal. In fact, one can show a more general result, for real powers as well.

Theorem 4.3 (Khare–Tao [89]). *Fix an integer $N \geq 2$ and real exponents $n_0 < \cdots < n_{N-1} < M$ in the set $\mathbb{Z}_+ \cup [N - 2, \infty)$. Suppose ρ, c_0, $\ldots$, $c_{N-1} > 0$ as above. Then there exists $c' < 0$ such that the function*

$$f(x) = c_0 x^{n_0} + \cdots + c_{N-1} x^{n_{N-1}} + c' x^M \qquad \big(x \in (0, \rho)\big)$$

preserves positivity entrywise on $\mathcal{P}_N\big((0, \rho)\big)$. [Here and below, we set $0^0 := 1$.]

The restriction of the n_j lying in $\mathbb{Z}_+ \cup [N-2, \infty)$ is a technical one that is explained in a later chapter on the study of entrywise powers preserving positivity on $\mathcal{P}_N\big((0, \infty)\big)$; see Theorem 6.1.

Remark 4.4. A stronger result, Theorem 4.15, which also applies to real powers, is stated below. We mention numerous ramifications of the results in this chapter following that result.

The proofs of the preceding two theorems crucially use type-A representation theory (specifically, a family of symmetric functions) that naturally emerges here via generalized Vandermonde determinants. These symmetric homogeneous polynomials are introduced and used in the next section.

For now, we explain how Theorem 4.3 helps achieve a complete classification of the sign patterns of a family of generalised power series, of the form

$$f(x) = \sum_{j=0}^{\infty} c_j x^{n_j}, \qquad n_j \in \mathbb{Z}_+ \cup [N-2, \infty) \text{ for all } j \geq 0,$$

but without the requirement that that exponents are non-decreasing. In this generality, one first notes that the Horn–Loewner-type Lemma 3.9 still applies: if some coefficient $c_{j_0} < 0$, then there must be at least N indices j such that $n_j < n_{j_0}$ and $c_j > 0$. The following result shows that once again, this necessary condition is the best possible.

Theorem 4.5 (Classification of sign patterns for real power-series preservers, Khare–Tao [89]). *Fix an integer $N \geq 2$, and distinct real exponents $n_0, n_1, \ldots$ in $\mathbb{Z}_+ \cup [N-2, \infty)$. Suppose $\epsilon_j \in \{0, \pm 1\}$ is a choice of sign for each $j \geq 0$, such that if $\epsilon_{j_0} = -1$ then $\epsilon_j = +1$ for at least N choices of j such that $n_j < n_{j_0}$. Given any $\rho > 0$, there exists a choice of coefficients c_j with sign ϵ_j such that*

$$f(x) := \sum_{j=0}^{\infty} c_j x^{n_j}$$

is convergent on $(0, \rho)$ and preserves positivity entrywise on $\mathcal{P}_N\big((0, \rho)\big)$.

Notice this result is strictly more general than Theorem 4.2, because the sequence $n_0, n_1, \ldots$ can contain an infinite decreasing sequence of positive non-integer powers, for example, all rational elements of $[N-2, \infty)$. Thus Theorem 4.5 covers a larger class of functions than even Hahn or Puiseux series.

Theorem 4.5 is derived from Theorem 4.3 in a similar fashion to the proof of Theorem 4.2, and we refer the reader to [89, Section 1] for the details.

4.2. **Schur polynomials; the sharp threshold bound for a single matrix.** We now explain how to prove Theorem 4.3. The present section will discuss the case of integer powers, and end by proving the theorem for a single 'generic' rank-one matrix. In the following section we show

how to extend the results to all rank-one matrices for integer powers. The subsequent section will complete the proof for real powers, and then for matrices of all ranks.

The key new tool that is indispensable to the following analysis is that of *Schur polynomials*. These can be defined in a number of equivalent ways; we refer the reader to [30] for more details, including the equivalence of these definitions shown using ideas of Karlin–Macgregor, Lindström, and Gessel–Viennot. For our purposes the definition of Cauchy is the most useful:

Definition 4.6. Given non-negative integers $N \geq 1$ and $n_0 < \cdots < n_{N-1}$, let

$$\mathbf{n} := (n_0, \ldots, n_{N-1})^T, \qquad \text{and} \quad \mathbf{n}_{\min} := (0, 1, \ldots, N-1)^T,$$

and define $V(\mathbf{n}) := \prod_{0 \leq i < j \leq N-1} (n_j - n_i)$.

Given a vector $\mathbf{u} = (u_1, \ldots, u_N)^T$ and a non-negative integer k, let $\mathbf{u}^{\circ k} := (u_1^k, \ldots, u_N^k)^T$, and let $\mathbf{u}^{\circ \mathbf{n}}$ be the $N \times N$ matrix with (j, k) entry $u_j^{n_{k-1}}$.

The *Schur polynomial* in variables $u_1, \ldots, u_N$ of *degree* $\mathbf{n}$ is given by

$$s_{\mathbf{n}}(\mathbf{u}) := \frac{\det \mathbf{u}^{\circ \mathbf{n}}}{\det \mathbf{u}^{\circ \mathbf{n}_{\min}}}. \tag{4.2}$$

Notice that the numerator is a generalized Vandermonde determinant, so a homogeneous and alternating polynomial, while the denominator is the usual Vandermonde determinant in the indeterminates $u_1, \ldots, u_N$. Hence their ratio $s_{\mathbf{n}}(\mathbf{u})$ is a homogeneous symmetric polynomial in $\mathbb{Z}[u_1, \ldots, u_N]$. It follows that Schur polynomials are well defined when working over any commutative unital ring.

Schur polynomials are an extremely well-studied family of symmetric functions. Their appeal lies in the important observation that they are the characters of all irreducible (finite-dimensional) polynomial representations of the complex Lie group $GL_n(\mathbb{C})$ (or of the Lie algebra $\mathfrak{sl}_{n+1}(\mathbb{C})$). In this setting, the definition of Cauchy is a special case of the Weyl character formula. Thus, its specialization yields the corresponding Weyl dimension formula, which will be of use below:

$$s_{\mathbf{n}}((1, \ldots, 1)^T) = \prod_{0 \leq i < j \leq N-1} \frac{n_j - n_i}{j - i} = \frac{V(\mathbf{n})}{V(\mathbf{n}_{\min})}. \tag{4.3}$$

An alternate proof of (4.3) comes from the *principal specialization formula*: for a variable q, one has that

$$s_{\mathbf{n}}((1, q, \ldots, q^{N-1})^T) = \prod_{0 \leq i < j \leq N-1} \frac{q^{n_j} - q^{n_i}}{q^j - q^i}; \tag{4.4}$$

this follows from (4.2) because now the numerator is also a standard Vandermonde determinant. We also refer the reader to [98] for many more results and properties of Schur polynomials.

Returning to polynomial positivity preservers, we wish to consider functions of the form

$$f(x) = c_0 x^{n_0} + \cdots + c_{N-1} x^{n_{N-1}} + c' x^M,$$

with non-negative integers $n_0 < \cdots < n_{N-1} < M$ and positive coefficients $c_0, \ldots, c_{N-1}$. We are interested in characterizing those $c' \in \mathbb{R}$ for which the entrywise map $f[-]$ preserves positivity on $\mathcal{P}_N\big((0, \rho)\big)$. By the Schur product theorem, this is equivalent to finding the smallest c' such that $f[-]$ is a preserver. We may assume that $c' < 0$, so we rescale by $t := |c'|^{-1}$ and define

$$p_t(x) := t \sum_{j=0}^{N-1} c_j x^{n_j} - x^M. \tag{4.5}$$

The goal now is to find the smallest $t > 0$ such that $p_t[-]$ preserves positivity on $\mathcal{P}_N\big((0, \rho)\big)$. We next achieve this goal for a single rank-one matrix.

Proposition 4.7. *With notation as above, define*

$$\mathbf{n}_j = (n_0, \ldots, n_{j-1}, \widehat{n_j}, n_{j+1}, \ldots, n_{N-1}, M)^T$$

for $0 \leq j \leq N - 1$. Given a vector $\mathbf{u} \in (0, \infty)^N$ with distinct coordinates, the following are equivalent.

(1) *The matrix $p_t[\mathbf{u}\mathbf{u}^T]$ is positive semidefinite.*
(2) $\det p_t[\mathbf{u}\mathbf{u}^T] \geq 0$.
(3) $t \geq \displaystyle\sum_{j=0}^{N-1} \frac{s_{\mathbf{n}_j}(\mathbf{u})^2}{c_j s_{\mathbf{n}}(\mathbf{u})^2}.$

In particular, this shows that for a generic rank-one matrix in $\mathcal{P}_N\big((0, \rho)\big)$, there does exist a positivity-preserving polynomial with a negative leading term.

In essence, the equivalences in Proposition 4.7 hold more generally; this is distilled into the following lemma.

Lemma 4.8 (Khare–Tao [90][5]). *Fix $\mathbf{w} \in \mathbb{R}^N$ and a positive-definite matrix H. Fix $t > 0$ and define $P_t := tH - \mathbf{w}\mathbf{w}^T$. The following are equivalent.*

(1) P_t *is positive semidefinite.*
(2) $\det P_t \geq 0$.
(3) $t \geq \mathbf{w}^T H^{-1} \mathbf{w} = 1 - \dfrac{\det(H - \mathbf{w}\mathbf{w}^T)}{\det H}.$

We refer the reader to [90] for the detailed proof of Lemma 4.8, remarking only that the equality in assertion (3) follows by using Schur complements in two different ways to expand the determinant of the matrix $\begin{bmatrix} H & \mathbf{w} \\ \mathbf{w}^T & 1 \end{bmatrix}$.

---

[5]The work [90] is an extended abstract of the paper [89], but some of the results in it have different proofs from [89].

Now Proposition 4.7 follows directly from Lemma 4.8, by setting

$$H = \sum_{j=0}^{N-1} c_j \mathbf{u}^{\circ n_j} (\mathbf{u}^{\circ n_j})^T \quad \text{and} \quad \mathbf{w} = \mathbf{u}^{\circ M},$$

where H is positive definite because of the following general matrix factorization (which is also used below).

Proposition 4.9. *Let $f(x) = \sum_{k=0}^{M} f_k x^k$ be a polynomial with coefficients in a commutative ring R. For any integer $N \geq 1$ and any vectors $\mathbf{u} = (u_1, \dots, u_N)^T$ and $\mathbf{v} = (v_1, \dots, v_N)^T \in R^N$, it holds that*

$$f[t\mathbf{u}\mathbf{v}^T] = \sum_{k=0}^{M} f_k t^k \mathbf{u}^{\circ k} (\mathbf{v}^{\circ k})^T \tag{4.6}$$

$$= \begin{bmatrix} 1 & u_1 & \cdots & u_1^M \\ 1 & u_2 & \cdots & u_2^M \\ \vdots & \vdots & \ddots & \vdots \\ 1 & u_N & \cdots & u_N^M \end{bmatrix} \begin{bmatrix} f_0 & 0 & \cdots & 0 \\ 0 & f_1 t & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f_M t^M \end{bmatrix} \begin{bmatrix} 1 & v_1 & \cdots & v_1^M \\ 1 & v_2 & \cdots & v_2^M \\ \vdots & \vdots & \ddots & \vdots \\ 1 & v_N & \cdots & v_N^M \end{bmatrix}^T,$$

where 1 is a multiplicative identity which is adjoined to R if necessary.

Now to adopt Lemma 4.8(3), this same equation and the Cauchy–Binet formula allow one to compute $\det(H - \mathbf{w}\mathbf{w}^T)$ in the present situation, and this yields precisely that $t \geq \sum_{j=0}^{N-1} \dfrac{s_{\mathbf{n}_j}(\mathbf{u})^2}{c_j s_{\mathbf{n}}(\mathbf{u})^2}$, as desired.

4.3. The threshold for all rank-one matrices: a Schur positivity result.

We continue toward a proof of Theorem 4.3. The next step is to use Proposition 4.7 to achieve an intermediate goal: a threshold bound for c' that works for all rank-one matrices in $\mathcal{P}_N\big((0, \rho)\big)$, still working with integer powers. Clearly, to do so one has to understand the supremum of each ratio $R_j := s_{\mathbf{n}_j}(\mathbf{u})^2 / s_{\mathbf{n}}(\mathbf{u})^2$, as $\mathbf{u}$ runs over vectors in $(0, \sqrt{\rho})^N$ with distinct coordinates. More precisely, one has to understand the supremum of the weighted sum $\sum_j R_j / c_j$.

This observation was first made in the work [11] for the case $n_j = j$, that is, $\mathbf{n} = \mathbf{n}_{\min}$. It led to the first proof of Theorem 4.3, with all of the denominators being the same: $s_{\mathbf{n}_{\min}}(\mathbf{u}) = 1$. We now use another equivalent definition of Schur polynomials, by Littlewood, realizing them as sums of monomials corresponding to certain Young tableaux. Every monomial has a non-negative integer coefficient. It follows by the continuity and homogeneity of $s_{\mathbf{n}_j}$ and the Weyl Dimension Formula (4.3), that the supremum in the previous paragraph equals the value at $(\sqrt{\rho}, \dots, \sqrt{\rho})^T$, namely

$$\sup_{\mathbf{u} \in (0, \sqrt{\rho})^N} s_{\mathbf{n}_j}(\mathbf{u})^2 = \frac{V(\mathbf{n}_j)^2}{V(\mathbf{n}_{\min})^2} \rho^{M - n_j}.$$

Since all of these suprema are attained at the same point $\sqrt{\rho}(1,\dots,1)^T$, the weighted sum in Proposition 4.7(3) also attains its supremum at the same point. Thus, we conclude using Proposition 4.7 that

$$f(x) = \sum_{j=0}^{N-1} c_j x^{n_j} + c' x^M$$

preserves positivity entrywise on all rank-one matrices $\mathbf{u}\mathbf{u}^T \in \mathcal{P}_N\big((0,\rho)\big)$ if and only if

$$c' \geq -\left(\sum_{j=0}^{N-1} \frac{V(\mathbf{n}_j)^2}{c_j V(\mathbf{n}_{\min})^2}\rho^{M-n_j}\right)^{-1}.$$

In fact, if $\mathbf{n} = \mathbf{n}_{\min}$ then the entire argument above goes through even when one changes the domain to the open complex disc $D(0,\rho)$, or any intermediate domain $(0,\rho) \subset D \subset D(0,\rho)$. This is precisely the content of the main result in [11].

Theorem 4.10 (Belton–Guillot–Khare–Putinar [11])**.** *Fix $\rho > 0$ and integers $M \geq N \geq 2$. Let*

$$f(z) = \sum_{j=0}^{N-1} c_j z^j + c' z^M, \qquad \text{where } c_0,\dots,c_{N-1}, c' \in \mathbb{R},$$

and let $I := \overline{D}(0,\rho)$ be the closed disc in the complex plane with centre 0 and radius ρ. The following are equivalent.

(1) *The entrywise map $f[-]$ preserves positivity on $\mathcal{P}_N(I)$.*
(2) *The entrywise map $f[-]$ preserves positivity on rank-one matrices in $\mathcal{P}_N\big((0,\rho)\big)$.*
(3) *Either c_0, ..., c_{N-1}, c' are all non-negative, or c_0, ..., c_{N-1} are positive and*

$$c' \geq -\left(\sum_{j=0}^{N-1} \frac{V(\mathbf{n}_j)^2}{c_j V(\mathbf{n}_{\min})^2}\rho^{M-j}\right)^{-1},$$

where $\mathbf{n}_j := (0,1,\dots,j-1,\widehat{j},j+1,\dots,N-1,M)^T$ for $0 \leq j \leq N-1$.

This theorem provides a complete understanding of which polynomials of degree at most N preserve positivity entrywise on $\mathcal{P}_N\big((0,\rho)\big)$ and, more generally, on any subset of $\mathcal{P}_N\big(\overline{D}(0,\rho)\big)$ that contains the rank-one matrices in $\mathcal{P}_N\big((0,\rho)\big)$.

Remark 4.11. Clearly (1) $\Longrightarrow$ (2) here, and the proof of (2) $\Longleftrightarrow$ (3) was outlined above via Proposition 4.7. We defer mentioning the proof strategy for (2) $\Longrightarrow$ (1), because we will later see a similar theorem over $I = (0,\rho)$ for more general powers n_j. The proof of that result, Theorem 4.15, will be outlined in some detail.

Having dealt with the base case of $\mathbf{n} = \mathbf{n}_{\min}$, as well as $\mathbf{n} = (k, k + 1, \ldots, k + N - 1)$ for any $k \in \mathbb{Z}_+$, which holds by the Schur product theorem, we now turn to the general case. In general, $s_{\mathbf{n}}(\mathbf{u})$ is no longer a monomial, and so it is no longer clear if and where the supremum of each ratio $s_{\mathbf{n}_j}(\mathbf{u})^2 / s_{\mathbf{n}}(\mathbf{u})^2$, or of their weighted sum, is attained for $\mathbf{u} \in (0, \sqrt{\rho})^N$. The threshold bound for all rank-one matrices itself is not apparent, and the bound for all matrices in $\mathcal{P}_N\big((0, \rho)\big)$ is even more inaccessible.

By a mathematical miracle, it turns out that the same phenomenon as in the base case holds in general. Namely, the ratio of each $s_{\mathbf{n}_j}$ and $s_{\mathbf{n}}$ attains its supremum at $\sqrt{\rho}(1, \ldots, 1)^T$. Hence one can proceed as above to obtain a uniform threshold for c', which works for all rank-one matrices in $\mathcal{P}_N\big((0, \rho)\big)$.

Example 4.12. To explain the ideas of the preceding paragraph, we present an example. Suppose

$$N = 3, \qquad \mathbf{n} = (0, 2, 3), \qquad M = 4, \quad \text{and} \quad \mathbf{u} = (u_1, u_2, u_3)^T.$$

Then

$$\mathbf{n}_3 = (0, 2, 4),$$
$$s_{\mathbf{n}}(\mathbf{u}) = u_1 u_2 + u_2 u_3 + u_3 u_1,$$
$$\text{and} \quad s_{\mathbf{n}_3}(\mathbf{u}) = (u_1 + u_2)(u_2 + u_3)(u_3 + u_1).$$

The claim is that $s_{\mathbf{n}_3}(\mathbf{u})/s_{\mathbf{n}}(\mathbf{u})$ is coordinatewise non-decreasing for $\mathbf{u} \in (0, \infty)^3$; the assertion about its supremum on $(0, \sqrt{\rho})^N$ immediately follows from this. It suffices by symmetry to show the claim only for one variable, say u_3. By the quotient rule,

$$s_{\mathbf{n}}(\mathbf{u}) \partial_{u_3} s_{\mathbf{m}}(\mathbf{u}) - s_{\mathbf{m}}(\mathbf{u}) \partial_{u_3} s_{\mathbf{n}}(\mathbf{u}) = (u_1 + u_2)(u_1 u_3 + 2u_1 u_2 + u_2 u_3)u_3,$$

and this is clearly non-negative on the positive orthant, proving the claim. As we see, the above expression is, in fact, monomial positive, from which numerical positivity follows immediately.

