



ACADEMIC  
PRESS

Available at  
**WWW.MATHEMATICSWEB.ORG**  
POWERED BY SCIENCE @ DIRECT®

Journal of Multivariate Analysis 85 (2003) 375–394

Journal of  
Multivariate  
Analysis

<http://www.elsevier.com/locate/jmva>

# Concentrated matrix Langevin distributions

Yasuko Chikuse

*Faculty of Engineering, Kagawa University, 2217-20 Hayashi-cho, Takamatsu, Kagawa 761-0396, Japan*

Received 28 November 2000

---

## Abstract

This paper concerns the matrix Langevin distributions, exponential-type distributions defined on the two manifolds of our interest, the Stiefel manifold  $V_{k,m}$  and the manifold  $P_{k,m-k}$  of  $m \times m$  orthogonal projection matrices idempotent of rank  $k$  which is equivalent to the Grassmann manifold  $G_{k,m-k}$ . Asymptotic theorems are derived when the concentration parameters of the distributions are large. We investigate the asymptotic behavior of distributions of some (matrix) statistics constructed based on the sample mean matrices in connection with testing hypotheses of the orientation parameters, and obtain asymptotic results in the estimation of large concentration parameters and in the classification of the matrix Langevin distributions.

© 2003 Elsevier Science (USA). All rights reserved.

*AMS 1991 subject classifications:* primary 62H11; 62E20; secondary 62F05; 62F12; 33C20

*Keywords:* Stiefel manifold; Grassmann manifold; Orthogonal projection matrix; Asymptotics for large concentrations; Matrix Langevin distributions; Hypergeometric functions with matrix arguments

---

## 1. Introduction

The Stiefel manifold  $V_{k,m}$  is the space a point of which is a set of  $k$  orthonormal vectors in  $R^m$  ( $k \leq m$ ), so that  $V_{k,m} = \{X(m \times k); X'X = I_k\}$ , where  $I_k$  is the  $k \times k$  identity matrix. The manifold  $V_{k,m}$  is an analytic manifold with dimension  $km - \frac{1}{2}k(k+1)$ . For  $m = k$ ,  $V_{k,m}$  is the orthogonal group  $O(m)$  of  $m \times m$  orthonormal matrices. A random matrix  $X$  on  $V_{k,m}$  is said to have the matrix Langevin (or von Mises–Fisher) distribution, denoted by  $L(m, k; F)$ , if its density function is

---

*E-mail address:* [chikuse@eng.kagawa-u.ac.jp](mailto:chikuse@eng.kagawa-u.ac.jp).

given by [14]

$$\text{etr}(F'X)/{}_0F_1(\tfrac{1}{2}m; \tfrac{1}{4}F'F) \quad \text{with } F \text{ an } m \times k \text{ matrix,} \quad (1.1)$$

where  $\text{etr}(A) = \exp(\text{tr } A)$ , and the  ${}_pF_q$  is a hypergeometric function with matrix argument due to e.g., [12,19,24]. Writing the (unique) singular value decomposition (s.v.d.) of  $F$  of rank  $p$  ( $p \leq k$ ) as

$$F = \Gamma A \Theta' \quad \text{with } \Gamma \in \tilde{V}_{p,m}, \Theta \in V_{p,k} \quad \text{and } A = \text{diag}(\lambda_1, \dots, \lambda_p), \\ \lambda_1 \geq \dots \geq \lambda_p > 0, \quad (1.2)$$

the  $\lambda_i$ 's control concentrations in the directions determined by the orientations  $\Gamma$  and  $\Theta$ . Here  $\tilde{V}_{p,m}$  denotes the  $2^{-p}$ th part of  $V_{p,m}$  consisting of matrices whose elements of the first row are positive. The distribution has the modal orientation  $\Gamma \Theta'$ , which is unique if and only if  $\lambda_k > 0$ .

The Grassmann manifold  $G_{k,m-k}$  is the space whose points are  $k$ -dimensional hyperplanes in  $R^m$  containing the origin, and is equivalent to the manifold  $P_{k,m-k}$ , the set of all  $m \times m$  orthogonal projection matrices idempotent of rank  $k$ . The manifold  $P_{k,m-k}$  is an analytic manifold with dimension  $km - k^2$ , on which we shall conduct statistical analysis.

For a random matrix  $P$  on  $P_{k,m-k}$ , the distribution having the density function

$$\text{etr}(BP)/{}_1F_1(\tfrac{1}{2}k; \tfrac{1}{2}m; B) \quad (1.3)$$

is a slight modification of the Downs' [14] distribution (1.1) on the Stiefel manifold, and may be called the matrix Langevin distribution on  $P_{k,m-k}$ , which is denoted by  $L^{(P)}(m, k; B)$ . Here  $B$  is an  $m \times m$  symmetric matrix having the (unique) spectral decomposition (s.d.) of rank  $p$ ,

$$B = \Gamma A \Gamma' \quad \text{with } \Gamma \in \tilde{V}_{p,m} \text{ and } A = \text{diag}(\lambda_1, \dots, \lambda_p), \lambda_1 \geq \dots \geq \lambda_p. \quad (1.4)$$

We impose identifiability restrictions on  $B$ ,

$$\text{tr } B = b \text{ being fixed} \quad (1.5)$$

or

$$\text{rank } B = p < m. \quad (1.6)$$

The distribution has the mode  $\Gamma_1 \Gamma_1'$ , for  $\Gamma = (\Gamma_1 \Gamma_2)$ , with  $\Gamma_1$  being  $m \times k$ , which is unique if and only if  $\text{rank } B \geq k$  and  $\lambda_k > \lambda_{k+1}$ . When  $B$  is positive semi-definite, the  $\lambda_i$ 's may control the concentrations of the distribution about the mode. Chikuse and Watson [11] defined and discussed distribution (1.3) as a special case of a more general family of distributions expressed in terms of zonal polynomials or hypergeometric functions with matrix argument.

Both of the  $L(m, k; F)$  and  $L^{(P)}(m, k; B)$  distributions are of exponential type and have been used most commonly on our manifolds  $V_{k,m}$  and  $P_{k,m-k}$ , respectively. The sample mean matrices  $\bar{X} = \sum_{i=1}^n X_i/n$  and  $\bar{P} = \sum_{i=1}^n P_i/n$  are sufficient statistics and useful for inferential problems, where  $X_1, \dots, X_n$  and  $P_1, \dots, P_n$  are random samples of size  $n$  from the respective distributions. The distributions of  $\bar{X}, \bar{P}$  and related statistics are expressed in integral forms involving hypergeometric functions

with matrix arguments, which seem to be intractable in themselves, and thus we must resort to asymptotic theory. Chikuse [5] and Chikuse and Watson [11] obtained some asymptotic results on distributional and inferential problems concerning the matrix Langevin distributions, for the cases of small concentrations (i.e., near the uniformities) and large sample sizes on  $V_{k,m}$  and  $P_{k,m-k}$ , and Chikuse [6] for the case of high dimension  $m$  on  $V_{k,m}$ . In this paper we discuss asymptotic theorems concerning the distributions for the cases when the concentration parameters  $\Lambda$ 's are large on both  $V_{k,m}$  and  $P_{k,m-k}$ .

For the special case  $k = 1$ , the observations from the unit hypersphere  $V_{1,m}$  are directed unit vectors, i.e., directions, and those from the real projective space  $G_{1,m-1}$  are axes or undirected lines through the origin, i.e., one-dimensional subspaces. There exists a large literature of applications of these directional statistics and its statistical analysis. Most directional statistics in practice occur in two or three dimensions, i.e., on the circle ( $m = 2$ ) and the sphere ( $m = 3$ ). Directional analysis of data played important roles in the Earth Sciences, Astrophysics, Biology, Meteorology, Animal Behavior and many other fields; see e.g., [16,26].

The analysis of data on the general Stiefel manifold  $V_{k,m}$  is required in particular for  $k \leq m \leq 3$  in practical applications in the Medical Sciences, Astronomy and other fields. See [14,16,20,27] for the analyses of the data of vectorcardiogram orientations and of measurements of orbits of comets. One is naturally interested in  $k$ -dimensional subspaces as observations from the general Grassmann manifold  $G_{k,m-k}$ . We note that if  $X$  is an observation on the Stiefel manifold  $V_{k,m}$ ,  $XX'$  is an observation on the manifold  $P_{k,m-k}$  equivalent to the Grassmann manifold  $G_{k,m-k}$ . Examples of observations on  $G_{k,m-k}$  arise in the signal processing of radar with  $m$  elements observing  $k$  targets. The Grassmann manifold is a rather new subject treated as a statistical sample space.

In Section 2 we investigate the asymptotic behavior of distributions of some (matrix) statistics constructed based on the sample mean matrices in connection with testing hypotheses of the orientation parameters  $\Gamma$  and  $\Theta$  of the  $L(m, k; F)$  distribution and of the orientation parameter  $\Gamma$  of the  $L^{(P)}(m, k; B)$  distribution, when  $\Lambda$ 's are known and large.

Section 3 discusses asymptotics in the estimation of large concentration parameters  $\Lambda$ 's and in the classification of the matrix Langevin distributions for large  $\Lambda$ 's. We deal with the maximum likelihood estimators (m.l.e.'s) and related maximum marginal and/or maximum semi-marginal likelihood estimators of  $\Lambda$ , which are given by the solutions of systems of partial differential equations involving hypergeometric functions with matrix arguments. We show methods to give approximate solutions for large  $\Lambda$ 's. Asymptotic distributions of the classification function are obtained in the problem of classification into one of two matrix Langevin distributions under the condition of *local closeness* of the two distributions on each of our manifolds.

Some more discussions of the  $L(m, k; F)$  and  $L^{(P)}(m, k; B)$  distributions and the related statistical analyses on the manifolds  $V_{k,m}$  and  $P_{k,m-k}$  may be found in e.g. [8,20,22,25] and [9,10], respectively.

Before closing this section, we present some distributional results which are relevant for the discussion in the following sections.

