# Complete Asymptotic Expansions for the Normalizing Constants of High-Dimensional Matrix Bingham and Matrix Langevin Distributions

*Armine BAGYAN and Donald RICHARDS*

*Department of Statistics, Pennsylvania State University, University Park, PA 16802, USA*
E-mail: *aub171@psu.edu, dsr11@psu.edu*

**Abstract.** For positive integers $d$ and $p$ such that $d \geq p$, let $\mathbb{R}^{d \times p}$ denote the set of $d \times p$ real matrices, $I_p$ be the identity matrix of order $p$, and $V_{d,p} = \{x \in \mathbb{R}^{d \times p} \mid x'x = I_p\}$ be the Stiefel manifold in $\mathbb{R}^{d \times p}$. Complete asymptotic expansions as $d \to \infty$ are obtained for the normalizing constants of the matrix Bingham and matrix Langevin probability distributions on $V_{d,p}$. The accuracy of each truncated expansion is strictly increasing in $d$; also, for sufficiently large $d$, the accuracy is strictly increasing in $m$, the number of terms in the truncated expansion. Lower bounds are obtained for the truncated expansions when the matrix parameters of the matrix Bingham distribution are positive definite and when the matrix parameter of the matrix Langevin distribution is of full rank. These results are applied to obtain the rates of convergence of the asymptotic expansions as both $d \to \infty$ and $p \to \infty$. Values of $d$ and $p$ arising in numerous data sets are used to illustrate the rate of convergence of the truncated approximations as $d$ or $m$ increases. These results extend recently-obtained asymptotic expansions for the normalizing constants of the high-dimensional Bingham distributions.

*Key words:* Frobenius norm; generalized hypergeometric function of matrix argument; Grassmann manifold; Hadamard's inequality; hippocampus; neural spike activity; Stiefel manifold; Super Chris (the rat); symmetric cone; zonal polynomial

*2020 Mathematics Subject Classification:* 60E05; 62H11; 62E20; 62R30

## 1 Introduction

For positive integers $d$ and $p$ such that $d \geq p$ let $\mathbb{R}^{d \times p}$ denote the set of $d \times p$ real matrices, $I_p$ be the identity matrix of order $p$, and $V_{d,p} = \{x \in \mathbb{R}^{d \times p} \mid x'x = I_p\}$ denote the *Stiefel manifold*. The manifold $V_{d,p}$ is endowed with a unique probability measure, denoted by $\mathrm{d}x$, that is invariant under the change of variables $x \mapsto HxK$, where $x \in V_{d,p}$, and $H$ and $K$ are orthogonal matrices of orders $d \times d$ and $p \times p$, respectively (see Chikuse [4, p. 16], Muirhead [30, p. 70]).

In this article, we first consider the matrix Bingham probability distributions on $V_{d,p}$. For non-zero symmetric matrices $A \in \mathbb{R}^{p \times p}$ and $\Sigma \in \mathbb{R}^{d \times d}$, a random matrix $X \in V_{d,p}$ is said to have a *(full) matrix Bingham distribution with parameters $A$ and $\Sigma$* if the probability density function of $X$, with respect to the invariant measure $\mathrm{d}x$, is

$$\phi(x; A, \Sigma) = [\Phi_{d,p}(A, \Sigma)]^{-1} \exp(\operatorname{tr} Ax'\Sigma x), \qquad x \in V_{d,p}, \tag{1.1}$$

where $\Phi_{d,p}(A, \Sigma)$ is the normalizing constant. The matrix Bingham family is now ubiquitous in the statistical analysis of directional data and shape analysis, and the extant literature now provides many properties of the family and its applications in fields including astronomy, biology,

---

computer vision, medicine, meteorology, physics and psychology; see, e.g., [3, 4, 7, 22, 23, 25, 24, 29, 33].

We also consider the matrix Langevin (or matrix von Mises–Fisher) family of probability distributions on $V_{d,p}$. For non-zero $B \in \mathbb{R}^{d \times p}$, a random matrix $X \in V_{d,p}$ is said to have a *matrix Langevin distribution with parameter $B$* if the probability density function of $X$, with respect to the invariant measure $\mathrm{d}x$, is

$$\psi(x; B) = [\Psi_{d,p}(B)]^{-1} \exp(\operatorname{tr} B'x), \qquad x \in V_{d,p}, \tag{1.2}$$

where $\Psi_{d,p}(B)$ is the normalizing constant. These distributions too are widely used in the statistical analysis of shape and directional data; see, e.g., Chikuse [4], Mardia and Jupp [29], Oualkacha and Rivest [31].

A major difficulty in statistical inference with the matrix Bingham and matrix Langevin distributions is the calculation of the constants $\Phi_{d,p}(A, \Sigma)$ and $\Psi_{d,p}(B)$. Some authors have described this issue as the "main obstacle" to maximum likelihood estimation because the normalizing constants, although known analytically, are difficult to compute in practice (Mantoux, et al. [28, pp. 4, 11]). For that reason, substantial attention has been paid to the problem of calculating those constants (see Kume and Wood [25], and Kume, Preston and Wood [24]).

In a recent article [1], we studied the Bingham distributions for the case in which $p = 1$. Under a growth condition on $\Sigma$ we obtained for $\Phi_{d,1}(1, \Sigma)$ a complete asymptotic series expansion in negative powers of $d$, and we derived a bound for the tail of that series upon truncation after, say, $m$ terms. In [1], we used an explicit formula for the zonal polynomials indexed by partitions with one part; however, as no explicit formula is known for the zonal polynomials in general, a more powerful approach is needed for general $p$, and the purpose of the present article is to provide such an approach.

In the present article, we obtain for all $d$ and $p$ complete asymptotic expansions of $\Phi_{d,p}(A, \Sigma)$ and $\Psi_{d,p}(B)$ in negative powers of $d$; these expansions are derived subject to growth conditions on the Frobenius norms of $\Sigma$ and $B$ as $d \to \infty$. We accomplish these results using detailed properties of the zonal polynomials, including an integral representation due to James [18].

We also derive an infinite series bound and a closed-form bound for the remainder term in each asymptotic expansion. These bounds, as they are valid for all $d \geq p$, extend the results in [1]. We show further that the new bounds decrease strictly to zero as $d$, or as $m$, the number of terms in the truncated expansion, increases; equivalently, the accuracy of each truncated asymptotic expansion is strictly increasing in $d$ or $m$. We devote close attention to the closed-form upper bound as it is likely to be more widely used than its infinite series counterpart; in particular, we derive conditions on the values of the parameters such that the closed-form bound is unimodal in $m$.

When the matrices $A$ and $\Sigma$ in (1.1) are positive definite, we obtain a lower bound on the truncated expansion for $\Phi_{d,p}(A, \Sigma)$. Further, we derive a lower bound for the truncated expansion of $\Psi_{d,p}(B)$ when $B$ is of full rank. These lower bounds can be expressed in closed-form, and they decrease to zero as $d \to \infty$ or $m \to \infty$.

The contents of the article are as follows. In Section 2, we obtain the asymptotic expansion for $\Phi_{d,p}(A, \Sigma)$ and we remark on the properties of the expansion, and in Section 3, we obtain a similar expansion for $\Psi_{d,p}(B)$. In Section 4, we survey the range of values of $d$ and $p$ that have appeared in the literature on analyses of data on Stiefel manifolds, describe the rate of convergence to zero of each bound as $d$ or $m$ increases, and conclude by noting that our results extend to more general classes of distributions on the Stiefel and Grassmann manifolds and to the complex analogs of the matrix Bingham and matrix Langevin distributions. We provide in Section 5.1 some necessary properties of the zonal polynomials, and all proofs are given in Sections 5.2–5.4.

## 2  Complete asymptotics for the normalizing constant of the high-dimensional matrix Bingham distribution

A *partition* $\kappa = (\kappa_1, \ldots, \kappa_d)$ is a vector of integers such that $\kappa_1 \geq \cdots \geq \kappa_d \geq 0$. The non-zero $\kappa_j$ are called the *parts* of $\kappa$. The *length* of $\kappa$, denoted by $\ell(\kappa)$, is the number of non-zero $\kappa_j$, $j = 1, \ldots, d$, and the *weight* of $\kappa$ is $|\kappa| = \kappa_1 + \cdots + \kappa_d$.

