Invariant kernels on the space of complex covariance matrices

Cyrus Mostajeran and Salem Said

Abstract

The present work sets forth the analytical tools which make it possible to construct and compute invariant kernels on the space of complex covariance matrices. The main result is the L^1-Godement theorem, which is used to obtain a general analytical expression that yields any invariant kernel which is, in a certain natural sense, also integrable. Using this expression, one can design and explore new families of invariant kernels, all while incurring a rather moderate computational effort. The expression comes in the form of a determinant (it is a determinantal expression), and is derived from the notion of spherical transform, which arises when the space of complex covariance matrices is considered as a Riemannian symmetric space.

Index Terms

kernel method, symmetric space, covariance matrix, Bochner's theorem, Godement's theorem, spherical function

I. Introduction

In order to successfully bring kernel learning methods to bear on data which belong to non-Euclidean spaces, it seems necessary to meet two objectives.

First, the ability to construct and compute sufficiently rich families of kernels. Preferably, one should not be limited to a handful of "canonical" kernels (such as the heat kernels and its various cousins), but should be able to design and explore new families of kernels, at will. In addition, evaluation of these new kernels should be numerically straightforward.

Second, it should be a requirement that these kernels have to be compatible with the fundamental geometric properties of the space at hand. For example, if the data belong to a symmetric (or just a homogeneous) space, then the kernels have to be invariant.

The present work sets out to achieve these two objectives, for the case of data which belong to the space of complex covariance matrices.

Section II provides general background used to formulate the main results. It identifies the space M of complex covariance matrices as a Riemannian symmetric space under the action of the group G of invertible complex matrices. The main objects of study being the invariant kernels $\mathcal{K} : M \times M \to \mathbb{C}$ (invariant means under the action of G), Proposition 1 shows that there is a straightforward equivalence between these invariant kernels and U-invariant positive definite functions $f : M \to \mathbb{C}$ (U-invariant means under the action of the group of unitary matrices, a subgroup of G). Thus, functions $\mathcal{K}$ of two variables are replaced with functions f of one variable.

Section III is concerned with the spherical transform. This provides an expansion of any integrable U-invariant function, in terms of so-called spherical functions (integrable means with respect to the Riemannian volume of M) [1] [2]. Spherical functions are eigenfunctions of the Laplace-Beltrami operator which arises from the Riemannian geometry of M. These functions generalise the Schur polynomials, which play a key role across group theory, multivariate statistics and random matrix theory [3] [4] [5]. The spherical transform of $f : M \to \mathbb{C}$ is a symmetric function $\hat{f} : \mathbb{R}^N \to \mathbb{C}$ (M is the space of $N \times N$ complex covariance matrices). To say that $\hat{f}$ is symmetric means $\hat{f}$ remains unchanged after any permutation of its N real arguments.

Proposition 2 shows that the spherical functions on M admit an explicit determinantal expression (Formula (9), here called the Gelfand-Naimark formula). Proposition 3 uses this expression to write down the spherical transform, as well as its inverse transform, in the form of a multivariate integral. A worked example is then presented, which shows how, in certain cases, the analytical expression of these integrals can be found using the Andréief identity, widely used in random matrix theory [6].

Section IV opens with Theorem 1, the L^1−Godement theorem. A somewhat more restricted statement of this theorem was given in [7]. It is here extended to the case at hand, of the space of complex covariance matrices. As explained in [7], the L^1−Godement theorem is based on the celebrated Godement theorem, which is a genralisation of the classical Bochner's theorem to the context of symmetric spaces [8].

According to Theorem 1, U-invariant positive definite functions can be obtained by taking inverse spherical transforms of positive symmetric functions. These inverse spherical transforms are evaluated in Proposition 4. As a result, one can generate new U-invariant positive definite functions, simply by choosing suitable positive symmetric functions and plugging them into Proposition 4. In fact, Theorem 1 sets up a bijective correspondence between the set of integrable U-invariant positive definite functions, and a certain set of positive symmetric functions. In turn, Proposition 4 gives a general analytical expression for any integrable U-invariant positive definite function. In this sense (via Proposition 1), it gives the general expression of any integrable invariant kernel.

Section IV closes with two worked examples, which showcase how Proposition 4 can be used in order to design and explore new families of invariant kernels, and also to recover any old ones (*e.g.* the heat kernel).

Section V provides final remarks, summarising the main results of the present work, and indicating their possible future extensions. In particular, it explains how the approach based on Theorem 1 and Proposition 4 can still be applied to other spaces of covariance matrices, not just complex, but possibly real, quaternion, or even block-Toeplitz. It also explains the general mathematical context which makes this approach feasible, in the first place. An explicit determinantal expression (similar to the Gelfand-Naimark formula (9)) for the spherical functions is readily available on any symmetric space whose group of isometries (here, the group G of invertible complex matrices) is a complex Lie group. Symmetric spaces with this property are called symmetric spaces of type IV [9]. The symmetric spaces of type IV are precisely (a) spaces of complex covariance matrices (b) spaces of complex covariance matrices which are also orthogonal (c) spaces of complex covariance matrices which are also symplectic (d) certain other so-called exceptional spaces. These are the non-compact analogues (the technical word is duals) of the compact Lie groups (unitary, orthogonal, symplectic, and exceptional).

II. General background

A. Positive definite functions

Denote by M the space of $N \times N$ complex covariance matrices. Specifically, these are $N \times N$ Hermitian positive definite matrices. Moreover, denote by G the group of $N \times N$ invertible complex matrices, and by U the group of $N \times N$ unitary matrices, a subgroup of G.

Recall that G acts transitively on M in the following way: $g \cdot x = gxg^\dagger$ for $g \in G$ and $x \in M$ (where $\dagger$ denotes the conjugate-transpose) [10] [11]. For this action, U is the stabiliser of the identity matrix $\mathrm{id} \in M$. In other words, $g \cdot \mathrm{id} = \mathrm{id}$ if and only if $g \in U$.

A kernel $\mathcal{K}$ is a continuous function $\mathcal{K} : M \times M \to \mathbb{C}$, such that for any $x_1, \ldots, x_n \in M$ (here, $n = 2, 3, \ldots$), the $n \times n$ matrix with elements $\mathcal{K}(x_i, x_j)$ is Hermitian positive semidefinite. The focus of the present work is on invariant kernels. These are kernels which satisfy $\mathcal{K}(g \cdot x, g \cdot y) = \mathcal{K}(x, y)$ for all $g \in G$ and all $x, y \in M$ [7] [12].