Let  $Z$  be an  $m \times k$  random matrix (of rank  $k$ ). For the (unique) polar decomposition of  $Z$ ,  $Z = H_Z T_Z^{1/2}$ , with  $H_Z = Z(Z'Z)^{-1/2} \in V_{k,m}$  and  $T_Z = Z'Z$ , we have (see [17, Lemma 1.4]; [24, Theorem 2.1.14])

$$(dZ) = [\pi^{km/2} / \Gamma_k(\frac{1}{2}m)] |T_Z|^{(m-k-1)/2} (dT_Z) [dH_Z], \quad (1.7)$$

where  $(dZ)$ ,  $(dT_Z)$  and  $[dH_Z]$  are the Lebesgue measures on the spaces of  $m \times k$  matrices and of  $k \times k$  symmetric matrices and the normalized invariant measure on  $V_{k,m}$  (see [15,18]), respectively. Here  $A^{1/2}$  denotes the unique square root of a positive semi-definite matrix  $A$ , and  $\Gamma_k(a) = \pi^{k(k-1)/4} \prod_{i=1}^k \Gamma[a - \frac{1}{2}(i-1)]$ . Throughout the paper, density functions of distributions are expressed with respect to the Lebesgue measures and the normalized invariant measures on the respective spaces and manifolds; for the invariant measures on  $G_{k,m-k}$  and  $P_{k,m-k}$ , see [11,18].

An  $m \times k$  random matrix  $Z$  is said to have the normal  $N_{m,k}(0, I_m \otimes I_k)$  distribution if the density function is given by

$$\varphi^{m,k}(Z) = (2\pi)^{-km/2} \text{etr}(-\frac{1}{2}Z'Z),$$

so that the characteristic function is given by

$$\Phi_Z(T) = E(\text{etr } T'Z) = \text{etr}(-\frac{1}{2}T'T),$$

for an  $m \times k$  matrix  $T$ . The random matrix  $Y = \Sigma_1^{1/2} Z \Sigma_2^{1/2} + M$ , for an  $m \times k$  matrix  $M$  and  $m \times m$  and  $k \times k$  positive definite matrices  $\Sigma_1$  and  $\Sigma_2$ , respectively, is said to have the normal  $N_{m,k}(M, \Sigma_1 \otimes \Sigma_2)$  distribution, whose density function is written as

$$\varphi^{(m,k)}(Y; M, \Sigma_1 \otimes \Sigma_2) = |\Sigma_1|^{-k/2} |\Sigma_2|^{-m/2} \varphi^{(m,k)}[\Sigma_1^{-1/2}(Y - M)\Sigma_2^{-1/2}].$$

## 2. Asymptotics in connection with testing problems

### 2.1. On the Stiefel manifold $V_{k,m}$

We are concerned with the matrix Langevin  $L(m, k; F)$  distribution (1.1) with the s.v.d. (1.2). When the concentration parameter  $\Lambda$  is known and large, we consider testing the null hypothesis

$$H_0: \Gamma = \Gamma_0 \quad \text{and} \quad \Theta = \Theta_0 \quad \text{with } \Gamma_0 \in V_{p,m} \text{ and } \Theta_0 \in V_{p,k} \text{ given}, \quad (2.1)$$

against a sequence of local alternative hypotheses

$$H_1: \begin{cases} \Gamma = (\Gamma_0 + \Psi_1 \Lambda^{-1/2})(I_p + \Lambda^{-1/2} \Psi_1' \Psi_1 \Lambda^{-1/2})^{-1/2} & \text{with } \Psi_1' \Gamma_0 = 0, \\ \Theta = (\Theta_0 + \Psi_2 \Lambda^{-1/2})(I_p + \Lambda^{-1/2} \Psi_2' \Psi_2 \Lambda^{-1/2})^{-1/2} & \text{with } \Psi_2' \Theta_0 = 0, \end{cases} \quad (2.2)$$

so that we can write

$$\Psi_1 = \Gamma_0^\perp \Phi_1 \quad \text{and} \quad \Psi_2 = \Theta_0^\perp \Phi_2,$$

for certain  $(m-p) \times p$  and  $(k-p) \times p$  matrices  $\Phi_1$  and  $\Phi_2$ , respectively, where  $\Gamma_0^\perp$  and  $\Theta_0^\perp$  are chosen such that  $(\Gamma_0 \ \Gamma_0^\perp) \in O(m)$  and  $(\Theta_0 \ \Theta_0^\perp) \in O(k)$ , respectively. Under the hypothesis  $H_1$  (2.2), we can write

$$\Gamma = \Gamma_0 + \Psi_1 A^{-1/2} - \frac{1}{2} \Gamma_0 A^{-1/2} \Psi_1' \Psi_1 A^{-1/2} + O(A^{-3/2})$$

and

$$\Theta = \Theta_0 + \Psi_2 A^{-1/2} - \frac{1}{2} \Theta_0 A^{-1/2} \Psi_2' \Psi_2 A^{-1/2} + O(A^{-3/2}),$$

where  $O(A^{-l})$  denotes the terms of order  $l$  in the elements of  $A^{-1}$ .

For the sample mean matrix  $\bar{X} = \sum_{i=1}^n X_i/n$  with a random sample  $X_1, \dots, X_n$  from the  $L(m, k; F)$  distribution, we shall investigate asymptotic behavior for large  $A$  of the following matrix statistics in connection with the above-mentioned test:

$$Y = n^{1/2} \begin{pmatrix} \Gamma_0' \\ \Gamma_0^{\perp'} \end{pmatrix} (\Gamma_0 \Theta_0' - \bar{X}) \Theta_0 A^{1/2} = n^{1/2} \begin{bmatrix} (I_p - \Gamma_0' \bar{X} \Theta_0) A^{1/2} \\ -\Gamma_0^\perp \bar{X} \Theta_0 A^{1/2} \end{bmatrix} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}, \quad (2.3)$$

$$\begin{aligned} Z &= n^{1/2} A^{1/2} \Gamma_0' (\Gamma_0 \Theta_0' - \bar{X}) (\Theta_0 \ \Theta_0^\perp) \\ &= n^{1/2} [A^{1/2} (I_p - \Gamma_0' \bar{X} \Theta_0) - A^{1/2} \Gamma_0' \bar{X} \Theta_0^\perp] = (Z_1 \ Z_2) \end{aligned} \quad (2.4)$$

and

$$V = 2n^{1/2} Y_1 A^{1/2}; \quad (2.5)$$

we note the similarity between (2.3) and (2.4) with the relationship  $Z_1 = A^{1/2} Y_1 A^{-1/2}$ .

The characteristic function of  $Y$  under the hypothesis  $H_1$  (2.2) is, for an  $m \times p$  matrix  $T = (T_1' \ T_2')'$  with  $T_1$  being  $p \times p$ ,

$$\Phi_Y(T) = E(\text{etr } iT'Y) = \text{etr}(in^{1/2} A^{1/2} T_1') Q^n, \quad (2.6)$$

where

$$Q = {}_0F_1(\tfrac{1}{2}m; \tfrac{1}{4}AA') / {}_0F_1(\tfrac{1}{2}m; \tfrac{1}{4}A^2)$$

with

$$\begin{aligned} A &= \Theta_0 A \Gamma_0' \\ &+ [-in^{-1/2} \Theta_0 A^{1/2} (T_1' \Gamma_0' + T_2' \Gamma_0^{\perp'}) + \Theta_0 A^{1/2} \Phi_1' \Gamma_0^{\perp'} + \Theta_0^\perp \Phi_2 A^{1/2} \Gamma_0'] \\ &+ [-\tfrac{1}{2} \Theta_0 A^{1/2} \Phi_1' \Phi_1 A^{-1/2} \Gamma_0' - \tfrac{1}{2} \Theta_0 A^{-1/2} \Phi_2' \Phi_2 A^{1/2} \Gamma_0' + \Theta_0^\perp \Phi_2 \Phi_1' \Gamma_0^{\perp'}] \\ &+ O(A^{-1/2}). \end{aligned}$$

We have

$$AA' = O_2(A^2) + O_{3/2}(A^{3/2}) + O_1(A) + O(A^{1/2}), \quad (2.7)$$

where

$$O_2(A^2) = \Theta_0 A^2 \Theta_0', \quad O_{3/2}(A^{3/2}) = O_{3/2} + O_{3/2}'$$

with

$$O_{3/2} = -in^{-1/2} \Theta_0 A T_1 A^{1/2} \Theta_0' + \Theta_0 A^{3/2} \Phi_2' \Theta_0^{\perp'}$$

and

$$O_1(A) = -n^{-1} \Theta_0 A^{1/2} (T_1' T_1 + T_2' T_2) A^{1/2} \Theta_0' + O_1 + O_1' + \Theta_0^{\perp} \Phi_2 A \Phi_2' \Theta_0^{\perp'} \quad (2.8)$$

with

$$O_1 = -\Theta_0 (in^{-1/2} A^{1/2} \Phi_1' T_2 A^{1/2} + \frac{1}{2} A^{3/2} \Phi_2' \Phi_2 A^{-1/2}) \Theta_0' \\ - in^{-1/2} \Theta_0 A^{1/2} T_1' A^{1/2} \Phi_2' \Theta_0^{\perp'}.$$

Now, when  $AA'$  is of rank  $p$  having nonzero latent roots  $a_1^2, \dots, a_p^2 (a_1^2 \geq \dots \geq a_p^2 > 0)$ , the function  ${}_0F_1(\frac{1}{2}m; \frac{1}{4}AA')$  is asymptotically, for large  $A$ , expressed as in [23, Section 5]. The latent roots  $a_i$  for the matrix  $AA'$  given by (2.7) may be obtained by using the *perturbation theory* for large  $A$ . The following lemma on the perturbation theory is essentially due to Bellman [3, Chapter 4].

