

# Diffusion Tensor Imaging and Deconvolution on Spaces of Positive Definite Symmetric Matrices

Peter Kim, Donald Richards

► **To cite this version:**

Peter Kim, Donald Richards. Diffusion Tensor Imaging and Deconvolution on Spaces of Positive Definite Symmetric Matrices. Xavier Pennec. 2nd MICCAI Workshop on Mathematical Foundations of Computational Anatomy, Oct 2008, New-York, United States. pp.140-149, 2008. <inria-00632882>

**HAL Id: inria-00632882**

**<https://hal.inria.fr/inria-00632882>**

Submitted on 16 Oct 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Diffusion Tensor Imaging and Deconvolution on Spaces of Positive Definite Symmetric Matrices

Peter T. Kim \* and Donald St. P. Richards \*\*

University of Guelph and Penn State University

**Abstract.** Diffusion tensor imaging can be studied as a deconvolution density estimation problem on the space of positive definite symmetric matrices. We develop a nonparametric estimator for the common 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 smoothness conditions on the density of the intrinsic error in the random sample, we derive bounds on the rates of convergence of our nonparametric estimator to the true density.

**Key words:** Deconvolution; diffusion tensor imaging; Harish-Chandra  $c$ -function; Helgason-Fourier transform; inverse problems; Laplace-Beltrami operator; magnetic resonance imaging; Sobolev spaces; Wishart distribution.

## 1 Introduction

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 purposes 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  $3 \times 3$  positive-definite covariance matrix of the local diffusion process at the given location (Fletcher and Joshi [3, 4]). 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 [9], Hasan, *et al.* [5]).

In DTI brain imaging, the diffusion of water molecules within and between *voxels*, the three-dimensional volume elements that constitute an image, reveal

---

\* Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario N1G 2W1, Canada. Research supported in part by the Natural Sciences and Engineering Research Council of Canada, grant DG-46204. Email: pkim@uoguelph.ca.

\*\* Department of Statistics, The Pennsylvania State University, University Park, PA 16802. Research supported in part by the U.S. National Science Foundation, grants DMS-0112069 and DMS-0705210. Email: richards@stat.psu.edu.

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.* [12], Rosenbloom, *et al.* [14], Pomara, *et al.* [13], Matthews and Arnold [10]). 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.* [8]; Schwartzman, *et al.* [15, 16]; Zhu, *et al.* [18]).

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 [15] 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. 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. 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 noncompact 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, 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, see for example Muirhead [11]. 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 [6], Terras [17]), 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  $\mathcal{P}_m$  and we provide the explicit details in the case of the Wishart distribution. All proofs can be found in Kim and Richards [7].

## 2 Preliminaries

Throughout the paper, we denote by  $G$  the general linear group  $\text{GL}(m, \mathbb{R})$ , of  $m \times m$  nonsingular real matrices and by  $K$  the group,  $\text{O}(m)$ , 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

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

$g \in G$ ,  $x \in \mathcal{P}_m$ , 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

$$K \backslash G \rightarrow \mathcal{P}_m, \quad Kg \mapsto g'g. \quad (1)$$

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 (1) 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  $\tilde{f}(g) = \tilde{f}(k_1gk_2)$  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 (1),  $K\tilde{X} \mapsto X$ , equivalently  $\tilde{X}'\tilde{X} = X$  and, similarly,  $K\tilde{\varepsilon} \mapsto \varepsilon$ . Then  $\tilde{X}\tilde{\varepsilon} \in G$  is mapped *via* (1) to

$$\tilde{X}\tilde{\varepsilon} \mapsto (\tilde{X}\tilde{\varepsilon})'(\tilde{X}\tilde{\varepsilon}) \equiv \tilde{\varepsilon}'\tilde{X}'\tilde{X}\tilde{\varepsilon} = \tilde{\varepsilon}'X\tilde{\varepsilon} \in \mathcal{P}_m. \quad (2)$$