It is convenient to study invariant kernels indirectly, by studying U-invariant positive definite functions [7]. A function $f : M \to \mathbb{C}$ is called U-invariant if $f(u \cdot x) = f(x)$ for all $u \in U$ and $x \in M$. This means that $f(x) = f_o(\rho)$ where f_o is a symmetric function and $\rho = (\rho_1, \ldots, \rho_N)$ are the eigenvalues of x: $f_o(\rho)$ remains unchanged after any permutation of $(\rho_1, \ldots, \rho_N)$.

If f is continuous, then it is called positive definite if, for any $x_1, \ldots, x_n \in M$ (where $n = 2, 3, \ldots$), the $n \times n$ matrix with elements $f(x_i^{-1/2} x_j x_i^{-1/2})$ is Hermitian positive semidefinite.

The two concepts (invariant kernel and U-invariant positive definite function) are equivalent.

Proposition 1: If $\mathcal{K}$ is an invariant kernel, then the function $f(x) = \mathcal{K}(x, \mathrm{id})$ is U-invariant and positive definite. If f is a U-invariant and positive definite function, $\mathcal{K}(x, y) = f(y^{-1/2} x y^{-1/2})$ defines an invariant kernel.

Proof: let $f(x) = \mathcal{K}(x, \mathrm{id})$. If $u \in U$, then $f(u \cdot x) = \mathcal{K}(u \cdot x, \mathrm{id}) = \mathcal{K}(u \cdot x, u \cdot \mathrm{id})$ because $u \cdot \mathrm{id} = \mathrm{id}$. However, if $\mathcal{K}$ is invariant, then $\mathcal{K}(u \cdot x, u \cdot \mathrm{id}) = \mathcal{K}(x, \mathrm{id})$. Therefore, $f(u \cdot x) = f(x)$ and f is U-invariant. To see that f is positive definite, it is enough to note that

$$f(y^{-1/2} x y^{-1/2}) = \mathcal{K}(y^{-1/2} \cdot x, \mathrm{id}) = \mathcal{K}(x, y^{1/2} \cdot \mathrm{id}) = \mathcal{K}(x, y)$$

where the first equality follows from the definition of $g \cdot x$, by taking $g = y^{-1/2}$, and the second equality because $\mathcal{K}$ is invariant. Thus, for any $x_1, \ldots, x_n \in M$, the matrix with elements $f(x_i^{-1/2} x_j x_i^{-1/2})$ is the same as the matrix with elements $\mathcal{K}(x_j, x_i)$, which is positive semidefinite because $\mathcal{K}$ is a kernel. In addition, continuity of f follows from continuity of $\mathcal{K}$, and this ensures f is positive definite.

Conversely, let f be U-invariant and positive definite. This clearly implies that $\mathcal{K}(x, y) = f(y^{-1/2} x y^{-1/2})$ is a kernel. To see that this $\mathcal{K}$ is invariant, note that $\mathcal{K}(x, y) = f_o(\rho)$ where $\rho = (\rho_1, \ldots, \rho_N)$ are the eigenvalues of $y^{-1/2} x y^{-1/2}$. These are the same as the eigenvalues of $y^{-1} x$ because the two matrices are similar: $y^{-1} x = y^{-1/2} (y^{-1/2} x y^{-1/2}) y^{1/2}$. By the same argument, $\mathcal{K}(g \cdot x, g \cdot y) = f_o(\rho')$ where $\rho' = (\rho'_1, \ldots, \rho'_N)$ are the eigenvalues of $(g \cdot y)^{-1}(g \cdot x)$. However, this last matrix is similar to $y^{-1} x$,

$$(g \cdot y)^{-1}(g \cdot x) = (g^\dagger)^{-1}(y^{-1} x)(g^\dagger)$$

Therefore, $\rho' = \rho$ and $\mathcal{K}(g \cdot x, g \cdot y) = \mathcal{K}(x, y)$, as required. ∎

B. Riemannian geometry

An explicit description of U-invariant positive definite functions relies on the Riemannian geometry of M. Precisely, it relies on the fact that M is a Riemannian symmetric space [9] [2].

Note that M is an open subset of H, the real vector space of $N \times N$ Hermitian matrices. Therefore, M is a differentiable manifold with its tangent space at any $x \in M$ naturally isomorphic to H. Now, with this in mind, consider the Riemannian metric on M,

$$\langle v, w\rangle_x = \mathrm{Re}\left[\mathrm{tr}(x^{-1}vx^{-1}w)\right] \qquad v, w \in H \tag{1}$$

where Re denotes the real part and tr the trace. This is the affine-invariant metric, which was made popular by [13].

For any $g \in G$, the map $x \mapsto g \cdot x$ is an isometry of the metric (1). The same is true for inversion $x \mapsto x^{-1}$. These two facts together show that M satisfies the definition of a Riemannian symmetric space (for details, see [11]).

The class of U-invariant functions behaves in a special way with respect to the metric (1). For example, let vol denote the Riemannian volume element of this metric. If $f : M \to \mathbb{C}$ is an integrable U-invariant function [11],

$$\int_M f(x)\mathrm{vol}(dx) = \frac{C_N}{N!}\int_{\mathbb{R}_+^N} f_o(\rho)(V(\rho))^2 \prod_{k=1}^{N} \rho_k^{-N} d\rho_k \tag{2}$$

where $f(x) = f_o(\rho)$ is a symmetric function of the eigenvalues $(\rho_1, \ldots, \rho_N)$ of x, and where V stands for the Vandermonde polynomial. Here, and throughout the following, C_N denotes a positive constant that only depends on N and whose value is allowed to differ from one formula to another.

Moreover, if L is the Laplace-Beltrami operator of the metric (1), and f is smooth and U-invariant, then [1] [2],

$$Lf = \sum_{k=1}^{N} \rho_k^2 \frac{\partial^2 f_o}{\partial \rho_k^2} + 2\sum_{k<\ell} \frac{\rho_k \rho_\ell}{\rho_k - \rho_\ell}\left(\frac{\partial f_o}{\partial \rho_k} - \frac{\partial f_o}{\partial \rho_\ell}\right) + N\sum_{k=1}^{N} \rho_k \frac{\partial f_o}{\partial \rho_k} \tag{3}$$

Formulas (2) and (3) arise systematically from the Riemannian geometry of M, but they are also familiar in certain problems of multivariate statistics and random matrix theory [14] [3] [4] (this is further discussed in Section III).

