

DECONVOLUTION DENSITY ESTIMATION ON SPACES OF POSITIVE DEFINITE SYMMETRIC MATRICES

BY PETER T. KIM* AND DONALD ST. P. RICHARDS†

University of Guelph, and Penn State University and SAMSI

Motivated by applications in microwave engineering and diffusion tensor imaging, we study the problem of deconvolution density estimation on the space of positive definite symmetric matrices. We develop a nonparametric estimator for the density function of a random sample of positive definite matrices. Our estimator is based on the Helgason-Fourier transform and its inversion, the natural tools for analysis of compositions of random positive definite matrices. Under several smoothness conditions on the density of the intrinsic error in the random sample, we derive upper bounds on the rates of convergence of our nonparametric estimator to the true density.

1. Introduction. In this paper, we are motivated by applications in microwave engineering and medical imaging to study the problem of deconvolution density estimation on the space of positive definite (symmetric) matrices.

In microwave engineering, researchers have studied problems in which random observations arise as positive definite matrices. Terras [38], p. 156 ff. reviews the problem of transmitting electrical signals over a long, lossless line containing random inhomogeneities, and shows how the problem may be modeled using mathematical analysis on the space of 2×2 positive definite unimodular matrices. All in all, Terras [38], sections 3.1-3.2 contains a fascinating account of the appearance of sequences of random positive definite matrices in a real-world engineering problem.

The appearance in medical imaging of sequences of random positive definite matrices has become commonplace due to developments in *diffusion tensor imaging* (DTI), a method of imaging based upon the observation that water molecules in biological tissue are always in motion. For the pur-

*Research supported in part by the Natural Sciences and Engineering Research Council of Canada, grant OGP46204.

†Research supported in part by the U.S. National Science Foundation, grants DMS-0112069 and DMS-0705210.

AMS 2000 subject classifications: Primary 62G20; Secondary 65R32.

Keywords and phrases: Diffusion tensor imaging, deconvolution, Harish-Chandra c -function, Helgason-Fourier transform, inverse problems, Laplace-Beltrami operator, magnetic resonance imaging, Sobolev spaces, Wishart distribution.

poses of mathematical modeling, it is generally assumed that the diffusion of water molecules at any given location in biological tissue follows a Brownian motion. A diffusion tensor image then is represented by the 3×3 positive-definite covariance matrix of the local diffusion process at the given location. DTI seeks to detect the diffusion of water protons between and within distinct tissue cells, and to derive estimates of the dominant orientation and direction of the Brownian motion (Le Bihan [24], Hasan, *et al.* [12]).

In DTI brain imaging, the diffusion of water molecules within and between *voxels*, the three-dimensional volume elements that constitute an image, reveal both the orientation of fibers comprising white-matter tracts in the brain and the coherence of fibers, the extent to which fibers are aligned together. DTI may be the only non-invasive, *in vivo* procedure which enables the study of deep brain white-matter fibers. Consequently, DTI has been found to be highly promising for comparing the human brain in normal states with abnormal states caused by strokes, epileptic seizures, tumors, white-matter abnormalities, multiple sclerosis lesions, HIV-infection, traumatic brain injuries, aging, Alzheimer's disease, alcoholism, and developmental disorders; and there are potential applications to psychiatric conditions including schizophrenia, autism, cognitive and learning disabilities (Neumann-Haefelin, *et al.* [29], Rosenbloom, *et al.* [34], Pomara, *et al.* [31], Matthews and Arnold [27]). In addition, DTI has been applied in research on the pathology of organ and tissue types such as the human breast, kidney, lingual, cardiac, skeletal muscles, and spinal cord (Damon, *et al.* [2]).

It is well-known that magnetic resonance imaging, from which diffusion tensors are derived, is endowed inherently with random noise. Hence, DTI data also contain noise (Basu, *et al.* [1]), and it is natural that statistical inferential issues arise in the analysis of DTI data (Koltchinskii, *et al.* [22]; Schwartzman, *et al.* [35, 36]; Zhu, *et al.* [42]).

In this paper, we study the problem of estimating the probability density function of a population of positive definite matrices based on a random sample from that population. An instance in which this problem arises may be obtained from Schwartzman [35] who studied the two-sample comparison of twelve children divided into two groups according to reading ability, where the issue is to compare physical characteristics of brain tissue of the two groups on the basis of DTI images. In addition to comparing the population parameters, it is natural to seek estimators of the underlying density functions, and then it will be important to estimate the rates of convergence of the density estimators.

The deconvolution density estimation problem has been widely studied on Euclidean spaces (Diggle and Hall [3], Fan [4], Koo [23], Mair and Ruym-

gaart [26]). In this classical setting, the commutative nature of the underlying mathematical operations renders the problem amenable to classical mathematical methods. The deconvolution problem has also been studied on certain compact manifolds (Healy *et al.* [14], Hendriks [16], Kim *et al.* [18, 19, 20, 21]); in that setting, the problem is solvable using well-known generalizations of classical Fourier analysis.

By contrast, \mathcal{P}_m , the space of $m \times m$ positive definite matrices, is a non-compact Riemannian manifold and has an intrinsic non-commutative nature, and it is natural to expect that the deconvolution problem will be more difficult in that setting. To the best of our knowledge, Pesenson [30] is the only author who has studied the deconvolution problem on \mathcal{P}_m , albeit under band-limiting restrictions on the underlying density functions. Thus, no results yet are available for the general deconvolution problem on \mathcal{P}_m .

To date, the primary statistical emphases regarding mathematical methods on \mathcal{P}_m are motivated by properties of the Wishart distribution (Muirhead [28], Letac and Massam [25], Richards [32, 33]), estimation problems associated with the Wishart distribution (Haff [9, 10, 11]), and high-dimensional random matrices and their eigenvalue distributions (Johnstone [17], Takemura and Sheena [37]). To solve the general deconvolution problem on \mathcal{P}_m , more advanced mathematical methods are required. While much is known in the mathematical literature about the necessary methods (Helgason [15], Terras [40]), virtually nothing about those methods has appeared in the statistical literature. Using these new methods, we develop nonparametric methods for solving the deconvolution problem on \mathcal{P}_m .

In summary, section 2 provides notation, and introduces the *Helgason-Fourier transform* and its inversion formula, involving the *Harish-Chandra c -function*; we have provided the necessary details of these concepts so as to make the paper fully accessible to readers who are new to this area. We formalize in section 3 the statistical procedure in terms of measurement errors on \mathcal{P}_m and present the main results on general deconvolution density estimation on \mathcal{P}_m . In section 4, we provide the explicit details in the case of the Wishart distribution. Finally, all proofs are provided in section 5.

2. Preliminaries. Throughout the paper, we denote by G the general linear group $\mathrm{GL}(m, \mathbb{R})$, of $m \times m$ nonsingular real matrices and by K the group, $\mathrm{O}(m)$, of $m \times m$ orthogonal matrices. The group G acts transitively on \mathcal{P}_m , the space of $m \times m$ positive definite matrices, by the action

$$(2.1) \quad G \times \mathcal{P}_m \rightarrow \mathcal{P}_m, \quad (g, x) \mapsto g'xg,$$

$g \in G$, $x \in \mathcal{P}_m$, where g' denotes the transpose of g . Under this action, the isotropy group of the identity in G is K , hence the homogeneous space $K \backslash G$

can be identified with \mathcal{P}_m by the “natural” mapping

$$(2.2) \quad K \backslash G \rightarrow \mathcal{P}_m, \quad Kg \mapsto g'g.$$

In distinguishing between left and right cosets, we place the quotient operation on the left and right of the group, respectively.