Define the *shifted factorial*, $(a)_k = a(a+1)(a+2)\cdots(a+k-1)$, where $a \in \mathbb{R}$ and $k = 0, 1, 2, \ldots$; in particular, $(a)_0 \equiv 1$ since the defining product is empty. For $a \in \mathbb{R}$ and a partition $\kappa = (\kappa_1, \ldots, \kappa_d)$, the *partitional shifted factorial* is

$$(a)_\kappa = \prod_{i=1}^{d} \left( a - \frac{1}{2}(i-1) \right)_{\kappa_i}. \tag{2.1}$$

Assume that $\mathbb{R}_{\mathrm{sym}}^{d \times d}$ denote the space of real symmetric $d \times d$ matrices. For a partition $\kappa = (\kappa_1, \ldots, \kappa_d)$ and $\Sigma \in \mathbb{R}_{\mathrm{sym}}^{d \times d}$, the *zonal polynomial* $C_\kappa(\Sigma)$ is an eigenfunction of the Laplace–Beltrami operator in the eigenvalues of $\Sigma$ (see [17], [30, Chapter 7] and [35]). Further, $C_\kappa(\Sigma)$ depends only on the eigenvalues of $\Sigma$, and is symmetric in those eigenvalues.

Tables of the polynomials $C_\kappa(\Sigma)$ for $|\kappa| \leq 12$ were provided by James [16, pp. 498–499] and Parkhurst and James [32]. Jiu and Koutschan [20] have provided software to calculate rapidly the zonal polynomials to high values of $|\kappa|$, so the asymptotic expansions obtained in this article can be calculated for practical values of $m$.

Let $X \in V_{d,p}$ be a random matrix having the matrix Bingham distribution with the density function $\phi(x; A, \Sigma)$ as in (1.1). It is well known that the normalizing constant $\Phi_{d,p}(A, \Sigma)$ can be expressed as a generalized hypergeometric function of two matrix arguments, denoted by ${}_0F_0(A, \Sigma)$, or by the zonal polynomial expansion of that function,

$$\Phi_{d,p}(A, \Sigma) = {}_0F_0(A, \Sigma) = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{|\kappa|=k} \frac{C_\kappa(A) C_\kappa(\Sigma)}{C_\kappa(I_d)}, \tag{2.2}$$

where the inner sum is over all partitions $\kappa$ such that $\ell(\kappa) \leq p$. This formula for $\Phi_{d,p}(A, \Sigma)$ in terms of ${}_0F_0(A, \Sigma)$ dates to De Waal [5]; cf. Bingham [2], Chikuse [4, p. 108]. By applying the methods of [11, Theorem 6.3], we find that the series (2.2) converges absolutely for all $(A, \Sigma)$ and uniformly on compact regions.

For the special case in which $A = I_p$, on applying to (2.2) the identity (5.7), written in the form $C_\kappa(I_p)/C_\kappa(I_d) = (p/2)_\kappa/(d/2)_\kappa$, we obtain the normalizing constant in terms of a confluent hypergeometric function of matrix argument,

$$\Phi_{d,p}(I_p, \Sigma) = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{|\kappa|=k} \frac{(p/2)_\kappa}{(d/2)_\kappa} C_\kappa(\Sigma) = {}_1F_1\left( \frac{1}{2}p; \frac{1}{2}d; \Sigma \right),$$

as shown by various authors, e.g., Bingham [2], Chikuse [4, p. 33] and Muirhead [30, p. 288].

Similar to [1, equation (2.4)], when $d$ is large, our asymptotic approximation to $\Phi_{d,p}(A, \Sigma)$ is the sum of the terms up to degree $m-1$ of the series (2.2), i.e.,

$$\Phi_{d,p}(A, \Sigma) \approx \sum_{k=0}^{m-1} \frac{1}{k!} \sum_{|\kappa|=k} \frac{C_\kappa(A) C_\kappa(\Sigma)}{C_\kappa(I_d)}, \tag{2.3}$$

$m \geq 2$, and now we study the remainder series

$$\Phi_{d,p;m}(A, \Sigma) = \sum_{k=m}^{\infty} \frac{1}{k!} \sum_{|\kappa|=k} \frac{C_\kappa(A) C_\kappa(\Sigma)}{C_\kappa(I_d)}. \tag{2.4}$$

In the sequel, we will encounter the constants $\alpha_p = (2\pi)^{-(p-1)/4p} p^{1/4p}$ and $\gamma_1 = (\sqrt{3}+1)/2$, arising in Lemmas 5.1 and 5.2, respectively. Also, if $a_1, \ldots, a_p$ are the eigenvalues of $A$, then we define $A_+ = \mathrm{diag}(|a_1|, \ldots, |a_p|)$, the diagonal matrix with diagonal entries $|a_1|, \ldots, |a_p|$. In the following result, the proof of which is given in Section 5.3, we extend a result in [1] by deriving an upper bound for $\Phi_{d,p;m}(A, \Sigma)$ in terms of the Frobenius norm, $\|\Sigma\| = [\mathrm{tr}(\Sigma^2)]^{1/2}$.

**Theorem 2.1.** *Let $A \in \mathbb{R}^{p \times p}_{\mathrm{sym}}$, $\Sigma \in \mathbb{R}^{d \times d}_{\mathrm{sym}}$, and suppose that $\|\Sigma\| \leq \gamma_0 d^{r/2}$, where $\gamma_0 > 0$ and $0 \leq r < 1$. Then for all $m \geq 2$,*

$$|\Phi_{d,p;m}(A,\Sigma)| \leq \alpha_p \sum_{k=m}^{\infty} \frac{\left[\gamma_0 \gamma_1 p^{1/2}(\mathrm{tr}\, A_+) d^{-(1-r)/2}\right]^k}{(k!)^{1/2}} \tag{2.5}$$

$$\leq \alpha_p (4e/\pi)^{1/4} (e/m)^{(m/2)-(1/4)} \left[\gamma_0 \gamma_1 p^{1/2}(\mathrm{tr}\, A_+) d^{-(1-r)/2}\right]^m$$

$$\times \exp\left(c_m \left[\gamma_0 \gamma_1 p^{1/2}(\mathrm{tr}\, A_+) d^{-(1-r)/2}\right]^2/2\right), \tag{2.6}$$

*where $c_m = (1+m^{-1})^{-m} e$. In particular, $\Phi_{d,p;m}(A, \Sigma) = O(d^{-(1-r)m/2})$ as $d \to \infty$.*

**Remark 2.2.**

(i) The growth condition, $\|\Sigma\| \leq \gamma_0 d^{r/2}$ with $\gamma_0 > 0$ and $r \in [0, 1)$, is the same as was assumed in [1] for the Bingham distribution, i.e., the case $p = 1$. Further, the bounds (2.5) and (2.6) provide a rate of convergence, viz., $\Phi_{d,p;m}(A, \Sigma) = O(d^{-m(1-r)/2})$ as $d \to \infty$, that is the same as we obtained in [1] for the case in which $p = 1$. On the other hand, unlike the bounds in [1], which are valid only for sufficiently large $d$, the bounds (2.5) and (2.6) are valid for all $d \geq p \geq 1$ and even for all $r \in \mathbb{R}$.

(ii) The bounds (2.5) and (2.6) each are increasing in $\gamma_0$ and in $r$; this is consistent with the restriction $\|\Sigma\| \leq \gamma_0 d^{r/2}$ representing a larger matrix region as $\gamma_0$ or $r$ increases.

(iii) Each bound is also strictly increasing in $p$ or $\mathrm{tr}(A_+)$ for fixed $d$, $m$, $\gamma_0$, and $r$. This follows from the fact that, in the $k$th summand in the bound in (2.5), the term $\alpha_p p^{k/2}$ is strictly increasing in $p$; therefore, the bound itself is strictly increasing. A similar argument also yields the same conclusion for the bound in (2.6).