**Lemma 2.1.** Let  $B_0, B_1$  and  $B_2$  be  $k \times k$  symmetric matrices,  $B_0$  have latent roots  $\xi_1, \dots, \xi_k$  and the corresponding latent vectors  $\mathbf{x}_1, \dots, \mathbf{x}_k$  of unit length, and  $\xi_i$  have multiplicity one for a fixed  $i$ . Let  $B_0$  be perturbed to

$$B = B_0 + \varepsilon B_1 + \varepsilon^2 B_2 + O(\varepsilon^3),$$

and assume that the corresponding perturbations of  $\xi_i$  and  $\mathbf{x}_i$  are

$$\mu_i = \xi_i + \varepsilon \xi_{i1} + \varepsilon^2 \xi_{i2} + O(\varepsilon^3) \quad \text{and} \quad \mathbf{y}_i = \mathbf{x}_i + \varepsilon \mathbf{x}_{i1} + \varepsilon^2 \mathbf{x}_{i2} + O(\varepsilon^3).$$

Then we have

$$\xi_{i1} = \mathbf{x}_i' B_1 \mathbf{x}_i, \quad (2.9)$$

$$\mathbf{x}_{i1} = \sum_{j=1, j \neq i}^k \frac{c_{j(i)}}{\xi_j - \xi_i} \mathbf{x}_j \quad \text{with} \quad c_{j(i)} = -\mathbf{x}_j' B_1 \mathbf{x}_i \quad (2.10)$$

and

$$\xi_{i2} = \mathbf{x}_i' B_1 \mathbf{x}_{i1} + \mathbf{x}_i' B_2 \mathbf{x}_i. \quad (2.11)$$

**Proof.** Equating the coefficients of  $\varepsilon$  and  $\varepsilon^2$  in  $B\mathbf{y}_i = \mu_i \mathbf{y}_i$  gives

$$(B_0 - \xi_i I_k) \mathbf{x}_{i1} = (\xi_{i1} I_k - B_1) \mathbf{x}_i \quad (2.12)$$

and

$$(B_0 - \xi_i I_k) \mathbf{x}_{i2} = (\xi_{i1} I_k - B_1) \mathbf{x}_{i1} + (\xi_{i2} I_k - B_2) \mathbf{x}_i, \quad (2.13)$$

and equating the coefficients of  $\varepsilon$  in  $\mathbf{y}'_i \mathbf{y}_i = 1$  gives

$$\mathbf{x}'_i \mathbf{x}_{i1} = 0. \quad (2.14)$$

Premultiplying (2.12) by  $\mathbf{x}'_i$  and using  $B_0 \mathbf{x}_i = \xi_i \mathbf{x}_i$  and  $\mathbf{x}'_i \mathbf{x}_i = 1$  gives the desired result (2.9). Since the right-hand side of (2.12) is expressed as  $\sum_{j=1}^k c_{j(i)} \mathbf{x}_j$ , with  $c_{j(i)}$  being given in (2.10) and  $c_{i(i)} = 0$ , substituting  $\mathbf{x}_{i1} = \sum_{j=1, j \neq i}^k b_{j(i)} \mathbf{x}_j$ , with  $b_{j(i)} = \mathbf{x}'_j \mathbf{x}_{i1}$ , into the left-hand side of (2.12), and using  $B_0 \mathbf{x}_j = \xi_j \mathbf{x}_j$ ,  $j = 1, \dots, k$ , we obtain  $b_{j(i)}(\xi_j - \xi_i) = c_{j(i)}$ , for  $j \neq i, j = 1, \dots, k$ , which yields the desired result (2.10). Similarly, premultiplying (2.13) by  $\mathbf{x}'_i$ , in view of (2.14), gives the desired result (2.11).  $\square$

Putting  $A = NA_0$ , with  $N(>0)$  large and  $A_0$  fixed, and  $\varepsilon = N^{-1/2}$ , we apply Lemma 2.1 to the matrix

$$B = \frac{AA'}{N^2} = O_2(A_0^2) + \varepsilon O_{3/2}(A_0^{3/2}) + \varepsilon^2 O_1(A_0) + O(\varepsilon^3),$$

where  $O_2(\cdot)$ ,  $O_{3/2}(\cdot)$ , and  $O_1(\cdot)$  are defined by (2.8). We obtain

$$\begin{aligned} \xi_i &= \lambda_{0i}^2, \quad \xi_{i1} = -\frac{2i}{n^{1/2}} \lambda_{0i}^{3/2} T_{1ii}, \\ \xi_{i2} &= -\lambda_{0i} \left[ \frac{1}{n} (T'_1 T_1 + T'_2 T_2) + \frac{2i}{n^{1/2}} \Phi'_1 T_2 + \Phi'_2 \Phi_2 \right]_{ii} \\ &\quad + \frac{\lambda_{0i}}{n} \sum_{j=1, j \neq i}^p \frac{\lambda_{0j}(\lambda_{0i}^{1/2} T_{1ij} + \lambda_{0j}^{1/2} T_{1ji})}{\lambda_{0j}^2 - \lambda_{0i}^2} + \lambda_{0i} \sum_{j=p+1}^k \Phi_{2j-p,i}^2, \quad \text{for } i = 1, \dots, p \end{aligned} \quad (2.15)$$

and

$$\xi_i = \xi_{i1} = \xi_{i2} = 0, \quad \text{for } i = p+1, \dots, k,$$

where  $\lambda_i = N\lambda_{0i}$ ,  $i = 1, \dots, p$ , and, in general,  $A_{ij}$  indicates the  $ij$ -element of a matrix  $A$ . We can express the latent roots  $a_i$  as

$$a_i = N\xi_i^{1/2} \left[ 1 + \varepsilon \frac{\xi_{i1}}{2\xi_i} + \varepsilon^2 \left( \frac{\xi_{i2}}{2\xi_i} - \frac{\xi_{i1}^2}{8\xi_i^2} \right) + O(\varepsilon^3) \right]. \quad (2.16)$$

Thus, substituting (2.15) into (2.16) and using Muirhead [23, Section 5] with large  $AA'$  and  $A^2$ , we obtain, from (2.6),

$$\begin{aligned} \Phi_Y(T) &= \exp \frac{1}{2} \left[ -\text{tr } T'_1 T_1 + \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^p \frac{(\lambda_i^{1/2} T_{1ij} + \lambda_j^{1/2} T_{1ji})^2}{\lambda_i + \lambda_j} \right. \\ &\quad \left. - \text{tr } T'_2 T_2 - 2in^{1/2} \text{tr } \Phi'_1 T_2 \right] [1 + O(A^{-1/2})]. \end{aligned}$$

A similar method is applied to  $Z$  defined by (2.4). Next, the characteristic function of the vector  $\mathbf{v} = (V_{11}, \dots, V_{pp})'$  of the diagonal elements of  $V = (V_{ij})$  defined by

(2.5), for a  $p \times 1$  vector  $\mathbf{t} = (t_1, \dots, t_p)'$ , is also expanded, by using the perturbation method, as

$$\begin{aligned}\Phi_{\mathbf{v}}(\mathbf{t}) &= E \exp(i\mathbf{t}'\mathbf{v}) = \exp\left(in \sum_{i=1}^p \frac{\Delta_{ii}t_i}{1-2it_i}\right) \prod_{i=1}^p (1-2it_i)^{-(m-p)n/2} \\ &\times \prod_{i < j}^p \left[1 - \frac{2i(\lambda_i t_i + \lambda_j t_j)}{\lambda_i + \lambda_j}\right]^{-n/2} [1 + O(\Lambda^{-1})] \\ &\text{with } \Delta = (\Delta_{ij}) = \sum_{i=1}^2 \Phi_i' \Phi_i.\end{aligned}$$

Hence, although the vector variate  $\mathbf{v}$  is correlated,  $v = \text{tr } V = \sum_{i=1}^p V_{ii}$  is shown to be asymptotically distributed as noncentral  $\chi^2_{r,n \text{ tr } \Delta}$  with  $r = \frac{1}{2}pn(2m-p-1)$ , for large  $\Lambda$ .

The above discussion is summarized in

**Theorem 2.2.** *Under the hypothesis  $H_1$  (2.2), for large  $\Lambda$ , the distributions of the matrix statistics  $Y$ ,  $Z$  and  $V$  defined in (2.3)–(2.5) have the following asymptotic properties:*

(i)  $Y_1$  and  $Y_2$  are independent, and  $Y_2$  is distributed as normal  $N_{m-p,p}(-n^{1/2}\Phi_1, I_{m-p} \otimes I_p)$ . The distribution of  $Y_1$  is such that the diagonal elements of  $Y_1$  are equal to zero almost everywhere, that  $(Y_{1ij}, Y_{1ji})'$ ,  $i \neq j$ , is bivariate normal with zero means and covariance matrix

$$\begin{pmatrix} g_{ij} & -h_{ij} \\ -h_{ji} & g_{ji} \end{pmatrix} \quad \text{with} \quad g_{ij} = \frac{\lambda_j}{\lambda_i + \lambda_j} \quad \text{and} \quad h_{ij} = h_{ji} = \frac{(\lambda_i \lambda_j)^{1/2}}{\lambda_i + \lambda_j}, \quad \text{for } i \neq j,$$

and that  $Y_{1ij}$  is independent of all the other elements  $Y_{1i'j'}[(i', j') \neq (j, i)]$  of  $Y_1$ . Here we assume the ratios  $\lambda_i/\lambda_j$  are constant for large  $\Lambda$ .

(ii)  $Z_1$  and  $Z_2$  are independent,  $Z_2$  is distributed as normal  $N_{p,k-p}(-n^{1/2}\Phi_2', I_p \otimes I_{k-p})$ , and the distribution of  $Z_1$  is similar to that of  $Y_1$  except that the variance of  $Z_{1ij}$ ,  $i \neq j$ , is  $g_{ij}^* = \lambda_i/(\lambda_i + \lambda_j)$  instead of  $g_{ij}$ .

(iii)  $Y_2$  and  $Z_2$  are independent.

(iv)  $v = \text{tr } V$  is distributed as noncentral  $\chi^2_{r,n \text{ tr } \Delta}$ , where  $r = \frac{1}{2}pn(2m-p-1)$  and  $\Delta = \sum_{i=1}^2 \Phi_i' \Phi_i$ .