A random matrix $X \in \mathcal{P}_m$ is said to be K -invariant if $X \stackrel{\mathcal{L}}{=} k'Xk$ for all $k \in K$, where “ $\stackrel{\mathcal{L}}{=}$ ” denotes equality in distribution. A function f on \mathcal{P}_m is called K -invariant if $f(k'xk) = f(x)$ for all $k \in K$, $x \in \mathcal{P}_m$; we will indicate that f is K -invariant by writing its domain as \mathcal{P}_m/K , with a similar notation for K -invariant positive definite random matrices.

By means of the relationship (2.2) between $K \backslash G$ and \mathcal{P}_m , we identify K -invariant functions on \mathcal{P}_m with K -biinvariant functions on G , i.e., functions $\tilde{f} : G \rightarrow \mathbb{C}$ which satisfy $\tilde{f}(g) = \tilde{f}(k_1 g k_2)$ for all $k_1, k_2 \in K$ and $g \in G$. In particular, $\mathcal{P}_m/K \simeq K \backslash G / K$ where “ \simeq ” denotes diffeomorphic equivalence.

Consider random matrices $X, \varepsilon \in \mathcal{P}_m$ with corresponding group elements $\tilde{X}, \tilde{\varepsilon} \in G$, respectively. By the natural map (2.2), $K\tilde{X} \mapsto X$, equivalently $\tilde{X}'\tilde{X} = X$ and, similarly, $K\tilde{\varepsilon} \mapsto \varepsilon$. Then $\tilde{X}\tilde{\varepsilon} \in G$ is mapped *via* (2.2) to

$$(2.3) \quad \tilde{X}\tilde{\varepsilon} \mapsto (\tilde{X}\tilde{\varepsilon})'(\tilde{X}\tilde{\varepsilon}) \equiv \tilde{\varepsilon}'\tilde{X}'\tilde{X}\tilde{\varepsilon} = \tilde{\varepsilon}'X\tilde{\varepsilon} \in \mathcal{P}_m.$$

If $\tilde{\varepsilon}$ is K -biinvariant, i.e., $\tilde{\varepsilon} \stackrel{\mathcal{L}}{=} k_1\tilde{\varepsilon}k_2$ for all $k_1, k_2 \in K$ then $\tilde{\varepsilon}'\tilde{\varepsilon} = \varepsilon = k'\Lambda k$ where $k \in K$ and $\Lambda \in \mathcal{P}_m$ is the diagonal matrix of eigenvalues of ε . Consequently, we define $\varepsilon^{1/2} = k'\Lambda^{1/2}k$ and note that, by K -biinvariance of $\tilde{\varepsilon}$, the relationship (2.3) on the group G corresponds in distribution to $\varepsilon^{1/2}X\varepsilon^{1/2} \in \mathcal{P}_m$, where $X \in \mathcal{P}_m$ and $\varepsilon \in \mathcal{P}_m/K$. Bearing this in mind, we formally make the following definition.

DEFINITION 2.1. Suppose that $X \in \mathcal{P}_m$ and $\varepsilon \in \mathcal{P}_m/K$ are random matrices. Then the *composition of X and ε* is

$$(2.4) \quad X \circ \varepsilon = \varepsilon^{1/2}X\varepsilon^{1/2}$$

where $\varepsilon^{1/2}$ is the positive definite square root of ε .

We remark that the composition operation has arisen in the context of central limit theorems on \mathcal{P}_m ; see Terras [38, 40], Richards [33], Graczyk [8].

There is an alternative approach leading to compositions (2.4). We begin by noting that \mathcal{P}_m is a Riemannian manifold, hence each pair $X, Y \in \mathcal{P}_m$ defines a unique geodesic path. Since \mathcal{P}_m is also a homogeneous space, *viz.*, $\mathcal{P}_m = K \backslash G$, then there exists $V \in G$ such that

$$(2.5) \quad Y = VXV'.$$

This model can be viewed as a multiplicative analog of classical regression analysis, with Y serving as the dependent variable, X as the independent variable, and V as the error variable. We can also study this relationship between X and Y using the logarithm map on \mathcal{P}_m . To that end, each observation on Y may be transformed into an observation on y where $\exp(y) = Y$ with $y \in T(\mathcal{P}_m)$, the tangent space of \mathcal{P}_m . Similarly, each measurement on X may be transformed into a measurement on $x \in T(\mathcal{P}_m)$ where $\exp(x) = X$. By postulating the existence of a “small” error $v \in T(\mathcal{P}_m)$ such that $y = x+v$ and, by exponentiating this linear “regression” relationship between y and x , we are led naturally to (2.5); see Terras [40], section 4.1-4.2 for details on the exponentiation of such linear relationships on $T(\mathcal{P}_m)$.

If we assume that the measurement error V is *isotropic* or has no preferred orientation, i.e., V has a K -invariant distribution, then the conclusion is the model (2.5) in which $V \in G/K$. As noted in [34], in the context of DTI, the assumption that water molecules diffuse isotropically is appropriate for regions such as the ventricles, which are large fluid-filled spaces deep in the brain. On the other hand, water molecules located in white-matter fiber are constrained by the axon sheath; this forces greater movement along the longitudinal axes of fibers than across the axes, and then diffusion may be isotropic only at sufficiently small scales.

Returning to (2.5), we apply polar coordinates on G , viz., $V = k\varepsilon^{1/2}$ where $k \in K$ and $\varepsilon \in \mathcal{P}_m$. Since V is K -invariant then $V \stackrel{\mathcal{L}}{=} k'V = k'k\varepsilon^{1/2} = \varepsilon^{1/2}$; hence (2.5) reduces to $Y \stackrel{\mathcal{L}}{=} \varepsilon^{1/2}X\varepsilon^{1/2} \equiv X \circ \varepsilon$, which agrees with the definition (2.1) since $\varepsilon = V'V$ is K -invariant. Thus, the problem is to estimate nonparametrically the density of X based on a random sample from $Y = X \circ \varepsilon$ where the error matrix ε is isotropic.

2.1. The Helgason-Fourier Transform. Let $C_c^\infty(\mathcal{P}_m)$ denote the collection of complex-valued, infinitely differentiable, compactly supported functions f on \mathcal{P}_m . For $w \in \mathcal{P}_m$ and $j = 1, \dots, m$, denote by $|w_j|$ the principal minor of order j of w . For $s = (s_1, \dots, s_m) \in \mathbb{C}^m$, the *power function* $p_s : \mathcal{P}_m \rightarrow \mathbb{C}$ is defined by

$$(2.6) \quad p_s(w) = \prod_{j=1}^m |w_j|^{s_j},$$

$w \in \mathcal{P}_m$.

Let dk denote the Haar measure on K , normalized to have total volume equal to one. Then

$$(2.7) \quad h_s(w) = \int_K p_s(k'wk) dk$$

$w \in \mathcal{P}_m$, $s \in \mathbb{C}^m$, is a *zonal spherical function* on \mathcal{P}_m . It is well-known that the spherical functions play a fundamental role in harmonic analysis on symmetric spaces; see Helgason [15]; in particular, (2.7) is a special case of Harish-Chandra's formula for the general spherical functions. If s_1, \dots, s_m are nonnegative integers then, except for a constant factor, (2.7) is an integral formula for the zonal polynomials which arise often in aspects of multivariate distribution theory; see Muirhead [28], pp. 231–232.

Let $w = (w_{ij}) \in \mathcal{P}_m$. Then, up to a constant factor, the unique G -invariant measure on \mathcal{P}_m is

$$(2.8) \quad d_*w = |w|^{-(m+1)/2} \prod_{1 \leq i \leq j \leq m} dw_{ij},$$

where $|w|$ is the determinant of w .