(iv) It is evident that the bound in (2.5), as a function of $d$ or $m$, decreases strictly to zero as $d \to \infty$ or $m \to \infty$. Equivalently, the accuracy of the truncated asymptotic expansion (2.3) is strictly increasing in $d$ or $m$. Further, it is evident that the closed-form bound in (2.6) is strictly decreasing in $d$.

(v) The bound (2.5) on $\Phi_{d,p;m}(A, \Sigma)$ also allows us to obtain asymptotics as both $d$ and $p$ tend to infinity. By (5.10), $\alpha_p \to (2\pi)^{-1/4}$ as $p \to \infty$, and therefore we deduce that $\Phi_{d,p;m}(A, \Sigma) = O(1)$ if $(\mathrm{tr}\, A_+)^2 p d^{-(1-r)} = O(1)$, i.e., if $(\mathrm{tr}\, A_+)^2 p = O(d^{1-r})$.

**Remark 2.3.**

(i) The behavior of the closed-form bound (2.6), as a function of $m$, is more intricate. We shall prove in Section 5.3 that the bound (2.6) is strictly decreasing in $m$ if $d$ is sufficiently large, precisely, if $d \geq [\gamma_0 \gamma_1 p^{1/2}(\mathrm{tr}\, A_+)]^{2/(1-r)}$.

(ii) For $d < [\gamma_0 \gamma_1 p^{1/2}(\mathrm{tr}\, A_+)]^{2/(1-r)}$, the bound (2.6), as a function of $m$, is unimodal. Therefore, for smaller values of $d$, the bound (2.6) may increase for small values of $m$, but it eventually attains a maximum value and then decreases strictly to zero thereafter. This shows again that if $d$ is small, then higher values of $m$ may need to be chosen in order for the approximation (2.3) to result in small errors, and the bounds (2.5) and (2.6) enable us to ascertain suitable values of $m$ in order for the approximation (2.3) to achieve any required level of accuracy.

To close this section, we obtain a lower bound for the remainder term $\Phi_{d,p;m}(A, \Sigma)$ for the case in which $A$ and $\Sigma$ are arbitrary positive definite matrices.

**Theorem 2.4.** *Let $A \in \mathbb{R}^{p \times p}_{\mathrm{sym}}$ and $\Sigma \in \mathbb{R}^{d \times d}_{\mathrm{sym}}$ be positive definite, let $\sigma_{(1)} \geq \cdots \geq \sigma_{(d)}$ denote the ordered eigenvalues of $\Sigma$, and set $N_d = \frac{1}{2}(d+2)(d-1)$. Then for all $m \geq 1$,*

$$\Phi_{d,p;m}(A, \Sigma) \geq (2+m)^{-mN_d}(1+m)^{-(1+m)N_d} \sum_{k=m}^{\infty} \left( \frac{1+m}{2+m} \right)^{kN_d} \frac{(\sigma_{(p)} \operatorname{tr} A)^k}{k!}. \tag{2.7}$$

**Remark 2.5.** Let $W$ be a discrete random variable having a Poisson distribution with parameter $\mu = \left( \frac{1+m}{2+m} \right)^{N_d} \sigma_{(p)} \operatorname{tr} A$. Then (2.7) can be written as

$$\Phi_{d,p;m}(A, \Sigma) \geq (2+m)^{mN_d}(1+m)^{-(1+m)N_d} \mathrm{e}^{\mu} \mathbb{P}(W \geq m), \tag{2.8}$$

and this result reveals several properties of the lower bound. First, by using the formula $\mathbb{P}(W \geq m) = 1 - \mathbb{P}(W < m)$, it is seen that the right-hand side of (2.8) can be expressed in closed-form. Second, as $m \to \infty$, it is simple to show that the right-hand side of (2.8) is asymptotically equal to $\mathrm{e}^{N_d + \mu} m^{-N_d} \mathbb{P}(W \geq m)$, where $W$ is Poisson-distributed with parameter $\mu \simeq \sigma_{(p)} \operatorname{tr} A$, and this asymptotic lower bound converges to 0 as $m \to \infty$.

Third, suppose that $\mu$ is large. By the central limit theorem, $(W - \mu)/\sqrt{\mu}$ is approximately distributed as $Z$, a random variable having the standard normal distribution (with mean 0 and standard deviation 1). Applying the usual correction for continuity that is used whenever, the distribution of a discrete random variable is approximated by the distribution of a continuous random variable, we obtain

$$\mathbb{P}(W \geq m) = \mathbb{P}\left( W \geq m - \frac{1}{2} \right) = \mathbb{P}\left( \frac{W - \mu}{\sqrt{\mu}} \geq \frac{m - \frac{1}{2} - \mu}{\sqrt{\mu}} \right) \simeq \mathbb{P}\left( Z \geq \frac{m - \frac{1}{2} - \mu}{\sqrt{\mu}} \right),$$

which can be evaluated numerically using the standard normal distribution function.

Fourth, it is evident that the lower bound in (2.7) decreases strictly to 0 as $d \to \infty$.

## 3 Complete asymptotics for the normalizing constant of the high-dimensional matrix Langevin distribution

Let $X \in V_{d,p}$ be a random matrix having the matrix Langevin distribution with the density function $\psi(x; A, \Sigma)$ given in (1.2). The normalizing constant $\Psi_{d,p}(B)$ can be expressed as a generalized hypergeometric function of matrix argument,

$$\Psi_{d,p}(B) = {}_0F_1\left( \frac{1}{2}; \frac{1}{4}B'B \right) = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{|\kappa|=k} \frac{C_\kappa\left( \frac{1}{4}B'B \right)}{(d/2)_\kappa},$$

cf. Chikuse [4, p. 31]. By applying [11, Theorem 6.3], we find that this series converges absolutely for all $B \in \mathbb{R}^{d \times p}$ and uniformly on compact regions.

Similar to the approach in Section 2, we use as an asymptotic approximation to $\Psi_{d,p}(B)$ the truncated series,

$$\Psi_{d,p}(B) \approx \sum_{k=0}^{m-1} \frac{1}{k!} \sum_{|\kappa|=k} \frac{C_\kappa\left( \frac{1}{4}B'B \right)}{(d/2)_\kappa},$$

$m \geq 2$, and we derive for the remainder series,

$$\Psi_{d,p;m}(B) = \sum_{k=m}^{\infty} \frac{1}{k!} \sum_{|\kappa|=k} \frac{C_\kappa\left( \frac{1}{4}B'B \right)}{(d/2)_\kappa}, \tag{3.1}$$

an upper bound that is a complete asymptotic expansion in negative powers of $d$. The proof of the following result is given in Section 5.4, and the constants $\alpha_p = (2\pi)^{-(p-1)/4p} p^{1/4p}$ and $\gamma_1 = (\sqrt{3}+1)/2$ again appear in the proof.

**Theorem 3.1.** *Let $B \in \mathbb{R}^{d \times p}$ be such that $\|B\| \leq 2\gamma_0^{1/2} d^{r/4}$, where $\gamma_0 > 0$ and $0 \leq r < 3$. Then for all $m \geq 2$, $\Psi_{d,p;m}(B) \geq 0$ and*

$$\Psi_{d,p;m}(B) \leq \alpha_p \sum_{k=m}^{\infty} \frac{\left[2\gamma_0\gamma_1 p^{5/2} d^{-(3-r)/2}\right]^k}{(k!)^{1/2}} \tag{3.2}$$

$$\leq \alpha_p (4\mathrm{e}/\pi)^{1/4} (\mathrm{e}/m)^{(m/2)-(1/4)} \left[2\gamma_0\gamma_1 p^{5/2} d^{-(3-r)/2}\right]^m$$

$$\times \exp\left(c_m \left[2\gamma_0\gamma_1 p^{5/2} d^{-(3-r)/2}\right]^2 / 2\right). \tag{3.3}$$

*In particular, $\Psi_{d,p;m}(B) = O\left(d^{-(3-r)m/2}\right)$ as $d \to \infty$.*

We note that each comment in Remark 2.2 has a direct analog for the matrix Langevin distribution. In particular, the bound in (3.2) is strictly decreasing in $d$ and $m$; strictly increasing in $\gamma_0$ and $r$; and strictly increasing in $p$ for fixed values of all other parameters. Also, the bound in (3.3) is strictly decreasing in $d$, and it is strictly decreasing in $m$ if $d$ is sufficiently large.