To close the present section, consider a special case of the integral formula (2). Assume that $f_o(\rho)$ factors into $f_o(\rho) = w(\rho_1)\ldots w(\rho_N)$ where w is an integrable function such that

$$\int_0^\infty |w(\rho)|\rho^N d\rho < \infty$$

Then, the volume integral in (2) is convergent and admits a determinantal expression

$$\int_M f(x)\mathrm{vol}(dx) = C_N \det\left[\int_0^\infty w(\rho)\rho^{k+\ell-N}d\rho\right]_{k,\ell=0}^{N-1} \tag{4}$$

This is an application of the Andréief identity, widely used in random matrix theory, and was pointed out in [15].

Example : the expression (4) can be used to compute the Gaussian integral of [16] [17]

$$Z(\sigma) = \int_M \exp\left[-\frac{d^2(x, \mathrm{id})}{2\sigma^2}\right]\mathrm{vol}(dx) \tag{5}$$

where $d(\cdot,\cdot)$ denotes the Riemannian distance induced on M by the metric (1). In (4), this integral corresponds to $w(\rho) = \exp[-\log^2(\rho)/2\sigma^2]$, which yields

$$Z(\sigma) = C_N \det\left[\sigma e^{(\sigma^2/2)(k+\ell-N+1)^2}\right]_{k,\ell=0}^{N-1} \tag{6}$$

a formula due to [15].

III. THE SPHERICAL TRANSFORM

The key ingredient which will be employed in constructing and computing U-invariant positive definite functions is the spherical transform. Roughly, this provides an expansion of any well-behaved U-invariant function, in terms of eigenfunctions of the Laplace-Beltrami operator (3), which are known as spherical functions [1] [2].

The set of all spherical functions is described as follows [2]. Consider first the power function, $\Delta_s : M \to \mathbb{C}$ where $s = (s_1, \ldots, s_N)$ belongs to $\mathbb{C}^N$. This is

$$\Delta_s(x) = (\Delta_1(x))^{s_1 - s_2}(\Delta_2(x))^{s_2 - s_3} \ldots (\Delta_N(x))^{s_N} \tag{7}$$

where $\Delta_k(x)$ is the k-th leading principal minor of $x \in M$. A spherical function is a function of the form

$$\Phi_\lambda(x) = \int_U \Delta_{\lambda+\delta}(u \cdot x)\, du \tag{8}$$

where $\lambda \in \mathbb{C}^N$ and $\delta_k = \frac{1}{2}(2k - N - 1)$, while du denotes the normalised Haar measure on the unitary group U. Two functions Φ_λ and $\Phi_{\lambda'}$ are identical if and only if $(\lambda_1, \ldots, \lambda_N)$ is a permutation of $(\lambda'_1, \ldots, \lambda'_N)$.

Each Φ_λ is U-invariant and an eigenfunction of the Laplace-Beltrami operator (3), with eigenvalue $(\lambda, \lambda) - (\delta, \delta)$. Here, $(\mu, \mu) = \sum_{k=1}^N \mu_k^2$ for $\mu \in \mathbb{C}^N$ [2] (Theorem XIV.3.1).

The first claim of the present section is that the spherical functions Φ_λ, while initially given by the integral formula (8), admit the following determinantal expression.

Proposition 2: If $\lambda \in \mathbb{C}^N$ and Φ_λ is given by (8), then

$$\Phi_\lambda(x) = \prod_{k=1}^{N-1} k! \times \frac{\det\left[\rho_k^{\lambda_\ell + (N-1)/2}\right]_{k,\ell=1}^N}{V(\lambda)V(\rho)} \tag{9}$$

where $(\rho_1, \ldots, \rho_N)$ are the eigenvalues of x and V stands for the Vandermonde polynomial.

The proof of Proposition 2 will be given in Appendix A. Formula (9) will be called the Gelfand-Naimark formula, as it is a generalisation of the formula introduced by Gelfand and Naimark [18], in 1950. The reason why the spherical functions (8) admit the determinantal expression (9) is that the group G is here a complex Lie group. This fact will be discussed again in Section V, and is the foundation of the proof in Appendix A.

Now, let $f : M \to \mathbb{C}$ be an integrable U-invariant function (integrable means with respect to vol, as in (2)). Its spherical transform is the function $\hat{f} : \mathbb{R}^N \to \mathbb{C}$,

$$\hat{f}(t) = \int_M f(x)\Phi_{-\mathrm{i}t}(x)\mathrm{vol}(dx) \tag{10}$$

where $\mathrm{i} = \sqrt{-1}$ [2]. An inversion theorem for the spherical transform (10) is given in [2] (Theorem XIV.5.3). Specifically, if $\hat{f}$ satisfies the integrability condition

$$\int_{\mathbb{R}^N} |\hat{f}(t)|(V(t))^2 dt < \infty \tag{11}$$

then the following inversion formula holds,

$$f(x) = C_N \int_{\mathbb{R}^N} \hat{f}(t)\Phi_{\mathrm{i}t}(x)(V(t))^2 dt \tag{12}$$

After substituting (2) and (9) into (10) and (12), the following is obtained.

Proposition 3: For the spherical transform pair (10)–(12),

$$\hat{f}(t) = \frac{C_N}{V(-\mathrm{i}t)} \times \frac{1}{N!}\int_{\mathbb{R}^N_+} f_o(\rho)V(\rho)\det\left[\rho_k^{-\mathrm{i}t_\ell-(N+1)/2}\right]d\rho \tag{13}$$

$$f(x) = \frac{C_N}{V(\mathrm{i}\rho)} \times \frac{1}{N!}\int_{\mathbb{R}^N} \hat{f}(t)V(t)\det\left[\rho_k^{\mathrm{i}t_\ell+(N-1)/2}\right]dt \tag{14}$$

where $f(x) = f_o(\rho)$ is a symmetric function of the eigenvalues $(\rho_1,\ldots,\rho_N)$ of x.