DEFINITION 2.2. For $s = (s_1, \dots, s_m) \in \mathbb{C}^m$ and $k \in K$, the *Helgason-Fourier transform* ([40], p. 87) of a function $f \in C_c^\infty(\mathcal{P}_m)$ is

$$(2.9) \quad \mathcal{H}f(s, k) = \int_{\mathcal{P}_m} f(w) \overline{p_s(k'wk)} d_*w,$$

where $\overline{p_s(k'wk)}$ denotes complex conjugation and d_*w is the G -invariant measure (2.8).

For the case in which $f \in C_c^\infty(\mathcal{P}_m/K)$, we make the change of variables $w \rightarrow k_1'wk_1$ in (2.9), $k_1 \in K$, and integrate with respect to the Haar measure dk_1 . Applying the invariance of f and the formula (2.7) for the zonal spherical function, we deduce that $\mathcal{H}f(s, k)$ does not depend on k ; specifically, $\mathcal{H}f(s, k) = \widehat{f}(s)$ where

$$(2.10) \quad \widehat{f}(s) = \int_{\mathcal{P}_m} f(w) \overline{h_s(w)} d_*w,$$

is the *zonal spherical transform* of f .

2.2. The Inversion Formula for the Helgason-Fourier Transform. Let

$$A = \{\text{diag}(a_1, \dots, a_m) : a_j > 0, j = 1, \dots, m\},$$

denote the group of diagonal positive definite matrices in G , and

$$N = \{n = (n_{ij}) \in G : n_{ij} = 0, 1 \leq j < i \leq m; n_{jj} = 1, j = 1, \dots, m\}$$

be the subgroup of G consisting of upper-triangular matrices with all diagonal entries equal to 1. It is well-known (see Terras [40], p. 20) that each

$g \in G$ can be decomposed uniquely as $g = kan$ where $k \in K$, $a \in A$, and $n \in N$; this result is the *Iwasawa decomposition*, and (k, a, n) are called the *Iwasawa coordinates* of g .

For $a, b \in \mathbb{C}$ with $\operatorname{Re}(a), \operatorname{Re}(b) > 0$, let

$$B(a, b) = \frac{\Gamma(a) \Gamma(b)}{\Gamma(a + b)}$$

denote the well-known beta function, where $\Gamma(\cdot)$ is the classical gamma function. For $s = (s_1, \dots, s_m) \in \mathbb{C}^m$, the *Harish-Chandra c -function* is

$$(2.11) \quad c_m(s) = \prod_{1 \leq i < j \leq m-1} \frac{B(\frac{1}{2}, s_i + \dots + s_j + \frac{1}{2}(j - i + 1))}{B(\frac{1}{2}, \frac{1}{2}(j - i + 1))}.$$

Let $\rho \equiv (\frac{1}{2}, \dots, \frac{1}{2}, \frac{1}{4}(1 - m))$ and set

$$(2.12) \quad \omega_m = \frac{\prod_{j=1}^m \Gamma(j/2)}{(2\pi i)^m \pi^{m(m+1)/4} m!}.$$

We shall use the notation

$$(2.13) \quad \mathbb{C}^m(\rho) = \{s \in \mathbb{C}^m : \operatorname{Re}(s) = -\rho\}$$

because this subset of \mathbb{C}^m arises frequently in the sequel, and we also define

$$(2.14) \quad d_*s = \omega_m |c_m(s)|^{-2} ds_1 \cdots ds_m$$

since this measure is ubiquitous in our development.

Let $M = \{\operatorname{diag}(\pm 1, \dots, \pm 1)\}$ be the collection of $m \times m$ diagonal matrices with entries ± 1 on the diagonal; then M is a subgroup of K and is of order 2^m . By factorizing the Haar measure dk on K , it may be shown ([40], p. 88) that there exists an invariant measure $d\bar{k}$ on the coset space K/M such that

$$\int_{\bar{k} \in K/M} d\bar{k} = 1.$$

In stating the inversion formula for the Helgason-Fourier transform, we make particular use of the notation (2.12)-(2.14). The inversion formula then is the following result.

THEOREM 2.3. (Helgason [15]) *For $f \in C_c^\infty(\mathcal{P}_m)$ and $w \in \mathcal{P}_m$,*

$$(2.15) \quad f(w) = \int_{\mathbb{C}^m(\rho)} \int_{\bar{k} \in K/M} \mathcal{H}f(s, k) p_s(k'wk) d\bar{k} d_*s.$$

In particular, if $f \in C_c^\infty(\mathcal{P}_m/K)$ then

$$(2.16) \quad f(w) = \int_{\mathbb{C}^m(\rho)} \widehat{f}(s) h_s(w) d_*s.$$

We refer to Terras [40], p. 87 ff. for a detailed treatment of this inversion formula and many references to the literature.

For the case in which $m = 1$, the Helgason-Fourier transform is the classical Mellin transform and (2.15) reduces to the corresponding classical inverse Mellin transform, *viz.*, if $f : \mathbb{R}_+ \rightarrow \mathbb{C}$ and

$$\widehat{f}(s) = \int_0^\infty t^{s-1} f(t) dt$$

then, formally,

$$f(t) = (2\pi i)^{-1} \int_{\operatorname{Re}(s)=\alpha} t^{-s} \widehat{f}(s) ds$$

where α is a constant.

2.3. The Convolution Property of the Helgason-Fourier Transform. For $f \in L^1(\mathcal{P}_m)$ and $h \in L^1(\mathcal{P}_m/K)$ we define $f * h$, the convolution of f and h , by

$$(2.17) \quad (f * h)(w) = \int_{\mathcal{P}_m} f(z) h(z^{-1/2} w z^{-1/2}) d_* z,$$

$w \in \mathcal{P}_m$. Thus, if f and h are the density functions of independent random matrices $X \in \mathcal{P}_m$ and $\varepsilon \in \mathcal{P}_m/K$, respectively, then $f * h$ is the density function of the composition $X \circ \varepsilon$.

For $f \in C_c^\infty(\mathcal{P}_m)$ and $h \in C_c^\infty(\mathcal{P}_m/K)$, the *convolution property of the Helgason-Fourier transform* is that

$$(2.18) \quad \mathcal{H}(f * h)(s, k) = \mathcal{H}f(s, k) \widehat{h}(s),$$

$s \in \mathbb{C}^m$, $k \in K$; see Terras [40], Theorem 1, p. 88.

2.4. Eigenvalues, the Laplacian, and Sobolev spaces. For $w = (w_{ij}) \in \mathcal{P}_m$, let $\partial/\partial w$ be the $m \times m$ matrix

$$\frac{\partial}{\partial w} = \left(\frac{1}{2}(1 + \delta_{ij}) \frac{\partial}{\partial w_{ij}} \right)$$

where δ_{ij} denotes Kronecker's delta. The Laplacian, Δ , on \mathcal{P}_m can be written in terms of the local coordinates w_{ij} as

$$\Delta = -\operatorname{tr} \left(\left(w \frac{\partial}{\partial w} \right)^2 \right);$$

see [40], p. 106. The power function p_s is an eigenfunction of Δ : Let

$$r_j = s_j + s_{j+1} + \cdots + s_m + \frac{1}{4}(m - 2j + 1),$$

$j = 1, \dots, m$; then

$$\Delta p_s(w) = \lambda_s p_s(w)$$

where

$$\lambda_s = -(r_1^2 + \dots + r_m^2 - \frac{1}{48}m(m^2 - 1)).$$

Since $\operatorname{Re}(s) = -\rho$ then each r_j , $j = 1, \dots, m$ is purely imaginary, so $\lambda_s > 0$. See [40], p. 49 for the explicit calculation of λ_s ; [28], p. 229 or [32], p. 283 also provide alternative approaches to that calculation.