As regards asymptotics for $d, p \to \infty$, by proceeding as in Remark 2.2, we deduce that $\Psi_{d,p;m}(B) = O(1)$ if $p^{5/2} d^{-(3-r)/2} = O(1)$, i.e., if $p = O\left(d^{(3-r)/5}\right)$.

In closing this section, we provide a lower bound for $\Psi_{d,p;m}(B)$.

**Theorem 3.2.** *Let $B \in \mathbb{R}^{d \times p}$ such that $B$ is of rank $p$. Let $\beta_{(1)} \geq \cdots \geq \beta_{(p)}$ denote the ordered eigenvalues of $\frac{1}{4}B'B$, $N_p = \frac{1}{2}(p+2)(p-1)$, and $\mu = \left(\frac{1+m}{2+m}\right)^{N_p} 2p\beta_{(p)}\mathrm{e}^{-1}$. Then for all $m \geq 1$,*

$$\Psi_{d,p;m}(B) \geq \frac{(1+m)^{-N_p}(d+m)^{-m}(2p\beta_{(p)})^m}{m!} {}_1F_2(1; m+1, d+1+m; \mu). \tag{3.4}$$

In proving this result, we will also deduce the weaker bound,

$$\Psi_{d,p;m}(B) \geq \frac{(1+m)^{-N_p}(d+m)^{-m}(2p\beta_{(p)})^m}{m!},$$

a bound that is still noteworthy as it is a single-term lower bound, it is in closed-form, and it converges to 0 as $d \to \infty$ or $m \to \infty$.

## 4 Implications for applications

The bounds derived in Sections 2 and 3 for the remainder terms in the corresponding asymptotic expansions appear to be the first such explicit bounds available in the literature. Therefore, we remark on the possible ramifications of those bounds for a variety of applications derived from a survey of the literature on applications of the matrix Bingham and matrix Langevin distributions.

The literature on such applications contains a broad range of reported values of $d$ and $p$. The case in which $(d, p) = (3, 2)$ dates to analyses of data, drawn from astronomy and vectorcardiography, as provided in the classic article of Jupp and Mardia [21]. The same value, $(d, p) = (3, 2)$ still appears in the recent literature, e.g., in data representing the orientation of two-dimensional elliptical paths of objects in near-Earth orbit (Jauch, Hoff and Dunson [19]) and various graph data sets (Mantoux et al. [28]).

For larger values of $d$ and $p$, a variety of applications have appeared in the literature. Hoff [12] applied the matrix Bingham distributions with $(d, p) = (270, 2)$ to analyze data on protein-protein interaction networks.

Mantoux et al. [28] provided illustrative examples with low- and high-dimensional synthetic data for $(d, p) = (3, 2)$, $(20, 5)$, $(20, 10)$, and $(40, 20)$; and with $d = 21$ and $p \in \{2, 5, 10\}$, they also analyzed resting-state functional magnetic resonance imaging (rs-fMRI) data from the UK Biobank repository of large-scale brain-imaging data set on tens of thousands of subjects (Sudlow et al. [36]).

Holbrook, Vandenberg-Rodes and Shahbaba [13] considered Stiefel manifold data with values of $d$ ranging from 18 to 53. Their work included a study of data, collected during a memory experiment on rodents, in which neural spike activity was recorded in the hippocampi of six rats. Especially interesting is the analysis in [13] of the 53-dimensional data on Super Chris, a particular rat whose neural spike data exhibited statistical features substantially distinct from the neural spike data of his fellow rats; cf. Granados-Garcia et al. [10].

Jauch et al. [19] analyzed with $(d, p) = (365, 3)$ a meteorological data that included average daily temperatures for a collection of Canadian weather stations; for the case in which $p \geq 2$, this value of $d$ was the largest that we encountered in the literature.

In summary, for $p \geq 2$, we found applications in the literature with values of $(d, p)$ where $d$ ranges between 3 and 365, and $p$ ranges between 2 and 20. Further, for $p = 1$, values of $d$ as large as $62,501$ were noted in [1, Section 4].

The computation of the bounds in Theorems 2.1 and 3.1 is straightforward as it reduces to calculating the function $R_m(t)$, or the upper bound on $R_m(t)$, in (5.36) and (5.47), respectively. As we noted earlier, these bounds are (strictly) decreasing in $d$ (and in $m$ if $t \leq 1$), increasing in $\gamma_0$ and $r$, and increasing in $p$. We reiterate that, in such calculations, the relationship between $t$ and $d$ is that

$$t = \begin{cases} \gamma_0\gamma_1 p^{1/2}(\operatorname{tr} A_+)d^{-(1-r)/2} & \text{(in Theorem 2.1)}, \\ 2\gamma_0\gamma_1 p^{5/2}d^{-(3-r)/2} & \text{(in Theorem 3.1)}, \end{cases}$$

equivalently,

$$d = \begin{cases} \left(\gamma_0\gamma_1 p^{1/2}(\operatorname{tr} A_+)/t\right)^{2/(1-r)} & \text{(in Theorem 2.1)}, \\ \left(2\gamma_0\gamma_1 p^{5/2}/t\right)^{2/(3-r)} & \text{(in Theorem 3.1)}. \end{cases}$$

For small values of $d$, e.g., $d = 3$, the resulting values of $t$ usually are large and the bounds also can be large, indicating that larger choices for $m$ may be needed. For large values of $d$, e.g., $d = 365$, the resulting values of $t$ typically are small, and the resulting bounds in terms of $R_m(t)$ quickly become negligible even for small values of $m$.

In summary, we expect that the results obtained here will find broad usage as they enable for the first time theoretical analyses and numerical calculations of bounds on the accuracy of the resulting expansions. In particular, these results enable calculation of the given expansions to any desired practical degree of accuracy.

## 5 Proofs

### 5.1 Properties of the zonal polynomials

James [18] obtained for the zonal polynomials $C_\kappa(\Sigma)$ a crucial integral formula, as we now explain. Let $\operatorname{Or}(d) = \{H \in \mathbb{R}^{d \times d} \mid H'H = I_d\}$ be the group of orthogonal matrices in $\mathbb{R}^{d \times d}$. For $H \in \operatorname{Or}(d)$, denote by $dH$ the Haar measure on $\operatorname{Or}(d)$, normalized to have total volume 1. Also let $\det_j(\Sigma)$ be the $j$th principal minor of $\Sigma$, $j = 1, \ldots, d$. For each partition $\kappa = (\kappa_1, \ldots, \kappa_d)$,

$$C_\kappa(\Sigma) = C_\kappa(I_d) \int_{\operatorname{Or}(d)} \prod_{j=1}^{\ell(\kappa)} \left[\det_j\left(H'\Sigma H\right)\right]^{\kappa_j - \kappa_{j+1}} dH, \tag{5.1}$$

where $\kappa_{d+1} \equiv 0$ and

$$C_\kappa(I_d) = 2^{2|\kappa|}|\kappa|!(d/2)_\kappa \frac{\displaystyle\prod_{1 \le i < j \le \ell(\kappa)} (2\kappa_i - 2\kappa_j - i + j)}{\displaystyle\prod_{i=1}^{\ell(\kappa)} (2\kappa_i + \ell(\kappa) - i)!}. \tag{5.2}$$

The formula (5.1) was first given by James [18, equation (10.4)], further applications of it were given in [26] and [34], and the formula was extended to a more general setting in [11, Theorem 4.8].