In fact, an even stronger fact holds. Viewed as a polynomial in u_3, every coefficient in the above expression is in fact *Schur positive*. In other words, the coefficient of each u_3^j is a non-negative combination of Schur polynomials in u_1 and u_2:

$$(u_1 + u_2)(u_1 u_3 + 2u_1 u_2 + u_2 u_3)u_3 = \sum_{j \geq 0} p_j(u_1, u_2)u_3^j,$$

where

$$p_j(u_1, u_2) = \begin{cases} 2s_{(1,3)}(u_1, u_2) & \text{if } j = 1, \\ s_{(0,3)}(u_1, u_2) + s_{(1,2)}(u_1, u_2) & \text{if } j = 2, \\ 0 & \text{otherwise.} \end{cases}$$

In particular, this implies that each coefficient is monomial positive, whence numerically positive. We recall here that the monomial positivity of Schur polynomials follows from the definition of $s_{\mathbf{n}}(\mathbf{u})$ using Young tableaux.

The miracle to which we alluded above, is that the Schur positivity in the preceding example in fact holds in general.

Theorem 4.13 (Khare–Tao [89])**.** *If $n_0 < \cdots < n_{N-1}$ and $m_0 < \cdots < m_{N-1}$ are N-tuples of non-negative integers such that $m_j \geq n_j$ for $j = 0$, $\ldots$, $N - 1$, then the function*

$$f_{\mathbf{m},\mathbf{n}} : (0, \infty)^N \to \mathbb{R}; \ \mathbf{u} \mapsto \frac{s_{\mathbf{m}}(\mathbf{u})}{s_{\mathbf{n}}(\mathbf{u})}$$

is non-decreasing in each coordinate. Furthermore, if

$$s_{\mathbf{n}}(\mathbf{u}) \partial_{u_N} s_{\mathbf{m}}(\mathbf{u}) - s_{\mathbf{m}}(\mathbf{u}) \partial_{u_N} s_{\mathbf{n}}(\mathbf{u}) \qquad (4.7)$$

is considered as a polynomial in u_N, then the coefficient of every monomial u_N^j is a Schur-positive polynomial in $u_1, \ldots, u_{N-1}$.

The second, stronger part of Theorem 4.13 follows from a deep and highly non-trivial result in symmetric function theory (or type-A representation theory) by Lam, Postnikov, and Pylyavskyy [92], following earlier results by Skandera. We refer the reader to this paper and to [89] for more details. Notice also that the first assertion in Theorem 4.13 only requires the numerical positivity of the expression (4.7). This is given a separate proof in [89], using the method of condensation due to Charles Lutwidge Dodgson [40].[6] In this context, we add for completeness that in [89] the authors also show a log-supermodularity (or FKG, or MTP_2) phenomenon for determinants of totally positive matrices.

4.4. **Real powers; the threshold works for all matrices.** We now return to the proof of Theorem 4.3, which holds for real powers. Our next step is to observe that the first part of Theorem 4.13 now holds for all real powers. Since one can no longer define Schur polynomials in this case, we work with generalized Vandermonde determinants instead:

Corollary 4.14. *Fix N-tuples of real powers $\mathbf{n} = (n_0 < \cdots < n_{N-1})$ and $\mathbf{m} = (m_0 < \cdots < m_{N-1})$, such that $n_j \leq m_j$ for all j. Letting $\mathbf{u}^{\circ \mathbf{n}} := [u_j^{n_{k-1}}]_{j,k=1}^N$ as above, the function*

$$f : \{\mathbf{u} \in (0, \infty)^N : u_i \neq u_j \ \text{if} \ i \neq j\} \to \mathbb{R}; \ \mathbf{u} \mapsto \frac{\det \mathbf{u}^{\circ \mathbf{m}}}{\det \mathbf{u}^{\circ \mathbf{n}}}$$

is non-decreasing in each coordinate.

We sketch here one proof. The version for integer powers, Theorem 4.13, gives the version for rational powers, by taking a "common denominator" $L \in \mathbb{Z}$ such that Lm_j and Ln_j are all integers, and using a change of variables $y_j := u_j^{1/L}$. The general version for real powers then follows by considering rational approximations and taking limits.

---

[6]This article by Dodgson immediately follows his better-known 1865 publication, *Alice's Adventures in Wonderland.*

Corollary 4.14 helps prove the real-power version of Theorem 4.3, just as Theorem 4.13 would have shown the integer-powers case of Theorem 4.3. Namely, first note that Proposition 4.7 holds even when the n_j are real powers; the only changes are (a) to assume that the coordinates of $\mathbf{u}$ are distinct, and (b) to rephrase the last assertion (3) to the following:

$$t \geq \sum_{j=0}^{N-1} \frac{(\det \mathbf{u}^{\circ \mathbf{n}_j})^2}{c_j (\det \mathbf{u}^{\circ \mathbf{n}})^2}.$$

These arguments help prove the first part of the following result, which is the culmination of these ideas.

Theorem 4.15 (Khare–Tao [89]). *Fix an integer $N \geq 1$ and real exponents $n_0 < \cdots < n_{N-1} < M$, as well as scalars $\rho > 0$ and $c_0, \ldots, c_{N-1}, c'$. Let*

$$f(x) := \sum_{j=0}^{N-1} c_j x^{n_j} + c' x^M.$$

The following are equivalent.

(1) *The function f preserves positivity entrywise on all rank-one matrices in $\mathcal{P}_N\big((0, \rho)\big)$.*

(2) *The function f preserves positivity entrywise on all Hankel rank-one matrices in $\mathcal{P}_N\big((0, \rho)\big)$.*

(3) *Either the coefficients $c_0, \ldots, c_{N-1}$ and c' are non-negative, or $c_0, \ldots, c_{N-1}$ are positive and*

$$c' \geq -\left(\sum_{j=0}^{N-1} \frac{V(\mathbf{n}_j)^2}{c_j V(\mathbf{n})^2} \rho^{M-n_j} \right)^{-1},$$

where $V(\mathbf{n})$ and $\mathbf{n}_j$ are as defined above.

If, moreover, the exponents n_j all lie in $\mathbb{Z}_+ \cup [N - 2, \infty)$, then these assertions are also equivalent to the following.

(4) *The function f preserves positivity entrywise on $\mathcal{P}_N\big((0, \rho)\big)$.*

Before sketching the proof, we note several ramifications of this result.

(1) The theorem completely characterizes linear combinations of up to $N + 1$ powers that entrywise preserve positivity on $\mathcal{P}_N\big((0, \rho)\big)$. The same is true for any subset of $\mathcal{P}_N\big((0, \rho)\big)$ that contains all rank-one positive semidefinite Hankel matrices.

(2) As discussed above, Theorem 4.15 implies Theorem 4.5, which helps in understanding which sign patterns correspond to countable sums of real powers that preserve positivity entrywise on $\mathcal{P}_N\big((0, \rho)\big)$ (or on the subset of rank-one matrices). In particular, the existence of sign patterns which are not all non-negative shows the existence of functions which preserve positivity on $\mathcal{P}_N$ but not on $\mathcal{P}_{N+1}$.

(3) Theorem 4.15 bounds $A^{\circ M}$ in terms of a multiple of $\sum_{j=0}^{N-1} c_j A^{\circ n_j}$. More generally, one can do this for an arbitrary convergent power series instead of a monomial, in the spirit of Theorem 4.2. Even more generally, one may work with Laplace transforms of measures; see Corollary 4.17 below.

For completeness, we also mention two developments related (somewhat more distantly) to the above results.

- A refinement of a conjecture of Cuttler, Greene, and Skandera (2011) and its proof; see [89] for more details. In particular, this approach assists with a novel characterization of weak majorization, using Schur polynomials.
- A related "Schubert cell-type" stratification of the cone $\mathcal{P}_N(\mathbb{C})$; see [11] for further details.

We conclude this section by outlining the proof of Theorem 4.15.

Proof. Clearly, (4) $\implies$ (1) $\implies$ (2). If (2) holds, then, by Corollary 3.11 at $a = 0$, either all the c_j and c' are non-negative, or c_j is positive for all j. Thus, we suppose that $c_j > 0 > c'$.

Note that if $\mathbf{u}(u_0) := (1, u_0, \ldots, u_0^{N-1})^T$ for some $u_0 \in (0, 1)$, then

$$A(u_0) := \rho u_0^2 \mathbf{u}(u_0) \mathbf{u}(u_0)^T$$

is a rank-one Hankel matrix and hence in our test set. Repeating the analysis in Section 4.2, using generalized Vandermonde determinants instead of Schur polynomials and rank-one Hankel matrices of the form $A(u_0)$,

$$|c'|^{-1} \geq \sup_{u_0 \in (0,1)} \sum_{j=0}^{N-1} \frac{(\det[\sqrt{\rho} u_0 \mathbf{u}(u_0)]^{\circ \mathbf{n}_j})^2}{c_j (\det[\sqrt{\rho} u_0 \mathbf{u}(u_0)]^{\circ \mathbf{n}})^2}$$

$$= \sum_{j=0}^{N-1} \lim_{u_0 \to 1^-} \sum_{j=0}^{N-1} \frac{(\det \mathbf{u}(u_0)^{\circ \mathbf{n}_j})^2}{c_j (\det \mathbf{u}(u_0)^{\circ \mathbf{n}})^2} (\rho u_0^2)^{M - n_j},$$

where the equality follows from Corollary 4.14 above. The real-exponent version of (4.4) holds if $q \in (0, \infty) \setminus \{1\}$ and the exponents n_j are real and non-decreasing:

$$\det \mathbf{u}(q)^{\circ \mathbf{n}} = \prod_{0 \leq i < k \leq N-1} (q^{n_k} - q^{n_i}) = V(q^{\circ \mathbf{n}}).$$

Applying this identity, the above computation yields

$$|c'|^{-1} \geq \lim_{u_0 \to 1^-} \sum_{j=0}^{N-1} \frac{V(u_0^{\circ \mathbf{n}_j})^2}{V(u_0^{\circ \mathbf{n}})^2} \frac{(\rho u_0^2)^{M - n_j}}{c_j} = \sum_{j=0}^{N-1} \frac{V(\mathbf{n}_j)^2}{c_j V(\mathbf{n})^2} \rho^{M - n_j}.$$

Thus (2) $\implies$ (3). Conversely, that (3) $\implies$ (1) follows by a similar analysis to that given above, using Corollary 4.14 and the density of matrices $\mathbf{u}\mathbf{u}^T$, where $\mathbf{u} \in (0, \sqrt{\rho})^N$ has distinct entries, in the set of all rank-one matrices in $\mathcal{P}_N((0, \rho))$.

It remains to show that (1) $\implies$ (4) if all the exponents $n_j \in \mathbb{Z}_+ \cup [N-2, \infty)$. We proceed by induction on N. The case $N = 1$ is immediate. For the inductive step, we apply the extension principle of the following Proposition 4.16 with $h = f$, which requires verification that $f'[-]$ preserves positivity on $\mathcal{P}_{N-1}$. This is a straightforward calculation via the induction hypothesis. $\square$

The following extension principle was inspired by work of FitzGerald and Horn [51].

Proposition 4.16 (Khare–Tao [89]). *Suppose $0 < \rho \leq \infty$, and $I = (0, \rho)$, $(-\rho, \rho)$ or the closure of one of these sets. Let $h : I \to \mathbb{R}$ be a continuously differentiable function on the interior of I. If $h'[-]$ preserves positivity entrywise on $\mathcal{P}_{N-1}(I)$ and $h[-]$ does so on the rank-one matrices in $\mathcal{P}_N(I)$, then $h[-]$ in fact preserves positivity on all of $\mathcal{P}_N(I)$.*[7]

Proposition 4.16 relies on two arguments found in [51]: (a) every matrix in $\mathcal{P}_N$ may be written as the sum of a rank-one matrix in $\mathcal{P}_N$, and a matrix in $\mathcal{P}_{N-1}$ with its last row and column both zero, and (b) applying the integral identity

$$h(x) - h(y) = \int_x^y h'(t)\,\mathrm{d}t = \int_0^1 (x-y)h'(\lambda x + (1-\lambda)y)\,\mathrm{d}\lambda$$

entrywise to this decomposition. See [89, Section 3] for more details. The original use of these arguments was when h is a power function; this is explained in Chapter 6 below.

4.5. **Power series preservers and beyond; unbounded domains.** In the remainder of this chapter, we use Theorem 4.15 to derive several corollaries; thus, we retain and use the notation of that theorem. As discussed following Theorem 4.15, the first consequence extends the theorem from bounding monomials $A^{\circ M} = (x^M)[A]$ by a multiple of $\sum_{j=0}^{N-1} c_j A^{\circ n_j}$, to bounding $f[A]$ for more general power series. Even more generally, one can work with Laplace transforms of real measures on $\mathbb{R}$.

Corollary 4.17 (Khare–Tao [89]). *Let the notation be as for Theorem 4.15, with $c_j > 0$ for all j. Suppose μ is a real measure supported on $[n_{N-1} + \epsilon, \infty)$ for some $\epsilon > 0$, and let*

$$g_\mu(x) := \int_{n_{N-1}+\epsilon}^\infty x^t \,\mathrm{d}\mu(t). \tag{4.8}$$

---

[7]An analogous version of this results holds for $I = D(0, \rho)$ or its closure in $\mathbb{C}$, with $h : I \to \mathbb{C}$ analytic. This is used to prove the corresponding implication in Theorem 4.10 above.

If g_μ is absolutely convergent at ρ, then there exists a finite threshold $t_\mu > 0$ such that, for all $A \in \mathcal{P}_N\big((0, \rho)\big)$, the matrix

$$t_\mu \sum_{j=0}^{N-1} c_j A^{\circ n_j} - g_\mu[A]$$

is positive semidefinite.

Proof. By Theorem 4.15 and the fact that $\mathcal{P}_N(\mathbb{R})$ is a closed convex cone, it suffices to show the finiteness of the quantity

$$\int_{n_{N-1}+\epsilon}^{\infty} \sum_{j=0}^{N-1} \frac{V(\mathbf{n}_j)^2}{c_j V(\mathbf{n})^2} \rho^{M-n_j} \, \mathrm{d}\mu_+(M),$$

where μ_+ is the positive part of μ. This follows from the hypotheses. $\square$

We now turn to the $\rho = \infty$ case, which was briefly alluded to above. In other words, the domain is now unbounded: $I = (0, \infty)$. As in the bounded-domain case, the question of interest is to classify all possible sign patterns of polynomial or power-series preservers on $\mathcal{P}_N(I)$ for a fixed integer N.

Similar to the above discussion for bounded I, the crucial step in classifying sign patterns of power series (or more general functions, as in Theorem 4.5) is to work with integer powers and precisely one coefficient that can be negative. Thus, one first observes that Lemma 3.9(2) holds in the unbounded-domain case $I = (0, \infty)$. Hence given a polynomial

$$f(x) = \sum_{j=0}^{2N-1} c_j x^{n_j} + c' x^M,$$

where

$$0 \le n_0 < \cdots < n_{N-1} < M < n_N < n_{N+1} \cdots < n_{2N-1},$$

if $f[-]$ preserves positivity on $\mathcal{P}_N\big((0, \infty)\big)$, then either all the coefficients $c_0, \ldots, c_{2N-1}, c'$ are non-negative, or $c_0, \ldots, c_{2N-1}$ are positive and c' can be negative. In this case, an explicit threshold is not known as it is in Theorem 4.15, but we now explain why such a threshold exists.

We start from (4.6) and repeat the subsequent analysis via the Cauchy–Binet formula. To find a uniform threshold for c' that works for all rank-one matrices in $\mathcal{P}_N\big((0, \infty)\big)$, it suffices to bound, uniformly from above, certain ratios of sums of squares of Schur polynomials. This may be done because of the following tight bounds.

Proposition 4.18 (Khare–Tao [89]). *If $\mathbf{n} := (n_0, \ldots, n_{N-1})$ and $\mathbf{u} := (u_1, \ldots, u_N)$, where $n_0 < \cdots < n_{N-1}$ are non-negative integers and $u_1 \le \cdots \le u_N$ are non-negative real numbers, then*

$$\mathbf{u^{n-n_{min}}} \le s_{\mathbf{n}}(\mathbf{u}) \le \frac{V(\mathbf{n})}{V(\mathbf{n}_{min})} \mathbf{u^{n-n_{min}}}, \tag{4.9}$$

where $\mathbf{n}_{\min} := (0, \ldots, n_{N-1})$. *The constants* 1 *and* $V(\mathbf{n})/V(\mathbf{n}_{\min})$ *on each side of (4.9) cannot be improved.*

We refer the reader to [89, Section 4] for further details, including how Proposition 4.18 implies the existence of preservers f as above for rank-one matrices with $c' < 0$. The extension from rank-one matrices to all of $\mathcal{P}_N\big((0, \infty)\big)$ is carried out using the extension principle in Proposition 4.16.

In a sense, Proposition 4.18 isolates the 'leading term' of every Schur polynomial. This calculation can be generalized to the case of non-integer powers,[8] which helps extend the above results for the unbounded domain $I = (0, \infty)$ to real powers. This yields the desired classification, similar to Theorem 4.5 in the bounded-domain case.

Theorem 4.19 (Khare–Tao [89]). *Let* $N \geq 2$, *and let* $\{\alpha_j : j \geq 0\} \subset \mathbb{Z}_+ \cup [N-2, \infty)$ *be a set of distinct real numbers. For each* $j \geq 0$, *let* $\epsilon_j \in \{0, \pm 1\}$ *be a sign and suppose that, whenever* $\epsilon_{j_0} = -1$, *then* $\epsilon_j = +1$ *for at least* N *choices of* j *such that* $\alpha_j < \alpha_{i_0}$ *and also for at least* N *choices of* j *such that* $\alpha_j > \alpha_{i_0}$. *There exists a series with real coefficients,*

$$f(x) = \sum_{j=0}^{\infty} c_j x^{\alpha_j}$$

which converges on $(0, \infty)$, *preserves positivity entrywise on* $\mathcal{P}_N\big((0, \infty)\big)$, *and is such that* c_j *has the same sign as* ϵ_j *for all* $j \geq 0$.

Note that, in particular, Theorem 4.19 reaffirms that the Horn–Loewner-type conditions in Lemma 3.9(2) are sharp.

4.6. **Digression: Schur polynomials from smooth functions, and new symmetric function identities.** Before proceeding to additional applications of Theorem 4.15 and related results, we take a brief detour to explain how Schur polynomials arise naturally from any sufficiently differentiable function.