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  $\varepsilon$ . Consequently, we define  $\varepsilon^{1/2} = k'\Lambda^{1/2}k$  and note that, by  $K$ -biinvariance of  $\tilde{\varepsilon}$ , the relationship (2) 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. Suppose that  $X \in \mathcal{P}_m$  and  $\varepsilon \in \mathcal{P}_m/K$  are random matrices. Then the composition of  $X$  and  $\varepsilon$  is

$$X \circ \varepsilon = \varepsilon^{1/2} X \varepsilon^{1/2} \quad (3)$$

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

There is an alternative approach leading to compositions (3). 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 a unique  $V \in G$  such that

$$Y = V X V'. \quad (4)$$

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 (4); see Terras [17], 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 (4) in which  $V \in G/K$ . As noted in [14], 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 (4), 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  $V \stackrel{L}{\equiv} k'V = k'k\varepsilon^{1/2} = \varepsilon^{1/2}$ ; hence (4) reduces to  $Y \stackrel{L}{\equiv} \varepsilon^{1/2} X \varepsilon^{1/2} \equiv X \circ \varepsilon$ , which agrees with (3) 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  $\varepsilon$  is isotropic.

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

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

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

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

$$h_s(w) = \int_K p_s(k'wk) dk \tag{5}$$

$w \in \mathcal{P}_m$ ,  $s \in \mathbb{C}^m$ , 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 [6]. In particular, (5) 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, (8) is an integral formula for the zonal polynomials which arise often in aspects of multivariate distribution theory; see Muirhead [11], 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

$$d_*w = |w|^{-(m+1)/2} \prod_{1 \leq i \leq j \leq m} dw_{ij}, \tag{6}$$

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

For  $s = (s_1, \dots, s_m) \in \mathbb{C}^m$  and  $k \in K$ , the *Helgason-Fourier transform* ([17], p. 87) of a function  $f \in C_c^\infty(\mathcal{P}_m)$  is

$$\mathcal{H}f(s, k) = \int_{\mathcal{P}_m} f(w) \overline{p_s(k'wk)} d_*w, \tag{7}$$

where  $\overline{p_s(k'wk)}$  denotes complex conjugation and  $d_*w$  is the  $G$ -invariant measure (6).

For the case in which  $f \in C_c^\infty(\mathcal{P}_m/K)$ , we make the change of variables  $w \rightarrow k'_1wk_1$  in (7),  $k_1 \in K$ , and integrate with respect to the Haar measure  $dk_1$ . Applying the invariance of  $f$  and the formula (5) 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

$$\widehat{f}(s) = \int_{\mathcal{P}_m} f(w) \overline{h_s(w)} d_*w, \tag{8}$$

is the *zonal spherical transform* of  $f$ .

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 [17], 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  $\text{Re}(a), \text{Re}(b) > 0$ , let  $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$  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

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

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

$$\omega_m = \frac{\prod_{j=1}^m \Gamma(j/2)}{(2\pi i)^m \pi^{m(m+1)/4} m!} \tag{9}$$

We shall use the notation

$$\mathbb{C}^m(\rho) = \{s \in \mathbb{C}^m : \text{Re}(s) = -\rho\} \tag{10}$$

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

$$d_*s = \omega_m |c_m(s)|^{-2} ds_1 \cdots ds_m \tag{11}$$

since this measure is ubiquitous in our development.

Let  $M = \{\text{diag}(\pm 1, \dots, \pm 1)\}$  be the collection of  $m \times m$  diagonal matrices with entries  $\pm 1$  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 ([17], 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 (9)-(11). The inversion formula then is the following result.

For  $f \in C_c^\infty(\mathcal{P}_m)$  and  $w \in \mathcal{P}_m$ ,

$$f(w) = \int_{\mathbb{C}^m(\rho)} \int_{\bar{k} \in K/M} \mathcal{H}f(s, k) p_s(k'wk) d\bar{k} d_*s. \tag{12}$$

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

$$f(w) = \int_{\mathbb{C}^m(\rho)} \hat{f}(s) h_s(w) d_*s. \tag{13}$$

We refer to Terras [17], p. 87 ff. for a detailed treatment of this inversion formula and many references to the literature.