**Remark 2.3.** We note that, with  $L_i$  being the density of the sample under the hypothesis  $H_i$ ,  $i = 0, 1$ ,  $\log(L_0/L_1) = -n(\text{tr } \Gamma_0^{-1} \bar{X} \Theta_0 \Lambda^{1/2} \Phi_1' + \text{tr } \Theta_0^{-1} \bar{X}' \Gamma_0 \Lambda^{1/2} \Phi_2')$ , ignoring the terms of  $O(1)$ . Therefore, it might be recommended to use statistics based on  $Y_2$  for the test on the parameter  $\Gamma$ ,  $Z_2$  on the parameter  $\Theta$ , and  $Y_2$  and  $Z_2$  combined or  $\text{tr } V$  on the parameters  $\Gamma$  and  $\Theta$  simultaneously, for large  $\Lambda$ .

**Corollary 2.4.** *Putting  $\Phi_1 = 0$  and  $\Phi_2 = 0$  in Theorem 2.2 yields the asymptotic properties under the null hypothesis  $H_0$  (2.1).*

For the rest of Section 2.1, we obtain some asymptotic results for large  $\Lambda$  concerning the  $L(m, k; F)$  distribution with the (unique) s.v.d.  $F = \Gamma \Lambda \Theta'$  of rank  $p = k$ , which can be derived from Theorem 2.2. Letting  $X$  be a random matrix distributed as  $L(m, k; F)$ , we can express

$$X = \Gamma \Theta' + \Gamma Y_1 \Lambda^{-1/2} \Theta' + \Gamma^\perp Y_2 \Lambda^{-1/2} \Theta' + O(\Lambda^{-1}), \quad \text{for large } \Lambda,$$

where  $Y_1$  and  $Y_2$  are independent random matrices distributed as described in (i) of Theorem 2.2 with  $\Phi_1 = 0$ ; indicating that  $X$  approaches the mode  $\Gamma \Theta'$  in probability for large  $\Lambda$ .

We consider the problem of testing the null hypothesis  $H_0 : \Gamma = \Gamma_0$  and  $\Theta = \Theta_0$  against the alternative hypothesis  $H_1 : \Gamma \neq \Gamma_0$  and  $\Theta \neq \Theta_0$ , when  $\Lambda$  is known and large. Putting the (unique) s.v.d. of the sample mean matrix  $\bar{X}$  of sample size  $n$  as  $\bar{X} = \bar{H}_1 \bar{X}_d \bar{H}_2'$ , where  $\bar{H}_1 \in \tilde{V}_{k,m}$ ,  $\bar{H}_2 \in O(k)$ , and  $\bar{X}_d = \text{diag}(x_1, \dots, x_k)$ ,  $x_1 > \dots > x_k > 0$ , we have the m.l.e.'s of  $\Gamma$  and  $\Theta$ ,  $\hat{\Gamma} = \bar{H}_1$  and  $\hat{\Theta} = \bar{H}_2$  (see Section 3.1), and the likelihood ratio test statistic is given by  $2n \text{tr}(\bar{X}_d - \Gamma_0' \bar{X} \Theta_0) \Lambda$ , being distributed asymptotically for large  $n$  and  $\Lambda$  as  $\chi_{k(m-k)}^2$  under  $H_0$ . We have already seen that  $\text{tr} V = 2n \text{tr}(I_k - \Gamma_0' \bar{X} \Theta_0) \Lambda$  is distributed asymptotically for large  $\Lambda$  as  $\chi_{nk(2m-k-1)/2}^2$  under  $H_0$ .

Now, interpreting the terms in the identity “ $n \text{tr}(I_k - \Gamma_0' \bar{X} \Theta_0) \Lambda = n \text{tr}(I_k - \bar{X}_d) \Lambda + n \text{tr}(\bar{X}_d - \Gamma_0' \bar{X} \Theta_0) \Lambda$ ”, as the dispersion of sample about true mode being equal to the sum of those of sample about sample mode and of sample mode about true mode, the argument of analysis of variance may lead us to use the F-test statistic

$$(f_2/f_1)[\text{tr}(\bar{X}_d - \Gamma_0' \bar{X} \Theta_0) \Lambda] / [\text{tr}(I_k - \bar{X}_d) \Lambda], \quad (2.17)$$

with  $f_1 = k(m-k)$  and  $f_2 = \frac{1}{2}nk(2m-k-1) - f_1$ , which is distributed asymptotically for large  $n$  and  $\Lambda$  as  $F_{f_1, f_2}$  under  $H_0$ ; statistic (2.17) becomes asymptotically noncentral  $F_{f_1, f_2; n \text{tr } \Lambda}$ , with  $\Lambda = \sum_{i=1}^2 \Phi_i' \Phi_i$  under the alternative hypothesis  $H_1$  (2.2).

Khatri and Mardia [21] considered some problems of the concentrated matrix Langevin  $L(m, k; F)$  distributions, e.g., the maximum likelihood estimator of  $\Lambda$  and some test statistics. Large-concentration asymptotics for the Langevin distribution with  $k = 1$  were obtained in connection with some test problems by Watson [27,28].

## 2.2. On the manifold $P_{k,m-k}$

We are concerned with the matrix Langevin  $L^{(P)}(m, k; B)$  distribution (1.3) with the s.d. (1.4) of rank  $p < m$  [see (1.6)], where  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$ ,  $\lambda_1 \geq \dots \geq \lambda_p > 0$ , is known and large. We consider testing the null hypothesis

$$H_0 : \Gamma = \Gamma_0 \quad \text{with } \Gamma_0 \in V_{p,m} \text{ given}, \quad (2.18)$$

against a sequence of local alternative hypotheses

$$\begin{aligned} H_1 : \Gamma &= (\Gamma_0 + \Psi A^{-1/2})(I_p + A^{-1/2} \Psi' \Psi A^{-1/2})^{-1/2} \\ &= \Gamma_0 + \Psi A^{-1/2} - \frac{1}{2} \Gamma_0 A^{-1/2} \Psi' \Psi A^{-1/2} + O(A^{-3/2}) \quad \text{with } \Psi' \Gamma_0 = 0, \end{aligned} \quad (2.19)$$

so that we can write  $\Psi = \Gamma_0^\perp \Phi$  for an  $(m-p) \times p$  matrix  $\Phi$ .

For the sample mean matrix  $\bar{P} = \sum_{i=1}^n P_i/n$  with a random sample  $P_1, \dots, P_n$  from the  $L^{(p)}(m, k; B)$  distribution, we shall investigate the asymptotic behavior for large  $A$  of the matrix statistic  $S$  in connection with the above-mentioned test, where

$$\begin{aligned} S &= 2n \begin{pmatrix} A^{1/2} & 0 \\ 0 & I_{m-p} \end{pmatrix} (\Gamma_0 \quad \Gamma_0^\perp)' (I_m - \bar{P}) (\Gamma_0 \quad \Gamma_0^\perp) \begin{pmatrix} A^{1/2} & 0 \\ 0 & I_{m-p} \end{pmatrix} \\ &= 2n \begin{pmatrix} A^{1/2} \Gamma_0' (I_m - \bar{P}) \Gamma_0 A^{1/2} & A^{1/2} \Gamma_0' (I_m - \bar{P}) \Gamma_0^\perp \\ \Gamma_0^\perp{}' (I_m - \bar{P}) \Gamma_0 A^{1/2} & \Gamma_0^\perp{}' (I_m - \bar{P}) \Gamma_0^\perp \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{pmatrix}. \end{aligned} \quad (2.20)$$

The characteristic function of  $S$  under the hypothesis  $H_1$  (2.19) is, for an  $m \times m$  symmetric matrix  $T = (T_{ij})$  having the partition in accordance with (2.20),

$$\Phi_S(T) = E(\text{etr } iTS) = \text{etr}[2in(AT_{11} + T_{22})]Q^n, \quad (2.21)$$

where

$$Q = {}_1F_1(\tfrac{1}{2}k; \tfrac{1}{2}m; A) / {}_1F_1(\tfrac{1}{2}k; \tfrac{1}{2}m; A)$$

with

$$\begin{aligned} A &= \begin{bmatrix} A^{1/2}(I_p - 2iT_{11} - \tfrac{1}{2}\Phi' \Phi A^{-1} - \tfrac{1}{2}A^{-1}\Phi' \Phi)A^{1/2} & A^{1/2}(\Phi' - 2iT_{12}) \\ (\Phi - 2iT_{12}')A^{1/2} & \Phi\Phi' - 2iT_{22} \end{bmatrix} \\ &\quad + O(A^{-1/2}) \end{aligned} \quad (2.22)$$

We use the asymptotic expansions for the  ${}_1F_1$  functions with the matrix argument  $A$  and with the matrix argument  $A$  of type (2.22) for large  $A$  (see [13, Theorems 3.2 and 3.1, respectively]). In view of the Kummer relationship for the  ${}_1F_1$  function (see [19, (51)]), we obtain, from (2.21),

$$\begin{aligned} \Phi_S(T) &= |I_p - 2iT_{11}|^{-(m-k)n/2} \{ {}_1F_1[\tfrac{1}{2}(m-k); \tfrac{1}{2}(m-p); \\ &\quad (\Phi' - 2iT_{12})'(I_p - 2iT_{11})^{-1}(\Phi' - 2iT_{12}) + 2iT_{22} - \Phi\Phi'] \}^n [1 + O(A^{-1/2})], \end{aligned} \quad (2.23)$$

from which  $S_{11}$ ,  $S_{12}$ , and  $S_{22}$  are seen to be, in general, dependent.

Now putting  $T_{11} = 0$  and  $T_{12} = 0$  in (2.23) gives the limit characteristic function of  $S_{22}$

$$\Phi_{S_{22}}^*(T_{22}) = \{ {}_1F_1[\tfrac{1}{2}(m-k); \tfrac{1}{2}(m-p); 2iT_{22}] \}^n. \quad (2.24)$$

When  $p < k$ , the right-hand side of (2.24) indicates the characteristic function of the matrix variate  $2 \sum_{i=1}^n R_i$ , where the  $R_i$  are independent and uniformly distributed on the manifold  $P_{m-k, (m-p)-(m-k)} = P_{m-k, k-p}$ . When  $p = k$ , it follows that we have  $S_{22} = 2nI_{m-k}$  (almost everywhere) which is independent of  $S_{11}$  and  $S_{22}$  in the limit.