Theorem 4.20 (Khare [88]). *Fix non-negative integers* $m_0 < m_1 < \cdots < m_{N-1}$, *as well as scalars* $\epsilon > 0$ *and* $a \in \mathbb{R}$. *Let* $M := m_0 + \cdots + m_{N-1}$ *and suppose the function* $f : [a, a+\epsilon) \to \mathbb{R}$ *is* M-*times differentiable at* a. *Given vectors* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N$, *define* $\Delta : [0, \epsilon') \to \mathbb{R}$ *for a sufficiently small* $\epsilon' \in (0, \epsilon)$ *by setting*

$$\Delta(t) := \det f[a\mathbf{1}_{N \times N} + t\mathbf{u}\mathbf{v}^T].$$

Then,

$$\Delta^{(M)}(0) = \sum_{\mathbf{m} \vdash M} \binom{M}{m_0, m_1, \ldots, m_{N-1}} V(\mathbf{u}) V(\mathbf{v}) s_{\mathbf{m}}(\mathbf{u}) s_{\mathbf{m}}(\mathbf{v}) \prod_{k=0}^{N-1} f^{(m_k)}(a),$$
$$(4.10)$$

---

[8] We refer the reader again to [89, Section 5] for the details, which use additional concepts from type-A representation theory: the Harish-Chandra–Itzykson–Zuber integral and Gelfand–Tsetlin patterns.

*where the first factor in the summand is a multinomial coefficient, and we
sum over all partitions* $\mathbf{m} = (m_0, \ldots, m_{N-1})$ *of* M *with unequal parts, that
is,* $M = m_0 + \cdots + m_{N-1}$ *and* $0 \le m_0 < \cdots < m_{N-1}$.

In particular, $\Delta(0) = \Delta'(0) = \cdots = \Delta^{(\binom{N}{2}-1)}(0) = 0$.

Remark 4.21. As a special case, if $f : \mathbb{R} \to \mathbb{R}$ is smooth at a, and $\mathbf{u}$,
$\mathbf{v} \in \mathbb{R}^N$, then defining $\Delta(t) := \det f[a\mathbf{1}_{N \times N} + t\mathbf{u}\mathbf{v}^T]$ gives a function Δ
which is smooth at 0, and Theorem 4.20 gives all of these derivatives via
the formula (4.10). The general version of Theorem 4.20 is a key ingredi-
ent in showing Theorem 3.10, which subsumes all known variants of Horn–
Loewner-type necessary conditions in fixed dimension.

The key determinant computation required to prove the original Horn–
Loewner necessary condition in fixed dimension (see Theorem 3.4) is the
special case of Theorem 4.20 where $\mathbf{u} = \mathbf{v}$ and $m_j = j$ for all j. In this
situation, $s_{\mathbf{m}}(\mathbf{u}) = s_{\mathbf{m}}(\mathbf{v}) = 1$, so Schur polynomials do not appear. The
general version of Theorem 4.20 decouples the vectors $\mathbf{u}$ and $\mathbf{v}$, and holds
for all $M > 0$ if f is smooth (as in Loewner's setting). Moreover, it reveals
the presence of Schur polynomials in every case other than the ones studied
by Loewner, that is, when $M > \binom{N}{2}$.

While Theorem 4.20 involves derivatives of a smooth function, the result
and its proof are, in fact, completely algebraic, and valid over any commu-
tative ring. To show this, an algebraic analogue of the differential operator
is required, with more structure than is given by a derivation. The precise
statement and its proof may be found in [88, Section 2].

We conclude this section by applying Theorem 4.20 and its algebraic
avatar to symmetric function theory. We begin by recalling the famous
Cauchy summation identity [98, Example I.4.6]: if $f_0(x) := 1 + x + x^2 + \cdots$
is the geometric series, viewed as a formal power series over a commutative
unital ring R, and $u_1, \ldots, u_N, v_1, \ldots, v_N$ are commuting variables, then

$$\det f_0[\mathbf{u}\mathbf{v}^T] = V(\mathbf{u})V(\mathbf{v}) \sum_{\mathbf{m}} s_{\mathbf{m}}(\mathbf{u})s_{\mathbf{m}}(\mathbf{v}), \qquad (4.11)$$

where the sum runs over all partitions $\mathbf{m}$ with at most N parts.[9]

A natural question is whether similar formulae hold when f_0 is replaced by
other formal power series. Very few such results were known; this includes
one due to Frobenius [55], for the function $f_c(x) := (1 - cx)/(1 - x)$ with c

---

[9]Usually one uses infinitely many indeterminates in symmetric function theory, but
given the connection to the entrywise calculus in a fixed dimension, we will restrict our
attention to u_j and v_j for $1 \le j \le N$.

an scalar. (This is also connected to theta functions and elliptic Frobenius–Stickelberger–Cauchy determinant identities.) For this function,

$$
\begin{aligned}
\det f_c[\mathbf{u}\mathbf{v}^T] &= \det \Big[\frac{1 - cu_j v_k}{1 - u_j v_k} \Big]_{j,k=1}^{N} \\
&= V(\mathbf{u})V(\mathbf{v})(1 - c)^{N-1} \\
&\quad \times \Big(\sum_{\mathbf{m}:m_0=0} s_{\mathbf{m}}(\mathbf{u})s_{\mathbf{m}}(\mathbf{v}) + (1 - c) \sum_{\mathbf{m}:m_0>0} s_{\mathbf{m}}(\mathbf{u})s_{\mathbf{m}}(\mathbf{v}) \Big).
\end{aligned}
$$
(4.12)

A third, obvious identity is if f is a 'fewnomial' with at most $N-1$ terms. In this case, $f[\mathbf{u}\mathbf{v}^T]$ is a sum of at most $N-1$ rank-one matrices, and so its determinant vanishes.

The following result extends all three of these cases to an arbitrary formal power series over an arbitrary commutative ring R, and with an additional $\mathbb{Z}_+$-grading.

Theorem 4.22 (Khare [88]). *Fix a commutative unital ring R and let t be an indeterminate. Let $f(t) := \sum_{M \geq 0} f_M t^M \in R[[t]]$ be an arbitrary formal power series. Given vectors $\mathbf{u}, \mathbf{v} \in R^N$, where $N \geq 1$, we have that*

$$
\det f[t\mathbf{u}\mathbf{v}^T] = V(\mathbf{u})V(\mathbf{v}) \sum_{M \geq \binom{N}{2}} t^M \sum_{\mathbf{m}=(m_{N-1},\ldots,m_0) \vdash M} s_{\mathbf{m}}(\mathbf{u})s_{\mathbf{m}}(\mathbf{v}) \prod_{k=0}^{N-1} f_{m_k}.
$$
(4.13)

The heart of the proof involves first computing, for each $M \geq 0$, the coefficient of t^M in $\det f[t\mathbf{u}\mathbf{v}^T]$, over the "universal ring"

$$
R' := \mathbb{Q}[u_1, \ldots, u_N, v_1, \ldots, v_N, f_0, f_1, \ldots],
$$

where u_j, v_k and f_m are algebraically independent over $\mathbb{Q}$. These coefficients are seen to equal $\Delta^{(M)}(0)/M!$, by the algebraic version of Theorem 4.20. Thus, (4.13) holds over R'. Then note that both sides of (4.13) lie in the subring $R_0 := \mathbb{Z}[u_1, \ldots, u_N, v_1, \ldots, v_N, f_0, f_1, \ldots]$, so the identity holds in R_0. Finally, it holds as claimed by specializing from R_0 to R.

An alternate approach to proving Theorem 4.22 is also provided in [88]. The identity (4.6) is applied, along with the Cauchy–Binet formula, to each truncated Taylor–Maclaurin polynomial $f_{\leq M}$ of $f(x)$. The result follows by taking limits in the t-adic topology, using the t-adic continuity of the determinant function.

4.7. **Further applications: linear matrix inequalities, Rayleigh quotients, and the cube problem.** This chapter ends with further ramifications and applications of the above results. First, notice that Theorem 4.15 implies the following linear matrix inequality version that is 'sharp' in more than one sense:

Corollary 4.23. *Fix $\rho > 0$, real exponents $n_0 < \cdots < n_{N-1} < M$ for some integer $N \geq 1$, and scalars $c_j > 0$ for all j. Then,*

$$A^{\circ M} \leq \mathcal{C}\Big(c_0 A^{\circ n_0} + \cdots + c_{N-1} A^{\circ n_{N-1}}\Big),$$

$$\text{where } \mathcal{C} = \sum_{j=0}^{N-1} \frac{V(\mathbf{n}_j)^2}{c_j V(\mathbf{n})^2} \rho^{M-n_j},$$

for all $A \in \mathcal{P}_N\big((0,\rho)\big)$ of rank one, or of all ranks if $n_0, \ldots, n_{N-1} \in \mathbb{Z}_+ \cup [N-2,\infty)$. Moreover, the constant $\mathcal{C}$ is the smallest possible, as is the number of terms N on the right-hand side.

Seeking a uniform threshold such as $\mathcal{C}$ in the preceding inequality can also be achieved (as explained above) by first working with a single positive matrix, then optimizing over all matrices. The first step here can be recast as an extremal problem that involves Rayleigh quotients:

Proposition 4.24 (see [11, 89]). *Fix an integer $N \geq 2$ and real exponents $n_0 < \cdots < n_{N-1} < M$, where each $n_j \in \mathbb{Z}_+ \cup [N-2,\infty)$. Given positive scalars $c_0, \ldots, c_{N-1}$, let*

$$h(x) := \sum_{j=0}^{N-1} c_j x^{n_j} \qquad \big(x \in (0,\infty)\big).$$

Then, for $0 < \rho < \infty$ and $A \in \mathcal{P}_N\big([0,\rho]\big)$,

$$t\, h[A] \geq A^{\circ M} \qquad \text{if and only if} \quad t \geq \varrho(h[A]^{\dagger/2} A^{\circ M} h[A]^{\dagger/2}), \qquad (4.14)$$

where $\varrho[B]$ and $B^{\dagger}$ denote the spectral radius and the Moore–Penrose pseudo-inverse of a square matrix B, respectively. Moreover, for every non-zero matrix $A \in \mathcal{P}_N\big([0,\rho]\big)$, the following variational formula holds:

$$\varrho(h[A]^{\dagger/2} A^{\circ M} h[A]^{\dagger/2}) = \sup_{\mathbf{u} \in (\ker h[A])^{\perp} \setminus \{\mathbf{0}\}} \frac{\mathbf{u}^T A^{\circ M} \mathbf{u}}{\mathbf{u}^T h[\mathbf{u}\mathbf{u}^T] \mathbf{u}} \leq \sum_{j=0}^{N-1} \frac{V(\mathbf{n}_j)^2}{V(\mathbf{n})^2} \frac{\rho^{M-n_j}}{c_j}.$$

Proposition 4.24 is shown using the Kronecker normal form for matrix pencils; see the treatment in [57, Section X.6]. When the matrix A is a generic rank-one matrix, the above generalized Rayleigh quotient has a closed-form expression, which features Schur polynomials for integer powers. This reveals connections between Rayleigh quotients, spectral radii, and symmetric functions.

Proposition 4.25. *Let the notation be as in Proposition 4.24, but now with n_j not necessarily in $\mathbb{Z}_+ \cup [N-2,\infty)$. If $A = \mathbf{u}\mathbf{u}^T$, where $\mathbf{u} \in (0,\infty)^N$ has distinct coordinates, then $h[A]$ is invertible, and the threshold bound*

$$\varrho(h[A]^{\dagger/2} A^{\circ M} h[A]^{\dagger/2}) = (\mathbf{u}^{\circ M})^T h[\mathbf{u}\mathbf{u}^T]^{-1} \mathbf{u}^{\circ M} = \sum_{j=0}^{N-1} \frac{(\det \mathbf{u}^{\circ \mathbf{n}_j})^2}{c_j (\det \mathbf{u}^{\circ \mathbf{n}})^2}. \qquad (4.15)$$

In fact, the proof of the final equality in (4.15) is completely algebraic, and reveals new determinantal identities that hold over any field $\mathbb{F}$ with at least N elements.

Proposition 4.26 (Khare–Tao [89]). *Suppose $N \geq 1$ and $0 \leq n_0 < \cdots < n_{N-1} < M$ are integers, and $\mathbf{u}, \mathbf{v} \in \mathbb{F}^N$ each have distinct coordinates. Let $c_j \in \mathbb{F}^\times$ and define $h(t) := \sum_{j=0}^{N-1} c_j t^{n_j}$. Then $h[\mathbf{uv}^T]$ is invertible, and*

$$(\mathbf{v}^{\circ M})^T h[\mathbf{uv}^T]^{-1} \mathbf{u}^{\circ M} = \sum_{j=0}^{N-1} \frac{\det \mathbf{u}^{\circ \mathbf{n}_j} \det \mathbf{v}^{\circ \mathbf{n}_j}}{c_j \det \mathbf{u}^{\circ \mathbf{n}} \det \mathbf{v}^{\circ \mathbf{n}}}.$$

The final result is a variant of the matrix-cube problem [104], and connects to spectrahedra [22, 135] and modern optimization theory. Given two or more real symmetric $N \times N$ matrices $A_0, \ldots, A_{M+1}$, the corresponding matrix cube of size $2\eta > 0$ is

$$\mathcal{U}[\eta] := \Big\{ A_0 + \sum_{m=1}^{M+1} u_m A_m : u_m \in [-\eta, \eta] \Big\}.$$

The matrix-cube problem is to find the largest $\eta > 0$ such that $\mathcal{U}[\eta] \subset \mathcal{P}_N(\mathbb{R})$. In the present setting of the entrywise calculus, the above results imply asymptotically matching upper and lower bounds for the size of the matrix cube.

Theorem 4.27 (see [11, 89]). *Suppose $M \geq 0$ and $0 \leq n_0 < n_1 < \cdots$ are integers. Fix positive scalars $\rho > 0$, $0 < \alpha_1 < \cdots < \alpha_{M+1}$, and $c_j > 0$ for all $j \geq 0$, and define for each $N \geq 1$ and each matrix $A \in \mathcal{P}_N\big([0, \rho]\big)$, the cube*

$$\mathcal{U}_A[\eta] := \left\{ \sum_{j=0}^{N-1} c_j A^{\circ n_j} + \sum_{m=1}^{M+1} u_m A^{\circ(n_{N-1}+\alpha_m)} : u_m \in [-\eta, \eta] \right\}. \qquad (4.16)$$

Also define for $N \geq 1$ and $\alpha > 0$:

$$\mathcal{K}_\alpha(N) := \sum_{j=0}^{N-1} \frac{V(\mathbf{n}_j(\alpha, N))^2}{V(\mathbf{n}(N))^2} \frac{\rho^{\alpha - n_j}}{c_j}, \qquad (4.17)$$

where $\mathbf{n}(N) := (n_0, \ldots, n_{N-1})^T$, and

$$\mathbf{n}_j(\alpha, N) := (n_0, \ldots, n_{j-1}, n_{j+1}, \ldots, n_{N-1}, n_{N-1} + \alpha).$$

Then for each fixed $N \geq 1$, we have the uniform upper and lower bounds:

$$\eta \leq \big(\mathcal{K}_{\alpha_1}(N) + \cdots + \mathcal{K}_{\alpha_{M+1}}(N)\big)^{-1} \implies \mathcal{U}_A[\eta] \subset \mathcal{P}_N \text{ for all } A \in \mathcal{P}_N\big([0, \rho]\big)$$
$$\implies \eta \leq \mathcal{K}_{\alpha_{M+1}}(N)^{-1}.$$
$$(4.18)$$

Moreover, if the n_j grow linearly, in that

$$\alpha_{M+1} - \alpha_M \geq n_{j+1} - n_j \qquad \text{for all } j \geq 0,$$

then the lower and upper bounds for $\eta = \eta_N$ in (4.18) are asymptotically equal as $N \to \infty$:

$$\lim_{N \to \infty} \mathcal{K}_{\alpha_{M+1}}(N)^{-1} \sum_{m=1}^{M+1} \mathcal{K}_{\alpha_m}(N) = 1.$$

5. Totally non-negative matrices and positivity preservers

In this chapter, we discuss variant notions of matrix positivity that are well studied in the literature, *total positivity* and *total non-negativity*, and characterize the maps which preserve these properties.

Definition 5.1. A real matrix A is said to be *totally non-negative* or *totally positive* if every minor of A is non-negative or positive, respectively. We will denote these matrices, as well as the property, by TN and TP.

In older texts, such matrices were called *totally positive* and *strictly totally positive*, respectively.

To introduce the theory of total positivity, we can do no better than quote from the preface of Karlin's magisterial book [85]: "Total positivity is a concept of considerable power that plays an important role in various domains of mathematics, statistics and mechanics". Karlin goes on to list "problems involving convexity, moment spaces, eigenvalues of integral operators, ... oscillation properties of solutions of linear differential equations ... the theory of approximations ... statistical decision procedures ... discerning uniformly most powerful tests for hypotheses ... ascertaining optimal policy for inventory and production processes ... analysis of diffusion-type stochastic processes, and ... coupled mechanical systems."

Perhaps the earliest result on total positivity is due to Fekete, in correspondence with Pólya [50] published in 1912 (see Lemma 5.10). Schoenberg observed the variation-diminishing properties of TP matrices in 1930 [119], and published a series of papers on Pólya frequency functions, which are defined in terms of total positivity, in the 1950s [127, 126, 128]. Independently of Schoenberg, Krein's investigation of ordinary differential equations led him to the total positivity of Green's functions for certain differential operators, and in the mid-1930s his works with Gantmacher looked at spectral and other properties of totally positive matrices and kernels; see [58] and [85, Section 10.6].

For more on these four authors, one may consult the afterwork of Pinkus's book on total positivity [106], which also contains a wealth of results on totally positive and totally non-negative matrices. For a modern collection of applications of the theory of total positivity, see the book edited by Gasca and Micchelli [60].

More recently, total positivity has had a major impact on Lie theory. Lusztig extended the theory of total positivity to the setting of linear algebraic groups; see [97] for an exposition of this work. This led Fomin and

Zelevinsky to investigate the combinatorics of Lusztig's theory [53] and resulted in the invention of cluster algebras [54]. These objects have generated an enormous amount of activity in a short period of time, with connections across a wide range of areas within representation theory, combinatorics, geometry, and mathematical physics. For the latter, we will mention only the totally non-negative Grassmannian [110], its connections with scattering amplitudes for quantum field theories [4], and the work by Kodama and Williams on regular soliton solutions of the Kadomtsev–Petviashvili equation [91].