By (5.1), $C_\kappa(\Sigma)$ is homogeneous of degree $|\kappa|$; moreover $C_\kappa(\Sigma)$ depends only on the $\ell(\kappa)$ non-zero parts of $\kappa$, so that $C_\kappa(\Sigma)$ is unchanged if zeros are appended to $\kappa$. Since any symmetric matrix can be diagonalized by an orthogonal transformation, then it also follows from (5.1) that $C_\kappa(\Sigma)$ is a symmetric function of the eigenvalues of $\Sigma$.

The importance of the normalization (5.2) is that the zonal polynomials satisfy the identity,

$$(\operatorname{tr} \Sigma)^k = \sum_{|\kappa|=k} C_\kappa(\Sigma), \tag{5.3}$$

for all $k = 0, 1, 2, \ldots$; see [16, p. 479, equation (21)], [30, p. 228, equation (iii)], [35, equation (35.4.6)]. Also noteworthy is the special case of (5.2) in which $\kappa = (k)$, a partition with one part

$$C_{(k)}(I_d) = \frac{(d/2)_k}{(1/2)_k}. \tag{5.4}$$

The (seemingly formidable) expression for $C_\kappa(I_d)$ in (5.2) arises from the representation theory of $\mathfrak{S}_k$, the group of permutations on $k$ symbols, and we comment on the connection as follows. It is a classical result that the collection of irreducible representations of $\mathfrak{S}_k$ is in one-to-one correspondence with the set of partitions of $k$ (see Ledermann [27, Chapter 4]). For a given partition $\kappa = (\kappa_1, \ldots, \kappa_d)$ we use the standard notation $\chi_{[2\kappa]}(1)$ for the dimension of the irreducible representation, corresponding to the partition $2\kappa = (2\kappa_1, \ldots, 2\kappa_d)$, of the symmetric group $\mathfrak{S}_{2|\kappa|}$ on $2|\kappa|$ symbols. James [16, pp. 478–479, equations (18) and (20)] proved that

$$C_\kappa(I_d) = \frac{2^{2|\kappa|}|\kappa|!(d/2)_\kappa}{(2|\kappa|)!} \chi_{[2\kappa]}(1) = \frac{(d/2)_\kappa}{(1/2)_{|\kappa|}} \chi_{[2\kappa]}(1), \tag{5.5}$$

and then (5.2) is derived from (5.5) by applying Frobenius' famous formula for $\chi_{[2\kappa]}(1)$ (Ledermann [27, p. 123]).

Another consequence of (5.5) is that

$$\frac{C_\kappa(I_d)}{(d/2)_\kappa} = \frac{\chi_{[2\kappa]}(1)}{(1/2)_{|\kappa|}}. \tag{5.6}$$

Since the dimension of an irreducible representation does not change if zeros are appended to the corresponding partition, then it follows that, for any partition $\kappa$ such that $\ell(\kappa) \le p$, the right-hand side of (5.6) remains unchanged if $d$ is replaced by $p$; hence

$$\frac{C_\kappa(I_d)}{(d/2)_\kappa} = \frac{C_\kappa(I_p)}{(p/2)_\kappa}. \tag{5.7}$$

We remark that the identity (5.7) can be verified directly, albeit laboriously, from (5.2). Further, (5.7) has been noted explicitly in the literature, e.g., by Chikuse [4, p. 30] and by Edelman et al. [8, p. 259]. Also, if $\ell(\kappa) = 1$, then (5.7) reduces to (5.4).

The zonal polynomials also have a combinatorial formulation. Recall that the set of all partitions is endowed with the *lexicographic ordering*: For partitions $\kappa = (\kappa_1, \ldots, \kappa_d)$ and $\lambda = (\lambda_1, \ldots, \lambda_d)$ such that $|\kappa| = |\lambda|$, we say that $\lambda$ *is less than* $\kappa$, written $\lambda < \kappa$, if there exists $i \in \{1, \ldots, d-1\}$ such that $(\lambda_1, \ldots, \lambda_{i-1}) = (\kappa_1, \ldots, \kappa_{i-1})$ and $\lambda_i < \kappa_i$.

For $\Sigma \in \mathbb{R}^{d \times d}_{\text{sym}}$, denote by $\sigma_1, \ldots, \sigma_d$ the eigenvalues of $\Sigma$. For each partition $\kappa = (\kappa_1, \ldots, \kappa_\ell)$ of length $\ell$, the *monomial symmetric function of $\Sigma$ corresponding to $\kappa$* is

$$M_\kappa(\Sigma) = \sum_{(j_1, \ldots, j_\ell)} \sigma_{j_1}^{\kappa_1} \cdots \sigma_{j_\ell}^{\kappa_\ell}, \tag{5.8}$$

where the sum is taken over all distinct permutations $(j_1, \ldots, j_\ell)$ of $\ell$ integers drawn from the set $\{1, \ldots, d\}$. James [17] (see also Muirhead [30, Section 7.2]) proved that

$$C_\kappa(\Sigma) = \sum_{\lambda \leq \kappa} c_{\kappa, \lambda} M_\lambda(\Sigma), \tag{5.9}$$

where the constants $c_{\kappa, \lambda}$ are *nonnegative* and the summation is over all partitions $\lambda$ such that $\lambda \leq \kappa$, i.e., $\lambda$ is less than or equal to $\kappa$ in the lexicographic ordering.

## 5.2 Inequalities for the partitional shifted factorials and the zonal polynomials

**Lemma 5.1.** *Let $a_1, \ldots, a_p$ be nonnegative integers such that $a_1 + \cdots + a_p \geq 2$, and define*

$$\alpha_p = (2\pi)^{-(p-1)/4p} p^{1/4p}. \tag{5.10}$$

*Then*

$$\prod_{i=1}^{p} (pa_i)! \leq \alpha_p^{2p} p^{p(a_1 + \cdots + a_p)} \left[ (a_1 + \cdots + a_p)! \right]^p. \tag{5.11}$$

**Proof.** Since the multinomial coefficient, $(a_1 + \cdots + a_p)/a_1! \cdots a_p!$, is a positive integer, then

$$\prod_{i=1}^{p} a_i! \leq (a_1 + \cdots + a_p)!. \tag{5.12}$$

It is well known that the classical gamma function is increasing on the interval $[2, \infty)$; therefore, $\Gamma\left(a + (i-1)p^{-1}\right) \leq \Gamma\left(a + (p-1)p^{-1}\right) = \Gamma\left(a + 1 - p^{-1}\right)$ for $a \geq 2$ and all $i = 1, \ldots, p$. Applying Gauss' multiplication formula for the gamma function, we obtain

$$\Gamma(pa) = (2\pi)^{-(p-1)/2} p^{pa-(1/2)} \prod_{i=1}^{p} \Gamma\left(a + (i-1)p^{-1}\right)$$

$$\leq (2\pi)^{-(p-1)/2} p^{pa-(1/2)} \left[ \Gamma\left(a + 1 - p^{-1}\right) \right]^p, \tag{5.13}$$

where the latter inequality holds for $a \geq 2$.