In the sequel, we assume that  $p = k$ . By the inversion of (2.23), we have the limit joint density function  $f_S^*(S)$  of  $S_{11}$ ,  $S_{12}$  and  $S_{22}$ , with respect to the measure  $(dS) = (dS_{11})(dS_{12})(dS_{22})$ ,

$$f_S^*(S) = a_m \int \text{etr}(-iST) \Phi_S^*(T) (dT) \quad \text{with } a_m = \frac{2^{m(m-1)/2}}{(2\pi)^{m(m+1)/2}}, \quad (2.25)$$

where

$$\begin{aligned} \Phi_S^*(T) &= |I_k - 2iT_{11}|^{-(m-k)n/2} \text{etr} n[(\Phi' - 2iT_{12})' \\ &\quad \times (I_k - 2iT_{11})^{-1} (\Phi' - 2iT_{12}) + 2iT_{22} - \Phi\Phi']. \end{aligned}$$

Evaluating (2.25) with  $T_{22} = 0$ , with respect to  $(dT_{12})$  and then  $(dT_{11})$ , we obtain the limit joint density function of  $S_{11}$  and  $S_{12}$

$$\begin{aligned} f_{S_{11}, S_{12}}^*(S_{11}, S_{12}) &= \{2^{k(m-k)(n-1)/2} (4n\pi)^{k(m-k)/2} \Gamma_k[\tfrac{1}{2}(m-k)(n-1)]\}^{-1} \\ &\quad \times |S_{11} - \tfrac{1}{2n} S_{12} S_{12}'|^{[(m-k)(n-1)-k-1]/2} \text{etr}(-\tfrac{1}{2} S_{11} - \Phi S_{12} - n\Phi'\Phi), \\ &\quad \text{for } S_{11} - S_{12} S_{12}'/2n > 0; \end{aligned} \quad (2.26)$$

that is,  $S_{12}$  is marginally distributed as normal  $N_{k, m-k}(-2n\Phi', 2nI_k \otimes I_{m-k})$ , and  $S_{11} - S_{12} S_{12}'/2n (> 0)$  is distributed as Wishart  $W_k((m-k)(n-1), I_k)$  conditionally with given  $S_{12}$ .

To obtain the limit marginal distribution of  $S_{11}$  from (2.26), we have only to evaluate the integral  $I$ , say, with respect to  $(dS_{12})$  for  $S_{11} > S_{12} S_{12}'/2n$  in (2.26). Making the transformation  $S_{12} \rightarrow S_{12}H$  for  $H \in O(m-k)$  and integrating over  $O(m-k)$  in view of the integral representation of the  ${}_0F_1$  function (see [19, (27)]), the term  $\text{etr}(-\Phi S_{12})$  can be replaced by  ${}_0F_1[\tfrac{1}{2}(m-k); \tfrac{1}{4}\Phi' \Phi S_{12} S_{12}']$  in the integral  $I$ . Next, transforming the resulting integral with respect to  $(dS_{12})$  to that with respect to  $(dU)$ , where  $U = S_{11}^{-1/2} S_{12} S_{12}' S_{11}^{-1/2}/2n$ , assuming  $k \leq m-k$ , in view of (1.7), the integral  $I$  becomes

$$\begin{aligned} I &= c |S_{11}|^{(m-k)/2} \int_0^{I_k} |U|^{(m-2k-1)/2} |I_k - U|^{[(m-k)(n-1)-k-1]/2} \\ &\quad \times {}_0F_1[\tfrac{1}{2}(m-k); \tfrac{1}{2}n S_{11}^{1/2} \Phi' \Phi S_{11}^{1/2} U] (dU), \end{aligned} \quad (2.27)$$

for a suitable constant  $c$ . Expressing the  ${}_0F_1$  function in (2.27) as the inverse Laplace transform of the  ${}_0F_0$  function (see [19, (29)]), integrating over  $0 < U < I_k$  in view of the integral representation of the  ${}_1F_1$  function (see [19, (47)]), and finally evaluating the inverse Laplace transform of the  ${}_1F_1$  function (see [19, (29)]), we are led to, for a suitable constant  $c'$ ,

$$I = c' |S_{11}|^{(m-k)/2} {}_0F_1[\tfrac{1}{2}(m-k)n; \tfrac{1}{2}n\Phi' \Phi S_{11}].$$

Summarizing the above results establishes

**Theorem 2.5.** *We assume that  $p = k$  and that the hypothesis  $H_1$  (2.19) holds. For large  $A$ , the distribution of the matrix statistic  $S = (S_{ij})$  defined by (2.20) has the following*

asymptotic properties:

(i)  $S_{22}$  is independent of  $S_{11}$  and  $S_{12}$ , and it holds that  $S_{22} = 2nI_{m-k}$  almost everywhere.

(ii)  $S_{12}$  has the marginal normal  $N_{k,m-k}(-2n\Phi', 2nI_k \otimes I_{m-k})$  distribution, and hence  $S_{12}S'_{12}$  follows the marginal noncentral Wishart  $W_k(m-k, 2nI_k; 4n^2\Phi'\Phi)$  distribution (assuming  $k \leq m-k$ ).  $S_{11} - S_{12}S'_{12}/2n$  is distributed as Wishart  $W_k[(m-k)(n-1), I_k]$  conditionally with given  $S_{12}$  (assuming  $S_{11} - S_{12}S'_{12}/2n > 0$ ).

(iii)  $S_{11}$  has the marginal noncentral Wishart  $W_k[(m-k)n, I_k; 2n\Phi'\Phi]$  distribution. The conditional density function of  $S_{12}$  given  $S_{11}$  is given by

$$\frac{\Gamma_k[\frac{1}{2}(m-k)n]}{\pi^{k(m-k)/2}\Gamma_k[\frac{1}{2}(m-k)(n-1)]} \frac{|I_k - (2nS_{11})^{-1}S_{12}S'_{12}|^{[(m-k)(n-1)-k-1]/2} \text{etr}(-\Phi S_{12})}{|2nS_{11}|^{(m-k)/2} {}_0F_1[\frac{1}{2}(m-k)n; \frac{1}{2}n\Phi'\Phi S_{11}]}$$

**Remark 2.6.** From Theorem 2.5, it might be suggested to use statistics based on  $S_{11}$  for the test on the parameter  $\Gamma$ , for large  $\Lambda$ . We note that, with  $L_i$  being the density of the sample under the hypothesis  $H_i$ ,  $i = 0, 1$ ,  $\log(L_0/L_1) = -n \text{tr}(\Lambda^{1/2}\Gamma'_0\bar{P}\Gamma_0^\perp\Phi - \Phi'\Gamma_0^\perp\bar{P}\Gamma_0\Lambda^{1/2})$ , ignoring the terms of  $O(1)$ . Therefore  $S_{12}$  (or  $S_{12}S'_{12}$ ) may be also a useful matrix statistic for the test on the parameter  $\Gamma$ , for large  $\Lambda$ . Further correction terms of  $O(\Lambda^{-1})$ , in terms of the noncentral Laguerre polynomials with matrix arguments, of the asymptotic expansion for the distribution of  $S_{11}$  have been given by Chikuse [7].

**Corollary 2.7.** Putting  $\Phi = 0$  in Theorems 2.5 yields the asymptotic properties under the null hypothesis  $H_0$  (2.18).

For the rest of Section 2.2, we discuss some asymptotic results for large  $\Lambda$  concerning the  $L^{(P)}(m, k; B)$  distribution with the (unique) s.d.  $B = \Gamma\Lambda\Gamma'$  of rank  $k < m$ , which can be derived from Theorem 2.5. Letting  $P$  be a random matrix distributed as  $L^{(P)}(m, k; B)$ , we can express

$$P = \Gamma\Gamma' + \frac{1}{2}\Gamma\Lambda^{-1/2}S_{12}\Gamma'^\perp + \frac{1}{2}\Gamma^\perp S'_{12}\Lambda^{-1/2}\Gamma' + O(\Lambda^{-1}) \quad \text{for large } \Lambda,$$

where  $S_{12}$  is distributed as  $N_{k,m-k}(0, 2I_k \otimes I_{m-k})$ ; indicating that  $P$  approaches the mode  $\Gamma\Gamma'$  in probability for large  $\Lambda$ .

We consider the problem of testing the null hypothesis  $H_0 : \Gamma = \Gamma_0$  against the alternative hypothesis  $H_1 : \Gamma \neq \Gamma_0$ , when  $\Lambda$  is known and large. Putting the (unique) s.d. of the sample mean matrix  $\bar{P}$  of sample size  $n$  as  $\bar{P} = \bar{H}\bar{P}_d\bar{H}'$ , where  $\bar{H} = (\bar{H}_1 \ \bar{H}_1^\perp) \in O(m)$  with  $\bar{H}_1$  being  $m \times k$ , and  $\bar{P}_d = \text{diag}(p_1, \dots, p_m) = \text{diag}(\bar{P}_{d(1)}, \bar{P}_{d(2)}, 1 > p_1 > \dots > p_m > 0$ , with  $\bar{P}_{d(1)}$  being  $k \times k$ , we have the m.l.e. of  $\Gamma$ ,  $\hat{\Gamma} = \bar{H}_1$  (see Section 3.1), and the likelihood ratio test statistic is given by  $2n \text{tr}(\bar{P}_{d(1)} - \Gamma'_0\bar{P}\Gamma_0)\Lambda$ , being distributed asymptotically for large  $n$  and  $\Lambda$  as  $\chi^2_{km-k(k+1)/2}$  under  $H_0$ . We use the argument of analysis of variance, where we

have the identity “ $n \operatorname{tr}(I_k - \Gamma'_0 \bar{P} \Gamma_0) A = n \operatorname{tr}(I_k - \bar{P}_{d(1)}) A + n \operatorname{tr}(\bar{P}_{d(1)} - \Gamma'_0 \bar{P} \Gamma_0) A$ ”, and we have already seen that  $2n \operatorname{tr}(I_k - \Gamma'_0 \bar{P} \Gamma_0) A$  is distributed asymptotically for large  $A$  as  $\chi^2_{nk(m-k)}$ . Thus we may use the F-test statistic

$$(f_2/f_1)[\operatorname{tr}(\bar{P}_{d(1)} - \Gamma'_0 \bar{P} \Gamma_0) A] / [\operatorname{tr}(I_k - \bar{P}_{d(1)}) A], \quad (2.28)$$

with  $f_1 = km - k(k+1)/2$  and  $f_2 = nk(m-k) - f_1$ , which is distributed asymptotically for large  $n$  and  $A$  as  $F_{f_1, f_2}$  under  $H_0$ , and as noncentral  $F_{f_1, f_2; 2n \operatorname{tr} \Phi' \Phi}$  under the alternative hypothesis  $H_1$  (2.19).