Example 5.2. Perhaps the most well-known class of totally positive matrices consists of the *(generalized) Vandermonde matrices*: for real numbers $0 < x_1 < \cdots < x_m$ and $\alpha_1 < \cdots < \alpha_n$, the $m \times n$ matrix

$$A := [x_j^{\alpha_k}]_{1 \le j \le m,\ 1 \le k \le n}$$

is totally positive. Indeed, it suffices to show the positivity of any such matrix determinant $\det A$ when $m = n$. That $\det A$ is non-zero follows from Laguerre's extension of Descartes' rule of signs (see [82]) and by fixing the x_j and considering a linear homotopy from $(0, 1, \ldots, n-1)$ to $(\alpha_1, \ldots, \alpha_n)$, one obtains a continuous non-vanishing function from the usual Vandermonde determinant $\prod_{1 \le j < k \le n}(x_k - x_j)$ (which is positive) to $\det A$.

Example 5.3. Another prominent class of symmetric totally positive matrices consists of the Hankel moment matrices $H_\mu := [s_{j+k}(\mu)]_{j,k \ge 0}$ corresponding to admissible measures μ; see Definition 3.16.

5.1. **Totally non-negative and totally positive kernels.** An important generalization of TN and TP matrices is given by the following functional form.

Definition 5.4. Let X and Y be totally ordered sets, and let $K : X \times Y \to \mathbb{R}$ be a kernel.

(1) The kernel K is *totally positive of order r*, denoted TP_r, if, for any n-tuples of points $x_1 < \cdots < x_n$ in X and $y_1 < \cdots < y_n$ in Y, where $1 \le n \le r$, the matrix

$$[K(x_j, y_k)]_{j,k=1}^n$$

has positive determinant.
(2) The kernel K is *totally positive* if K is TP_r for all $r \ge 1$.
(3) Similarly, one defines TN_r kernels and totally non-negative kernels by replacing the word "positive" in the above by "non-negative."

If $X = \{1, \ldots, m\}$ and $Y = \{1, \ldots, n\}$, we recover the earlier notions of totally positive and totally non-negative matrices. When X and Y are taken to be real intervals, TN and TP kernels can be thought of as continuous analogues of TN and TP matrices. In fact, one has a continuous analogue of the Cauchy–Binet formula, which generalizes its traditional version.

Theorem 5.5 (Basic Composition Lemma, see e.g. [85, 86]). *Suppose X, Y, $Z \subset \mathbb{R}$ and let μ be a non-negative Borel measure on Y. Suppose $K : X \times Y \to \mathbb{R}$ and $L : Y \times Z \to \mathbb{R}$ are pointwise Borel measurable with respect to Y, and let*

$$M : X \times Z \to \mathbb{R}; \ (x, z) \mapsto \int_Y K(x, y) L(y, z) \, d\mu(y).$$

If M is well defined on the whole of $X \times Z$, then

$$\det \begin{bmatrix} M(x_1, z_1) & \ldots & M(x_1, z_m) \\ \vdots & \ddots & \vdots \\ M(x_m, z_1) & \ldots & M(x_m, z_m) \end{bmatrix}$$

$$= \int \cdots \int_{y_1 < y_2 < \cdots < y_m \in Y} \det[K(x_i, y_j)]_{i,j=1}^m \det[L(y_j, z_k)]_{j,k=1}^m \prod_{j=1}^m d\mu(y_j).$$

As an immediate consequence, we have the following corollary.

Corollary 5.6. *In the setting of Theorem 5.5, if the kernels K and L are both TN_r or TP_r for some $r \geq 1$, then M has the same property. In particular, if K and L are both TN or TP, then so is M.*

We conclude this part with an observation of Pólya that connects to a class of well-studied functions, and also implies the positive definiteness of the Gaussian kernel. Recall from the proof of Theorem 2.4 above that this latter property was crucially used by Schoenberg in characterizing metric space embeddings into Hilbert space; however, its proof above was only outlined (via the more sophisticated machinery of Fourier analysis and Bochner's theorem).

Lemma 5.7 (Pólya). *The Gaussian kernel $K : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by $K(x, y) := \exp(-(x - y)^2)$ is totally positive.*

Proof. It suffices to show that every square matrix generated from the kernel has positive determinant. Given real numbers $x_1 < \cdots < x_n$ and $y_1 < \cdots < y_n$, we observe the following factorization:

$$[\exp(-(x_j - y_k)^2)]_{j,k=1}^n$$
$$= \operatorname{diag}[\exp(-x_j^2)]_{j=1}^n [\exp(2 x_j y_k)]_{j,k=1}^n \operatorname{diag}[\exp(-y_k^2)]_{k=1}^n.$$

The proof concludes by observing that all three matrices on the right-hand side have positive determinants, the second because it is a Vandermonde matrix $[p_j^{\alpha_k}]$ with $p_j = \exp(2 x_j)$ and $\alpha_k = y_k$. $\qquad\square$

Example 5.8. The Gaussian function $f(x) = \exp(-x^2)$ is thus an example of a *Pólya frequency function*, that is, one for which $f(x - y)$ is a TP kernel on $\mathbb{R} \times \mathbb{R}$. As noted above, these functions were intensively studied by Schoenberg, and continue to be much studied in mathematics and statistics; two of the classic references are [29, 43].

The case of the multivariate Gaussian kernel follows immediately from the one-dimensional version.

Corollary 5.9. *For all $d \geq 1$, the Gaussian kernel*

$$\mathbb{R}^d \times \mathbb{R}^d \to (0, \infty); \ (\mathbf{x}, \mathbf{y}) \mapsto K(\mathbf{x}, \mathbf{y}) := \exp(-\|\mathbf{x} - \mathbf{y}\|^2)$$

is positive semidefinite on $\mathbb{R}^d \times \mathbb{R}^d$. In other words, the matrix $[\exp(-\|\mathbf{x}_j - \mathbf{x}_k\|^2)]_{j,k=1}^n$ is positive semidefinite for all $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^d$.

Proof. The $d = 1$ case is a direct consequence of Lemma 5.7, and the case of general d follows from this by using the Schur product theorem. $\square$

5.2. **Entrywise preservers of totally non-negative Hankel matrices.** In the recent article [48] by Fallat, Johnson, and Sokal, the authors study when various classes of totally non-negative (TN) matrices are closed under taking sums or Schur products. As they observe, the set of all TN matrices is not closed under these operations; for example, the 3×3 identity matrix and the all-ones matrix $\mathbf{1}_{3 \times 3}$ are both TN but their sum is not.

It is of interest to isolate a class of TN matrices that is a closed convex cone, and is furthermore closed under taking Schur products. Indeed, it is under these conditions that the observation of Pólya–Szegö (see Section 3.1) holds, leading to large classes of TN preservers.

Such a class of matrices has been identified in both the dimension-free as well as fixed-dimension settings. It consists of the *TN Hankel matrices.* In a fixed dimension, there is the following classical result from 1912.

Lemma 5.10 (Fekete [50])**.** *Let A be a possibly rectangular real Hankel matrix such that all of its contiguous minors are positive. Then A is totally positive.*

Recall that a minor is said to be *contiguous* if it is obtained from successive rows and successive columns of A.

If A is a square Hankel matrix, let $A^{(1)}$ be the square submatrix of A obtained by removing the first row and the last column. Notice that every contiguous minor of A is a principal minor of either A or $A^{(1)}$. Combined with Fekete's lemma, these observations help show another folklore result.

Theorem 5.11. *Let A be a square real Hankel matrix. Then A is TN or TP if and only if both A and $A^{(1)}$ are positive semidefinite or positive definite, respectively.*

Theorem 5.11 is a very useful bridge between matrix positivity and total non-negativity. A related dimension-free variant (see [2, 59]) concerns the Stieltjes moment problem: a sequence $(s_0, s_1, \ldots,)$ is the moment sequence of an admissible measure on $\mathbb{R}_+$ (see Definition 3.16) if and only if the Hankel matrices $H := (s_{j+k})_{j,k \geq 0}$ and $H^{(1)}$ (obtained by excising the first row of H, or equivalently, the first column) are both positive semidefinite. By Theorem 5.11, this is equivalent to saying that H is totally non-negative.

With Theorem 5.11 in hand, one can easily show several basic facts about TN Hankel matrices; we collect these in the following result for convenience.

Lemma 5.12. *For an integer $N \geq 1$ and a set $I \subset \mathbb{R}_+$, let $HTN_N(I)$ denote the set of $N \times N$ TN Hankel matrices with entries in I. For brevity, we let $HTN_N := HTN_N(\mathbb{R}_+)$.*

(1) *The family HTN_N is closed under taking sums and non-negative scalar multiples, or more generally, integrals against non-negative measures (as long as these exist).*

(2) *In particular, if μ is an admissible measure supported on $\mathbb{R}_+$, then its moment matrix $H_\mu := \left(s_{j+k}(\mu)\right)_{j,k=0}^{\infty}$ is totally non-negative.*

(3) *HTN_N is closed under taking entrywise products.*

(4) *If the power series $f(x) = \sum_{k \geq 0} c_k x^k$ is convergent on $I \subset \mathbb{R}_+$, with $c_k \geq 0$ for all $k \geq 0$, then the entrywise map $f[-]$ preserves total non-negativity on $HTN_N(I)$, for all $N \geq 1$.*

Given Lemma 5.12(4), which is identical to the start of the story for positivity preservers, it is natural to expect parallels between the two settings. For example, one can ask if a Schoenberg-type phenomenon also holds for preservers of total non-negativity on $\bigcup_{N \geq 1} HTN_N\left([0, \rho)\right)$ with $0 < \rho \leq \infty$. As we now explain, this is indeed the case; we will set $\rho = \infty$ for ease of exposition. From Theorem 3.14 and the subsequent discussion, it follows via Hamburger's theorem that the class of functions $\sum_{k \geq 0} c_k x^k$ with all $c_k \geq 0$ characterizes the entrywise maps preserving the set of moment sequences of admissible measures supported on $[-1, 1]$. By the above discussion, in considering the family of matrices HTN_N for all $N \geq 1$, we are studying moment sequences of admissible measures supported on $I = \mathbb{R}_+$, or the related Hausdorff moment problem for $I = [0, 1]$. In this case, one also has a Schoenberg-like characterization, outside of the origin.

Theorem 5.13 (Belton–Guillot–Khare–Putinar [12]). *Let $f : \mathbb{R}_+ \to \mathbb{R}$. The following are equivalent.*

(1) *Applied entrywise, the map f preserves the set HTN_N for all $N \geq 1$.*

(2) *Applied entrywise, the map f preserves positive semidefiniteness on HTN_N for all $N \geq 1$.*

(3) *Applied entrywise, the map f preserves the set of moment sequences of admissible measures supported on $\mathbb{R}_+$.*

(4) *Applied entrywise, the map f preserves the set of moment sequences of admissible measures supported on $[0, 1]$.*

(5) *The function f agrees on $(0, \infty)$ with an absolutely monotonic entire function, hence is non-decreasing, and $0 \leq f(0) \leq \lim_{\epsilon \to 0^+} f(\epsilon)$.*

Remark 5.14. If we work only with $f : (0, \infty) \to \mathbb{R}$, then we are interested in matrices in HTN_N with positive entries. Since the only matrices in HTN_N with a zero entry are scalar multiples of the elementary square matrices E_{11} or E_{NN} (equivalently, the only admissible measures supported

in $\mathbb{R}_+$ with a zero moment are of the form $c\delta_0$), the test set does not really reduce, and hence the preceding theorem still holds in essence: we must replace HTN_N by $HTN_N\big((0,\infty)\big)$ in (1) and (2), reduce the class of admissible measures to those that are not of the form $c\delta_0$ in (3) and (4), and end (5) at 'entire function'. These five modified statements are, once again, equivalent, and provide further equivalent conditions to those of Vasudeva (Theorems 3.3 and 3.18).

In a similar vein, we now present the classification of sign patterns of polynomial or power-series functions that preserve TN entrywise in a fixed dimension on Hankel matrices. This too turns out to be exactly the same as for positivity preservers.

Theorem 5.15 (Khare–Tao [89])**.** *Fix $\rho > 0$ and real exponents $n_0 < \cdots < n_{N-1} < M$. For any real coefficients $c_0, \ldots, c_{N-1}, c'$, let*

$$f(x) := \sum_{j=0}^{N-1} c_j x^{n_j} + c' x^M. \tag{5.1}$$

The following are equivalent.

(1) *The entrywise map $f[-]$ preserves TN on the rank-one matrices in $HTN_N\big((0,\rho)\big)$.*

(2) *The entrywise map $f[-]$ preserves positivity on the rank-one matrices in $HTN_N\big((0,\rho)\big)$.*

(3) *Either all the coefficients $c_0, \ldots, c_{N-1}, c'$ are non-negative, or $c_0, \ldots, c_{N-1}$ are positive and $c' \geq -\mathcal{C}^{-1}$, where*

$$\mathcal{C} = \sum_{j=0}^{N-1} \frac{V(\mathbf{n}_j)^2}{V(\mathbf{n})^2} \frac{\rho^{M-n_j}}{c_j}. \tag{5.2}$$

If $n_j \in \mathbb{Z}_+ \cup [N-2, \infty)$ for $j = 0, \ldots, N-1$, then conditions (1), (2) and (3) are further equivalent to the following.

(4) *The entrywise map $f[-]$ preserves TN on $HTN_N\big([0,\rho]\big)$.*

In particular, this produces further equivalent conditions to Theorem 4.15. Notice that assertion (2) here is valid because the rank-one matrices used in proving Theorem 4.15 are of the form $c\mathbf{u}\mathbf{u}^T$, where $\mathbf{u} = (1, u_0, \ldots, u_0^{N-1})^T$, $u_0 \in (0, 1)$, and $c \in (0, \rho)$, so that $c\mathbf{u}\mathbf{u}^T \in HTN_N\big((0,\rho)\big)$.

The consequences of Theorem 4.15 also carry over for TN preservers. For instance, one can bound Laplace transforms analogously to Corollary 4.17, by replacing the words "positive semidefinite" by "totally non-negative" and the set $\mathcal{P}_N\big((0,\rho)\big)$ by $HTN_N\big((0,\rho)\big)$. Similarly, one can completely classify the sign patterns of power series that preserve TN entrywise on Hankel matrices of a fixed size:

Theorem 5.16 (Khare–Tao [89])**.** *Theorems 4.5 and 4.19 hold upon replacing the phrase "preserves positivity entrywise on $\mathcal{P}_N\big((0,\rho)\big)$" with "preserves TN entrywise on $HTN_N\big((0,\rho)\big)$", for both $\rho < \infty$ and for $\rho = \infty$.*

We point the reader to [89, End of Section 9] for details.

To conclude, it is natural to seek a general result that relates the positivity preservers on $\mathcal{P}_N(I)$ and TN preservers on the set $HTN_N(I)$ for domains $I \subset \mathbb{R}_+$. Here is one variant which helps prove the above theorems, and which essentially follows from Theorem 5.11.

Proposition 5.17 (Khare–Tao [89]). *Fix integers $1 \leq k \leq N$ and a scalar $0 < \rho \leq \infty$. Suppose $f : [0, \rho) \to \mathbb{R}$ is such that the entrywise map $f[-]$ preserves positivity on $\mathcal{P}_N^k([0, \rho))$, the set of matrices in $\mathcal{P}_N([0, \rho))$ with rank no more than k. Then $f[-]$ preserves total non-negativity on $HTN_N([0, \rho)) \cap \mathcal{P}_N^k([0, \rho))$.*

5.3. **Entrywise preservers of totally non-negative matrices.** The TN property is very rigid when it comes to entrywise operations, as the following result makes clear.

Theorem 5.18 ([13, Theorem 2.1]). *Let $F : \mathbb{R}_+ \to \mathbb{R}$ be a function and let $d := \min(m, n)$, where m and n are positive integers. The following are equivalent.*

(1) *F preserves TN entrywise on $m \times n$ matrices.*
(2) *F preserves TN entrywise on $d \times d$ matrices.*
(3) *F is either a non-negative constant or*
 (a) *$(d = 1)$ $F(x) \geq 0$;*
 (b) *$(d = 2)$ $F(x) = cx^\alpha$ for some $c > 0$ and some $\alpha \geq 0$;*
 (c) *$(d = 3)$ $F(x) = cx^\alpha$ for some $c > 0$ and some $\alpha \geq 1$;*
 (d) *$(d \geq 4)$ $F(x) = cx$ for some $c > 0$.*

Proof. That (1) $\iff$ (2) is immediate, as is the equivalence of (2) and (3) when $d = 1$. For larger values of d, we sketch the implication (2) $\implies$ (3).

For $d = 2$, let the totally non-negative matrices

$$A(x, y) := \begin{bmatrix} x & xy \\ 1 & y \end{bmatrix} \quad \text{and} \quad B(x, y) := \begin{bmatrix} xy & x \\ y & 1 \end{bmatrix} \quad (x, y \geq 0). \quad (5.3)$$

If the non-constant function F preserves TN entrywise for 2×2 matrices, then the non-negativity of the determinants of $F[A(x, y)]$ and $F[B(x, y)]$ gives that

$$F(xy)F(1) = F(x)F(y) \qquad \text{for all } x, y \geq 0. \quad (5.4)$$

It follows that F is strictly positive. Applying Vasudeva's argument, as set out before Proposition 3.8, now implies that F is continuous on $(0, \infty)$. Since the identity (5.4) shows that $x \mapsto F(x)/F(1)$ is multiplicative, there exists an exponent $\alpha \in \mathbb{R}_+$ such that $F(x) = F(1)x^\alpha$ for all $x > 0$. The final details are left as an exercise.

For $d = 3$, note that the 3×3 matrix $A \oplus 0$ is totally non-negative if and only if the 2×2 matrix A is. Hence the previous working gives that $F(x) = cx^\alpha$ for some $c > 0$ and $\alpha \geq 0$. Looking at $\det F[C]$ for the totally

non-negative matrix

$$C := \begin{bmatrix} 1 & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1 \end{bmatrix} \tag{5.5}$$

shows that we must have $\alpha \geq 1$.

The argument to rule out the possibility that $\alpha \in [1,2)$ when $d \geq 4$ is more involved, but makes use of an example of Fallat, Johnson and Sokal [48, Example 5.8]. Full details are provided in [13]. □

If our totally non-negative matrices are also required to be symmetric, and so positive semidefinite, then the classes of preservers are enlarged somewhat, but still fairly restrictive.

Theorem 5.19 ([13, Theorem 2.3]). *Let $F : \mathbb{R}_+ \to \mathbb{R}$ and let d be a positive integer. The following are equivalent.*

(1) *F preserves TN entrywise on symmetric $d \times d$ matrices.*
(2) *F is either a non-negative constant or*
  (a) *$(d = 1)$ $F \geq 0$;*
  (b) *$(d = 2)$ F is non-negative, non-decreasing, and multiplicatively mid-convex, that is, $F(\sqrt{xy})^2 \leq F(x)F(y)$ for all x, $y \in [0, \infty)$, so continuous;*
  (c) *$(d = 3)$ $F(x) = cx^\alpha$ for some $c > 0$ and some $\alpha \geq 1$;*
  (d) *$(d = 4)$ $F(x) = cx^\alpha$ for some $c > 0$ and some $\alpha \in \{1\} \cup [2, \infty)$;*
  (e) *$(d \geq 5)$ $F(x) = cx$ for some $c > 0$.*

5.4. **Entrywise preservers of totally positive matrices.** In moving from total non-negativity to total positivity, we face two significant technical challenges. Firstly, the idea of realizing totally non-negative $d \times d$ matrices as submatrices of totally non-negative $(d + 1) \times (d + 1)$ matrices, by padding with zeros, does not transfer to the TP setting. Secondly, it is no longer possible to use Vasudeva's idea to establish multiplicative midpoint convexity, since the test matrices used for this are not always totally positive.