Setting $a = a_1 + \cdots + a_p + p^{-1}$ in (5.13), we obtain

$$(p(a_1 + \cdots + a_p))! = \Gamma(p(a_1 + \cdots + a_p) + 1)$$

$$\leq (2\pi)^{-(p-1)/2} p^{p(a_1 + \cdots + a_p) + (1/2)} [\Gamma(a_1 + \cdots + a_p + 1)]^p$$

$$= (2\pi)^{-(p-1)/2} p^{p(a_1 + \cdots + a_p) + (1/2)} [(a_1 + \cdots + a_p)!]^p. \tag{5.14}$$

Applying (5.14) to (5.12), we find that

$$\prod_{i=1}^{p}(pa_i)! \le \left(p(a_1 + \cdots + a_p)\right)! \le (2\pi)^{-(p-1)/2}p^{p(a_1+\cdots+a_p)+(1/2)}\left((a_1 + \cdots + a_p)!\right)^p$$

$$\equiv \alpha_p^{2p}p^{p(a_1+\cdots+a_p)}[(a_1 + \cdots + a_p)!]^p.$$

This completes the proof.                                                                                               ■

**Lemma 5.2.** *Let* $d \ge p$ *and let* $\kappa = (\kappa_1, \ldots, \kappa_p)$ *be a partition of length* $\ell(\kappa) \le p$ *and weight* $|\kappa| \ge 2$. *Then*

$$(2p)^{-|\kappa|}d^{|\kappa|} \le (d/2)_\kappa \le 2^{-|\kappa|}(d + |\kappa|)^{|\kappa|}. \tag{5.15}$$

*Also, with* $\gamma_1 = \left(\sqrt{3} + 1\right)/2 \simeq 1.366025$, *we have*

$$\prod_{i=1}^{p} \frac{\left(d^{1/2}/2\right)_{p\kappa_i}}{(d/2)_{p\kappa_i}} \le \left[\alpha_p(|\kappa|!)^{1/2}\gamma_1^{|\kappa|}p^{|\kappa|/2}d^{-|\kappa|/2}\right]^p. \tag{5.16}$$

**Proof.** By (2.1), we have

$$(d/2)_\kappa = \prod_{j=1}^{p}\left(\frac{1}{2}d - \frac{1}{2}(j-1)\right)_{\kappa_j} = \prod_{j=1}^{p}\prod_{i=1}^{\kappa_j}\left(\frac{1}{2}d - \frac{1}{2}(j-1) + i - 1\right). \tag{5.17}$$

For all $i \ge 1$, $j \le p$, and $d \ge p$, we also have

$$\frac{1}{2}d - \frac{1}{2}(j-1) + i - 1 \ge \frac{1}{2}d - \frac{1}{2}(p-1) = \frac{1}{2}\left(1 - \frac{p-1}{d}\right)d \ge \frac{1}{2}\left(1 - \frac{p-1}{p}\right)d = (2p)^{-1}d.$$

Therefore, by (5.17),

$$(d/2)_\kappa \ge \prod_{j=1}^{p}\prod_{i=1}^{\kappa_j}[(2p)^{-1}d] = (2p)^{|\kappa|}d^{-|\kappa|}.$$

This establishes the lower bound in (5.15).

Next, we apply to the product in (5.17) the arithmetic-geometric mean inequality. Then we obtain

$$(d/2)_\kappa \le \left(\frac{1}{|\kappa|}\sum_{j=1}^{p}\sum_{i=1}^{\kappa_j}\left(\frac{1}{2}d - \frac{1}{2}(j-1) + i - 1\right)\right)^{|\kappa|}$$

$$= \left(\frac{1}{2|\kappa|}\left(d|\kappa| - \sum_{j=1}^{p}j\kappa_j + \sum_{j=1}^{p}\kappa_j^2\right)\right)^{|\kappa|} \le \left(\frac{1}{2|\kappa|}\left(d|\kappa| + \left(\sum_{j=1}^{p}\kappa_j\right)^2\right)\right)^{|\kappa|}$$

$$= 2^{-|\kappa|}(d + |\kappa|)^{|\kappa|}.$$

This proves the upper bound in (5.15).

Further, it is known from [1, Lemma A.2] that, for all $r \ge 1$, $\frac{\left(d^{1/2}/2\right)_r}{(d/2)_r} \le [(r-1)!]^{1/2}\gamma_1^{r-1}d^{-r/2}$, and since $\gamma_1 > 1$, then we also have

$$\frac{\left(d^{1/2}/2\right)_r}{(d/2)_r} \le [(r-1)!]^{1/2}\gamma_1^{r-1}d^{-r/2} \cdot \gamma_1 r^{1/2} = (r!)^{1/2}\gamma_1^r d^{-r/2}.$$

Since the latter inequality is valid for $r = 0$, then we have shown that, for all $r \geq 0$,

$$\frac{\left(d^{1/2}/2\right)_r}{(d/2)_r} \leq (r!)^{1/2}\gamma_1^r d^{-r/2}. \tag{5.18}$$

Now we apply (5.18) to obtain

$$\prod_{i=1}^p \frac{\left(d^{1/2}/2\right)_{p\kappa_i}}{(d/2)_{p\kappa_i}} \leq \prod_{i=1}^p [(p\kappa_i)!]^{1/2}\gamma_1^{p\kappa_i} d^{-p\kappa_i/2} = \gamma_1^{p|\kappa|} d^{-p|\kappa|/2} \prod_{i=1}^p [(p\kappa_i)!]^{1/2}. \tag{5.19}$$

By (5.11) in Lemma 5.1 with $a_i = \kappa_i$, $i = 1, \ldots, p$, we also have

$$\prod_{i=1}^p (p\kappa_i)! \leq \alpha_p^{2p} p^{p|\kappa|}(|\kappa|!)^p,$$

and by substituting this bound into (5.19), we obtain (5.16). ∎

For $\Sigma \in \mathbb{R}_{\text{sym}}^{d \times d}$, let $\sigma_1, \ldots, \sigma_d$ denote the eigenvalues of $\Sigma$ and define the diagonal matrix $\Sigma_+ = \text{diag}(|\sigma_1|, \ldots, |\sigma_d|)$. The following result is a consequence of (5.9), the combinatorial formula for the zonal polynomials.

**Lemma 5.3.** *Let $\kappa$ be a partition and $\Sigma \in \mathbb{R}_{\text{sym}}^{d \times d}$. Then $|C_\kappa(\Sigma)| \leq C_\kappa(\Sigma_+)$.*

**Proof.** By (5.9), the nonnegativity of the constants $c_{\kappa,\lambda}$, and the triangle inequality, we have

$$|C_\kappa(\Sigma)| \leq \left|\sum_{\lambda \leq \kappa} c_{\kappa,\lambda} M_\lambda(\Sigma)\right| \leq \sum_{\lambda \leq \kappa} c_{\kappa,\lambda}|M_\lambda(\Sigma)|.$$

By (5.8), for all partitions $\lambda$,

$$|M_\lambda(\Sigma)| = \left|\sum \sigma_{j_1}^{\lambda_1} \cdots \sigma_{j_\ell}^{\lambda_\ell}\right| \leq \sum |\sigma_{j_1}|^{\lambda_1} \cdots |\sigma_{j_\ell}|^{\lambda_\ell} = M_\lambda(\Sigma_+).$$

Therefore,

$$|C_\kappa(\Sigma)| \leq \sum_{\lambda \leq \kappa} c_{\kappa,\lambda} M_\lambda(\Sigma_+) = C_\kappa(\Sigma_+).$$

The proof now is complete. ∎

The following result will be crucial for deriving bounds on the tails of the asymptotic series for the normalizing constants of the matrix Bingham and matrix Langevin distributions.

**Proposition 5.4.** *Let $\kappa$ be a partition such that $\ell(\kappa) \leq p$, where $p \leq d$, and let $\Sigma \in \mathbb{R}_{\text{sym}}^{d \times d}$. Then*

$$|C_\kappa(\Sigma)| \leq \left(\prod_{i=1}^p \frac{\left(d^{1/2}/2\right)_{p\kappa_i}}{(d/2)_{p\kappa_i}}\right)^{1/p} C_\kappa(I_d)\|\Sigma\|^{|\kappa|}. \tag{5.20}$$

**Proof.** By Lemma 5.3, $|C_\kappa(\Sigma)| \leq C_\kappa(\Sigma_+)$. Since $\|\Sigma\| = \|\Sigma_+\|$, then it suffices to assume, without loss of generality, that $\Sigma = \Sigma_+$, i.e., that $\Sigma$ is positive semidefinite and diagonal, with diagonal entries $\sigma_1, \ldots, \sigma_d$.

Since $\Sigma$ is positive semidefinite, then so is every principal submatrix of $H'\Sigma H$, for all $H \in \text{Or}(d)$. Therefore, $\det_j(H'\Sigma H) \geq 0$ for all $j = 1, \ldots, d$.