The proof of this proposition will not be given in detail, as it merely consists in performing straightforward algebraic simplifications.

It is remarkable that the spherical transform does not involve all the spherical functions Φ_λ but only the functions $\Phi_{\mathrm{i}t}$ where $t \in \mathbb{R}^N$. These functions have in common the property that they correspond to real, negative eigenvalues of the Laplace-Beltrami operator (3) : each $\Phi_{\mathrm{i}t}$ corresponds to the eigenvalue $-(t,t)-(\delta,\delta)$. The spherical functions that do not appear in the spherical transform are interesting in their own right. For instance, if $\lambda+\delta = m$ where $(m_1,\ldots,m_N)$ are positive integers arranged in decreasing order, then one has

$$\Phi_\lambda(x) = \prod_{k=1}^{N-1} k! \times S_m(\rho)/V(\lambda) \tag{15}$$

where S_m denotes the Schur polynomial corresponding to $(m_1,\ldots,m_N)$. Schur polynomials are very important in the study of circular and unitary-invariant random matrix ensembles [3] [4]. Mathematically, this is because they provide the irreducible characters of the unitary group [4] [5].

Example : continuing the previous example (at the end of Paragraph II-B), consider the integrals

$$Z(\sigma,\lambda) = \int_M \exp\left[-\frac{d^2(x,\mathrm{id})}{2\sigma^2}\right]\Phi_\lambda(x)\,\mathrm{vol}(dx) \tag{16}$$

where Φ_λ was defined in (8). If $\lambda = -\delta$ then $Z(\sigma,\lambda)$ is just $Z(\sigma)$ from (5). On the other hand, note that $Z(\sigma,-\mathrm{i}t) = \hat{f}(t)$ where $f(x) = \exp[-d^2(x,\mathrm{id})/2\sigma^2]$. As in the previous example, $f_o(\rho) = w(\rho_1)\ldots w(\rho_N)$ where $w(\rho) = \exp[-\log^2(\rho)/2\sigma^2]$. Replacing this into (13), it is possible to apply the Andréief identity (as stated in [6], Chapter 11), which yields

$$\hat{f}(t) = \frac{C_N}{V(-\mathrm{i}t)} \times \det\left[\int_0^\infty w(\rho)\rho^{\delta_k-\mathrm{i}t_\ell-1}\,d\rho\right]_{k,\ell=1}^N \tag{17}$$

In particular, recalling that $w(\rho)$ is here given by a log-normal function,

$$Z(\sigma,-\mathrm{i}t) = \frac{C_N}{V(-\mathrm{i}t)} \times \det\left[\sigma\exp\left((\sigma^2/2)(\delta_k-\mathrm{i}t_\ell)^2\right)\right]_{k,\ell=1}^N \tag{18}$$

In fact, all of the integrals $Z(\sigma,\lambda)$ in (16) are convergent, and can be expressed by analytic continuation of (18) (simply by replacing $-\mathrm{i}t_\ell$ with λ_ℓ).

IV. Constructing invariant kernels

The L^1–Godement theorem was introduced in [7]. Roughly, this theorem shows that U-invariant positive definite functions can be obtained by taking inverse spherical transforms of positive symmetric functions.

Theorem 1: Let $f : M \to \mathbb{C}$ be an integrable U-invariant function (integrable means with respect to vol). Then, f is positive definite if and only if

$$f(x) = \int_{\mathbb{R}^N} g(t)\Phi_{\mathrm{i}t}(x)(V(t))^2 dt \tag{19}$$

where the function $g : \mathbb{R}^N \to \mathbb{R}$ is positive ($g(t) \geq 0$ for all t), symmetric, and satisfies the integrability condition

$$\int_{\mathbb{R}^N} g(t)\,(V(t))^2 dt < \infty \tag{20}$$

Moreover, this function g is then unique — recall that g is said to be symmetric if $g(t)$ remains unchanged after any permutation of $(t_1, \ldots, t_N)$.

The proof of Theorem 1 will be given in Appendix B.

As explained in [7], the L^1–Godement theorem is based on the celebrated Godement theorem, which generalises Bochner's theorem to the context of symmetric spaces [8]. In [7], the only-if part of this theorem was used to check whether a given function f is positive definite or not. Here, the if part will be used to construct and compute positive definite functions.

To do so, note that (19) is essentially an inverse spherical transform as in (12), with $g(t)$ instead of $\hat{f}(t)$. Therefore, just as in (14), it is possible to rewrite (19),

$$f(x) = \frac{1}{V(\mathrm{i}\rho)} \times \frac{1}{N!} \int_{\mathbb{R}^N} g(t) V(t) \det\left[\rho_k^{\mathrm{i}t_\ell + (N-1)/2}\right] dt \tag{21}$$

where $(\rho_1, \ldots, \rho_N)$ are the eigenvalues of x. To obtain a positive definite function f, it is then enough to choose a suitable positive function g and then evaluate the integral (21). This is considered in the following proposition.

Proposition 4: Let $g : \mathbb{R}^N \to \mathbb{R}$ satisfy the conditions of Theorem 1.

(a) assume that $g(t)$ factors into $g(t) = \gamma(t_1) \ldots \gamma(t_N)$, where γ is a positive function. It follows from (21) that

$$f(x) = \frac{(\det(x))^{(N-1)/2}}{V(\mathrm{i}\rho)} \times \det\left[\int_{\mathbb{R}} \gamma(t) t^{k-1} e^{\mathrm{i}ts_\ell}\, dt\right] \tag{22}$$

whenever the integrals under the determinant exist. Here, $s_\ell = \log(\rho_\ell)$ for $\ell = 1, \ldots, N$.

(b) assume that the inverse Fourier transform

$$\tilde{g}(s) = \int_{\mathbb{R}^N} g(t) e^{\mathrm{i}(s,t)}\, dt \qquad (s_1, \ldots, s_N) \in \mathbb{R}^N$$

is smooth. It follows from (21) that

$$f(x) = \frac{(\det(x))^{(N-1)/2}}{V(\rho)} \times V\left(-\frac{\partial}{\partial s}\right) \tilde{g}(s)\Big|_{s_\ell = \log(\rho_\ell)} \tag{23}$$

where $V(\partial/\partial s)$ is the Vandermonde operator $V(\partial/\partial s) = \prod_{i<j}(\partial/\partial s_j - \partial/\partial s_i)$.