The first issue leads us into the domain of *totally positive completion problems* [47]. It is possible to do this generality, using parametrizations of TP matrices [53] or exterior bordering [46, Chapter 9], but the following result has the advantage of providing an explicit embedding into a well-known class of matrices.

Lemma 5.20 ([13, Lemma 3.2]). *Any totally positive 2×2 matrix may be realized as the leading principal submatrix of a positive multiple of a rectangular totally positive generalized Vandermonde matrix of any larger size.*

Remark 5.21 ([13, Remark 3.4]). Lemma 5.20 can be strengthened to the following completion result: given integers m, $n \geq 2$, an arbitrary 2×2 matrix A occurs as a minor in a totally positive $m \times n$ matrix at any given

position (that is, in a specified pair of rows and pair of columns) if and only if A is totally positive.

The other tool which will be vital to our deliberations is the following result of Whitney.

Theorem 5.22 ([139, Theorem 1]). *The set of totally positive $m \times n$ matrices is dense in the set of totally non-negative $m \times n$ matrices.*

With these tools in hand, we are able to provide a complete classification of the entrywise TP preservers of each fixed size, akin to the results in the preceding section.

Theorem 5.23 ([13, Theorem 3.1]). *Let $F : (0, \infty) \to \mathbb{R}$ be a function and let $d := \min(m, n)$, where m and n are positive integers. The following are equivalent.*

  (1) *F preserves total positivity entrywise on $m \times n$ matrices.*
  (2) *F preserves total positivity entrywise on $d \times d$ matrices.*
  (3) *The function F satisfies*
      (a) *$(d = 1)$ $F(x) > 0$;*
      (b) *$(d = 2)$ $F(x) = cx^\alpha$ for some $c > 0$ and some $\alpha > 0$;*
      (c) *$(d = 3)$ $F(x) = cx^\alpha$ for some $c > 0$ and some $\alpha \geq 1$;*
      (d) *$(d \geq 4)$ $F(x) = cx$ for some $c > 0$.*

Proof. We sketch the proof that $(2) \implies (3)$ when $d = 2$ and $d \geq 3$. For the first case, working with the matrix

$$\begin{bmatrix} y & x \\ x & y \end{bmatrix} \qquad (y > x > 0)$$

shows that F takes positive values and is increasing, so is Borel measurable and continuous except on a countable set. We now fix a point of continuity a and use the totally positive matrices

$$A(x, y, \epsilon) := \begin{bmatrix} ax & axy \\ a - \epsilon & ay \end{bmatrix} \quad \text{and} \quad B(x, y, \epsilon) := \begin{bmatrix} axy & ax \\ ay & a + \epsilon \end{bmatrix}$$

to show that

$$0 \leq \lim_{\epsilon \to 0^+} \det F[A(x, y, \epsilon)] = F(ax)F(ay) - F(axy)F(a)$$

$$\text{and} \quad 0 \leq \lim_{\epsilon \to 0^+} \det F[B(x, y, \epsilon)] = F(a)F(axy) - F(ax)F(ay)$$

for all $x, y > 0$. Hence $G : x \mapsto F(ax)/F(a)$ is such that

$$G(xy) = G(x)G(y) \qquad \text{for all } x, y > 0,$$

so G is a measurable solution of the Cauchy functional equation. It follows that $G(x) = x^\alpha$ for some $\alpha \in \mathbb{R}$. As F, and so G, is increasing, we must have $\alpha > 0$.

Finally, if $d \geq 3$, then the embedding of Lemma 5.20 and the previous working give positive constants c and α such that $F(x) = cx^\alpha$. In particular,

the function F admits a continuous extension $\tilde{F}$ to $\mathbb{R}_+$. The density of TP in TN, that is, Theorem 5.22, implies that $\tilde{F}$ preserves TN entrywise on $d \times d$ matrices. Theorem 5.18 now establishes the form of $\tilde{F}$, and so of F. $\qquad \square$

We may consider a version of the previous theorem which restricts to the case of totally positive matrices which are symmetric. A moment's thought leads to the consideration of a symmetric version of the matrix completion problem.

Lemma 5.24 ([13, Lemma 3.7]). *Any symmetric totally positive 2×2 matrix occurs as the leading principal submatrix of a totally positive $d \times d$ Hankel matrix, where $d \geq 2$ can be taken arbitrary large.*

Proof. It suffices to embed the matrix

$$\begin{bmatrix} 1 & a \\ a & b \end{bmatrix} \qquad (0 < a < \sqrt{b})$$

into such a Hankel matrix. It is an exercise to prove the existence of a continuous function $f : [0,1] \to \mathbb{R}_+;\ x \mapsto cx^s$ such that

$$\int_0^1 f(x)\,\mathrm{d}x = a \qquad \text{and} \qquad \int_0^1 f(x)^2\,\mathrm{d}x = b,$$

and then setting

$$a_{jk} := \int_0^1 f(x)^{j+k}\,\mathrm{d}x \qquad (j, k \geq 0)$$

gives a Hankel matrix A as required. The verification of total positivity may be made with the help of Andréief's identity,

$$\det \left[\int \phi_i(x)\psi_j(x)\,\mathrm{d}x \right]_{i,j=1}^k$$
$$= \frac{1}{k!} \int \cdots \int \det(\phi_i(x_j))_{i,j=1}^k \det(\psi_i(x_j))_{i,j=1}^k\,\mathrm{d}x_1 \cdots \mathrm{d}x_k,$$

where $\phi_i(x) = f(x)^{\alpha_i - 1}$ and $\psi_j(x) = f(x)^{\beta_j - 1}$, with

$$1 \leq \alpha_1 < \cdots < \alpha_k \leq d \qquad \text{and} \qquad 1 \leq \beta_1 < \cdots < \beta_k \leq d,$$

together with the total positivity of generalized Vandermonde matrices. $\quad \square$

We remark here that the preceding result can be further strengthened to have the symmetric TP 2×2 matrix occur in any "symmetric" position inside a larger square symmetric TP Hankel matrix, in the spirit of Remark 5.21. See [13, Theorem 3.9] for details.

We now state the symmetric version of Theorem 5.23.

Theorem 5.25 ([13, Theorem 3.6]). *Let $F : (0, \infty) \to \mathbb{R}$ and let d be a positive integer. The following are equivalent.*

    (1) *F preserves total positivity entrywise on symmetric $d \times d$ matrices.*
    (2) *The function F satisfies*

(a) $(d = 1)$ $F(x) > 0$;
(b) $(d = 2)$ F is positive, increasing, and multiplicatively mid-convex, that is, $F(\sqrt{xy})^2 \leq F(x)F(y)$ for all x, $y \in (0, \infty)$, so continuous;
(c) $(d = 3)$ $F(x) = cx^\alpha$ for some $c > 0$ and some $\alpha \geq 1$;
(d) $(d = 4)$ $F(x) = cx^\alpha$ for some $c > 0$ and some $\alpha \in \{1\} \cup [2, \infty)$.
(e) $(d \geq 5)$ $F(x) = cx$ for some $c > 0$.

Although we have developed the key ingredients to prove this theorem, we content ourselves with referring the interested reader to [13].

6. POWER FUNCTIONS

A natural approach to tackle the problem of characterizing entrywise preservers in fixed dimension is to examine if some natural simple functions preserve positivity. One such family is the collection of power functions, $f(x) = x^\alpha$ for $\alpha > 0$. Characterizing which fractional powers preserve positivity entrywise has recently received much attention in the literature. One of the first results in this area reads as follows.

Theorem 6.1 (FitzGerald and Horn [51, Theorem 2.2]). *Let $N \geq 2$ and let $A = [a_{jk}] \in \mathcal{P}_N(\mathbb{R}_+)$. For any real number $\alpha \geq N - 2$, the matrix $A^{\circ\alpha} := [a_{jk}^\alpha]$ is positive semidefinite. If $0 < \alpha < N - 2$ and α is not an integer, then there exists a matrix $A \in \mathcal{P}_N((0, \infty))$ such that $A^{\circ\alpha}$ is not positive semidefinite.*

Theorem 6.1 shows that every real power $\alpha \geq N - 2$ entrywise preserves positivity, while no non-integers in $(0, N - 2)$ do. This surprising "phase transition" phenomenon at the integer $N - 2$ is referred to as the "critical exponent" for preserving positivity. Studying which powers entrywise preserve positivity is a very natural and interesting problem. It also often provides insights to determine which general functions preserve positivity. For example, Theorem 6.1 suggests that functions that entrywise preserve positivity on $\mathcal{P}_N$ should have a certain number of non-negative derivatives, which is indeed the case by Theorem 3.4.

Outline of the proof. The first part of Theorem 6.1 relies on an ingenious idea that we now sketch. The result is obvious for $N = 2$. Let us assume it holds for some $N - 1 \geq 2$, let $A \in \mathcal{P}_N(\mathbb{R}_+)$, and let $\alpha \geq N - 2$. Write A in block form,

$$A = \begin{bmatrix} B & \xi \\ \xi^T & a_{NN} \end{bmatrix},$$

where B has dimension $(N - 1) \times (N - 1)$ and $\xi \in \mathbb{R}^{N-1}$. Assume without loss of generality that $a_{NN} \neq 0$ (as the case where $a_{NN} = 0$ follows from the induction hypothesis) and let $\zeta := (\xi^T, a_{NN})^T / \sqrt{a_{NN}}$. Then $A - \zeta\zeta^T = (B - \xi\xi^T)/a_{NN} \oplus 0$, where $(B - \xi\xi^T)/a_{NN}$ is the Schur complement of a_{NN}

in A. Hence $A - \zeta\zeta^T$ is positive semidefinite. By the fundamental theorem of calculus, for any $x, y \in \mathbb{R}$,

$$x^\alpha = y^\alpha + \alpha \int_0^1 (x - y)(\lambda x + (1 - \lambda)y)^{\alpha - 1} \, d\lambda.$$

Using the above expression entrywise, we obtain

$$A^{\circ\alpha} = \zeta^{\circ\alpha}(\zeta^{\circ\alpha})^T + \int_0^1 (A - \zeta\zeta^T) \circ (\lambda A + (1 - \lambda)\zeta\zeta^T)^{\circ(\alpha - 1)} \, d\lambda.$$

Observe that the entries of the last row and column of the matrix $A - \zeta\zeta^T$ are all zero. Using the induction hypothesis and the Schur product theorem, it follows that the integrand is positive semidefinite, and therefore so is $A^{\circ\alpha}$.

The converse implication in Theorem 6.1 is shown by considering a matrix of the form $a\mathbf{1}_{N\times N} + t\mathbf{u}\mathbf{u}^T$, where $a, t > 0$, the coordinates of $\mathbf{u}$ are distinct, and $t1$ is small. Recall this is the exact same class of matrices that was useful in proving the Horn–Loewner theorem 3.4 as well as its strengthening in Theorem 3.10. The original proof, by FitzGerald and Horn [51], used $\mathbf{u} = (1, 2, \ldots, N)^T$, while a later proof by Fallat, Johnson and Sokal [48] used the same argument, now with $\mathbf{u} = (1, u_0, \ldots, u_0^{N-1})^T$; the motivation in [48] was to work with Hankel matrices, and the matrix $a\mathbf{1}_{N\times N} + t\mathbf{u}\mathbf{u}^T$ is indeed Hankel. That said, the argument of FitzGerald and Horn works more generally than both of these proofs, to show that, for any non-integral power $\alpha \in (0, N-2)$, $a > 0$, and vector $\mathbf{u} \in (0, \infty)^N$ with distinct coordinates, there exists $t > 0$ such that $(a\mathbf{1}_{N\times N} + t\mathbf{u}\mathbf{u}^T)^{\circ\alpha}$ is not positive semidefinite. $\square$

In her 2017 paper [81], Jain provided a remarkable strengthening of the result mentioned at the end of the previous proof, which removes the dependence on t entirely.

Theorem 6.2 (Jain [81]). *Let*

$$A := [1 + u_j u_k]_{j,k=1}^N = \mathbf{1}_{N\times N} + \mathbf{u}\mathbf{u}^T,$$

where $N \geq 2$ and $\mathbf{u} = (u_1, \ldots, u_N)^T \in (0, \infty)^N$ has distinct entries. Then $A^{\circ\alpha}$ is positive semidefinite for $\alpha \in \mathbb{R}$ if and only if $\alpha \in \mathbb{Z}_+ \cup [N - 2, \infty)$.

Jain's result identifies a family of rank-two positive semidefinite matrices, every one of which encodes the classification of powers preserving positivity over all of $\mathcal{P}_N\big((0, \infty)\big)$. In a sense, her rank-two family is the culmination of previous work on positivity preserving powers for $\mathcal{P}_N\big((0, \infty)\big)$, since for rank-one matrices, every entrywise power preserves positivity: $(\mathbf{u}\mathbf{u}^T)^{\circ\alpha} = \mathbf{u}^{\circ\alpha}(\mathbf{u}^{\circ\alpha})^T$.

An immediate consequence of these results is the classification of the entrywise powers preserving positivity on the $N \times N$ TN Hankel matrices. Recall from the results in Section 5.2 (including Lemma 5.12(4)) that there is to be expected a strong correlation between this classification and the one in Theorem 6.1.

Corollary 6.3. *Given $N \geq 2$, the following are equivalent for an exponent $\alpha \in \mathbb{R}$.*

(1) *The entrywise power function $x \mapsto x^{\alpha}$ preserves total non-negativity on HTN_N (see Lemma 5.12).*
(2) *The entrywise map $x \mapsto x^{\alpha}$ preserves positivity on HTN_N.*
(3) *The entrywise map $x \mapsto x^{\alpha}$ preserves positivity on the matrices in $HTN_N\big((0, \infty)\big)$ of rank at most two.*
(4) *The exponent $\alpha \in \mathbb{Z}_+ \cup [N - 2, \infty)$.*

Proof. That $(4) \implies (2)$ and $(2) \implies (1)$ follow from Theorems 6.1 and 5.11, respectively. That $(1) \implies (2)$ and $(2) \implies (3)$ are obvious, and Jain's Theorem 6.2 shows that $(3) \implies (4)$. □

A problem related to the above study of entrywise powers preserving positivity, is to characterize infinitely divisible matrices. This problem was also considered by Horn in [80]. Recall that a complex $N \times N$ matrix is said to be *infinitely divisible* if $A^{\circ \alpha} \in \mathcal{P}_N$ for all $\alpha \in \mathbb{R}_+$. Denote the *incidence matrix* of A by $M(A)$:

$$M(A)_{jk} = m_{jk} := \begin{cases} 0 & \text{if } a_{jk} = 0 \\ 1 & \text{otherwise.} \end{cases}$$

Also, let

$$L(A) := \{\mathbf{x} \in \mathbb{C}^N : \sum_{j,k=1}^{N} m_{jk} x_j \overline{x_k} = 0\},$$

and note that $L(A)$ is the kernel of $M(A)$ if $M(A)$ is positive semidefinite.

Assuming the arguments of the entries are chosen in a *consistent way* [80], we let

$$\log^{\#} A := M(A) \circ \log[A] = [\mu_{jk} \log a_{jk}]_{j,k=1}^{N},$$

with the usual convention $0 \log 0 = 0$.

Theorem 6.4 (Horn [80, Theorem 1.4]). *An $N \times N$ matrix A is infinitely divisible if and only if (a) A is Hermitian, with $a_{jj} \geq 0$ for all j, (b) $M(A) \in \mathcal{P}_N$, and (c) $\log^{\#} A$ is positive semidefinite on $L(A)$.*

6.1. **Sparsity constraints.** Theorem 6.1 was recently extended to more structured matrices. Given $I \subset \mathbb{R}$ and a graph $G = (V, E)$ on the finite vertex set $V = \{1, \dots, N\}$, we define the cone of positive-semidefinite matrices with *zeros according to G*:

$$\mathcal{P}_G(I) := \{A = [a_{jk}] \in \mathcal{P}_N(I) : a_{jk} = 0 \text{ if } (j, k) \notin E \text{ and } i \neq j\}. \quad (6.1)$$

Note that if $(j, k) \in E$, then the entry a_{jk} is unconstrained; in particular, it is allowed to be 0. Consequently, the cone $\mathcal{P}_G := \mathcal{P}_G(\mathbb{R})$ is a closed subset of $\mathcal{P}_N$.

A natural refinement of Theorem 6.1 involves studying powers that entrywise preserve positivity on $\mathcal{P}_G$. In that case, the flavor of the problem changes significantly, with the discrete structure of the graph playing a prominent role.

Definition 6.5 (Guillot–Khare–Rajaratnam [69])**.** Given a simple graph $G = (V, E)$, let

$$\mathcal{H}_G := \{\alpha \in \mathbb{R} : A^{\circ\alpha} \in \mathcal{P}_G \text{ for all } A \in \mathcal{P}_G(\mathbb{R}_+)\}. \tag{6.2}$$

Define the *Hadamard critical exponent of G* to be

$$CE(G) := \min\{\alpha \in \mathbb{R} : [\alpha, \infty) \subset \mathcal{H}_G\}. \tag{6.3}$$

Notice that, by Theorem 6.1, for every graph $G = (V, E)$, the critical exponent $CE(G)$ exists, and lies in $[\omega(G) - 2, |V| - 2]$, where $\omega(G)$ is the size of the largest complete subgraph of G, that is, the clique number. To compute such critical exponents is natural and highly non-trivial.

FitzGerald and Horn proved that $CE(K_n) = n - 2$ for all $n \geq 2$ (Theorem 6.1), while it follows from [70, Proposition 4.2] that $CE(T) = 1$ for every tree T. For a general graph, it is not *a priori* clear what the critical exponent is or how to compute it. A natural family of graphs that encompasses both complete graphs and trees is that of chordal graphs. Recall that a graph is *chordal* if it does not contain an induced cycle of length 4 or more. Chordal graphs feature extensively in many areas, such as the theory of graphical models [93], and in problems involving positive-definite completions (see [130]). Examples of important chordal graphs include trees, complete graphs, Apollonian graphs, band graphs, and split graphs.

Recently, Guillot, Khare, and Rajaratnam [69] were able to compute the complete set of entrywise powers preserving positivity on $\mathcal{P}_G$ for all chordal graphs G. Here, the critical exponent can be described purely combinatorially.