### 3. Other asymptotics

#### 3.1. Estimation of large concentration parameters

##### 3.1.1. On the Stiefel manifold $V_{k,m}$

Let  $X_1, \dots, X_n$  be a random sample from the matrix Langevin distribution (1.1) where  $\lambda_1 > \dots > \lambda_p > 0$  are distinct, and put  $\bar{X} = \sum_{i=1}^n X_i/n$ . The m.l.e.'s  $\hat{\Gamma}$ ,  $\hat{\Theta}$ , and  $\hat{A} = \operatorname{diag}(\hat{\lambda}_1, \dots, \hat{\lambda}_p)$  of  $\Gamma$ ,  $\Theta$ , and  $A$  are given by the matrix of the latent vectors of  $\bar{X}\bar{X}'$ , that of  $\bar{X}'\bar{X}$ , corresponding to the first  $p$  largest ones of the latent roots  $x_1^2, \dots, x_k^2$  of  $\bar{X}'\bar{X}$  ( $x_1 > \dots > x_k > 0$ ), and the solution of

$$\partial \log_0 F_1(\tfrac{1}{2}m; \tfrac{1}{4}\hat{A}^2) / \partial \hat{\lambda}_i = x_i, \quad i = 1, \dots, p, \quad (3.1)$$

respectively (see [21]).

The matrix variate  $\bar{X}$  is a sufficient statistic, and its density function  $f_{\bar{X}}(\bar{X}; F)$  was obtained by Chikuse [4]. It is seen that  $f_{\bar{X}}(\bar{X}; F)$  is invariant both under the simultaneous transformation

$$\hat{\Gamma} \rightarrow H_1 \hat{\Gamma} \quad \text{and} \quad \Gamma \rightarrow H_1 \Gamma, \quad \text{for } H_1 \in O(m) \quad (3.2)$$

[and  $\hat{\Gamma}^\perp \rightarrow H_1 \hat{\Gamma}^\perp$ , where  $(\hat{\Gamma} \hat{\Gamma}^\perp) \in O(m)$  is the matrix of the latent vectors of  $\bar{X}\bar{X}'$ ], and under the simultaneous transformation

$$\hat{\Theta} \rightarrow H_2 \hat{\Theta} \quad \text{and} \quad \Theta \rightarrow H_2 \Theta, \quad \text{for } H_2 \in O(k) \quad (3.3)$$

[and  $\hat{\Theta}^\perp \rightarrow H_2 \hat{\Theta}^\perp$ , where  $(\hat{\Theta} \hat{\Theta}^\perp) \in O(k)$  is the matrix of the latent vectors of  $\bar{X}'\bar{X}$ ]. Therefore,  $\bar{X}_d = \operatorname{diag}(x_1, \dots, x_k)$  is quasi-sufficient for  $A$  in the sense of Barnard [2], and it is suggested to estimate  $A$  based on the marginal distribution of  $\bar{X}_d$ . It is seen that this maximum marginal likelihood estimator (m. marginal l.e.)  $\hat{\hat{A}} = \operatorname{diag}(\hat{\hat{\lambda}}_1, \dots, \hat{\hat{\lambda}}_p)$  of  $A$  is obtained as the Bayes estimator of  $A$  by employing the uniform prior distributions for  $\Gamma$  and  $\Theta$ . Then the marginal likelihood function is given by

$${}_0F_1^{(k)}(\tfrac{1}{2}m; \tfrac{1}{4}A^2, n^2 \bar{X}_d^2) / [{}_0F_1(\tfrac{1}{2}m; \tfrac{1}{4}A^2)]^n,$$

which leads to the m. marginal l.e.  $\hat{\Lambda}$  given by the solution of

$$\partial[-n \log {}_0F_1(\tfrac{1}{2}m; \tfrac{1}{4}\hat{\Lambda}^2) + \log {}_0F_1^{(k)}(\tfrac{1}{2}m; \tfrac{1}{4}\hat{\Lambda}^2, n^2 \bar{X}_d^2)] / \partial \hat{\lambda}_i = 0, \quad i = 1, \dots, p. \quad (3.4)$$

Here the  ${}_pF_q^{(k)}$  is a hypergeometric function with two matrix arguments (see e.g., [19]).

Next, taking the simultaneous transformation (3.2) only into consideration,  $\bar{X}'\bar{X}$  is quasi-sufficient for  $\Theta$  and  $\Lambda$ . The maximum semi-marginal likelihood estimator (m. semi-marginal l.e.)  $\hat{\Lambda}^* = \text{diag}(\hat{\lambda}_1^*, \dots, \hat{\lambda}_p^*)$  of  $\Lambda$  is obtained from the distributions of  $\bar{X}'\bar{X}$ , and is equivalent to the Bayes estimator of  $\Lambda$  obtained by employing the uniform prior distribution for  $\Gamma$ . Then the semi-marginal likelihood function is given by

$${}_0F_1(\tfrac{1}{2}m; \tfrac{1}{4}n^2 \Theta \Lambda^2 \Theta' \bar{X}'\bar{X}) / [{}_0F_1(\tfrac{1}{2}m; \tfrac{1}{4}\Lambda^2)]^n.$$

It is seen that the m. semi-marginal l.e. of  $\Theta$  is equal to the m.l.e.  $\hat{\Theta}$ , and the m. semi-marginal l.e.  $\hat{\Lambda}^*$  is given by the solution of

$$\partial[-n \log {}_0F_1(\tfrac{1}{2}m; \tfrac{1}{4}\hat{\Lambda}^{*2}) + \log {}_0F_1(\tfrac{1}{2}m; \tfrac{1}{4}n^2 \hat{\Lambda}^{*2} \bar{X}_{d(1)}^2)] / \partial \hat{\lambda}_i^* = 0, \quad i = 1, \dots, p, \quad (3.5)$$

where  $\bar{X}_{d(1)} = \text{diag}(x_1, \dots, x_p)$ .

When we take the simultaneous transformation (3.3) only into consideration,  $\bar{X}\bar{X}'$  is quasi-sufficient for  $\Gamma$  and  $\Lambda$ , and a similar method leads to the m. semi-marginal l.e.'s of  $\Gamma$  and  $\Lambda$  which are equal to the m.l.e  $\hat{\Gamma}$  and the m. semi-marginal l.e.  $\hat{\Lambda}^*$  given by (3.5), respectively.

Systems (3.1), (3.4), and (3.5) of partial differential equations can be solved approximately for small  $\Lambda$  (see [5]). Here we shall show a method to evaluate the approximate estimators of large  $\Lambda$ . Using the asymptotic expansion for  ${}_0F_1(\tfrac{1}{2}m; \tfrac{1}{4}\hat{\Lambda}^2)$  for large  $\hat{\Lambda}$  (see [23, Section 5]), (3.1) becomes

$$1 - \frac{m-p}{2\hat{\lambda}_i} - \frac{1}{2} \sum_{j=1, j \neq i}^p \frac{1}{\hat{\lambda}_i + \hat{\lambda}_j} + O(\hat{\Lambda}^{-2}) = x_i, \quad i = 1, \dots, p, \quad (3.6)$$

and, similarly for large  $\hat{\Lambda}^*$ , (3.5) becomes

$$\begin{aligned} (1 - x_i) - \frac{(m-p)(n-1)}{2n\hat{\lambda}_i^*} - \frac{1}{2} \sum_{j=1, j \neq i}^p \frac{1}{\hat{\lambda}_i^* + \hat{\lambda}_j^*} \\ + \frac{1}{2n} \sum_{j=1, j \neq i}^p \frac{x_i}{x_i\hat{\lambda}_i^* + x_j\hat{\lambda}_j^*} + O(\hat{\Lambda}^{*-2}) = 0, \quad i = 1, \dots, p. \end{aligned} \quad (3.7)$$

Using the asymptotic expansion for  ${}_0F_1^{(k)}(\frac{1}{2}m; \frac{1}{4}\hat{A}^2, n^2\bar{X}_d^2)$  for large  $\hat{A}$  (see [23, (6.4)]), (3.4) becomes

$$(1 - x_i) - \frac{(m-p)(n-1) - k + p}{2n\hat{\lambda}_i} - \frac{1}{2} \sum_{j=1, j \neq i}^p \frac{1}{\hat{\lambda}_i + \hat{\lambda}_j} + \frac{1}{n} \sum_{j=1, j \neq i}^p \frac{\hat{\lambda}_i}{\hat{\lambda}_i^2 - \hat{\lambda}_j^2} + O(\hat{A}^{-2}) = 0, \quad i = 1, \dots, p. \quad (3.8)$$

It is seen that the above three kinds of estimators are asymptotically equivalent to one another for large sample size  $n$ . Asymptotic solutions (3.6)–(3.8) may be obtained by iterative methods. For a simple comparison, let us consider the case  $p = 1$ . We have

$$\hat{\lambda}_1 = (m-1)/[2(1-x_1)] + O(\lambda_1^{-1}),$$

$$\hat{\lambda}_1^* = \hat{\lambda}_1 - [(m-1)/2n(1-x_1)] + O(\lambda_1^{-1})$$

and

$$\hat{\lambda}_1 = \hat{\lambda}_1^* - [(k-1)/2n(1-x_1)] + O(\lambda_1^{-1}),$$

and hence,  $\hat{\lambda}_1 \geq \hat{\lambda}_1^* \geq \hat{\lambda}_1$ , ignoring the terms of  $O(\lambda_1^{-1})$ , where the first and second equalities hold if and only if  $m = 1$  and  $k = 1$ , respectively.