For $\ell(\kappa) \leq p$, it follows from (5.1) that

$$C_\kappa(\Sigma) = C_\kappa(I_d) \int_{\mathrm{Or}(d)} \prod_{j=1}^{p} [\det_j(H'\Sigma H)]^{\kappa_j - \kappa_{j+1}} \, \mathrm{d}H. \tag{5.21}$$

Denote by $(H'\Sigma H)_{ii}$ the $i$th diagonal entry of $H'\Sigma H$, $i = 1, \ldots, d$. By Hadamard's inequality for the principal minors of a positive semidefinite matrix [14, p. 505],

$$\det_j(H'\Sigma H) \leq \prod_{i=1}^{j} (H'\Sigma H)_{ii}$$

for all $j = 1, \ldots, d$; therefore,

$$\prod_{j=1}^{p} [\det_j(H'\Sigma H)]^{\kappa_j - \kappa_{j+1}} \leq \prod_{j=1}^{p} \prod_{i=1}^{j} [(H'\Sigma H)_{ii}]^{\kappa_j - \kappa_{j+1}}$$

$$= \prod_{i=1}^{p} \prod_{j=i}^{p} [(H'\Sigma H)_{ii}]^{\kappa_j - \kappa_{j+1}} = \prod_{i=1}^{p} [(H'\Sigma H)_{ii}]^{\kappa_i}.$$

Inserting this bound into (5.21) and then applying Hölder's inequality, we obtain

$$C_\kappa(\Sigma) \leq C_\kappa(I_d) \int_{\mathrm{Or}(d)} \prod_{i=1}^{p} [(H'\Sigma H)_{ii}]^{\kappa_i} \, \mathrm{d}H$$

$$\leq C_\kappa(I_d) \left( \prod_{i=1}^{p} \int_{\mathrm{Or}(d)} [(H'\Sigma H)_{ii}]^{p\kappa_i} \, \mathrm{d}H \right)^{1/p}. \tag{5.22}$$

Denote by $h_{i,j}$ the $(i,j)$th entry of $H$, then $h_i = (h_{1,i}, \ldots, h_{d,i})'$ is the $i$th column of $H$, $i = 1, \ldots, d$. Since $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_d)$, then for all $i = 1, \ldots, d$,

$$(H'\Sigma H)_{ii} = \sum_{l=1}^{d} h_{l,i}^2 \sigma_l = h_i'\Sigma h_i.$$

Now regard $H$ as a random matrix having the uniform distribution (Haar measure) on $Or(d)$. Since the Haar measure is orthogonally invariant, and therefore invariant under permutation of the columns of $H$, it follows that the marginal distribution of $h_i$ does not depend on $i$. Denoting equality in distribution by $\overset{\mathcal{L}}{=}$, we obtain

$$(H'\Sigma H)_{ii} = h_i'\Sigma h_i \overset{\mathcal{L}}{=} h_1'\Sigma h_1 = (H'\Sigma H)_{11},$$

for all $i = 1, \ldots, d$. Consequently,

$$\int_{\mathrm{Or}(d)} [(H'\Sigma H)_{ii}]^{p\kappa_i} \, \mathrm{d}H = \int_{\mathrm{Or}(d)} [(H'\Sigma H)_{11}]^{p\kappa_i} \, \mathrm{d}H = \frac{C_{(p\kappa_i)}(\Sigma)}{C_{(p\kappa_i)}(I_d)}, \tag{5.23}$$

the second equality following from (5.1). Applying (5.23) to (5.22), we obtain

$$C_\kappa(\Sigma) \leq C_\kappa(I_d) \left( \prod_{i=1}^{p} \frac{C_{(p\kappa_i)}(\Sigma)}{C_{(p\kappa_i)}(I_d)} \right)^{1/p}. \tag{5.24}$$

By [1, equation (A.10)],

$$C_{(p\kappa_i)}(\Sigma) \leq \frac{(d^{1/2}/2)_{p\kappa_i}}{(1/2)_{p\kappa_i}} \|\Sigma\|^{p\kappa_i};$$

also, by (5.4),

$$C_{(p\kappa_i)}(I_d) = \frac{(d/2)_{p\kappa_i}}{(1/2)_{p\kappa_i}}.$$

Substituting the two latter results into (5.24), we obtain

$$C_\kappa(\Sigma) \leq C_\kappa(I_d) \left( \prod_{i=1}^p \frac{(d^{1/2}/2)_{p\kappa_i}}{(1/2)_{p\kappa_i}} \|\Sigma\|^{p\kappa_i} \cdot \frac{(1/2)_{p\kappa_i}}{(d/2)_{p\kappa_i}} \right)^{1/p}$$

$$= C_\kappa(I_d) \|\Sigma\|^{|\kappa|} \left( \prod_{i=1}^p \frac{(d^{1/2}/2)_{p\kappa_i}}{(d/2)_{p\kappa_i}} \right)^{1/p},$$

and this establishes (5.20). ∎

**Remark 5.5.** Suppose we evaluate both sides of the inequality (5.20) at $\Sigma = cI_d$, $c > 0$. Since $C_\kappa(cI_d) = c^{|\kappa|} C_\kappa(I_d)$ and $\|cI_d\|^{|\kappa|} = c^{|\kappa|} d^{|\kappa|/2}$, then the inequality reduces to

$$1 \leq \left( \prod_{i=1}^p \frac{(d^{1/2}/2)_{p\kappa_i}}{(d/2)_{p\kappa_i}} \right)^{1/p} d^{|\kappa|/2}. \tag{5.25}$$

For $d \to \infty$, we have

$$\frac{(d^{1/2}/2)_{p\kappa_i}}{(d/2)_{p\kappa_i}} \sim \frac{(d^{1/2}/2)^{p\kappa_i}}{(d/2)^{p\kappa_i}} = d^{-p\kappa_i/2},$$

therefore

$$\left( \prod_{i=1}^p \frac{(d^{1/2}/2)_{p\kappa_i}}{(d/2)_{p\kappa_i}} \right)^{1/p} d^{|\kappa|/2} \sim \left( \prod_{i=1}^p d^{-p\kappa_i/2} \right)^{1/p} d^{|\kappa|/2} = 1.$$

Hence for scalar matrices $\Sigma$, both sides of the inequality (5.20) are asymptotically equal as $d \to \infty$. This indicates that, in high dimensions, the inequality is close to asymptotically tight if $\Sigma$ is within a small neighborhood of any scalar matrix.

**Remark 5.6.** In the setting of general symmetric cones, another inequality for the zonal (spherical) polynomials was given by Faraut and Korányi [9, Theorem XII.1.1] (see also the proof of Lemma 6.5 in [11], where the same inequality can be derived for the matrix cones). In the notation used here, the inequality in [9, loc. cit.] states that if $\kappa = (\kappa_1, \ldots, \kappa_p)$ is a partition of length $p$, $\Sigma$ is a $d \times d$ positive definite matrix, and $\sigma_{(1)} \geq \cdots \geq \sigma_{(d)}$ are the ordered eigenvalues of $\Sigma$, then

$$C_\kappa(\Sigma) \leq C_\kappa(I_d) \prod_{j=1}^p \sigma_{(j)}^{\kappa_j}. \tag{5.26}$$

There is also the issue of which of the inequalities (5.20) and (5.26) is sharper. We now provide two examples to show that the answer depends on $d$ and on $\Sigma$.

For the case in which $\Sigma = cI_d$ where $c$ does not depend on $d$, we find that (5.26) is sharper than (5.20) since, by (5.25),

$$\prod_{i=1}^{p} \frac{\left(d^{1/2}/2\right)_{p\kappa_i}}{(d/2)_{p\kappa_i}} \geq d^{-p|\kappa|/2}.$$

More generally, it can be seen that if $\Sigma$ is "close" to a scalar matrix and $d$ is small, then we can expect (5.26) to be sharper than (5.20), however both bounds are asymptotic to $d^{-p|\kappa|/2}C_\kappa(I_d)$ for large $d$.