Theorem 6.6 (Guillot–Khare–Rajaratnam [69])**.** *Let $K_r^{(1)}$ denote the complete graph with one edge removed, and let G be a finite simple connected chordal graph. The critical exponent for entrywise powers preserving positivity on $\mathcal{P}_G$ is $r - 2$, where r is the largest integer such that K_r or $K_r^{(1)}$ is an induced subgraph of G. More precisely, the set of entrywise powers preserving $\mathcal{P}_G$ is $\mathcal{H}_G = \mathbb{Z}_+ \cup [r - 2, \infty)$, with r as before.*

The set of entrywise powers preserving positivity was also computed in [69] for cycles and bipartite graphs.

Theorem 6.7 (Guillot–Khare–Rajaratnam [69])**.** *The critical exponent of cycles and bipartite graphs is 1.*

Surprisingly, the critical exponent does not depend on the size of the graph for cycles and bipartite graphs. In particular, it is striking that any power greater than 1 preserves positivity for families of dense graphs such as bipartite graphs. Such a result is in sharp contrast to the general case, where

there is no underlying structure of zeros. That small powers can preserve positivity is important for applications, since such entrywise procedures are often used to regularize positive definite matrices, such as covariance or correlation matrices, where the goal is to minimally modify the entries of the original matrix (see [94, 143] and Chapter 7 below).

For a general graph, the problem of computing the set $\mathcal{H}_G$ or the critical exponent $CE(G)$ remains open. We now outline some other natural open problems in the area.

Problems.

(1) In every currently known case (Theorems 6.6, 6.7), $CE(G)$ is equal to $r - 2$, where r is the largest integer such that K_r or $K_r^{(1)}$ is an induced subgraph of G. Is the same true for every graph G?

(2) Is $CE(G)$ always an integer? Can this be proved without computing $CE(G)$ explicitly?

(3) Recall that every chordal graph is perfect. Can the critical exponent be calculated for other broad families of graphs such as the family of perfect graphs?

6.2. **Rank constraints and other Loewner properties.** Another approach to generalize Theorem 6.1 is to examine other properties of entrywise functions such as monotonicity, convexity, and super-additivity (with respect to the Loewner semidefinite ordering) [78, 68]. Given a set $V \subset \mathcal{P}_N(I)$, recall that a function $f : I \to \mathbb{R}$ is

- *positive on V* with respect to the Loewner ordering if $f[A] \geq 0$ for all $0 \leq A \in V$;
- *monotone on V* with respect to the Loewner ordering if $f[A] \geq f[B]$ for all $A, B \in V$ such that $A \geq B \geq 0$;
- *convex on V* with respect to the Loewner ordering if $f[\lambda A + (1 - \lambda)B] \leq \lambda f[A] + (1 - \lambda)f[B]$ for all $\lambda \in [0, 1]$ and all $A, B \in V$ such that $A \geq B \geq 0$;
- *super-additive on V* with respect to the Loewner ordering if $f[A + B] \geq f[A] + f[B]$ for all $A, B \in V$ for which $f[A + B]$ is defined.

The following relations between the first three notions were obtained by Hiai.

Theorem 6.8 (Hiai [78, Theorem 3.2])**.** *Let $I = (-\rho, \rho)$ for some $\rho > 0$.*

(1) *For each $N \geq 3$, the function f is monotone on $\mathcal{P}_N(I)$ if and only if f is differentiable on I and f' is positive on $\mathcal{P}_N(I)$.*

(2) *For each $N \geq 2$, the function f is convex on $\mathcal{P}_N(I)$ if and only if f is differentiable on I and f' is monotone on $\mathcal{P}_N(I)$.*

Power functions satisfying any of the above four properties have been characterized by various authors. In recent work, Hiai [78] has extended Theorem 6.1 by considering the odd and even extensions of the power functions to $\mathbb{R}$. For $\alpha > 0$, the even and odd extensions to $\mathbb{R}$ of the power function $f_\alpha(x) := x^\alpha$ are defined to be $\phi_\alpha(x) := |x|^\alpha$ and $\psi_\alpha(x) := \text{sign}(x)|x|^\alpha$.

The first study of powers $\alpha > 0$ for which ϕ_α preserves positivity entrywise on $\mathcal{P}_N(\mathbb{R})$ was carried out by Bhatia and Elsner [18]. Subsequently, Hiai studied the power functions ϕ_α and ψ_α that preserve Loewner positivity, monotonicity, and convexity entrywise, and showed for positivity preservers that the same phase transition occurs at $n - 2$ for ϕ_α and ψ_α, as demonstrated in [51]. The work was generalized in [68] to matrices satisfying rank constraints.

Definition 6.9. Fix non-negative integers $n \geq 2$ and $n \geq k$, and a set $I \subset \mathbb{R}$. Let $\mathcal{P}_n^k(I)$ denote the subset of matrices in $\mathcal{P}_n(I)$ that have rank at most k, and let

$$\mathcal{H}_{\mathrm{pos}}(n, k) := \{\alpha > 0 : x^\alpha \text{ preserves positivity on } \mathcal{P}_n^k(\mathbb{R}_+)\},$$

$$\mathcal{H}_{\mathrm{pos}}^\phi(n, k) := \{\alpha > 0 : \phi_\alpha \text{ preserves positivity on } \mathcal{P}_n^k(\mathbb{R})\}, \qquad (6.4)$$

$$\mathcal{H}_{\mathrm{pos}}^\psi(n, k) := \{\alpha > 0 : \psi_\alpha \text{ preserves positivity on } \mathcal{P}_n^k(\mathbb{R})\}.$$

Similarly, let $\mathcal{H}_J(n, k)$, $\mathcal{H}_J^\phi(n, k)$ and $\mathcal{H}_J^\psi(n, k)$ denote sets of the entrywise powers preserving Loewner properties on $\mathcal{P}_n^k(\mathbb{R}_+)$ or $\mathcal{P}_n^k(\mathbb{R})$, where $J \in \{\text{monotonicity, convexity, super-additivity}\}$.

The set of entrywise powers preserving the above notions are given in the table below (see [68, Theorem 1.2]).

7. Motivation from statistics

The study of entrywise functions preserving positivity has recently attracted renewed attraction due to its importance in the estimation and regularization of covariance/correlation matrices. Recall that the covariance between two random variables X_j and X_k is given by

$$\sigma_{jk} = \mathrm{Cov}(X_j, X_k) = E\big[(X_j - E[X_j])(X_k - E[X_k])\big],$$

where $E[X_j]$ denotes the expectation of X_j. In particular, $\mathrm{Cov}(X_j, X_j) = \mathrm{Var}(X_j)$, the variance of X_j. The *covariance matrix* of a random vector $\mathbf{X} := (X_1, \ldots, X_m)$, is the matrix $\Sigma := [\mathrm{Cov}(X_j, X_k)]_{j,k=1}^m$. Covariance matrices are a fundamental tool that measure linear dependencies between random variables. In order to discover relations between variables in data, statisticians and applied scientists need to obtain estimates of the covariance matrix Σ from observations $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^m$ of $\mathbf{X}$. A traditional estimator of Σ is the *sample covariance matrix S* given by

$$S = [s_{jk}]_{j,k=1}^m = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^T, \qquad (7.1)$$

where $\overline{\mathbf{x}} := \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i$ is the average of the observations. In the case where the random vector $\mathbf{X}$ has a multivariate normal distribution with mean μ and covariance matrix Σ, one can show that $\overline{\mathbf{x}}$ and $\frac{n-1}{n} S$ are the maximum likelihood estimators of μ and Σ, respectively [3, Chapter 3]. It is not difficult to show that S is an unbiased estimator of Σ. More generally, under weak

J	$\mathcal{H}_J(n,k)$	$\mathcal{H}_J^\phi(n,k)$	$\mathcal{H}_J^\psi(n,k)$
Positivity			
$k=1$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$
	G–K–R	G–K–R	G–K–R
$2 \le k \le n$	$\mathbb{N} \cup [n-2, \infty)$	$2\mathbb{N} \cup [n-2, \infty)$	$(-1+2\mathbb{N}) \cup [n-2, \infty)$
	FitzGerald–Horn	Hiai, Bhatia–Elsner,	Hiai, G–K–R
		G–K–R	
Monotonicity			
$k=1$	$\mathbb{R}_+$	$\mathbb{R}_+$	$\mathbb{R}_+$
	G–K–R	G–K–R	G–K–R
$2 \le k \le n$	$\mathbb{N} \cup [n-1, \infty)$	$2\mathbb{N} \cup [n-1, \infty)$	$(-1+2\mathbb{N}) \cup [n-1, \infty)$
	FitzGerald–Horn	Hiai, G–K–R	Hiai, G–K–R
Convexity			
$k=1$	$[1, \infty)$	$[1, \infty)$	$[1, \infty)$
	G–K–R	G–K–R	G–K–R
$2 \le k \le n$	$\mathbb{N} \cup [n, \infty)$	$2\mathbb{N} \cup [n, \infty)$	$(-1+2\mathbb{N}) \cup [n, \infty)$
	Hiai, G–K–R	Hiai, G–K–R	Hiai, G–K–R
Super-additivity			
$1 \le k \le n$	$\mathbb{N} \cup [n, \infty)$	$2\mathbb{N} \cup [n, \infty)$	$(-1+2\mathbb{N}) \cup [n, \infty)$
	G–K–R	G–K–R	G–K–R

TABLE 1. Summary of real Hadamard powers preserving Loewner properties, with additional rank constraints. See Bhatia–Elsner [18], FitzGerald–Horn [51], Guillot–Khare–Rajaratnam [68], and Hiai [78].

assumptions, one can show that the distribution of $\sqrt{n}(S - \Sigma)$ is asymptotically normal as $n \to \infty$. The exact description of the limiting distribution depends on the moments and the cumulants of $\mathbf{X}$ (see [20, Chapter 6.3]). For example, in the two-dimensional case, we have the following result.

Let $N_m(\mu, \Sigma)$ denote the m-dimensional normal distribution with mean μ and covariance matrix Σ.

Proposition 7.1 (see [20, Example 6.4]). *Let* $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^2$ *be an independent and identically distributed sample from a bivariate vector* $\mathbf{X} = (X_1, X_2)$ *with mean* $\mu = (\mu_1, \mu_2)$ *and finite fourth-order moments, and let* S *be as in Equation (7.1). Then*

$$\sqrt{n} \left[\begin{bmatrix} s_1^2 \\ s_{12} \\ s_2^2 \end{bmatrix} - \begin{bmatrix} \sigma_1^2 \\ \sigma_{12} \\ \sigma_2^2 \end{bmatrix} \right] \xrightarrow{\mathrm{d}} N_3(\mathbf{0}, \Omega),$$

where Ω *is the symmetric* 3×3 *matrix*

$$\Omega = \begin{bmatrix} \mu_4^1 - (\mu_2^1)^2 & \mu_{31}^{12} - \mu_{11}^{12}\mu_2^1 & \mu_{22}^{12} - \mu_2^1\mu_2^2 \\ \mu_{31}^{12} - \mu_{11}^{12}\mu_2^1 & \mu_{22}^{12} - (\mu_{11}^{12})^2 & \mu_{13}^{12} - \mu_{11}^{12}\mu_2^2 \\ \mu_{22}^{12} - \mu_2^1\mu_2^2 & \mu_{31}^{12} - \mu_{11}^{12}\mu_2^1 & \mu_4^2 - (\mu_2^2)^2 \end{bmatrix},$$

and $\mu_k^i = E[(X_i - \mu_i)^k]$ *and* $\mu_{kl}^{ij} = E[(X_i - \mu_i)^k(X_j - \mu_j)^l]$.

In traditional statistics, one usually assumes the number of samples n is large enough for asymptotic results such as the one above to apply. In covariance estimation, one typically requires a sample size at least a few times the number of variables m for that to apply. In such a case, the sample covariance matrix provides a good approximation of the true covariance matrix Σ. However, this ideal setting is rarely seen nowadays. Indeed, our systematic and automated way of collecting data today yields datasets where the number of variables is often orders of magnitude larger than the number of instances available for study [41]. Classical statistical methods were not designed and are not suitable for analyzing data in such settings. Developing new methodologies that are adapted to modern high-dimensional problems is the object of active research. In the case of covariance estimation, several strategies have been proposed to replace the traditional sample covariance matrix estimator S. These approaches typically leverage low-dimensional structures in the data (low rank, sparsity, ...) to obtain reasonable covariance estimates, even when the sample size is small compared to the dimension of the problem (see [111] for a detailed description of such techniques). One such approach involves applying functions to the entries of sample covariance matrices to improve their properties (see [6, 19, 44, 75, 76, 94, 114, 143]). For example, *hard thresholding* a matrix entails setting to zero the entries of the matrix that are smaller in absolute value than a prescribed value $\epsilon > 0$ (thinking the corresponding variables are independent, for example). Letting

$$f_\epsilon^H(x) = \begin{cases} x & \text{if } |x| > \epsilon, \\ 0 & \text{otherwise,} \end{cases} \tag{7.2}$$

hard thresholding is equivalent to applying the function f_ϵ^H entrywise to the entries of the matrix. Another popular example that was first studied in the context of wavelet shrinkage [42] is *soft thresholding*, where f_ϵ^H is replaced by

$$f_\epsilon^S : x \mapsto \text{sign}(x)\big(|x| - \epsilon\big)_+ \qquad \text{with } y_+ := \max\{y, 0\}.$$

Soft thresholding not only sets small entries to zero, it also shrinks all the other entries continuously towards zero. Several other thresholding and shrinkage procedures were also recently proposed in the context of covariance estimation (see [49] and the references therein).

Compared to other techniques, the above procedure has several advantages. Firstly, the resulting estimators are often significantly more precise than the sample covariance matrices. Secondly, applying a function to the entries of a matrix is very simple and not computationally intensive. The procedure can therefore be performed in very high dimensions and in real-time applications. This is in contrast to several other techniques that require solving optimization problems and often become too intensive to be used in modern applications. A downside of the entrywise calculus, however, is that the positive definiteness of the resulting matrices is not guaranteed. As the

parameter space of covariance matrices is the cone of positive definite matrices, it is critical that the resulting matrices be positive definite for the technique to be useful and widely applicable. The problem of characterizing positivity preservers thus has an immediate impact in the area of covariance estimation by providing useful functions that can be applied entrywise to covariance estimates in order to regularize them.

Several characterizations of when thresholding procedures preserve positivity have recently been obtained.

7.1. **Thresholding with respect to a graph.** In [72], the concept of thresholding with respect to a graph was examined. In this context, the elements to threshold are encoded in a graph $G = (V, E)$ with $V = \{1, \ldots, p\}$. If $A = (a_{jk})$ is a $p \times p$ matrix, we denote by A_G the matrix with entries

$$(A_G)_{jk} = \begin{cases} a_{jk} & \text{if } (j,k) \in E \text{ or } j = k, \\ 0 & \text{otherwise.} \end{cases}$$

We say that A_G is the matrix obtain by thresholding A with respect to the graph G. The main result of [72] characterizes the graphs G for which the corresponding thresholding procedure preserves positivity. Denote by $\mathcal{P}_N^+$ the set of real symmetric $N \times N$ positive definite matrices and by $\mathcal{P}_G^+$ the subset of positive definite matrices contained in $\mathcal{P}_G$ (see Equation (6.1)).

Theorem 7.2 (Guillot–Rajaratnam [72, Theorem 3.1]). *The following are equivalent:*

(1) $A_G \in \mathcal{P}_N^+$ *for all* $A \in \mathcal{P}_N^+$;
(2) $G = \bigcup_{i=1}^d G_i$, *where* G_1, ..., G_d *are disconnected and complete components of* G.

The implication (2) $\implies$ (1) of the theorem is intuitive and straightforward, since principal submatrices of positive definite matrices are positive definite. That (1) $\implies$ (2) may come as a surprise though, and shows that indiscriminate or arbitrary thresholding of a positive definite matrix can quickly lead to loss of positive definiteness.

Theorem 7.2 also generalizes to matrices that already have zero entries. In that case, the characterization of the positivity preservers remains essentially the same.

Theorem 7.3 (Guillot–Rajaratnam [72, Theorem 3.3]). *Let* $G = (V, E)$ *be an undirected graph and let* $H = (V, E')$ *be a subgraph of* G, *so that* $E' \subset E$. *Then* A_H *is positive definite for every* $A \in \mathcal{P}_G^+$ *if and only if* $H = G_1 \cup \cdots \cup G_k$, *where* G_1, ..., G_k *are disconnected induced subgraphs of* G.

7.2. **Hard and soft thresholding.** Theorems 7.2 and 7.3 address the case where positive definite matrices are thresholded with respect to a given pattern of entries, regardless of the magnitude of the entries of the original matrix. The more natural case where the entries are hard or soft thresholded

was studied in [72, 73]. In applications, it is uncommon to threshold the diagonal entries of estimated covariance matrices, as the diagonal contains the variance of the underlying variables. Hence, for a given function $f : \mathbb{R} \to \mathbb{R}$ and a real matrix $A = [a_{jk}]$, we let the matrix $f^*[A]$ be defined by setting

$$f^*[A]_{jk} := \begin{cases} f(a_{jk}) & \text{if } j \neq k, \\ a_{jk} & \text{otherwise.} \end{cases}$$

Theorem 7.4 (Guillot–Rajaratnam [72, Theorem 3.6]). *Let G be a connected undirected graph with $n \geq 3$ vertices. The following are equivalent.*

(1) *There exists $\epsilon > 0$ such that, for every $A \in \mathcal{P}_G^+$, we have $(f_\epsilon^H)^*[A] \in \mathcal{P}_n^+$.*
(2) *For every $\epsilon > 0$ and every $A \in \mathcal{P}_G^+$, we have $f_\epsilon^H[A] \in \mathcal{P}_n^+$.*
(3) *G is a tree.*

The case of soft thresholding was considered in [73]. Surprisingly, the characterization of the thresholding levels that preserve positivity is exactly the same as in the case of hard thresholding.

Theorem 7.5 (Guillot–Rajaratnam [73, Theorem 3.2]). *Let $G = (V, E)$ be a connected graph with $n \geq 3$ vertices. Then the following are equivalent:*

(1) *There exists $\epsilon > 0$ such that for every $A \in \mathcal{P}_G^+$, we have $(f_\epsilon^S)^*[A] \in \mathcal{P}_n^+$.*
(2) *For every $\epsilon > 0$ and every $A \in \mathcal{P}_G^+$, we have $f_\epsilon^S[A] \in \mathcal{P}_n^+$.*
(3) *G is a tree.*

An extension of Schoenberg's theorem (Theorem 2.12) to the case where the function f is only applied to the off-diagonal entries of the matrix was also obtained in [73].