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

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

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

$$\mathcal{H}(f * h)(s, k) = \mathcal{H}f(s, k) \hat{h}(s), \tag{14}$$

$s \in \mathbb{C}^m$ ,  $k \in K$ ; see Terras [17], Theorem 1, p. 88.

### 3 Deconvolution density estimation on $\mathcal{P}_m$

On the group  $G$ , the deconvolution problem arises from the statistical model

$$\tilde{Y} \stackrel{\mathcal{L}}{=} \tilde{X}\tilde{\varepsilon} \tag{15}$$

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  $\varepsilon$  also are independent, and (15) implies that

$$Y \stackrel{\mathcal{L}}{=} \varepsilon^{1/2} X \varepsilon^{1/2}; \tag{16}$$

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

Applying to (16) the convolution property (14) of the Helgason-Fourier transform, we obtain

$$\mathcal{H}f_Y(s, k) = \mathcal{H}f_X(s, k) \widehat{f}_{\varepsilon}(s), \tag{17}$$

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

$$\mathcal{H}_n f_Y(s, k) = \frac{1}{n} \sum_{\ell=1}^n \overline{p_s(k'Y_{\ell}k)}. \tag{18}$$

Substituting (18) in (17), together with the assumption that  $\widehat{f}_{\varepsilon}(s) \neq 0$ ,  $s \in \mathbb{C}^m$ , we obtain

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

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

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 (12) using a spectral cut-off.

We introduce the notation

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

where  $\mathbb{C}^m(\rho)$  is defined in (10). As an estimator of the population density  $f_X$ , we define the density estimator  $f_X^n$  given by

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

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



Let  $\{a_n\}$  and  $\{b_n\}$  be sequences of real numbers. We write  $a_n \ll b_n$  to mean  $a_n \leq Cb_n$  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

$$\int_{\mathcal{P}_m} |w_1|^{-1} \cdots |w_{m-1}|^{-1} |w|^{(m-1)/2} f_Y(w) d_*w < \infty, \quad (19)$$

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.

Estimation will proceed in the *Sobolev class* of functions,

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

where  $2\sigma > \dim \mathcal{P}_m = m(m+1)/2$ ,  $\Delta$  is the Laplacian on  $\mathcal{P}_m$  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,

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

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

**Theorem 1.** *Suppose there exists  $\beta \geq 0$  such that*

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

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  $n \rightarrow \infty$ ,

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

**Corollary 1.** *Suppose the distribution of  $\varepsilon$  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 in Theorem 1 is replaced by an exponential bound. In such a situation, we have the following result.

**Theorem 2.** *Suppose there exists  $\beta, \gamma > 0$  such that*

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

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

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

$s \in \mathbb{C}^m$ , where  $\gamma > 0$  is a scale parameter. Again by the inversion formula (13), the underlying density function is

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

$w \in \mathcal{P}_m$ . 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  $\mathcal{P}_m$ , see Terras [17], pp. 106-107. As a consequence, we obtain the following result.

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

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

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,

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

where  $\text{Re}(s_j + \dots + s_m) > (j-1)/2$ ,  $j = 1, \dots, m$ . Relative to the invariant measure  $d_*w$  in (6), the probability density function of the standard Wishart distribution  $W_m(N, \mathbf{I}_m)$  is

$$f_\varepsilon(w) = \frac{1}{\Gamma_m(0, \dots, 0, N/2)} |\frac{1}{2}w|^{N/2} \exp(-\frac{1}{2}\text{tr } w), \tag{20}$$

$w \in \mathcal{P}_m$ . We note that (20) is  $K$ -invariant and its Helgason-Fourier transform is well-known (Muirhead [11], p. 248; Terras [17], pp. 85-86),