### 3.1.2. On the manifold $P_{k,m-k}$

Let  $P_1, \dots, P_n$  be a random sample from the matrix Langevin distribution (1.3), where  $\lambda_1 > \dots > \lambda_p$  are distinct with  $p < m$  [see (1.6)], and put  $\bar{P} = \sum_{i=1}^n P_i/n$ . Let the s.d. of  $\bar{P}$  be  $\bar{P} = \bar{H}\bar{P}_d\bar{H}'$ , where  $\bar{H} \in O(m)$ , and  $\bar{P}_d = \text{diag}(p_1, \dots, p_m)$ ,  $1 > p_1 > \dots > p_m > 0$ . Due to Chikuse and Watson [11], the m.l.e.  $\hat{\Gamma}$  of  $\Gamma$  is seen to be  $\bar{H}_1$ , where  $\bar{H} = (\bar{H}_1 \bar{H}_1^\perp)$  with  $\bar{H}_1$  being  $m \times p$ , and the m.l.e.  $\hat{A} = \text{diag}(\hat{\lambda}_1, \dots, \hat{\lambda}_p)$  of  $A$  is given by the solution of

$$\partial \log {}_1F_1(\frac{1}{2}k; \frac{1}{2}m; \hat{A}) / \partial \hat{\lambda}_i = p_i, \quad i = 1, \dots, p. \quad (3.9)$$

The matrix variate  $\bar{P}$  is a sufficient statistic, and its density function is seen to be invariant under the simultaneous transformation

$$\Gamma \rightarrow H\Gamma \quad \text{and} \quad \hat{\Gamma} \rightarrow H\hat{\Gamma}, \quad \text{for } H \in O(m) \quad (3.10)$$

(and  $\hat{\Gamma}^\perp \rightarrow H\hat{\Gamma}^\perp$ , where  $\hat{\Gamma}^\perp = \bar{H}_1^\perp$ ).  $\bar{P}_d$  is quasi-sufficient for  $A$  in the sense of Barnard [2], and we shall estimate  $A$  based on the marginal distribution of  $\bar{P}_d$  by the Bayes method employing the uniform prior distribution for  $\Gamma$ . Then the marginal likelihood function is seen to be

$${}_0F_0^{(m)}(A, n\bar{P}_d) / [{}_1F_1(\frac{1}{2}k; \frac{1}{2}m; A)]^n,$$

which leads to the m. marginal i.e.  $\hat{\Lambda} = \text{diag}(\hat{\lambda}_1, \dots, \hat{\lambda}_p)$  given by the solution of

$$\partial[-n \log {}_1F_1(\tfrac{1}{2}k; \tfrac{1}{2}m; \hat{\Lambda}) + \log {}_0F_0^{(m)}(\hat{\Lambda}, n\bar{P}_d)] / \partial \hat{\lambda}_i = 0, \quad i = 1, \dots, p. \quad (3.11)$$

The approximate evaluation of systems (3.9) and (3.11) of partial differential equations for small  $\Lambda$  was given (see [11]), and here we shall solve the systems for large  $\Lambda$ . Using the asymptotic expansion for  ${}_1F_1(\tfrac{1}{2}k; \tfrac{1}{2}m; \hat{\Lambda})$  for large  $\hat{\Lambda}$  (see [13, Theorem 3.2]), (3.9) becomes

$$1 - [(m - k)/2\hat{\lambda}_i] + O(\hat{\Lambda}^{-2}) = p_i, \quad i = 1, \dots, p, \quad (3.12)$$

that is,

$$\hat{\lambda}_i = (m - k)/[2(1 - p_i)] + O(\Lambda^{-1}), \quad i = 1, \dots, p. \quad (3.13)$$

Similarly, using the asymptotic expansion for  ${}_0F_0^{(m)}(\hat{\Lambda}, n\bar{P}_d)$  for large  $\hat{\Lambda}$  (see [13, Section 5]), (3.11) becomes

$$(1 - p_i) - \frac{(m - k)n - m + p}{2n\hat{\lambda}_i} + \frac{1}{2n} \sum_{j=1, j \neq i}^p \frac{1}{\hat{\lambda}_i - \hat{\lambda}_j} + O(\hat{\Lambda}^{-2}) = 0, \quad i = 1, \dots, p. \quad (3.14)$$

It is seen that the above two kinds of estimators are asymptotically equivalent to one another for large sample size  $n$ . An asymptotic solution given by (3.14) may be obtained iteratively. For the simple case  $p = 1$ , we have

$$\hat{\lambda}_1 = (m - k)/[2(1 - p_1)] + O(\lambda_1^{-1})$$

and

$$\hat{\lambda}_1 = \hat{\lambda}_1 - [(m - 1)/2n(1 - p_1)] + O(\lambda_1^{-1}).$$

Therefore, ignoring the terms of  $O(\lambda_1^{-1})$ , we have  $\hat{\lambda}_1 \geq \hat{\lambda}_1$ , where the equality holds if and only if  $m = 1$ .

### 3.2. Classification of the matrix Langevin distributions

#### 3.2.1. On the Stiefel manifold $V_{k,m}$

We consider the problem of classification into one of two (known) matrix Langevin distributions  $\pi_j: L(m, k; F_j)$ , with the s.v.d.'s  $F_j = \Gamma_j \Lambda_j \Theta_j'$ ,  $j = 1, 2$ . For the classification analysis the reader may be referred to a standard textbook of multivariate statistical analysis (e.g. [1]). Taking the ratio of the density functions, the best classification rule is such that, given an observation  $X \in V_{k,m}$ , we choose

$$\pi_1 \quad \text{if } u(X) = \text{tr}(F_1 - F_2)'X \geq c$$

and

$$\pi_2 \quad \text{if } u(X) < c, \quad (3.15)$$

for a suitably chosen constant  $c$ . The misclassification probabilities are given by  $\Pr(u(X) < c \mid \pi_1)$  and  $\Pr(u(X) \geq c \mid \pi_2)$ . We investigate the asymptotic behavior of the classification function  $u(X)$  when  $A_1$  and  $A_2$  are both large.

We may be interested in the situation where the two populations are *locally close* in the following sense. Assuming that  $A_1 = A_2 = A$  being of rank  $p$ , for simplicity of argument, we consider the case where the following conditions are satisfied:

$$\Gamma_2 = (\Gamma_1 + \Psi_1 A^{-1/2})(I_p + A^{-1/2} \Psi_1' \Psi_1 A^{-1/2})^{-1/2} \quad \text{with } \Psi_1' \Gamma_1 = 0$$

and

$$\Theta_2 = (\Theta_1 + \Psi_2 A^{-1/2})(I_p + A^{-1/2} \Psi_2' \Psi_2 A^{-1/2})^{-1/2} \quad \text{with } \Psi_2' \Theta_1 = 0, \quad (3.16)$$

so that we can write  $\Psi_1 = \Gamma_1^\perp \Phi_1$  and  $\Psi_2 = \Theta_1^\perp \Phi_2$  for certain  $(m-p) \times p$  and  $(k-p) \times p$  matrices  $\Phi_1$  and  $\Phi_2$ , respectively.

The characteristic function of  $u(X)$  when  $\pi_1$  is true is

$$\Phi^{(\pi_1)}(t) = E_{\pi_1} e^{itu(X)} = {}_0F_1\left(\frac{1}{2}m; \frac{1}{4}AA'\right) / {}_0F_1\left(\frac{1}{2}m; \frac{1}{4}A^2\right),$$

with

$$A = (1 + it)\Theta_1 A \Gamma_1' - it\Theta_2 A \Gamma_2'. \quad (3.17)$$

Under condition (3.16), we have

$$AA' = O_2(A^2) + O_{3/2}(A^{3/2}) + O_1(A) + O(A^{1/2}),$$

where

$$O_2(A^2) = \Theta_1 A^2 \Theta_1', \quad O_{3/2}(A^{3/2}) = -it(\Theta_1 A^{3/2} \Phi_2' \Theta_1^{\perp'} + \Theta_1^\perp \Phi_2 A^{3/2} \Theta_1')$$

and

$$O_1(A) = it\Theta_1(A^{1/2} \Phi_1' \Phi_1 A^{1/2} + \frac{1}{2}A^{3/2} \Phi_2' \Phi_2 A^{-1/2} + \frac{1}{2}A^{-1/2} \Phi_2' \Phi_2 A^{3/2})\Theta_1' \\ + (it)^2(\Theta_1 A^{1/2} \Phi_1' \Phi_1 A^{1/2} \Theta_1' + \Theta_1^\perp \Phi_2 A \Phi_2' \Theta_1^{\perp'}).$$

Putting  $A = NA_0$  with  $N (>0)$  large and  $A_0$  fixed, and  $\varepsilon = N^{-1/2}$ , we apply the perturbation method to the matrix  $B = AA'/N^2$  as in Section 2. Writing the latent roots  $\mu_i$  of  $B$  as  $\mu_i = \xi_i + \varepsilon \xi_{i1} + \varepsilon^2 \xi_{i2} + O(\varepsilon^3)$ , we obtain

$$\xi_i = \lambda_{0i}^2, \quad \xi_{i1} = 0,$$

$$\xi_{i2} = it(1 + it)\lambda_{0i} \left( \sum_{j=1}^2 \Phi_j' \Phi_j \right)_{ii}, \quad \text{for } i = 1, \dots, p,$$

$$\xi_i = \xi_{i1} = \xi_{i2} = 0, \quad \text{for } i = p+1, \dots, k,$$

where  $\lambda_i = N\lambda_{0i}$ ,  $i = 1, \dots, p$ , leading to

$$\Phi^{(\pi_1)}(t) = \exp \left[ -\frac{1}{2} \delta (t^2 - it) \right] [1 + O(A^{-1})], \quad \text{with } \delta = \text{tr} \left( \sum_{j=1}^2 \Phi_j' \Phi_j \right).$$