The Helgason-Fourier transform has the property of changing the effect of invariant differential operators on functions to simple pointwise multiplication. That is, for $f \in C_c^\infty(\mathcal{P}_m)$,

$$(2.19) \quad \mathcal{H}(\Delta f)(s, k) = \lambda_s \mathcal{H}f(s, k),$$

$s \in \mathbb{C}^m$, $k \in K$ (see [40], p. 88). For $\sigma > 0$, we therefore define $\Delta^{\sigma/2}$, the $\sigma/2$ -fractional power of Δ , as the operator that satisfies the identity

$$(2.20) \quad \mathcal{H}(\Delta^{\sigma/2} f)(s, k) = \lambda_s^{\sigma/2} \mathcal{H}f(s, k),$$

for all $f \in C_c^\infty$, $s \in \mathbb{C}^m$, $k \in K$. In turn, we define the *Sobolev class* of functions,

$$(2.21) \quad H_\sigma(\mathcal{P}_m) = \{f \in C^\infty(\mathcal{P}_m) : \|\Delta^{\sigma/2} f\|^2 < \infty\},$$

where $2\sigma > \dim \mathcal{P}_m = m(m+1)/2$ and, for $f \in C^\infty(\mathcal{P}_m)$,

$$\|f\| = \left(\int_{\mathcal{P}_m} |f(w)|^2 d_* w \right)^{1/2}$$

denotes the $L^2(\mathcal{P}_m)$ -norm with respect to the invariant measure $d_* w$. For $Q > 0$, we also define the bounded Sobolev class,

$$(2.22) \quad H_\sigma(\mathcal{P}_m, Q) = \{f \in C^\infty(\mathcal{P}_m) : \|\Delta^{\sigma/2} f\|^2 < Q\},$$

where $2\sigma > \dim \mathcal{P}_m = m(m+1)/2$.

2.5. *The Plancherel Formula for the Helgason-Fourier Transform.* The Plancherel formula ([40], p. 88, Theorem 1) is that for $f \in C_c^\infty(\mathcal{P}_m)$,

$$(2.23) \quad \int_{\mathcal{P}_m} |f(w)|^2 d_* w = \int_{\mathbb{C}^m(\rho)} \int_{\bar{k} \in K/M} |\mathcal{H}f(s, \bar{k})|^2 d\bar{k} d_* s.$$

For the case in which $f \in C_c^\infty(\mathcal{P}_m/K)$, (2.23) simplifies *via* (2.10) to

$$(2.24) \quad \int_{\mathcal{P}_m} |f(w)|^2 d_* w = \int_{\mathbb{C}^m(\rho)} |\widehat{f}(s)|^2 d_* s.$$

3. Deconvolution density estimation on \mathcal{P}_m . On the group G , the deconvolution problem arises from the statistical model

$$(3.1) \quad \tilde{Y} \stackrel{\mathcal{L}}{=} \tilde{X}\tilde{\varepsilon}$$

where \tilde{X} is a random unobservable, $\tilde{\varepsilon}$ is an independent random error, and \tilde{Y} is the observed random measurement. We assume that $f_{\tilde{\varepsilon}}$, the density of $\tilde{\varepsilon}$, is known and K -biinvariant and that the unknown densities $f_{\tilde{X}}$ and $f_{\tilde{Y}}$ of \tilde{X} and \tilde{Y} , respectively, are K -invariant. Under the equivalence $K \backslash G \simeq \mathcal{P}_m$, we have $\tilde{\varepsilon} \mapsto \varepsilon$, $\tilde{X} \mapsto X$, and $\tilde{Y} \mapsto Y$ together with the identification $f_{\varepsilon}(\tilde{g}) = f_{\varepsilon}(\tilde{g}'\tilde{g})$, $f_X(\tilde{g}) = f_X(\tilde{g}'\tilde{g})$ and $f_Y(\tilde{g}) = f_Y(\tilde{g}'\tilde{g})$, $g \in G$. Since \tilde{X} and $\tilde{\varepsilon}$ are independent then X and ε also are independent, and (3.1) implies that

$$(3.2) \quad Y \stackrel{\mathcal{L}}{=} \varepsilon^{1/2} X \varepsilon^{1/2};$$

hence $f_Y = f_X * f_{\varepsilon}$.

Applying to (3.2) the convolution property (2.18) of the Helgason-Fourier transform, we obtain

$$(3.3) \quad \mathcal{H}f_Y(s, k) = \mathcal{H}f_X(s, k) \widehat{f}_{\varepsilon}(s),$$

$s \in \mathbb{C}^m$, $k \in K$. Given a random sample Y_1, \dots, Y_n from Y , we estimate the density function f_X as follows. We form $\mathcal{H}_n f_Y$, the *empirical Helgason-Fourier transform*,

$$(3.4) \quad \mathcal{H}_n f_Y(s, k) = \frac{1}{n} \sum_{\ell=1}^n \overline{p_s(k' Y_{\ell} k)}.$$

Substituting (3.4) in (3.3), together with the assumption that $\widehat{f}_{\varepsilon}(s) \neq 0$, $s \in \mathbb{C}^m$, we obtain

$$(3.5) \quad \mathcal{H}_n f_X(s, k) = \frac{\mathcal{H}_n f_Y(s, k)}{\widehat{f}_{\varepsilon}(s)},$$

$s \in \mathbb{C}^m$, $k \in K$.

Two elementary facts about $\mathcal{H}_n f_Y$ are:

$$(3.6) \quad \mathbb{E} \mathcal{H}_n f_Y(s, k) = \mathcal{H}f_Y(s, k),$$

i.e., $\mathcal{H}_n f_Y$ is an unbiased estimator of $\mathcal{H}f_Y$; and,

$$(3.7) \quad \text{Var}(\mathcal{H}_n f_Y(s, k)) = \frac{1}{n} \left(\mathcal{H}f_Y(-2\rho, k) - |\mathcal{H}f_Y(s, k)|^2 \right)$$

when $\operatorname{Re}(s) = -\rho$.

In analogy with classical Euclidean deconvolution, we introduce a smoothing parameter $T = T(n)$ where $T(n) \rightarrow \infty$ as $n \rightarrow \infty$, and then we apply the Helgason-Fourier inversion formula (2.15) using a spectral cut-off.

We introduce the notation

$$(3.8) \quad \mathbb{C}^m(\rho, T) = \{s \in \mathbb{C}^m(\rho) : \lambda_s < T\}$$

where $\mathbb{C}^m(\rho)$ is defined in (2.13). As an estimator of the population density f_X , we define the density estimator f_X^n given by

$$(3.9) \quad f_X^n(w) = \int_{\mathbb{C}^m(\rho, T)} \int_{\bar{k} \in K/M} \frac{\mathcal{H}_n f_Y(s, \bar{k})}{\widehat{f}_\varepsilon(s)} p_s(\bar{k}' w \bar{k}) d\bar{k} d_* s,$$

$w \in \mathcal{P}_m$. The estimator f_X^n will serve as our nonparametric deconvolution estimator of the density f_X .

We note that

$$(3.10) \quad \begin{aligned} \mathbb{E} f_X^n(w) &= \int_{\mathbb{C}^m(\rho, T)} \int_{\bar{k} \in K/M} \frac{\mathbb{E} \mathcal{H}_n f_Y(s, \bar{k})}{\widehat{f}_\varepsilon(s)} p_s(\bar{k}' w \bar{k}) d\bar{k} d_* s \\ &= \int_{\mathbb{C}^m(\rho, T)} \int_{\bar{k} \in K/M} \frac{\mathcal{H} f_Y(s, \bar{k})}{\widehat{f}_\varepsilon(s)} p_s(\bar{k}' w \bar{k}) d\bar{k} d_* s \\ &= \int_{\mathbb{C}^m(\rho, T)} \int_{\bar{k} \in K/M} \mathcal{H} f_X(s, \bar{k}) p_s(\bar{k}' w \bar{k}) d\bar{k} d_* s. \end{aligned}$$

Under certain smoothness assumptions on f_ε , (3.10) converges to $f_X(w)$ as $T \rightarrow \infty$.