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

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

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 3.** *Suppose that  $\varepsilon$  follows the Wishart distribution (20) with  $N > m - 1$ . If  $f_X \in H_\sigma(\mathcal{P}_m, Q)$  with  $\sigma > \dim \mathcal{P}_m/2$  then, as  $n \rightarrow \infty$ ,*

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

## References

1. Basu, S., Fletcher, T., Whitaker, S.: Rician noise removal in diffusion tensor MRI. In: Larsen, R., Nielsen, M., Sparring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 117–125. Springer, Heidelberg, (2006)
2. Damon, B. M., Ding, Z., Anderson, A. W., Freyer, A. S., Gore, J. C.: Validation of diffusion tensor MRI-based muscle fiber tracking. *Magn. Reson. Med.* 48(1), 97–104 (2002)
3. Fletcher, P. T., Joshi, S.: Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In: Šonka, M., Kakadiaris, I. A., Kybic, J. (eds.) ECCV 2004 Workshop CVAMIA. LNCS, vol. 3117, pp. 87–98. Springer, Heidelberg, (2004)
4. Fletcher, P. T. and Joshi, S.: Riemannian geometry for the statistical analysis of diffusion tensor data. *Signal Processing* 87(2) 250–262 (2007)
5. Hasan, K. M., Basser, P. J., Parker, D. L., Alexander, A. L.: Analytical computation of the eigenvalues and eigenvectors in DT-MRI. *J. Magn. Res.* 152(1), 41–47 (2001)
6. Helgason, S.: *Groups and Geometric Analysis*. Academic Press, Orlando, (1984)
7. Kim, P. T. and Richards, D. St. P.: Deconvolution density estimation on spaces of positive definite symmetric matrices. Preprint (2007)
8. Koltchinskii, V., Sakhnenko, L., Cao, S.: Integral curves of noisy vector fields and statistical problems in diffusion tensor imaging: Nonparametric kernel estimation and hypotheses testing. *Ann. Statist.* 35(4), 1576–1607 (2007)
9. Le Bihan, D.: *Diffusion and Perfusion Magnetic Resonance Imaging: Applications to Functional MRI*. Raven Press, New York, (1995)
10. Matthews, P. M., Arnold, D. L.: Magnetic resonance imaging of multiple sclerosis: new insights linking pathology to clinical evolution. *Curr. Opin. Neurol.* 14(3), 279–287 (2001)
11. Muirhead, R. J.: *Aspects of Multivariate Statistics*. Wiley, New York, (1978)
12. Neumann-Haefelin T., Moseley M. E., Albers, G. W.: New magnetic resonance imaging methods for cerebrovascular disease. *Ann. Neurol.* 47(5), 559–570 (2000)
13. Pomara, N., Crandalla, D. T., Choia, S. J., Johnson, G. Lim, K. O.: White-matter abnormalities in HIV-1 infection: A diffusion tensor imaging study. *Psychiatry Res.: Neuroimaging* 106(1), 15–24 (2001)
14. Rosenbloom, M., Sullivan, J. V., Pfefferbaum, A.: Using magnetic resonance imaging and diffusion tensor imaging to assess brain damage in alcoholics. *Alcohol Research & Health* 27(2), 146–152 (2003)
15. Schwartzman, A.: *Random Ellipsoids and False Discovery Rates: Statistics for Diffusion Tensor Imaging Data*. Ph.D. dissertation, Stanford University (2006)
16. Schwartzman, A., Dougherty, R. F., Taylor, J. E.: Cross-subject comparison of principal diffusion direction maps. *Magn. Reson. Med.* 53(6), 1423–1431 (2005)
17. Terras, A.: *Harmonic Analysis on Symmetric Spaces and Applications, II*. Springer, New York, (1988).
18. Zhu, H., Zhang, H., Ibrahim, J. G., Peterson, B. G.: Statistical analysis of diffusion tensors in diffusion-weighted magnetic image resonance data (with discussion). *J. Amer. Statist. Assoc.* 102, 1085–1102 (2007)