The proof of Proposition 4 is given in Appendix C.

Example: as an application of Proposition 4 (a), choose $g(t) = \gamma(t_1) \ldots \gamma(t_N)$ where $\gamma(t) = (\kappa/2)\exp(-\kappa|t|)$ for $\kappa > 0$. Replacing into (22) and using elementary properties of the Fourier transform,

$$f(x) = \frac{(\det(x))^{(N-1)/2}}{V(\rho)} \times \det\left[-\tilde{\gamma}^{(k-1)}(\log(\rho_\ell))\right] \tag{24}$$

where $\tilde{\gamma}^{(k-1)}$ is the $(k-1)$-th derivative of $\tilde{\gamma}(s) = (\kappa^2 + s^2)^{-1}$. Theorem 1 now says that this f is a positive definite function. In fact, (24) provides a whole family of positive definite functions, parameterised by κ.

Example: while it looks a bit too complicated, Proposition 4 (b) can be used to compute the heat kernel of M in closed form [1] [19]. This corresponds to $g(t) = \exp[-\kappa((t,t) + (\delta,\delta))]$ with $\kappa > 0$. Now, for this choice of $g(t)$, one has the following inverse Fourier transform

$$\tilde{g}(s) = \left((\pi/\kappa)^{N/2} e^{-\kappa(\delta,\delta)}\right) \exp\left[-(s,s)/4\kappa\right]$$

and one may use a beautiful identity from [19] (Chapter XII, Page 405)

$$V\left(-\frac{\partial}{\partial s}\right) \exp\left[-(s,s)/4\kappa\right] = (1/2\kappa)^{N(N-1)/2} V(s) \exp\left[-(s,s)/4\kappa\right] \tag{25}$$

Replacing this into (23) yields the heat kernel (rather, $f(x) = \mathcal{K}(x, \mathrm{id})$ where $\mathcal{K}$ is the heat kernel)

$$f(x) = C_\kappa \left(\det(x)\right)^{(N-1)/2} \times \left(V(\log(\rho))/V(\rho)\right) \exp\left[-|\log(\rho)|^2/4\kappa\right] \tag{26}$$

where $C_\kappa > 0$ is a constant and $|\log(\rho)|^2 = (\log(\rho), \log(\rho))$. Of course, this f is a positive definite function for each $\kappa > 0$, thanks to Theorem 1.

V. Final remarks

The motivation of the present work was to introduce a set of analytical tools which would make it possible to construct and compute invariant kernels on the space of complex covariance matrices. These tools are afforded by Theorem 1 and Proposition 4, stated in the previous Section IV.

Recall from Proposition 1 that there is a straightforward equivalence between invariant kernels and U-invariant positive definite functions. Theorem 1 sets up a bijective correspondence between the set of integrable U-invariant positive definite functions, and the set of positive symmetric functions which satisfy the integrability condition (11). Proposition 4 then makes it possible to explore the whole set of integrable U-invariant positive definite functions, simply by plugging suitable positive symmetric functions into (22) or (23).

The invariant kernels obtained from (22) or (23) can still be used on other spaces of covariance matrices (real, quaternion, block-Toeplitz, *etc.*). A space of covariance matrices may always be embedded into a space of complex covariance matrices (eventually of larger matrix size), in a way which preserves all the fundamental symmetry and invariance properties. It is then enough to compose kernels obtained from (22) or (23) with the correct embedding into a space of complex covariance matrices (in general, this embedding is rather easy to compute).

Theorem 1 and Proposition 4 rely heavily on Proposition 2, which provides the determinantal expression (9) for the spherical functions (8). In general, this kind of determinantal expression is readily available on any symmetric space whose group of isometries (that is the group G in the notation of Sections II and III) is a complex Lie group (see Formula (32), Appendix A). All the results obtained in the above can accordingly be generalised to symmetric spaces which have this property, called symmetric spaces of type IV in [9]. The symmetric spaces of type IV are precisely (a) spaces of complex covariance matrices (b) spaces of complex covariance matrices which are also orthogonal (c) spaces of complex covariance matrices which are also symplectic (d) certain other so-called exceptional spaces. These are the non-compact analogues (the technical word is duals) of the compact Lie groups (unitary, orthogonal, symplectic, and exceptional).

References

[1] S. Helgason, *Groups and geometric analysis : integral geometry, invariant differential operators, and spherical functions*. Academic Press, 1984.

[2] J. Faraut and A. Korányi, *Analysis on symmetric cones*. Claredon Press, 1994.

[3] P. Mergny, "Spherical integrals and their applications to random matrix theory," Ph.D. dissertation, Université Paris-Saclay, 2022.

[4] B. Jonnadula, J. P. Keating, and F. Mezzadri, "Symmetric function theory and unitary invariant ensembles," *Journal of Mathematical Physics*, vol. 62, 2021.

[5] E. S. Meckes, *The random matrix theory of the classical compact groups*. Cambridge University Press, 2019.

[6] G. Livan, M. Novaes, and P. Vivo, *Introduction to random matrices: theory and practice*. Springer Cham, 2018.

[7] N. Da Costa, C. Mostajeran, J. P. Ortega, and S. Said, "Invariant kernels on Riemannian symmetric spaces: a harmonic-analytic approach," Oct. 2023. [Online]. Available: arxiv:2310.19270

[8] R. Godement, "Introduction aux travaux de Selberg," *Séminaire Bourbaki, (no. 144)*, pp. 95–110, 1958.

[9] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*. Americal Mathematical Society, 2001.

[10] N. Guigui, N. Miolane, and X. Pennec, *Introduction to Riemannian Geometry and Geometric Statistics*. Now Publishers Inc., 2023.

[11] S. Said, "Statistical models and probabilistic methods on Riemannian manifolds," 2021, Université de Bordeaux, Habilitation thesis (HDR).

[12] A. Azangulov, A. Smolensky, A. Terenin, and S. Borovitsky, "Stationary kernels and Gaussian processes on Lie groups and their homogeneous spaces II: non-compact symmetric spaces," Apr. 2023. [Online]. Available: arxiv:2301.13088

[13] X. Pennec, "Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements," *Journal of Mathematical Imaging and Vision*, vol. 25, no. 1, pp. 127–154, 2006.