Let $\{a_n\}$ and $\{b_n\}$ be sequences of real numbers. We write $a_n \ll b_n$ to mean $a_n \leq C b_n$ for some constant $C > 0$, as $n \rightarrow \infty$ (the Vinogradov notation). We use the notation $a_n = o(b_n)$ to mean $a_n/b_n \rightarrow 0$, as $n \rightarrow \infty$. We also write $a_n \asymp b_n$ if both $a_n \ll b_n$ and $b_n \ll a_n$; and we write $a_n \sim b_n$ if $a_n/b_n \rightarrow 1$ as $n \rightarrow \infty$.

For technical reasons, we will also assume the moment condition

$$(3.11) \quad \int_{\mathcal{P}_m} |w_1|^{-1} \cdots |w_{m-1}|^{-1} |w|^{(m-1)/2} f_Y(w) d_* w < \infty,$$

on the principal minors $|y_1|, \dots, |y_m|$ of $y \in \mathcal{P}_m$. This assumption will be maintained throughout the rest of the paper.

THEOREM 3.1. *Suppose there exists $\beta \geq 0$ such that*

$$(3.12) \quad |\widehat{f}_\varepsilon(s)|^{-2} \ll T^\beta$$

as $T \rightarrow \infty$, for all $s \in \mathbb{C}^m(\rho, T)$. If $f_X \in H_\sigma(\mathcal{P}_m, Q)$ and $\sigma > \frac{1}{2} \dim \mathcal{P}_m \equiv m(m+1)/4$ then, as $n \rightarrow \infty$,

$$\mathbb{E} \|f_X^n - f_X\|^2 \ll n^{-2\sigma/(2\sigma+2\beta+\dim \mathcal{P}_m)}.$$

As an example, consider the special case in which

$$(3.13) \quad \widehat{f}_\varepsilon(s) = (1 + \gamma \lambda_s)^{-\beta},$$

$s \in \mathbb{C}^m$, where $\gamma > 0$ is a scale parameter. Since \widehat{f}_ε does not depend on $k \in K$ then the underlying density function f is K -invariant. By the Helgason-Fourier inversion formula (2.16),

$$(3.14) \quad f_\varepsilon(w) = \int_{\mathbb{C}^m(\rho)} (1 + \gamma \lambda_s)^{-\beta} h_s(w) d_*s,$$

$w \in \mathcal{P}_m$. Letting $\beta \rightarrow 0$, we obtain $\widehat{f}_\varepsilon(s) \rightarrow 1$, so the underlying probability distribution approaches the Dirac measure concentrated at \mathbf{I}_m , the identity matrix in \mathcal{P}_m , and the interpretation here is that observations are made without error. This result corresponds to the special case of Theorem 3.1 in which $\beta = 0$, and hence we obtain the following result.

COROLLARY 3.2. *Suppose the distribution of ε is concentrated at \mathbf{I}_m . If $f_X \in H_\sigma(\mathcal{P}_m, Q)$ where $\sigma > \frac{1}{2} \dim \mathcal{P}_m$ then as $n \rightarrow \infty$,*

$$\mathbb{E} \|f_X^n - f_X\|^2 \ll n^{-2\sigma/(2\sigma+\dim \mathcal{P}_m)}.$$

We shall also obtain a result for the situation in which the hypothesis (3.12) in Theorem 3.1 is replaced by an exponential bound. In such a situation, we have the following result.

THEOREM 3.3. *Suppose there exists $\beta, \gamma > 0$ such that*

$$(3.15) \quad |\widehat{f}_\varepsilon(s)|^{-2} \ll \exp(T^\beta/\gamma),$$

as $T \rightarrow \infty$, for all $s \in \mathbb{C}^m(\rho, T)$. If $f_X \in H_\sigma(\mathcal{P}_m, Q)$ with $\sigma > \frac{1}{2} \dim \mathcal{P}_m$ then, as $n \rightarrow \infty$,

$$\mathbb{E} \|f_X^n - f_X\|^2 \ll (\log n)^{-\sigma/\beta}.$$

In this situation, we consider the special case in which

$$(3.16) \quad \widehat{f}_\varepsilon(s) = \exp(-\gamma^{-1} \lambda_s^\beta),$$

$s \in \mathbb{C}^m$, where $\gamma > 0$ is a scale parameter. Again by the inversion formula (2.16), the underlying density function is

$$(3.17) \quad f_\varepsilon(w) = \int_{\mathbb{C}^m(\rho)} \exp(-\gamma^{-1} \lambda_s^\beta) h_s(w) d_* s,$$

$w \in \mathcal{P}_m$. The case in which $\beta = 1$ is particularly important and is called the *heat* or *Gaussian kernel*, since the latter is the fundamental solution to the heat equation on \mathcal{P}_m : $(\Delta - (\partial/\partial t))\phi = 0$, $t > 0$; see Terras [40], pp. 106-107. As a consequence, we obtain the following result.

COROLLARY 3.4. *Suppose that f_ε is Gaussian. If $f_X \in H_\sigma(\mathcal{P}_m, Q)$ where $\sigma > \frac{1}{2} \dim \mathcal{P}_m$ then, as $n \rightarrow \infty$,*

$$\mathbb{E} \|f_X^n - f_X\|^2 \ll (\log n)^{-\sigma}.$$

4. The Wishart distribution. A case which is familiar in multivariate statistics is that of the Wishart distribution, $W_m(N, \Sigma)$, where $\Sigma \in \mathcal{P}_m$ and $N > m - 1$. For $s = (s_1, \dots, s_m) \in \mathbb{C}^m$ define the multivariate gamma function,

$$(4.1) \quad \Gamma_m(s_1, \dots, s_m) = \pi^{m(m-1)/4} \prod_{j=1}^m \Gamma(s_j + \dots + s_m - \frac{1}{2}(j-1)),$$

where $\operatorname{Re}(s_j + \dots + s_m) > (j-1)/2$, $j = 1, \dots, m$. Relative to the invariant measure $d_* w$ in (2.8), the probability density function of the standard Wishart distribution $W_m(N, \mathbf{I}_m)$ is

$$(4.2) \quad f_\varepsilon(w) = \frac{1}{\Gamma_m(0, \dots, 0, N/2)} |\frac{1}{2}w|^{N/2} \exp(-\frac{1}{2}\operatorname{tr} w),$$

$w \in \mathcal{P}_m$. We note that (4.2) is K -invariant and its Helgason-Fourier transform is well-known (Muirhead [28], p. 248; Terras [40], pp. 85-86),

$$(4.3) \quad \widehat{f}_\varepsilon(s) = \frac{\Gamma_m((0, \dots, 0, N/2) + s^*)}{\Gamma_m(0, \dots, 0, N/2)} h_s(\frac{1}{2}\mathbf{I}_m)$$

where $s^* = (s_{m-1}, s_{m-2}, \dots, s_2, s_1, -(s_1 + \dots + s_m))$.

We have the following result.

LEMMA 4.1. *For $N > m - 1$, the Wishart distribution $W_m(N, \mathbf{I}_m)$ satisfies*

$$|\widehat{f}_\varepsilon(s)|^{-2} \ll \exp(\pi T^{1/2}),$$

as $T \rightarrow \infty$, where $s \in \mathbb{C}^m(\rho, T)$.