Theorem 7.6 (Guillot–Rajaratnam [73, Theorem 4.21]). *Let $0 < \rho \leq \infty$ and $f : (-\rho, \rho) \to \mathbb{R}$. The matrix $f^*[A]$ is positive semidefinite for all $A \in \mathcal{P}_n\big((-\rho, \rho)\big)$ and all $n \geq 1$ if and only if $f(x) = xg(x)$, where*

(1) *g is analytic on the disc $D(0, \rho)$;*
(2) *$\|g\|_\infty \leq 1$;*
(3) *g is absolutely monotonic on $(0, \rho)$.*

When $\rho = \infty$, the only functions satisfying the above conditions are the affine functions $f(x) = ax$ for $0 \leq a \leq 1$.

7.3. **Rank and sparsity constraints.** An explicit and useful characterization of entrywise functions preserving positivity on $\mathcal{P}_N$ for a fixed N still remains out of reach as of today. Motivated by applications in statistics, the authors in [70, 71] examined the cases where the matrices in $\mathcal{P}_N$ satisfy supplementary rank and sparsity constraints that are common in applications.

Observe that the sample covariance matrix (Equation (7.1)) has rank at most n, where n is the number of samples used to compute it. Moreover, as explained at the start of this Chapter, it is common in modern applications

that n is much smaller than the dimension p. Hence, when studying the regularization approach described above, it is natural to consider positive semidefinite matrices with bounded rank.

An immediate application of Schoenberg's theorem on spheres (see Equation (2.3)) provides a characterization of entrywise positivity preservers of correlation matrices of all dimensions, with rank bounded by n. Recall that a correlation matrix is the covariance matrix of a random vector where each variable has variance 1, so is a positive semidefinite matrix with diagonal entries equal to 1. As in Equation (2.3), we denote the ultraspherical orthogonal polynomials by $P_k^{(\lambda)}$.

Theorem 7.7 (Reformulation of [125, Theorem 1]). *Let $n \in \mathbb{N}$ and let $f : [-1, 1] \to \mathbb{R}$. The following are equivalent.*

(1) $f[A] \in \mathcal{P}_N$ for all correlation matrices $A \in \mathcal{P}_N\big([-1, 1]\big)$ with rank no more than n and all $N \geq 1$.
(2) $f(x) = \sum_{j=0}^{\infty} a_j P_j^{(\lambda)}(x)$ with $a_j \geq 0$ for all $j \geq 0$ and $\lambda = (n-1)/2$.

Proof. The result follows from [125, Theorem 1] and the observation that correlation matrices of rank at most n are in correspondence with Gram matrices of vectors in S^{n-1}. $\qquad\square$

In order to approach the case of matrices of a fixed dimension, we introduce some notation.

Definition 7.8. Let $I \subset \mathbb{R}$. Define $\mathcal{S}_n(I)$ to be the set of $n \times n$ symmetric matrices with entries in I. Let $\operatorname{rank} A$ denote the rank of a matrix A. We define:

$$\mathcal{S}_n^k(I) := \{A \in \mathcal{S}_n(I) : \operatorname{rank} A \leq k\},$$
$$\mathcal{P}_n^k(I) := \{A \in \mathcal{P}_n(I) : \operatorname{rank} A \leq k\}.$$

The main result in [71] provides a characterization of entrywise functions mapping $\mathcal{P}_n^l$ into $\mathcal{P}_n^k$.

Theorem 7.9 (Guillot–Khare–Rajaratnam [71, Theorem B]). *Let $0 < R \leq \infty$ and $I = [0, R)$ or $(-R, R)$. Fix integers $n \geq 2$, $1 \leq k < n - 1$, and $2 \leq l \leq n$. Suppose $f \in C^k(I)$. The following are equivalent.*

(1) $f[A] \in \mathcal{S}_n^k$ for all $A \in \mathcal{P}_n^l(I)$;
(2) $f(x) = \sum_{t=1}^{r} c_t x^{i_t}$ for some $c_t \in \mathbb{R}$ and some $i_t \in \mathbb{N}$ such that

$$\sum_{t=1}^{r} \binom{i_t + l - 1}{l - 1} \leq k. \tag{7.3}$$

Similarly, $f[-] : \mathcal{P}_n^l(I) \to \mathcal{P}_n^k$ if and only if f satisfies (2) and $c_t \geq 0$ for all t. Moreover, if $I = [0, R)$ and $k \leq n - 3$, then the assumption that $f \in C^k(I)$ is not required.

Notice that Theorem 7.9 is a fixed-dimension result with rank constraints. This may be considered a refinement of a similar, dimension-free result with rank constraints shown in [5], in which the authors arrive at the same conclusion as in part (2) above. We compare the two settings: in [5], (a) the hypotheses held for all dimensions N rather than in a fixed dimension; (b) the test matrices were a larger set in each dimension, compared to just the positive matrices considered in Theorem 7.9; (c) the test matrices did not consist only of rank-one matrices, similar to Theorem 7.9; and (d) the test functions f in the dimension-free case were assumed to be measurable, rather than C^k as in the fixed-dimension case. Thus, Theorem 7.9 is (a refinement of) the fixed-dimension case of the first main result in [5].[10]

The (2) $\implies$ (1) implication in Theorem 7.9 is clear. Indeed, let $i \geq 0$ and $A = \sum_{j=1}^{l} u_j u_j^T \in \mathcal{P}_n^l(I)$. Then

$$A^{\circ i} = \sum_{m_1 + \cdots + m_l = i} \binom{i}{m_1, \ldots, m_l} \mathbf{w_m} \mathbf{w_m}^T \qquad \text{where } \mathbf{w_m} := u_1^{\circ m_1} \circ \cdots \circ u_l^{\circ m_l}$$

and $\binom{i}{m_1, \ldots, m_l}$ is a multinomial coefficient. Note that there are exactly $\binom{i+l-1}{l-1}$ terms in the previous summation. Therefore rank $A^{\circ i} \leq \binom{i+l-1}{l-1}$, and so (1) easily follows from (2). The proof that (1) $\implies$ (2) is much more challenging; see [71] for details.

In [70], the authors focus on the case where sparsity constraints are imposed to the matrices instead of rank constraints. Positive semidefinite matrices with zeros according to graphs arise naturally in many applications. For example, in the theory of Markov random fields in probability theory ([93, 140]), the nodes of a graph G represent components of a random vector, and edges represent the dependency structure between nodes. Thus, absence of an edge implies marginal or conditional independence between the corresponding random variables, and leads to zeros in the associated covariance or correlation matrix (or its inverse). Such models therefore yield parsimonious representations of dependency structures. Characterizing entrywise functions preserving positivity for matrices with zeros according to a graph is thus of tremendous interest for modern applications. Obtaining such characterizations is, however, much more involved than the original problem considered by Schoenberg as one has to enforce and maintain the sparsity constraint. The problem of characterizing functions preserving positivity for sparse matrices is also intimately linked to problems in spectral graph theory and many other problems (see e.g. [79, 1, 105, 31]).

---

[10]We also point out the second main result in *loc. cit.*, that is, [5, Theorem 2], which classifies all continuous entrywise maps $f : \mathbb{C} \to \mathbb{C}$ that obey similar rank constraints in all dimensions. Such maps are necessarily of the form $g(z) = \sum_{j=1}^{p} \beta_j z^{m_j} (\bar{z})^{n_j}$, where the exponents m_j and n_j are non-negative integers. This should immediately remind the reader of Rudin's conjecture in the 'dimension-free' case, and its resolution by Herz; see Theorem 3.2.

As before, for a given graph $G = (V, E)$ on the finite vertex set $V = \{1, \ldots, N\}$, we denote by $\mathcal{P}_G(I)$ the set of positive-semidefinite matrices with entries in I and zeros according to G, as in (6.1). Given a function $f : \mathbb{R} \to \mathbb{R}$ and $A \in \mathcal{S}_{|G|}(\mathbb{R})$, denote by $f_G[A]$ the matrix such that

$$f_G[A]_{jk} := \begin{cases} f(a_{jk}) & \text{if } (j, k) \in E \text{ or } j = k, \\ 0 & \text{otherwise.} \end{cases}$$

The first main result in [70] is an explicit characterization of the entrywise positive preservers of $\mathcal{P}_G$ for any collection of trees (other than copies of K_2). Following Vasudeva's classification for $\mathcal{P}_{K_2}$ in Theorem 4.1, trees are the only other graphs for which such a classification is currently known.

Theorem 7.10 (Guillot–Khare–Rajaratnam [70, Theorem A]). *Suppose $I = [0, R)$ for some $0 < R \leq \infty$, and $f : I \to \mathbb{R}_+$. Let G be a tree with at least 3 vertices, and let A_3 denote the path graph on 3 vertices. The following are equivalent.*

(1) $f_G[A] \in \mathcal{P}_G$ *for every* $A \in \mathcal{P}_G(I)$;
(2) $f_T[A] \in \mathcal{P}_T$ *for all trees* T *and all matrices* $A \in \mathcal{P}_T(I)$;
(3) $f_{A_3}[A] \in \mathcal{P}_{A_3}$ *for every* $A \in \mathcal{P}_{A_3}(I)$;
(4) *The function* f *satisfies*

$$f\left(\sqrt{xy}\right)^2 \leq f(x)f(y) \qquad \text{for all } x, y \in I \tag{7.4}$$

and is super-additive on I, *that is,*

$$f(x + y) \geq f(x) + f(y) \qquad \text{whenever } x, y, x + y \in I. \tag{7.5}$$

The implication (4) $\implies$ (1) was further extended to all chordal graphs: it is the following result with $c = 2$ and $d = 1$.

Theorem 7.11 (Guillot–Khare–Rajaratnam [69]). *Let G be a chordal graph with a perfect elimination ordering of its vertices $\{v_1, \ldots, v_n\}$. For all $1 \leq k \leq n$, denote by G_k the induced subgraph on G formed by $\{v_1, \ldots, v_k\}$, so that the neighbors of v_k in G_k form a clique. Define $c = \omega(G)$ to be the clique number of G, and let*

$$d := \max\{\deg_{G_k}(v_k) : k = 1, \ldots, n\}.$$

If $f : \mathbb{R} \to \mathbb{R}$ is any function such that $f[-]$ preserves positivity on $\mathcal{P}_c^1(\mathbb{R})$ and $f[M + N] \geq f[M] + f[N]$ for all $M \in \mathcal{P}_d$ and $N \in \mathcal{P}_d^1$, then $f[-]$ preserves positivity on $\mathcal{P}_G(\mathbb{R})$. [Here, $\mathcal{P}_d^1$ denotes the matrices in $\mathcal{P}_d$ of rank at most one.]

See [69] for other sufficient conditions for a general entrywise function to preserve positivity on $\mathcal{P}_G$ for G chordal.

To state the final result in this section, recall that Schoenberg's theorem (Theorem 2.12) shows that entrywise functions preserving positivity for all matrices (that is, according to the family of complete graphs K_n for $n \geq 1$)

are absolutely monotonic on the positive axis. It is not clear if functions satisfying (7.4) and (7.5) in Theorem 7.10 are necessarily absolutely monotonic, or even analytic. As shown in [70, Proposition 4.2], the critical exponent (see Definition 6.5) of every tree is 1. Hence, functions satisfying (7.4) and (7.5) do not need to be analytic. The second main result in [70] demonstrates that even if the function is analytic, it can in fact have arbitrarily long strings of negative Taylor coefficients.

Theorem 7.12 (Guillot–Khare–Rajaratnam [70, Theorem B]). *There exists an entire function $f(z) = \sum_{n=0}^{\infty} a_n z^n$ such that*

(1) $a_n \in [-1, 1]$ *for every $n \geq 0$;*
(2) *The sequence $(a_n)_{n \geq 0}$ contains arbitrarily long strings of negative numbers;*
(3) *For every tree G, $f_G[A] \in \mathcal{P}_G$ for every $A \in \mathcal{P}_G(\mathbb{R}_+)$.*

In particular, if $\Delta(G)$ denotes the maximum degree of the vertices of G, then there exists a family G_n of graphs and an entire function f that is not absolutely monotonic, such that

(1) $\sup_{n \geq 1} \Delta(G_n) = \infty$;
(2) $f_{G_n}[A] \in \mathcal{P}_{G_n}$ *for every $A \in \mathcal{P}_{G_n}(\mathbb{R}_+)$.*

References

[1] Jim Agler, J. William Helton, Scott McCullough, and Leiba Rodman. Positive semidefinite matrices with a given sparsity pattern. In *Proceedings of the Victoria Conference on Combinatorial Matrix Analysis (Victoria, BC, 1987)*, volume 107, pages 101–149, 1988.

[2] Naum Ilyich Akhiezer. *The classical moment problem and some related questions in analysis.* Translated by N. Kemmer. Hafner Publishing Co., New York, 1965.

[3] Theodore W. Anderson. *An introduction to multivariate statistical analysis.* Wiley Series in Probability and Statistics. Wiley-Interscience (John Wiley & Sons), Hoboken, third edition, 2003.

[4] Nima Arkani-Hamed, Jacob L. Bourjaily, Freddy Cachazo, Alexander B. Goncharov, Alexander Postnikov, and Jaroslav Trnka. Scattering amplitudes and the positive Grassmannian. *Preprint*, available at http://arxiv.org/abs/1212.5605, 2012.

[5] Aharon Atzmon and Allan Pinkus. Rank restricting functions. *Linear Algebra Appl.*, 372:305–323, 2003.

[6] Zhi Dong Bai and Li-Xin Zhang. Semicircle law for Hadamard products. *SIAM J. Matrix Anal. Appl.*, 29(2):473–495, 2007.

[7] Victor S. Barbosa and Valdir Antonio Menegatto. Strictly positive definite kernels on compact two-point homogeneous spaces. *Math. Inequal. Appl.*, 19(2):743–756, 2016.

[8] Victor S. Barbosa and Valdir Antonio Menegatto. Strict positive definiteness on products of compact two-point homogeneous spaces. *Integral Transforms Spec. Funct.*, 28(1):56–73, 2017.

[9] Rick K. Beatson and Wolfgang zu Castell. Dimension hopping and families of strictly positive definite zonal basis functions on spheres. *J. Approx. Theory*, 221(C):22–37, 2017.

[10] Rick K. Beatson, Wolfgang zu Castell, and Yuan Xu. A Pólya criterion for (strict) positive-definiteness on the sphere. *IMA J. Numer. Anal.*, 34(2):550–568, 2014.

[11] Alexander Belton, Dominique Guillot, Apoorva Khare, and Mihai Putinar. Matrix positivity preservers in fixed dimension. I. *Adv. Math.*, 298:325–368, 2016.

[12] Alexander Belton, Dominique Guillot, Apoorva Khare, and Mihai Putinar. Moment-sequence transforms. *Preprint*, available at http://arxiv.org/abs/1610.05740, 2016.

[13] Alexander Belton, Dominique Guillot, Apoorva Khare, and Mihai Putinar. Total-positivity preservers. *Preprint*, available at http://arxiv.org/abs/1711.10468, 2017.

[14] Christian Berg, Jens Peter Reus Christensen, and Paul Ressel. Positive definite functions on abelian semigroups. *Math. Ann.*, 223(3):253–274, 1976.

[15] Christian Berg, Ana Paula Peron, and Emilio Porcu. Schoenberg's theorem for real and complex Hilbert spheres revisited. *J. Approx. Theory*, 228:58–78, 2018.

[16] Christian Berg and Emilio Porcu. From Schoenberg coefficients to Schoenberg functions. *Constr. Approx.*, 45(2):217–241, 2017.

[17] Serge Bernstein. Sur les fonctions absolument monotones. *Acta Math.*, 52(1):1–66, 1929.

[18] Rajendra Bhatia and Ludwig Elsner. Positivity preserving Hadamard matrix functions. *Positivity*, 11(4):583–588, 2007.

[19] Peter J. Bickel and Elizaveta Levina. Covariance regularization by thresholding. *Ann. Statist.*, 36(6):2577–2604, 2008.

[20] Martin Bilodeau and David Brenner. *Theory of multivariate statistics*. Springer Texts in Statistics. Springer-Verlag, New York, 1999.

[21] Garrett Birkhoff. Metric foundations of geometry. I. *Trans. Amer. Math. Soc.*, 55:465–492, 1944.

[22] Grigoriy Blekherman, Pablo A. Parrilo, and Rekha R. Thomas, editors. *Semidefinite optimization and convex algebraic geometry*, volume 13 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM); Mathematical Optimization Society, Philadelphia, 2013.

[23] Leonard Mascot Blumenthal. *Theory and applications of distance geometry*. Second edition. Chelsea Publishing Co., New York, 1970.

[24] Ralph P. Boas, Jr. and David Vernon Widder. Functions with positive differences. *Duke Math. J.*, 7(1):496–503, 1940.

[25] Salomon Bochner. Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse. *Math. Ann.*, 108(1):378–410, 1933.

[26] Salomon Bochner. Hilbert distances and positive definite functions. *Ann. of Math. (2)*, 42(3):647–656, 1941.

[27] Salomon Bochner. *Harmonic analysis and the theory of probability*. University of California Press, Berkeley, 1955.

[28] Rafaela Neves Bonfim, Jean Carlo Guella, and Valdir Antonio Menegatto. Strictly positive definite functions on compact two-point homogeneous spaces: the product alternative. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 14:112, 14 pages, 2018.

[29] Francesco Brenti. Unimodal, log-concave and Pólya frequency sequences in combinatorics. *Mem. Amer. Math. Soc.*, 81(413):viii+106, 1989.

[30] David Marius Bressoud and Shi-Yuan Wei. Combinatorial equivalence of definitions of the Schur function. In *A tribute to Emil Grosswald: number theory and related analysis*, volume 143 of *Contemp. Math.*, pages 59–64. American Mathematical Society, Providence, 1993.

[31] Richard A. Brualdi. *The mutually beneficial relationship of graphs and matrices*, volume 115. American Mathematical Society, Providence, 2011.

[32] Debao Chen, Valdir Antonio Menegatto, and Xingping Sun. A necessary and sufficient condition for strictly positive definite functions on spheres. *Proc. Amer. Math. Soc.*, 131(9):2733–2740, 2003.

[33] Jens Peter Reus Christensen and Paul Ressel. Functions operating on positive definite matrices and a theorem of Schoenberg. *Trans. Amer. Math. Soc.*, 243:89–95, 1978.

[34] Jens Peter Reus Christensen and Paul Ressel. Positive definite kernels on the complex Hilbert sphere. *Math. Z.*, 180(2):193–201, 1982.

[35] Henry Cohn and Matthew de Courcy-Ireland. The Gaussian core model in high dimensions. *Duke Math. J.*, 167(13):2417–2455, 2018.

[36] Henry Cohn and Abhinav Kumar. Universally optimal distribution of points on spheres. *J. Amer. Math. Soc.*, 20(1):99–148, 2007.