[14] R. I. Muirhead, *Aspects of multivariate statistical theory*. Wiley Interscience, 1982.

[15] L. Santilli and M. Tierz, "Riemannian Gaussian distributions, random matrix ensembles and diffusion kernels," *Nucl. Phys. B*, vol. 973, p. 115582, 2021.

[16] S. Said, H. Hajri, L. Bombrun, and B. C. Vemuri, "Gaussian distributions on Riemannian symmetric spaces : statistical learning with structured covariance matrices," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 752–772, 2018.

[17] S. Said, S. Heuveline, and C. Mostajeran, "Riemannian statistics meets random matrix theory: toward learning from high-dimensional covariance matrices," *IEEE Transactions on Information Theory*, vol. 69, no. 1, pp. 472–481, 2023.

[18] I. M. Gelfand and M. A. Naimark, "Unitary representations of the classical groups," *Trudy Mat. Inst. Steklov*, vol. 36, pp. 3–288, 1950.

[19] S. Lang and J. Jorgenson, *Spherical inversion on* $\mathrm{SL}_n(\mathrm{R})$. Springer-Verlag, 2001.

APPENDIX A
PROOF OF PROPOSITION 2

The proof starts from (8) and shows that it is equivalent to (9). The first step is to reduce to the case where x has unit determinant. Let $x = r\bar{x}$ where $\det(\bar{x}) = 1$. By a direct calculation, it follows from (7) that $\Delta_s(x) = r^{(1,s)}\Delta_s(\bar{x})$ where $(1,s) = \sum_{k=1}^{N} s_k$. In particular, $s = \lambda + \delta$ implies that $\Delta_s(x) = r^{(1,\lambda)}\Delta_s(\bar{x})$, because $(1,\delta) = 0$. Replacing into (8) yields the following identity

$$\Phi_\lambda(x) = r^{(1,\lambda)}\,\Phi_\lambda(\bar{x}) \tag{27}$$

Now, returning to (8), note that

$$\Phi_\lambda(\bar{x}) = \int_U \Delta_{\lambda+\delta}(u\cdot\bar{x})\,du \tag{28}$$

This integral can be restricted to SU, the special unitary subgroup of U (SU is the set of $u \in U$ with $\det(u) = 1$). Indeed, replacing u by $e^{\mathrm{i}\theta}u$ (with θ real) does not change $u\cdot\bar{x}$. Moreover, the normalised Haar measure on U descends to the normalised Haar measure on SU [**?**] (see Theorem 8.32). Therefore, (28) is equivalent to

$$\Phi_\lambda(\bar{x}) = \int_{SU} \Delta_{\lambda+\delta}(u\cdot\bar{x})\,du \tag{29}$$

The next step of the proof is to show that (29) is the same as the following Harish-Chandra integral [1] (Page 418)

$$\Phi_\lambda(\bar{x}) = \int_{SU} \exp\left[(2\lambda - 2\delta, \log a(u\bar{\rho}^{1/2}))\right]du \tag{30}$$

Here, $\bar{\rho} = (\bar{\rho}_1,\ldots,\bar{\rho}_N)$ are the eigenvalues of $\bar{x}$, and $a(u\bar{\rho}^{1/2})$ is the diagonal matrix with positive entries, such that $u\bar{\rho}^{1/2} = na(u\bar{\rho}^{1/2})h$ where the matrix n is upper triangular with ones on its diagonal, and h belongs to SU (vectors such as $\bar{\rho}$ will be identified with diagonal matrices, in a self-evident way, whenever that is convenient).

Because Φ_λ is U-invariant, it is enough to prove (30) when $\bar{x} = \bar{\rho}$. It will be helpfull to apply the identity $\Phi_\lambda(x) = \Phi_{-\lambda}(x^{-1})$ [2] (Theorem XIV.3.1). For short, let $a = a(u\bar{\rho}^{1/2})$, so $u\cdot\bar{\rho} = (u\bar{\rho}^{1/2})\cdot\mathrm{id}$ is equal to $na^2n^\dagger$. Using the identity just mentioned,

$$\Phi_\lambda(\bar{\rho}) = \Phi_{-\lambda}(\bar{\rho}^{-1}) = \int_{SU} \Delta_{-\lambda+\delta}\left(u\cdot\bar{\rho}^{-1}\right)du = \int_{SU} \Delta_{-\lambda+\delta}\left((n^\dagger)^{-1}\cdot a^{-2}\right)du \tag{31}$$

where the second equality follows from (29) and the third equality holds because $u\cdot\bar{\rho}^{-1} = (u\cdot\bar{\rho})^{-1}$ and $u\cdot\bar{\rho} = na^2n^\dagger$. However, since $(n^\dagger)^{-1}$ is lower triangular with ones on its diagonal, it is easy to see that $\Delta_{-\lambda+\delta}\left((n^\dagger)^{-1}\cdot a^{-2}\right) = \Delta_{-\lambda+\delta}\left(a^{-2}\right)$. Then, from (7) and the fact that a is diagonal,

$$\Delta_{-\lambda+\delta}\left((n^\dagger)^{-1}\cdot a^{-2}\right) = \prod_{k=1}^{N} a_k^{2(\lambda_k-\delta_k)} = \exp\left[(2\lambda-2\delta,\log a)\right]$$

and (30) follows immediately by replacing this into (31). The final step of the proof exploits the fact that G is a complex Lie group. In this case, the Harish-Chandra integral (30) admits a closed-form expression [1] (Page 432),

$$\Phi_\lambda(\bar{x}) = \frac{\Pi(-\delta)}{\Pi(\lambda)} \times \frac{\sum_{w\in S_N}\varepsilon(w)e^{(\lambda,w\bar{\rho})}}{\sum_{w\in S_N}\varepsilon(w)e^{(\delta,w\bar{\rho})}} \tag{32}$$

Here, $\Pi(\mu) = \prod_{k<\ell}(\mu_k - \mu_\ell)$ for $\mu \in \mathbb{C}^N$, S_N is the symmetric group (group of permutations of N objects), and $\varepsilon(w)$ is the signature of the permutation w, while $w\bar{\rho}$ denotes the action of that permutation on $(\bar{\rho}_1,\ldots,\bar{\rho}_N)$.