Consequently we deduce the following result.

THEOREM 4.2. *Suppose that ε follows the Wishart distribution (4.2) with $N > m - 1$. If $f_X \in H_\sigma(\mathcal{P}_m, Q)$ with $\sigma > \dim \mathcal{P}_m/2$ then, as $n \rightarrow \infty$,*

$$\mathbb{E}\|f_X^n - f_X\|^2 \ll (\log n)^{-2\sigma}.$$

From this result, one can see that the Wishart distribution has faster convergence in its Helgason-Fourier transform as compared to the Gaussian distribution. This of course results in a slower recovery in the corresponding deconvolution problem.

5. Proofs. The strategy of the proofs is, first, to decompose the integrated mean-squared error into its variance and bias components, *viz.*,

$$(5.1) \quad \begin{aligned} \mathbb{E}\|f_X^n - f_X\|^2 &= \mathbb{E}\|(f_X^n - \mathbb{E}f_X^n) + (\mathbb{E}f_X^n - f_X)\|^2 \\ &= \mathbb{E}\|f_X^n - \mathbb{E}f_X^n\|^2 + \|\mathbb{E}f_X^n - f_X\|^2, \end{aligned}$$

and, last, to estimate each component separately using estimates based on the Plancherel formula and the inversion formula for the Helgason-Fourier transform.

5.1. The Integrated Bias.

LEMMA 5.1. *Suppose that $f_X \in H_\sigma(\mathcal{P}_m, Q)$ and $\sigma > \dim \mathcal{P}_m/2$. Then*

$$\|\mathbb{E}f_X^n - f_X\|^2 \ll T^{-\sigma}.$$

PROOF. By (3.10), we have for $w \in \mathcal{P}_m$,

$$(5.2) \quad \begin{aligned} \mathbb{E}f_X^n(w) - f_X(w) &= \int_{\mathbb{C}^m(\rho, T)} \int_{\bar{k} \in K/M} \mathcal{H}f_X(s, \bar{k}) p_s(\bar{k}' w \bar{k}) d\bar{k} d_*s \\ &\quad - \int_{\mathbb{C}^m(\rho)} \int_{\bar{k} \in K/M} \mathcal{H}f_X(s, \bar{k}) p_s(\bar{k}' w \bar{k}) d\bar{k} d_*s \\ &= - \int_{\lambda_s > T, \operatorname{Re}(s) = -\rho} \int_{\bar{k} \in K/M} \mathcal{H}f_X(s, \bar{k}) p_s(\bar{k}' w \bar{k}) d\bar{k} d_*s \end{aligned}$$

Applying the Plancherel formula, (2.23), to (5.2), we obtain

$$\|\mathbb{E}f_X^n - f_X\|^2 = \int_{\lambda_s > T, \operatorname{Re}(s) = -\rho} \int_{K/M} |\mathcal{H}(f_X)(s, \bar{k})|^2 d\bar{k} d_*s.$$

Consequently,

$$\begin{aligned}
 \|\mathbb{E}f_X^n - f_X\|^2 &= \int_{\lambda_s \geq T, \operatorname{Re}(s) = -\rho} \int_{K/M} |\mathcal{H}f_X(s, \bar{k})|^2 d\bar{k} d_*s \\
 (5.3) \quad &\leq T^{-\sigma} \int_{\lambda_s \geq T, \operatorname{Re}(s) = -\rho} \int_{K/M} \lambda_s^\sigma |\mathcal{H}f_X(s, \bar{k})|^2 d\bar{k} d_*s \\
 &\leq T^{-\sigma} \int_{\mathbb{C}^m(\rho)} \int_{K/M} \lambda_s^\sigma |\mathcal{H}f_X(s, \bar{k})|^2 d\bar{k} d_*s.
 \end{aligned}$$

By (2.20),

$$\lambda_s^\sigma |\mathcal{H}f_X(s, k)|^2 \equiv |\lambda_s^{\sigma/2} \mathcal{H}f_X(s, k)|^2 = |\mathcal{H}(\Delta^{\sigma/2} f_X)(s, k)|^2,$$

therefore

$$\begin{aligned}
 \|\mathbb{E}f_X^n - f_X\|^2 &\leq T^{-\sigma} \int_{\mathbb{C}^m(\rho)} \int_{\bar{k} \in K/M} |\mathcal{H}(\Delta^{\sigma/2} f_X)(s, k)|^2 d\bar{k} d_*s \\
 &= T^{-\sigma} \int_{\mathcal{P}_m} |\Delta^{\sigma/2} f_X(w)|^2 d_*w,
 \end{aligned}$$

where the equality follows from the Plancherel formula. By assumption, $f_X \in H_\sigma(\mathcal{P}_m, Q)$ and $\sigma > \frac{1}{2} \dim \mathcal{P}_m$; therefore the latter integral is bounded above by Q , so we obtain

$$\|\mathbb{E}f_X^n - f_X\|^2 \leq Q T^{-\sigma},$$

and the proof is complete. \square

5.2. The Integrated Variance. To obtain bounds for the integrated variance, several preliminary calculations are needed. In particular, we begin with the variance calculation of the empirical Helgason-Fourier transform, which has similarities to the usual empirical characteristic function; see Feuerverger and Mureika [5].

LEMMA 5.2. *For $s \in \mathbb{C}^m(\rho)$ and $k \in K/M$,*

$$\mathbb{E}|\mathcal{H}_n f_Y(s, k) - \mathbb{E}\mathcal{H}_n f_Y(s, k)|^2 = \frac{1}{n} (|\mathcal{H}f_Y(-2\rho, k)|^2 - |\mathcal{H}f_Y(s, k)|^2).$$

PROOF. By (3.4),

$$\begin{aligned}
 |\mathcal{H}_n f_Y(s, k)|^2 &= \overline{\mathcal{H}_n f_Y(s, k)} \mathcal{H}_n f_Y(s, k) \\
 &= \frac{1}{n^2} \sum_{j, \ell=1}^n \overline{p_s(k'Y_j k)} p_s(k'Y_\ell k) \\
 (5.4) \quad &= \frac{1}{n^2} \left\{ \sum_{j=1}^n |p_s(k'Y_j k)|^2 + \sum_{j \neq \ell} \overline{p_s(k'Y_j k)} p_s(k'Y_\ell k) \right\}.
 \end{aligned}$$

Observe also that

$$\begin{aligned} p_s(w)\overline{p_s(w)} &= |w_1|^{s_1} \cdots |w_m|^{s_m} \overline{|w_1|^{s_1} \cdots |w_m|^{s_m}} \\ &= |w_1|^{2\operatorname{Re}(s_1)} \cdots |w_m|^{2\operatorname{Re}(s_m)} \\ &= p_{-2\rho}(w), \end{aligned}$$

since $\operatorname{Re}(s) = -\rho$. Applying this result to (5.4) and taking expectations, we obtain

$$\begin{aligned} \mathbb{E}|\mathcal{H}_n f_Y(s, k)|^2 &= \frac{1}{n^2} \mathbb{E} \left\{ \sum_{j=1}^n |p_s(k'Y_j k)|^2 + \sum_{j \neq \ell} \overline{p_s(k'Y_j k)} p_s(k'Y_\ell k) \right\} \\ &= \frac{1}{n^2} \left\{ \sum_{j=1}^n \mathbb{E} p_{-2\rho}(k'Y_j k) + \sum_{j \neq \ell} \mathbb{E} \overline{p_s(k'Y_j k)} \mathbb{E} p_s(k'Y_\ell k) \right\} \\ &= \frac{1}{n} \mathcal{H} f_Y(-2\rho, k) + \frac{n-1}{n} |\mathcal{H} f_Y(s, k)|^2, \end{aligned}$$

where the last equality follows from the fact that Y_1, \dots, Y_n are independent and identically distributed as Y , and because $\mathbb{E} p_s(k'Y k) = \mathcal{H} f_Y(s, k)$ by (3.6). \square

This allows us to establish the following bound.