The characteristic function  $\Phi^{(\pi_2)}(t)$  of  $u(X)$  when  $\pi_2$  is true is given in the form (3.17) with  $A = it\Theta_1\Lambda\Gamma'_1 + (1 - it)\Theta_2\Lambda\Gamma'_2$ . Interchanging “ $it$ ” with “ $1 + it$ ” in the previous discussion for  $\Phi^{(\pi_1)}(t)$  may yield

$$\Phi^{(\pi_2)}(t) = \exp[-\frac{1}{2}\delta(t^2 + it)][1 + O(\Lambda^{-1})].$$

Summarizing the above discussion yields

**Theorem 3.1.** *The classification function  $u(X)$  defined by (3.15) for the problem of classification is distributed asymptotically, for large  $\Lambda$ , as normal  $N(\frac{1}{2}\delta, \delta)$  when  $\pi_1$  is true, and  $N(-\frac{1}{2}\delta, \delta)$  when  $\pi_2$  is true, where  $\delta = \text{tr}(\sum_{j=1}^2 \Phi'_j \Phi_j)$  may indicate a local distance between the two sets of orientations  $(\Gamma_j, \Theta_j)$ ,  $j = 1, 2$ . The misclassification probabilities are given by*

$$P(u(X) < c \mid \pi_1) = (2\pi\delta)^{-1/2} \int_{-\infty}^c \exp\left[-\frac{1}{2\delta}\left(u - \frac{1}{2}\delta\right)^2\right] du$$

and

$$P(u(X) \geq c \mid \pi_2) = (2\pi\delta)^{-1/2} \int_c^{\infty} \exp\left[-\frac{1}{2\delta}\left(u + \frac{1}{2}\delta\right)^2\right] du.$$

### 3.2.2. On the manifold $P_{k,m-k}$

We shall consider a similar classification problem into one of two (known) matrix Langevin distributions  $\pi_j : L^{(P)}(m, k; B_j)$ , with the s.d.'s  $B_j = \Gamma_j \Lambda \Gamma'_j$ ,  $j = 1, 2$ , with  $\Lambda$  being of rank  $p < m$ . The best classification rule is such that, given an observation  $P \in P_{k,m-k}$ , we choose

$$\pi_1 \quad \text{if } u(P) = \text{tr}(B_1 - B_2)P \geq c$$

and

$$\pi_2 \quad \text{if } u(P) < c, \quad (3.18)$$

for a suitably chosen constant  $c$ . We may be interested in the case where the two populations are locally close, that is, when

$$\Gamma_2 = (\Gamma_1 + \Psi \Lambda^{-1/2})(I_p + \Lambda^{-1/2} \Psi' \Psi \Lambda^{-1/2})^{-1/2} \quad \text{with } \Psi' \Gamma_1 = 0,$$

so that we can write  $\Psi = \Gamma_1^\perp \Phi$  for a certain  $(m-p) \times p$  matrix  $\Phi$ .

The characteristic function of  $u(P)$  when  $\pi_1$  is true is

$$\Phi^{(\pi_1)}(t) = E_{\pi_1} e^{itu(P)} = {}_1F_1\left(\frac{1}{2}k; \frac{1}{2}m; A\right) / {}_1F_1\left(\frac{1}{2}k; \frac{1}{2}m; A\right), \quad (3.19)$$

where

$$A = \begin{bmatrix} \Lambda^{1/2} [I_p + \frac{1}{2}it(\Phi' \Phi \Lambda^{-1} + \Lambda^{-1} \Phi' \Phi)] \Lambda^{1/2} & -it\Lambda^{1/2} \Phi' \\ -it\Phi \Lambda^{1/2} & -it\Phi \Phi' \end{bmatrix} + O(\Lambda^{-1/2}). \quad (3.20)$$

Using the asymptotic expansions for the  ${}_1F_1$  functions with the matrix argument  $\Lambda$  and with the matrix argument of type (3.20) for large  $\Lambda$  (see [13, Theorems 3.2 and 3.1, respectively]), and assuming  $p = k$ , we obtain

$$\Phi^{(\pi_1)}(t) = \exp[-\frac{1}{2}\delta(t^2 - it)][1 + O(\Lambda^{-1/2})] \quad \text{with } \delta = 2 \operatorname{tr} \Phi' \Phi.$$

The characteristic function of  $\Phi^{(\pi_2)}(t)$  of  $u(P)$  when  $\pi_2$  is true is given by (3.19) together with (3.20) with “ $it$ ” being interchanged with “ $1 + it$ ”, and hence

$$\Phi^{(\pi_2)}(t) = \exp[-\frac{1}{2}\delta(t^2 + it)][1 + O(\Lambda^{-1/2})].$$

Thus summarizing the above discussion yields the following.

**Theorem 3.2.** *The classification function  $u(P)$  defined by (3.18) for the problem of classification is distributed asymptotically, for large  $\Lambda$ , as normal  $N(\frac{1}{2}\delta, \delta)$  when  $\pi_1$  is true, and  $N(-\frac{1}{2}\delta, \delta)$  when  $\pi_2$  is true, where  $\delta = 2 \operatorname{tr} \Phi' \Phi$  may indicate a local distance between the orientations  $\Gamma_1$  and  $\Gamma_2$ . The misclassification probabilities are given by*

$$P(u(P) < c \mid \pi_1) = (2\pi\delta)^{-1/2} \int_{-\infty}^c \exp\left[-\frac{1}{2\delta}\left(u - \frac{1}{2}\delta\right)^2\right] du$$

and

$$P(u(P) \geq c \mid \pi_2) = (2\pi\delta)^{-1/2} \int_c^{\infty} \exp\left[-\frac{1}{2\delta}\left(u + \frac{1}{2}\delta\right)^2\right] du.$$

## References

- [1] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, 2nd Edition, Wiley, New York, 1984.
- [2] G.A. Barnard, Some logical aspects of the fiducial argument, J. Roy. Statist. Soc. Ser. B 25 (1963) 111–114.
- [3] R. Bellman, Introduction to Matrix Analysis, 2nd Edition, McGraw-Hill, New York, 1970.
- [4] Y. Chikuse, Distributions of orientations on Stiefel manifolds, J. Multivariate Anal. 33 (1990) 247–264.
- [5] Y. Chikuse, Asymptotic expansions for distributions of the large sample matrix resultant and related statistics on the Stiefel manifold, J. Multivariate Anal. 39 (1991) 270–283.
- [6] Y. Chikuse, High dimensional asymptotic expansions for the matrix Langevin distributions on the Stiefel manifold, J. Multivariate Anal. 44 (1993) 82–101.
- [7] Y. Chikuse, Asymptotic theory for the concentrated matrix Langevin distributions on the Grassmann manifold, in: K. Matsusita, M.L. Puri, T. Hayakawa (Eds.), Statistical Sciences and Data Analysis, VSP, Utrecht, 1993, pp. 237–245.
- [8] Y. Chikuse, Density estimation on the Stiefel manifold, J. Multivariate Anal. 66 (1998) 188–206.
- [9] Y. Chikuse, Procrustes analysis on some special manifolds, Comm. Statist.-Theory Methods 28 (1999) 885–903.
- [10] Y. Chikuse, Orthogonal projective Gaussian distributions on the Grassmann manifold, being submitted.
- [11] Y. Chikuse, G.S. Watson, Large sample asymptotic theory of tests for uniformity on the Grassmann manifold, J. Multivariate Anal. 54 (1995) 18–31.
- [12] A.G. Constantine, Some non-central distribution problems in multivariate analysis, Ann. Math. Statist. 34 (1963) 1270–1285.

- [13] A.G. Constantine, R.J. Muirhead, Asymptotic expansions for distributions of latent roots in multivariate analysis, *J. Multivariate Anal.* 6 (1976) 369–391.
- [14] T.D. Downs, Orientation statistics, *Biometrika* 59 (1972) 665–676.
- [15] R.H. Farrell, *Multivariate Calculation*, Springer, New York, 1985.
- [16] N.I. Fisher, T. Lewis, B.J.J. Embleton, *Statistical Analysis of Spherical Data*, Cambridge University Press, London and New York, 1987.
- [17] C.S. Herz, Bessel functions of matrix argument, *Ann. Math.* 61 (1955) 474–523.
- [18] A.T. James, Normal multivariate analysis and the orthogonal group, *Ann. Math. Statist.* 25 (1954) 40–75.
- [19] A.T. James, Distributions of matrix variates and latent roots derived from normal samples, *Ann. Math. Statist.* 35 (1964) 475–501.
- [20] P.E. Jupp, K.V. Mardia, Maximum likelihood estimators for the matrix von Mises–Fisher and Bingham distributions, *Ann. Statist.* 7 (1979) 599–606.
- [21] C.G. Khatri, K.V. Mardia, The von Mises–Fisher matrix distribution in orientation statistics, *J. Roy. Statist. Soc. Ser. B* 39 (1977) 95–106.
- [22] K.V. Mardia, C.G. Khatri, Uniform distribution on a Stiefel manifold, *J. Multivariate Anal.* 7 (1977) 468–473.
- [23] R.J. Muirhead, Latent roots and matrix variates: a review of some asymptotic results, *Ann. Statist.* 6 (1978) 5–33.
- [24] R.J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley, New York, 1982.
- [25] M.J. Prentice, Orientation statistics without parametric assumptions, *J. Roy. Statist. Soc. Ser. B* 48 (1986) 214–222.
- [26] G.S. Watson, Orientation statistics in the earth sciences, *Bull. Geol. Inst. Univ. Uppsala* 2 (1970) 73–89.
- [27] G.S. Watson, *Statistics on Spheres*, in: *Lecture Notes in Mathematics* (University of Arkansas), Vol. 6, Wiley, New York, 1983.
- [28] G.S. Watson, The theory of concentrated Langevin distributions, *J. Multivariate Anal.* 14 (1984) 74–82.