Clearly, the polynomial Π is the Vandermonde polynomial, up to sign, while the sums in the second fraction are Leibniz expansions of certain determinants. Using these observations, and performing some basic simplifications,

$$\Phi_\lambda(\bar{x}) = \frac{\prod_{k=1}^{N-1} k!}{V(-\lambda)} \times \frac{\det\left[\bar{\rho}_k^{\lambda_\ell}\right]}{\det\left[\bar{\rho}_k^{N-\ell}\right]} = \frac{\prod_{k=1}^{N-1} k!}{V(-\lambda)} \times \frac{\det\left[\bar{\rho}_k^{\lambda_\ell}\right]}{V(-\bar{\rho})} = \prod_{k=1}^{N-1} k! \times \frac{\det\left[\bar{\rho}_k^{\lambda_\ell}\right]}{V(\lambda)V(\bar{\rho})} \tag{33}$$

Now, (9) can be retrieved from (27) and (33). To do so, it is enough to note that $\rho = r\bar{\rho}$ and that this implies (recall $\rho = (\rho_1, \ldots, \rho_N)$ are the eigenvalues of x)

$$r^{(1,\lambda)} \det\left[\bar{\rho}_k^{\lambda_\ell}\right] = \det\left[\rho_k^{\lambda_\ell}\right] \qquad \text{and} \qquad V(\bar{\rho}) = r^{-N(N-1)/2}V(\rho)$$

Then, since $r^N = \det(x)$ (which is the product of the eigenvalues ρ_k),

$$r^{(1,\lambda)} \frac{\det\left[\bar{\rho}_k^{\lambda_\ell}\right]}{V(\bar{\rho})} = r^{N(N-1)/2} \frac{\det\left[\rho_k^{\lambda_\ell}\right]}{V(\rho)} = \frac{\det\left[\rho_k^{\lambda_\ell + (N-1)/2}\right]}{V(\rho)} \tag{34}$$

Therefore, performing the multiplication in (27), with the help of (33) and (34), yields the required (9).

Appendix B
Proof of Theorem 1

if part: roughly, the idea of the proof is to show that the spherical functions $\Phi_{\mathrm{i}t}$ are positive definite functions. Then, (19) says that the function f is a positive linear combination of these $\Phi_{\mathrm{i}t}$ and is therefore positive definite. For $x \in M$, let $x = \exp(s)\bar{x}$ where $s = \log\det(x)$ so that $\det(\bar{x}) = 1$. From (27) (putting $\lambda = \mathrm{i}t$ and $\tau = (1,t)$),

$$\Phi_{\mathrm{i}t}(x) = e^{\mathrm{i}\tau s}\Phi_{\mathrm{i}t}(\bar{x}) \tag{35}$$

Thinking of s and $\bar{x}$ as functions of x, let $\varphi^\tau(x) = e^{\mathrm{i}\tau s}$ and $\bar{\varphi}(x) = \Phi_{\mathrm{i}t}(\bar{x})$, so that $\Phi_{\mathrm{i}t}(x) = \varphi^\tau(x)\bar{\varphi}(x)$. Now, recalling the well-known fact that a product of positive definite functions is positive definite, it is enough to show that φ^τ and $\bar{\varphi}$ are both positive definite. For any $x_1, \ldots, x_n \in M$, note that

$$\varphi^\tau(x_i^{-1/2}x_jx_i^{-1/2}) = e^{\mathrm{i}\tau(s_j - s_i)} \tag{36}$$

where $s_i = \log\det(x_i)$. Therefore, the matrix with elements $\varphi^\tau(x_i^{-1/2}x_jx_i^{-1/2})$ is the same as the matrix with elements $e^{\mathrm{i}\tau(s_j-s_i)}$, which is Hermitian positive semidefinite (of rank 1). This shows that φ^τ is positive definite. To see that $\bar{\varphi}$ is also positive definite, note that

$$\bar{\varphi}(x_i^{-1/2}x_jx_i^{-1/2}) = \Phi_{\mathrm{i}t}(\bar{x}_i^{-1/2}\bar{x}_j\bar{x}_i^{-1/2}) \tag{37}$$

However, according to [1] (Page 484), the restriction of $\Phi_{\mathrm{i}t}$ to the unit-determinant hypersurface (that is to the set of $x \in M$ with $\det(x) = 1$), which is given by the Harish-Chandra integral (30), is a positive definite function. In particular, the matrix whose elements appear in (37) is Hermitian positive semidefinite. This shows that $\bar{\varphi}$ is positive definite. Thus, being a product of positive definite functions, $\Phi_{\mathrm{i}t}$ is positive definite (for any t) as required. To conclude, recall Godement's theorem [7] [8] (in particular, Formula (2.4) in [7]). This gives a rigorous justification of the claim that f is positive definite because it is a positive linear combination of positive definite functions. The last step of the proof is thus a direct application of Godement's theorem.

only-if part : as stated in [7], the L^1– Godement theorem says that an integrable function f is positive definite if and only if its spherical transform is positive and integrable (details can be found in [7], Section 2 and Appendix A). In [7], it is required from the outset that the underlying symmetric space M should be of non-compact type (in particular, the group G should be semisimple). This requirement is clearly not satisfied, in the present context. The aim here is to explain that the L^1– Godement theorem can still be applied.

In [7], the requirement that M should be of non-compact type was introduced only in order to ensure that spherical functions on M are given by Harish-Chandra integrals (integrals of the form (30) in the proof of Proposition 2). The proof of the L^1– Godement theorem (see [7], Appendix A) relies on Godement's (much older) theorem [8], which applies to any symmetric space and in particular to the space M of complex covariance matrices. Specifically, the convolution product of two compactly-supported continuous U-invariant functions on M is commutative [2] (Proposition XIV.4.1), and this is the only hypothesis needed for Godement's theorem.

With this in mind, the proof is just an application of the L^1– Godement theorem as stated in [7]. Precisely, f is positive definite only if $\hat{f}$ is positive and integrable (which here means it satisfies the integrability condition (11)). Then, f is given by the inversion formula (12), which is the same as (19) with $g = C_{N}\,\hat{f}$.

To finish the proof, note that uniqueness of g follows by injectivity of the inverse spherical transform (the linear map that takes $\hat{f}$ to f according to (12)).