[37] Henry Cohn and Jeechul Woo. Three-point bounds for energy minimization. *J. Amer. Math. Soc.*, 25(4):929–958, 2012.

[38] Henry Cohn and Yufei Zhao. Sphere packing bounds via spherical codes. *Duke Math. J.*, 163(10):1965–2002, 2014.

[39] Jon Dattorro. Equality relating Euclidean distance cone to positive semidefinite cone. *Linear Algebra Appl.*, 428(11-12):2597–2600, 2008.

[40] Charles Lutwidge Dodgson. Condensation of determinants, being a new and brief method for computing their arithmetical values. *Proc. R. Soc. Lond.*, 15:150–155, 1866.

[41] David L. Donoho. High-dimensional data analysis: The curses and blessings of dimensionality. *AMS Math Challenges Lecture*, 1:32 pp., 2000.

[42] David L. Donoho and Iain M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.

[43] Bradley Efron. Increasing properties of Pólya frequency functions. *Ann. Math. Statist.*, 36(1):272–279, 1965.

[44] Noureddine El Karoui. Operator norm consistent estimation of large-dimensional sparse covariance matrices. *Ann. Statist.*, 36(6):2717–2756, 2008.

[45] Jan Emonds and Hartmut Führ. Strictly positive definite functions on compact abelian groups. *Proc. Amer. Math. Soc.*, 139(3):1105–1113, 2011.

[46] Shaun M. Fallat and Charles R. Johnson. *Totally nonnegative matrices*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, 2011.

[47] Shaun M. Fallat, Charles R. Johnson, and Ronald L. Smith. The general totally positive matrix completion problem with few unspecified entries. *Electron. J. Linear Algebra*, 7:1–20, 2000.

[48] Shaun M. Fallat, Charles R. Johnson, and Alan D. Sokal. Total positivity of sums, Hadamard products and Hadamard powers: results and counterexamples. *Linear Algebra Appl.*, 520:242–259, 2017.

[49] Jianqing Fan, Yuan Liao, and Han Liu. An overview of the estimation of large covariance and precision matrices. *The Econometrics Journal*, 19(1):C1–C32, 2016.

[50] Michael Fekete. Über ein problem von Laguerre. *Rend. Circ. Math. Palermo*, 34:89–120, 1912.

[51] Carl H. FitzGerald and Roger A. Horn. On fractional Hadamard powers of positive definite matrices. *J. Math. Anal. Appl.*, 61(3):633–642, 1977.

[52] Carl H. FitzGerald, Charles A. Micchelli, and Allan Pinkus. Functions that preserve families of positive semidefinite matrices. *Linear Algebra Appl.*, 221:83–102, 1995.

[53] Sergey Fomin and Andrei Zelevinsky. Total positivity: tests and parametrizations. *Math. Intelligencer*, 22(1):23–33, 2000.

[54] Sergey Fomin and Andrei Zelevinsky. Cluster algebras. I. Foundations. *J. Amer. Math. Soc.*, 15(2):497–529, 2002.

[55] Ferdinand Georg Frobenius. Über die elliptischen Funktionen zweiter Art. *J. reine angew. Math.*, 93:53–68, 1882.

[56] Ramesh Gangolli. Positive definite kernels on homogeneous spaces and certain stochastic processes related to Lévy's Brownian motion of several parameters. *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 3:121–226, 1967.

[57] Feliks Ruvimovich Gantmacher. *The theory of matrices. Vols. 1, 2.* Translated by K.A. Hirsch. Chelsea Publishing Co., New York, 1959.

[58] Feliks Ruvimovich Gantmacher and Mark Grigor'evich Krein. *Oscillation matrices and kernels and small vibrations of mechanical systems.* AMS Chelsea Publishing, Providence, revised edition, 2002. Edited and with a preface by Alex Eremenko.

[59] Feliks Ruvimovich Gantmacher and Mark Grigorievich Krein. Sur les matrices complètement non négatives et oscillatoires. *Compos. Math.*, 4:445–476, 1937.

[60] Mariano Gasca and Charles A. Micchelli, editors. *Total Positivity and Its Applications*, volume 359 of *Mathematics and Its Applications*. Springer, Utrecht, 1996.

[61] Israel Moiseevich Gelfand. Normierte Ringe. *Rec. Math. [Mat. Sbornik] N. S.*, 9(51):3–24, 1941.

[62] Tillmann Gneiting. Strictly and non-strictly positive definite functions on spheres. *Bernoulli*, 19(4):1327–1349, 2013.

[63] Jean Carlo Guella and Valdir Antonio Menegatto. Strictly positive definite kernels on a product of spheres. *J. Math. Anal. Appl.*, 435(1):286–301, 2016.

[64] Jean Carlo Guella and Valdir Antonio Menegatto. Strictly positive definite kernels on the torus. *Constr. Approx.*, 46(2):271–284, 2017.

[65] Jean Carlo Guella, Valdir Antonio Menegatto, and Ana Paula Peron. An extension of a theorem of Schoenberg to products of spheres. *Banach J. Math. Anal.*, 10(4):671–685, 2016.

[66] Jean Carlo Guella, Valdir Antonio Menegatto, and Ana Paula Peron. Strictly positive definite kernels on a product of spheres II. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 12:Paper No. 103, 15, 2016.

[67] Jean Carlo Guella, Valdir Antonio Menegatto, and Ana Paula Peron. Strictly positive definite kernels on a product of circles. *Positivity*, 21(1):329–342, 2017.

[68] Dominique Guillot, Apoorva Khare, and Bala Rajaratnam. Complete characterization of Hadamard powers preserving Loewner positivity, monotonicity, and convexity. *J. Math. Anal. Appl.*, 425(1):489–507, 2015.

[69] Dominique Guillot, Apoorva Khare, and Bala Rajaratnam. Critical exponents of graphs. *J. Combin. Theory Ser. A*, 139:30–58, 2016.

[70] Dominique Guillot, Apoorva Khare, and Bala Rajaratnam. Preserving positivity for matrices with sparsity constraints. *Trans. Amer. Math. Soc.*, 368(12):8929–8953, 2016.

[71] Dominique Guillot, Apoorva Khare, and Bala Rajaratnam. Preserving positivity for rank-constrained matrices. *Trans. Amer. Math. Soc.*, 369(9):6105–6145, 2017.

[72] Dominique Guillot and Bala Rajaratnam. Retaining positive definiteness in thresholded matrices. *Linear Algebra Appl.*, 436(11):4143–4160, 2012.

[73] Dominique Guillot and Bala Rajaratnam. Functions preserving positive definiteness for sparse matrices. *Trans. Amer. Math. Soc.*, 367(1):627–649, 2015.

[74] Henry Helson, Jean-Pierre Kahane, Yitzhak Katznelson, and Walter Rudin. The functions which operate on Fourier transforms. *Acta Math.*, 102(1-2):135–157, 1959.

[75] Alfred Hero and Bala Rajaratnam. Large-scale correlation screening. *J. Amer. Statist. Assoc.*, 106(496):1540–1552, 2011.

[76] Alfred Hero and Bala Rajaratnam. Hub discovery in partial correlation graphs. *IEEE Trans. Inform. Theory*, 58(9):6064–6078, 2012.

[77] Carl S. Herz. Fonctions opérant sur les fonctions définies-positives. *Ann. Inst. Fourier (Grenoble)*, 13(1):161–180, 1963.

[78] Fumio Hiai. Monotonicity for entrywise functions of matrices. *Linear Algebra Appl.*, 431(8):1125–1146, 2009.

[79] Leslie Hogben. Spectral graph theory and the inverse eigenvalue problem of a graph. *Electron. J. Linear Algebra*, 14:12–31 (electronic), 2005.

[80] Roger A. Horn. The theory of infinitely divisible matrices and kernels. *Trans. Amer. Math. Soc.*, 136:269–286, 1969.

[81] Tanvi Jain. Hadamard powers of some positive matrices. *Linear Algebra Appl.*, 528:147–158, 2017.

[82] Graham J. O. Jameson. Counting zeros of generalised polynomials: Descartes' rule of signs and Laguerre's extensions. *The Mathematical Gazette*, 90(518):223–234, 2006.

[83] Jean-Pierre Kahane. Sur un théorème de Wiener–Lévy. *C. R. Acad. Sci. Paris*, 246:1949–1951, 1958.

[84] Jean-Pierre Kahane and Walter Rudin. Caractérisation des fonctions qui opèrent sur les coefficients de Fourier-Stieltjes. *C. R. Acad. Sci. Paris*, 247:773–775, 1958.

[85] Samuel Karlin. *Total positivity. Vol. I*. Stanford University Press, Stanford, 1968.

[86] Samuel Karlin and Yosef Rinott. A generalized Cauchy–Binet formula and applications to total positivity and majorization. *J. Multivariate Anal.*, 27(1):284–299, 1988.

[87] Yitzhak Katznelson. Sur les fonctions opérant sur l'algèbre des séries de Fourier absolument convergentes. *C. R. Acad. Sci. Paris*, 247:404–406, 1958.

[88] Apoorva Khare. Smooth entrywise positivity preservers, a Horn–Loewner master theorem, and Schur polynomials. *Preprint*, available at http://arxiv.org/abs/1809.01823, 2018.

[89] Apoorva Khare and Terence Tao. On the sign patterns of entrywise positivity preservers in fixed dimension. *Preprint*, available at http://arxiv.org/abs/1708.05197, 2017.

[90] Apoorva Khare and Terence Tao. Schur polynomials, entrywise positivity preservers, and weak majorization. *Sem. Lothar. Combin.*, 80B:#14 (12 pp.), 2018.

[91] Yuji Kodama and Lauren Williams. KP solitons and total positivity for the Grassmannian. *Invent. Math.*, 198(3):637–699, 2014.

[92] Thomas Lam, Alexander Postnikov, and Pavlo Pylyavskyy. Schur positivity and Schur log-concavity. *Amer. J. Math.*, 129(6):1611–1622, 2007.

[93] Steffen L. Lauritzen. *Graphical models*, volume 17 of *Oxford Statistical Science Series*. The Clarendon Press Oxford University Press, New York, 1996. Oxford Science Publications.

[94] Ai Li and Steve Horvath. Network neighborhood analysis with the multi-node topological overlap measure. *Bioinformatics*, 23(2):222–231, 2007.

[95] Leo Liberti, Carlile Lavor, Nelson Maculan, and Antonio Mucherino. Euclidean distance geometry and applications. *SIAM Rev.*, 56(1):3–69, 2014.

[96] Lee Lorch and Donald Joseph Newman. On the composition of completely monotonic functions and completely monotonic sequences and related questions. *J. London Math. Soc. (2)*, 28(1):31–45, 1983.

[97] George Lusztig. Introduction to total positivity. In Joachim Hilgert, Jimmie D. Lawson, Karl-Hermann Neeb, and Ernest B. Vinberg, editors, *Positivity in Lie theory: open problems*, volume 26 of *De Gruyter Expositions in Mathematics*, pages 133–145. Walter de Gruyter & Co., Berlin, 1998.

[98] Ian G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, second edition, 1995. With contributions by A. Zelevinsky, Oxford Science Publications.

[99] Karl Menger. Untersuchungen über allgemeine Metrik. *Math. Ann.*, 100(1):75–163, 1928.

[100] Karl Menger. New Foundation of Euclidean Geometry. *Amer. J. Math.*, 53(4):721–745, 1931.

[101] Jesper Møller, Morten Nielsen, Emilio Porcu, and Ege Rubak. Determinantal point process models on the sphere. *Bernoulli*, 24(2):1171–1201, 2018.

[102] Oleg R. Musin. The kissing number in four dimensions. *Ann. of Math. (2)*, 168(1):1–32, 2008.

[103] Oleg R. Musin. Multivariate positive definite functions on spheres. In *Discrete geometry and algebraic combinatorics*, volume 625 of *Contemp. Math.*, pages 177–190. American Mathematical Society, Providence, 2014.

[104] Arkadi Nemirovski. Advances in convex optimization: conic programming. In *International Congress of Mathematicians. Vol. I*, pages 413–444. European Mathematical Society, Zürich, 2007.

[105] Allan Pinkus. Strictly positive definite functions on a real inner product space. *Adv. Comput. Math.*, 20(4):263–271, 2004.

[106] Allan Pinkus. *Totally positive matrices*, volume 181 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2010.

[107] Georg Pólya and Gabor Szegő. *Aufgaben und Lehrsätze aus der Analysis. Band II: Funktionentheorie, Nullstellen, Polynome Determinanten, Zahlentheorie*. Springer-Verlag, Berlin, 1971.

[108] Emilio Porcu, Alfredo Alegria, and Reinhard Furrer. Modeling temporally evolving and spatially globally dependent data. *Internat. Statist. Review*, 86(2), 2018.

[109] Emilio Porcu, Moreno Bevilacqua, and Marc G. Genton. Spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere. *J. Amer. Statist. Assoc.*, 111(514):888–898, 2016.

[110] Alexander Postnikov, David Speyer, and Lauren Williams. Matching polytopes, toric geometry, and the totally non-negative Grassmannian. *J. Algebraic Combin.*, 30(2):173–191, 2009.

[111] Mohsen Pourahmadi. *High-dimensional covariance estimation: with high-dimensional data*, volume 882. John Wiley & Sons, 2013.

[112] Mihai Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.*, 42(3):969–984, 1993.

[113] A. Wayne Roberts and Dale E. Varberg. *Convex functions*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York–London, 1973. Pure and Applied Mathematics, Vol. 57.

[114] Adam J. Rothman, Elizaveta Levina, and Ji Zhu. Generalized thresholding of large covariance matrices. *J. Amer. Statist. Assoc.*, 104(485):177–186, 2009.

[115] Walter Rudin. Transformations des coefficients de Fourier. *C. R. Acad. Sci. Paris*, 243:638–640, 1956.

[116] Walter Rudin. Positive definite sequences and absolutely monotonic functions. *Duke Math. J.*, 26(4):617–622, 1959.

[117] Walter Rudin. Some theorems on Fourier coefficients. *Proc. Amer. Math. Soc.*, 10(6):855–859, 1959.

[118] Walter Rudin. A strong converse of the Wiener-Levy theorem. *Canad. J. Math.*, 14(4):694–701, 1962.

[119] Isaac Jacob Schoenberg. Über variationsvermindernde lineare Transformationen. *Math. Z.*, 32(1):321–328, 1930.

[120] Isaac Jacob Schoenberg. Remarks to Maurice Fréchet's article "Sur la définition axiomatique d'une classe d'espace distanciés vectoriellement applicable sur l'espace de Hilbert". *Ann. of Math. (2)*, 36(3):724–732, 1935.

[121] Isaac Jacob Schoenberg. On certain metric spaces arising from Euclidean spaces by a change of metric and their imbedding in Hilbert space. *Ann. of Math. (2)*, 38(4):787–793, 1937.

[122] Isaac Jacob Schoenberg. Metric spaces and completely monotone functions. *Ann. of Math. (2)*, 39(4):811–841, 1938.

[123] Isaac Jacob Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.*, 44(3):522–536, 1938.

[124] Isaac Jacob Schoenberg. On metric arcs of vanishing Menger curvature. *Ann. of Math. (2)*, 41(4):715–726, 1940.

[125] Isaac Jacob Schoenberg. Positive definite functions on spheres. *Duke Math. J.*, 9(1):96–108, 1942.

[126] Isaac Jacob Schoenberg. On Pólya frequency functions. II. Variation-diminishing integral operators of the convolution type. *Acta Sci. Math. Szeged*, 12:97–106, 1950.

[127] Isaac Jacob Schoenberg. On Pólya frequency functions. I. The totally positive functions and their Laplace transforms. *J. Analyse Math.*, 1(1):331–374, 1951.

[128] Isaac Jacob Schoenberg and Anne M. Whitney. On Pólya frequency functions. III. The positivity of translation determinants with an application to the interpolation problem by spline curves. *Trans. Amer. Math. Soc.*, 74(2):246–259, 1953.

[129] Issai Schur. Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen. *J. reine angew. Math.*, 140:1–28, 1911.

[130] Ronald L. Smith. The positive definite completion problem revisited. *Linear Algebra Appl.*, 429(7):1442–1452, 2008.

[131] Ingo Steinwart. On the influence of the kernel on the consistency of support vector machines. *J. Mach. Learn. Res.*, 2(1):67–93, 2002.

[132] James Stewart. Positive definite functions and generalizations, an historical survey. *Rocky Mountain J. Math.*, 6(3):409–434, 1976.

[133] Vladimir Naumovich Vapnik. *The nature of statistical learning theory*. Statistics for Engineering and Information Science. Springer, New York, second edition, 2000.

[134] Harkrishan Lal Vasudeva. Positive definite matrices and absolutely monotonic functions. *Indian J. Pure Appl. Math.*, 10(7):854–858, 1979.

[135] Cynthia Vinzant. What is ... a spectrahedron? *Notices Amer. Math. Soc.*, 61(5):492–494, 2014.

[136] John von Neumann and Isaac Jacob Schoenberg. Fourier integrals and metric geometry. *Trans. Amer. Math. Soc.*, 50(2):226–251, 1941.

[137] Hsien-Chung Wang. Two-point homogeneous spaces. *Ann. of Math. (2)*, 55(1):177–191, 1952.

[138] Hermann Weyl. Harmonics on homogeneous manifolds. *Ann. of Math. (2)*, 35(3):486–499, 1934.

[139] Anne M. Whitney. A reduction theorem for totally positive matrices. *J. Analyse Math.*, 2(1):88–92, 1952.

[140] Joe Whittaker. *Graphical models in applied multivariate statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Ltd., Chichester, 1990.

[141] Yuan Xu. Positive definite functions on the unit sphere and integrals of Jacobi polynomials. *Proc. Amer. Math. Soc.*, 146(5):2039–2048, 2018.

[142] Yuan Xu and Elliott Ward Cheney. Strictly positive definite functions on spheres. *Proc. Amer. Math. Soc.*, 116(4):977–981, 1992.

[143] Bin Zhang and Steve Horvath. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.*, 4:Art. 17, 45 pp. (electronic), 2005.

[144] Johanna Ziegel. Convolution roots and differentiability of isotropic positive definite functions on spheres. *Proc. Amer. Math. Soc.*, 142(6):2063–2077, 2014.

(A. Belton) Department of Mathematics and Statistics, Lancaster University, Lancaster, UK
E-mail address: `a.belton@lancaster.ac.uk`

(D. Guillot) University of Delaware, Newark, DE, USA
E-mail address: `dguillot@udel.edu`

(A. Khare) Department of Mathematics, Indian Institute of Science; Analysis and Probability Research Group; Bangalore, India
E-mail address: `khare@iisc.ac.in`

(M. Putinar) University of California at Santa Barbara, CA, USA and Newcastle University, Newcastle upon Tyne, UK
E-mail address: `mputinar@math.ucsb.edu, mihai.putinar@ncl.ac.uk`