LEMMA 5.3. *As $T \rightarrow \infty$,*

$$\mathbb{E} \|f_X^n - \mathbb{E} f_X^n\|^2 \ll \sup_{s \in \mathbb{C}^m(\rho)} |\widehat{f}_\varepsilon(s)|^{-2} \frac{T^{\dim \mathcal{P}_m/2}}{n}.$$

PROOF. By the Plancherel formula, (2.23),

$$\begin{aligned} \mathbb{E} \|f_X^n - \mathbb{E} f_X^n\|^2 &= \int_{\mathbb{C}^m(\rho, T)} \int_{K/M} \mathbb{E} |\mathcal{H}_n f_X(s, k) - \mathbb{E} \mathcal{H}_n f_X(s, k)|^2 d\bar{k} d_* s \\ &= \int_{\mathbb{C}^m(\rho, T)} \int_{K/M} \mathbb{E} |\mathcal{H}_n f_Y(s, k) - \mathbb{E} \mathcal{H}_n f_Y(s, k)|^2 d\bar{k} |\widehat{f}_\varepsilon(s)|^{-2} d_* s \\ &\leq \frac{1}{n} \sup_{\lambda_s < T, \operatorname{Re}(s) = -\rho} |\widehat{f}_\varepsilon(s)|^{-2} \int_{K/M} |\mathcal{H} f_Y(-2\rho, \bar{k})|^2 d\bar{k} \int_{\mathbb{C}^m(\rho, T)} d_* s \\ &\ll \sup_{\lambda_s < T, \operatorname{Re}(s) = -\rho} |\widehat{f}_\varepsilon(s)|^{-2} \frac{T^{\dim \mathcal{P}_m/2}}{n}, \end{aligned}$$

as $T \rightarrow \infty$.

In the above calculation, we used the moment bound (3.11) together with the fact that $|\mathcal{H}f_Y(-2\rho, \bar{k})|^2$ varies continuously with respect to $\bar{k} \in K/M$, with the latter being compact. In addition, we used the fact that, as $T \rightarrow \infty$,

$$\sup_{\mathbb{C}^m(\rho, T)} |c_m(s)|^{-2} \ll T^{m(m-1)/4},$$

a result which follows from Helgason [15], p. 450, Proposition 7.2. \square

The proofs of all the theorems now are obtained by applying Lemma 5.1 and Lemma 5.3 to (5.1).

5.3. *Proof of Theorem 3.1.* By the condition of Theorem 3.1, (5.1) reduces to

$$\mathbb{E} \|f_X^n - f_X\|^2 \ll n^{-1} T^{\beta + \dim \mathcal{P}_m/2} + T^{-\sigma},$$

as $T \rightarrow \infty$. Thus, we choose $T \asymp T^{2/(2\beta + 2\sigma + \dim \mathcal{P}_m)}$ to optimize the upper bound.

5.4. *Proof of Theorem 3.3.* By the condition of Theorem 3.3, (5.1) reduces to

$$\mathbb{E} \|f_X^n - f_X\|^2 \ll n^{-1} T^{\dim \mathcal{P}_m/2} \exp(\gamma^{-1} T^\beta) + T^{-\sigma},$$

as $T \rightarrow \infty$. Thus, we choose $T \asymp (\log n)^{1/\beta}$ to optimize the upper bound.

5.5. *Proof of Lemma 4.1.* Let $s_j = -(1/2) + i\sigma_j$, for $j = 1, \dots, m-1$, and $s_m = ((m-1)/4) + i\sigma_m$, where $i^2 = -1$ and $\sigma_1, \dots, \sigma_m \in \mathbb{R}$. Define $s_j^* = s_{m-j}$ for $j = 1, \dots, m-1$, and $s_m^* = -(s_1 + \dots + s_m)$. Consequently,

$$(5.5) \quad s_j^* + \dots + s_m^* + \frac{N-j+1}{2} = \frac{2N-m+1}{4} - i(\sigma_{m-j+1} + \dots + \sigma_m).$$

By Stirling's approximation,

$$(5.6) \quad \Gamma(z) \sim z^{z-1/2} e^{-z} \sqrt{2\pi},$$

for $z \in \mathbb{C}$ with $0 < a \leq \operatorname{Re}(z) < b$, as $|z| \rightarrow \infty$; therefore

$$\begin{aligned} & \Gamma\left(s_j^* + \dots + s_m^* + \frac{N-j+1}{2}\right) \\ & \sim \sqrt{2\pi} \left(\frac{2N-m+1}{4} - i(\sigma_{m-j+1} + \dots + \sigma_m) \right)^{\frac{2N-m-1}{4} - i(\sigma_{m-j+1} + \dots + \sigma_m)} \\ & \quad \times \exp \left\{ \frac{2N-m+1}{4} - i(\sigma_{m-j+1} + \dots + \sigma_m) \right\} \end{aligned}$$

as $|\sigma_{m-j+1} + \cdots + \sigma_m| \rightarrow \infty$, for $j = 1, \dots, m$. Consequently,

$$\begin{aligned}
 & \left| \Gamma \left(s_j^* + \cdots + s_m^* + \frac{2N - m + 1}{4} \right) \right|^2 \\
 (5.7) \quad & \sim 2\pi \left(\left(\frac{2N - m + 1}{4} \right)^2 + (\sigma_{m-j+1} + \cdots + \sigma_m)^2 \right)^{(2N-m-1)/2} \\
 & \times \exp \left(2(\sigma_{m-j+1} + \cdots + \sigma_m) \arctan \left(\frac{4(\sigma_{m-j+1} + \cdots + \sigma_m)}{2N - m + 1} \right) \right) \\
 & \times \exp \left(\frac{2N - m + 1}{2} \right) \text{ and}
 \end{aligned}$$

as $|\sigma_{m-j+1} + \cdots + \sigma_m| \rightarrow \infty$, for $j = 1, \dots, m$.

By the condition that $\lambda_s < T$, it follows that there exists a positive constant C , independent of $\sigma_1, \dots, \sigma_m$, such that

$$(5.8) \quad \left| \Gamma_m(s_1^*, \dots, s_{m-1}^*, s_m^* + N/2) \right|^{-2} \leq C \exp(\pi T^{1/2})$$

as $T \rightarrow \infty$. Since $|h_s(\frac{1}{2}\mathbf{I}_m)|^2 = 1$ then it follows that as $T \rightarrow \infty$,

$$|\hat{f}_\varepsilon(s)|^{-2} \ll \exp(\pi T^{1/2}).$$

Acknowledgments. Portions of this research was done during fall, 2006 while D.R. was a Research Fellow at the Statistical and Applied Mathematical Sciences Institute, Research Triangle Park, NC. We are grateful to the entire SAMSI community for the warm and generous hospitality accorded to him during his stay.