APPENDIX C
PROOF OF PROPOSITION 4

Part (a) : in order to prove (22), it is enough to show that, on the right-hand side of (21),

$$\frac{1}{N!}\int_{\mathbb{R}^N} g(t)V(t)\det\left[\rho_k^{\mathrm{i}t_\ell+(N-1)/2}\right]dt = (\det(x))^{(N-1)/2}\,\det\left[\int_{\mathbb{R}}\gamma(t)t^{k-1}\,e^{\mathrm{i}ts_\ell}\,dt\right] \tag{38}$$

where $g(t) = \gamma(t_1)\ldots\gamma(t_N)$ and $s_\ell = \log(\rho_\ell)$. Note first that

$$\det\left[\rho_k^{\mathrm{i}t_\ell+(N-1)/2}\right] = \prod_{k=1}^{N}\rho_k^{(N-1)/2}\times\det\left[\rho_k^{\mathrm{i}t_\ell}\right] = (\det(x))^{(N-1)/2}\times\det\left[\rho_k^{\mathrm{i}t_\ell}\right]$$

This implies that (38) is equivalent to

$$\frac{1}{N!}\int_{\mathbb{R}^N} g(t)V(t)\det\left[\rho_k^{\mathrm{i}t_\ell}\right]dt = \det\left[\int_{\mathbb{R}}\gamma(t)t^{k-1}\,e^{\mathrm{i}ts_\ell}\,dt\right]$$

or, what is the same if $V(t)$ is expressed as a determinant,

$$\frac{1}{N!}\int_{\mathbb{R}^N} g(t)\det\left[t_\ell^{k-1}\right]\det\left[\rho_k^{\mathrm{i}t_\ell}\right]dt = \det\left[\int_{\mathbb{R}}\gamma(t)t^{k-1}\,e^{\mathrm{i}ts_\ell}\,dt\right] \tag{39}$$

Here, using the Andréief identity [6] (Chapter 11, Page 75), the left-hand side is equal to

$$\frac{1}{N!}\int_{\mathbb{R}^N}\det\left[t_\ell^{k-1}\right]\det\left[\rho_k^{\mathrm{i}t_\ell}\right]\prod_{\ell=1}^{N}\gamma(t_\ell)dt_\ell = \det\left[\int_{\mathbb{R}}\gamma(t)t^{k-1}\,\rho_\ell^{\mathrm{i}t}\,dt\right]$$

which is the same as the right-hand side (by definition of s_ℓ). Thus, (39) (equivalent to (38)) has been proven true.

Part (b) : comparing (21) and (23), it becomes clear that one must show

$$\frac{1}{N!}\int_{\mathbb{R}^N} g(t)V(t)\det\left[\rho_k^{\mathrm{i}t_\ell+(N-1)/2}\right]dt = (\det(x))^{(N-1)/2}\;\mathrm{i}^{N(N-1)/2}\;V\left(-\partial/\partial s\right)\tilde{g}(s) \tag{40}$$

However, as in the proof of Part (a), this is equivalent to

$$\frac{1}{N!}\int_{\mathbb{R}^N} g(t)V(t)\det\left[\rho_k^{\mathrm{i}t_\ell}\right]dt = \mathrm{i}^{N(N-1)/2}\, V\left(-\partial/\partial s\right)\tilde{g}(s) \tag{41}$$

After writing the Leibniz expansion of the determinant, the left-hand side is equal to

$$\frac{1}{N!}\sum_{w\in S_N}\varepsilon(w)\int_{\mathbb{R}^N} g(t)V(t)\prod_{k=1}^{N}\rho_k^{\mathrm{i}t_{w(k)}}\,dt = \frac{1}{N!}\sum_{w\in S_N}\varepsilon(w)\int_{\mathbb{R}^N} g(t)V(t)e^{\mathrm{i}(s,wt)}\,dt$$

Here, S_N is the group of permutations of $\{1,\ldots,N\}$ and $\varepsilon(w)$ is the signature of the permutation w. Moreover, on the right-hand side $s_\ell = \log(\rho_\ell)$ and wt denotes the action of the permutation w on $(t_1,\ldots,t_N)$. By introducing a new variable of integration $u = wt$ in each one of the integrals under the sum,

$$\frac{1}{N!}\sum_{w\in S_N}\varepsilon(w)\int_{\mathbb{R}^N} g(t)V(t)e^{\mathrm{i}(s,wt)}\,dt = \frac{1}{N!}\sum_{w\in S_N}\varepsilon(w)\int_{\mathbb{R}^N} g(w^{-1}u)V(w^{-1}u)e^{\mathrm{i}(s,u)}\,du$$

But the function g is symmetric, while the Vandermonde polynomial V is alternating — for any permutation w, $g(wu) = g(u)$ and $V(wu) = \varepsilon(w)V(u)$. Therefore, the above sum is equal to

$$\frac{1}{N!}\sum_{w\in S_N}\int_{\mathbb{R}^N} g(u)V(u)e^{\mathrm{i}(s,u)}\,du = \int_{\mathbb{R}^N} g(u)V(u)e^{\mathrm{i}(s,u)}\,du$$

and it now follows that the left-hand side of (41) is

$$\frac{1}{N!}\int_{\mathbb{R}^N} g(t)V(t)\det\left[\rho_k^{\mathrm{i}t_\ell}\right]dt = \int_{\mathbb{R}^N} g(u)V(u)e^{\mathrm{i}(s,u)}\,du \tag{42}$$

Finally, recalling the definition of the inverse Fourier transform $\tilde{g}(s)$, and differentiating under the integral, one has

$$V(\partial/\partial s)\tilde{g}(s) = \mathrm{i}^{N(N-1)/2}\int_{\mathbb{R}^N} g(u)V(u)e^{\mathrm{i}(s,u)}\,du$$

which can be replaced back into (42) to obtain

$$\frac{1}{N!}\int_{\mathbb{R}^N} g(t)V(t)\det\left[\rho_k^{\mathrm{i}t_\ell}\right]dt = (-\mathrm{i})^{N(N-1)/2}\,V(\partial/\partial s)\tilde{g}(s) = \mathrm{i}^{N(N-1)/2}\,V(-\partial/\partial s)\tilde{g}(s)$$

which is identical to (41), as required.