REFERENCES

- [1] Basu, S., Fletcher, T. and Whitaker, S. (2006). Rician noise removal in diffusion tensor MRI. In: *Medical Image Computing and Computer-Assisted Intervention, MICCAI 2006*, Lect. Notes Computer Science, **4190**, 117–125, New York: Springer.
- [2] Damon, B. M., Ding, Z., Anderson, A. W., Freyer, A. S. and Gore, J. C. (2002). Validation of diffusion tensor MRI-based muscle fiber tracking. *Magn. Reson. Med.*, **48**, 97–104
- [3] Diggle, P. J. and Hall, P. (1993). A Fourier approach to nonparametric deconvolution of a density estimate. *J. Roy. Statist. Soc., Ser. B*, **55**, 523–531.
- [4] Fan, J. (1992). Global behaviour of deconvolution kernel estimates. *Statist. Sinica*, **1**, 541–551.
- [5] Feuerverger, A. and Mureika, R. A. (1977). The empirical characteristic function and its applications. *Ann. Statist.*, **5**, 88–97.
- [6] Fletcher, P. T. and Joshi, S. (2004). Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In: *Proc. ECCV 2004 Workshop on Computer Vision Approaches to Medical Image Analysis*, Lect. Notes Comp. Sci., **3117**, 87–98, New York: Springer.

- [7] Fletcher, P. T. and Joshi, S. (2007). Riemannian geometry for the statistical analysis of diffusion tensor data. *Signal Processing*, **87**, 250–262.
- [8] Graczyk, P. (1992). A Central Limit Theorem on the space of positive definite symmetric matrices. *Ann. Inst. Fourier*, **42**, 857–874.
- [9] Haff, L. R. (1979). Estimation of the inverse covariance matrix: Random mixtures of the inverse Wishart matrix and the identity. *Ann. Statist.*, **7**, 1264–1276.
- [10] Haff, L. R. (1980). Empirical Bayes estimation of the multivariate normal covariance matrix. *Ann. Statist.*, **8**, 586–597.
- [11] Haff, L. R. (1991). The variational form of certain Bayes estimators. *Ann. Statist.*, **19**, 1163–1190.
- [12] Hasan, K. M., Basser, P. J., Parker, D. L. and Alexander, A. L. (2001). Analytical computation of the eigenvalues and eigenvectors in DT-MRI. *J. Magn. Res.*, **152**, 41–47.
- [13] Healy, D. M. and Kim, P. T. (1996). An empirical Bayes approach to directional data and efficient computation on the sphere. *Ann. Statist.*, **24**, 232–254.
- [14] Healy, D. M., Hendriks, H. and Kim, P. T. (1998). Spherical deconvolution. *J. Multivariate Anal.*, **67**, 1–22.
- [15] Helgason, S. (1984). *Groups and Geometric Analysis*. Orlando: Academic Press.
- [16] Hendriks, H. (1990). Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansions. *Ann. Statist.*, **18**, 832–849.
- [17] Johnstone, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.*, **29**, 295–327.
- [18] Kim, P. T. (1998). Deconvolution density estimation on $SO(N)$. *Ann. Statist.*, **26**, 1083–1102.
- [19] Kim, P. T. and Koo, J.-Y. (2000). Directional mixture models and optimal estimation of the mixing density. *Canad. J. Statist.*, **28**, 383–398.
- [20] Kim, P. T. and Koo, J.-Y. (2002). Optimal spherical deconvolution. *J. Multivariate Anal.*, **80**, 21–42.
- [21] Kim, P. T. and Richards, D. St. P. (2001). Deconvolution density estimation on compact Lie groups. *Contemp. Math.*, **287**, 155–171.
- [22] Koltchinskii, V., Sakhanenko, L. and Cao, S. (2007). Integral curves of noisy vector fields and statistical problems in diffusion tensor imaging: Nonparametric kernel estimation and hypotheses testing. *Ann. Statist.*, to appear.
- [23] Koo, J.-Y. (1993). Optimal rates of convergence for nonparametric statistical inverse problems. *Ann. Statist.*, **21**, 590–599.
- [24] Le Bihan, D. (1995). *Diffusion and Perfusion Magnetic Resonance Imaging: Applications to Functional MRI*. New York: Raven Press.
- [25] Letac, G. and Massam, H. (2004). All invariant moments of the Wishart distribution. *Scand. J. Statist.*, **31**, 295–318.
- [26] Mair, B. A. and Ruymgaart, F. H. (1996). Statistical inverse estimation in Hilbert scales. *SIAM J. Appl. Math.*, **56**, 1424–1444.
- [27] Matthews, P. M. and Arnold, D. L. (2001). Magnetic resonance imaging of multiple sclerosis: new insights linking pathology to clinical evolution. *Current Opinion in Neurology*, **14**, 279–287.
- [28] Muirhead, R. J. (1978). *Aspects of Multivariate Statistics*. Wiley, New York.
- [29] Neumann-Haefelin T., Moseley M. E. and Albers, G. W. (2000). New magnetic resonance imaging methods for cerebrovascular disease. *Ann. Neurol.*, **47**, 559–570.
- [30] Pesenson, I. (2006). Deconvolution of band-limited functions on non-compact symmetric spaces. *Houston J. Math.*, **32**, 183–204.
- [31] Pomara, N., Crandalla, D. T., Choia, S. J., Johnson, G. and Lim, K. O. (2001).

- White-matter abnormalities in HIV-1 infection: A diffusion tensor imaging study. *Psychiatry Research: Neuroimaging*, **106**, 15–24.
- [32] Richards, D. St. P. (1985). Applications of invariant differential operators to multivariate distribution theory. *SIAM J. Appl. Math.*, **45**, 280–288.
- [33] Richards, D. St. P. (1989). The Central Limit Theorem on spaces of positive definite matrices. *J. Multivariate Anal.*, **29**, 326–332.
- [34] Rosenbloom, M., Sullivan, J. V. and Pfefferbaum, A. (2003). Using magnetic resonance imaging and diffusion tensor imaging to assess brain damage in alcoholics. *Alcohol Research & Health*, **27**, 146–152.
- [35] Schwartzman, A. (2006). *Random Ellipsoids and False Discovery Rates: Statistics for Diffusion Tensor Imaging Data*. Ph.D. dissertation, Stanford University.
- [36] Schwartzman, A., Dougherty, R. F. and Taylor, J. E. (2005). Cross-subject comparison of principal diffusion direction maps. *Magn. Reson. Med.*, **53**, 1423–1431.
- [37] Takemura, A. and Sheena, Y. (2005). Distribution of eigenvalues and eigenvectors of Wishart matrix when the population eigenvalues are infinitely dispersed and its application to minimax estimation of covariance matrix. *J. Multivariate Anal.*, **94**, 271–299.
- [38] Terras, A. (1985). *Harmonic Analysis on Symmetric Spaces and Applications, I*. Springer-Verlag, New York.
- [39] Terras, A. (1987). Asymptotics of special functions and the Central Limit Theorem on the space \mathcal{P}_n of positive $n \times n$ matrices. *J. Multivariate Anal.*, **23**, 13–36.
- [40] Terras, A. (1988). *Harmonic Analysis on Symmetric Spaces and Applications, II*. Springer-Verlag, New York.
- [41] Zhang, C. H. (1990). Fourier methods for estimating mixing densities and distributions. *Ann. Statist.*, **18**, 806–830.
- [42] Zhu, H., Zhang, H., Ibrahim, J. G. and Peterson, B. G. (2007). Statistical analysis of diffusion tensors in diffusion-weighted magnetic image resonance data. *J. Amer. Statist. Assoc.* (with discussion), to appear.

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF GUELPH
GUELPH, ONTARIO N1G 2W1, CANADA
E-MAIL: pkim@uoguelph.ca

DEPARTMENT OF STATISTICS
THE PENNSYLVANIA STATE UNIVERSITY
UNIVERSITY PARK, PA 16802
AND
THE STATISTICAL AND APPLIED
MATHEMATICAL SCIENCES INSTITUTE
RESEARCH TRIANGLE PARK, NC 27709, USA
E-MAIL: richards@stat.psu.edu