On the other hand, suppose that $\Sigma$ is "far" from the scalar matrices $cI_d$ in the sense that $c$ does not depend on $d$ and $\|I_d\|/\|\Sigma\| \to \infty$ as $d \to \infty$. As an example, for constants $\sigma, \rho > 0$ and $r \in [0, 1)$, let the eigenvalues of $\Sigma$ be

$$\sigma_j = \begin{cases} \sigma d^{r/2}, & j = 1, \ldots, p, \\ \rho d^{-(d-p+1)/2}, & j = p+1, \ldots, d. \end{cases}$$

Then $\Sigma$ is non-singular and

$$\|\Sigma\| = \left(\sum_{j=1}^{d} \sigma_j^2\right)^{1/2} = \left[p\sigma^2 d^r + \rho^2(d-p)d^{-(d-p+1)}\right]^{1/2} \sim p^{1/2}\sigma d^{r/2},$$

as $d \to \infty$, so $\Sigma$ satisfies the hypotheses of Theorem 2.1. Further, $\Sigma$ is far from the scalar matrices $cI_d$ since $\|I_d\|/\|\Sigma\| \sim p^{-1/2}\sigma^{-1}d^{(1-r)/2} \to \infty$ as $d \to \infty$.

For this choice of $\Sigma$ and for a partition $\kappa$ of length $p$, the ratio of the right-hand sides of (5.20) and (5.26) equals

$$\left(\prod_{i=1}^{p} \frac{\left(d^{1/2}/2\right)_{p\kappa_i}}{(d/2)_{p\kappa_i}}\right)^{1/p} \frac{\|\Sigma\|^{|\kappa|}}{\prod_{j=1}^{p} \sigma_{(j)}^{\kappa_j}} \sim \frac{\left(d^{1/2}/2\right)^{|\kappa|}}{(d/2)^{|\kappa|}} \frac{\left(p^{1/2}\sigma d^{r/2}\right)^{|\kappa|}}{\left(\sigma d^{r/2}\right)^{|\kappa|}} = d^{-|\kappa|/2}p^{|\kappa|/2} \to 0$$

as $d \to \infty$. In such instances, the bound in (5.20) is sharper than the bound in (5.26) for sufficiently large $d$, and it becomes increasingly sharper afterwards as $d$ increases.

Next, we obtain a lower bound for the zonal polynomials. This bound tightens an inequality stated in [9, Theorem XII.1.1 (ii)]. (Readers who wish to match the notation used in [9] with the notation in the present article will find it helpful to review [6, Section 1.12], where the notation used for general irreducible symmetric cones is made explicit in the case of each of the five classes of cones.)

**Proposition 5.7.** *Let* $\kappa = (\kappa_1, \ldots, \kappa_p)$ *be a partition of length* $p$, $\Sigma$ *be a* $d \times d$ *positive definite matrix with ordered eigenvalues* $\sigma_{(1)} \geq \cdots \geq \sigma_{(d)}$, *and* $N_d = \frac{1}{2}(d+2)(d-1)$. *Then*

$$C_\kappa(\Sigma) \geq (1 + |\kappa|)^{-N_d} C_\kappa(I_d)\sigma_{(1)}^{\kappa_1} \cdots \sigma_{(p)}^{\kappa_p}. \tag{5.27}$$

**Proof.** By [9, p. 241, line 4],

$$\frac{C_\kappa(\Sigma)}{C_\kappa(I_d)} \geq \frac{T_1}{T_2}\sigma_{(1)}^{\kappa_1} \cdots \sigma_{(p)}^{\kappa_p}, \tag{5.28}$$

where $T_1 = \frac{1}{d_\kappa}$, with $d_\kappa$ being the dimension of a certain vector space of polynomials that are homogeneous of degree $k = |\kappa|$, and

$$T_2 = \prod_{1 \leq i < j \leq p} \frac{B\left(\kappa_i - \kappa_j + \frac{1}{2}(j-i-1) + 1, \frac{1}{2}\right)}{B\left(\frac{1}{2}(j-i-1) + 1, \frac{1}{2}\right)}. \tag{5.29}$$

In [9, p. 242, line 9], it is shown that

$$d_\kappa \le \binom{N_d + k}{N_d} = \frac{(N_d + k)!}{N_d! k!} = \prod_{j=1}^{N_d} \frac{j+k}{j} \le (1+k)^{N_d}$$

since $(j+k)/j \le 1 + k$, for all $j = 1, \ldots, N_d$. Therefore,

$$T_1 = \frac{1}{d_\kappa} \ge (1 + |\kappa|)^{-N_d}, \tag{5.30}$$

As regards an upper bound for $T_2$ we now prove, using an elementary probabilistic approach, that

$$T_2 \le 1. \tag{5.31}$$

For $b, c > 0$, let $V$ denote a random variable having a beta distribution with probability density function

$$\frac{v^{b-1}(1-v)^{c-1}}{B(b,c)}, \qquad 0 < v < 1.$$

Then for $\alpha \ge 0$,

$$\frac{B(\alpha + b, c)}{B(b,c)} = \frac{1}{B(b,c)} \int_0^1 v^{\alpha+b-1}(1-v)^{c-1} \, dv = \mathbb{E}(V^\alpha),$$

and it is trivial that $\mathbb{E}(V^\alpha) \le 1$ since $\alpha \ge 0$ and $0 < V < 1$, almost surely. Setting $\alpha = \kappa_i - \kappa_j$, $b = \frac{1}{2}(j - i - 1) + 1$, and $c = \frac{1}{2}$, we obtain

$$\frac{B\left(\kappa_i - \kappa_j + \frac{1}{2}(j-i-1) + 1, \frac{1}{2}\right)}{B\left(\frac{1}{2}(j-i-1) + 1, \frac{1}{2}\right)} \le 1 \tag{5.32}$$

for all $1 \le i < j \le p$. Applying (5.32) to each term in (5.29), we obtain (5.31), and by applying (5.30) and (5.31) to (5.28), then we obtain (5.27). ∎

### 5.3 The proofs for Section 2

**Proof of Theorem 2.1.** We begin by applying to (2.4) the triangle inequality, the upper bound from Proposition 5.4, and the inequality $|C_\kappa(A)| \le C_\kappa(A_+)$. Then we obtain

$$|\Phi_{d,p;m}(A, \Sigma)| \le \sum_{k=m}^\infty \frac{1}{k!} \sum_{|\kappa|=k} \frac{|C_\kappa(A)|}{C_\kappa(I_d)} \left( \prod_{i=1}^p \frac{(d^{1/2}/2)_{p\kappa_i}}{(d/2)_{p\kappa_i}} \right)^{1/p} C_\kappa(I_d) \|\Sigma\|^{|\kappa|}$$

$$\le \sum_{k=m}^\infty \frac{\|\Sigma\|^k}{k!} \sum_{|\kappa|=k} C_\kappa(A_+) \left( \prod_{i=1}^p \frac{(d^{1/2}/2)_{p\kappa_i}}{(d/2)_{p\kappa_i}} \right)^{1/p}. \tag{5.33}$$

By (5.16),

$$\left( \prod_{i=1}^p \frac{(d^{1/2}/2)_{p\kappa_i}}{(d/2)_{p\kappa_i}} \right)^{1/p} \le \alpha_p (|\kappa|!)^{1/2} \gamma_1^{|\kappa|} p^{|\kappa|/2} d^{-|\kappa|/2},$$

and by inserting this bound into (5.33), we obtain

$$\Phi_{d,p;m}(A,\Sigma)| \le \sum_{k=m}^{\infty} \frac{\|\Sigma\|^k}{k!} \sum_{|\kappa|=k} C_\kappa(A_+)\alpha_p(|\kappa|!)^{1/2}\gamma_1^{|\kappa|}p^{|\kappa|/2}d^{-|\kappa|/2}$$

$$= \alpha_p \sum_{k=m}^{\infty} \frac{\gamma_1^k p^{k/2}d^{-k/2}\|\Sigma\|^k}{(k!)^{1/2}} \sum_{|\kappa|=k} C_\kappa(A_+). \tag{5.34}$$

By (5.3), $\sum_{|\kappa|=k} C_\kappa(A_+) = (\operatorname{tr} A_+)^k$, and, by assumption, $\|\Sigma\| \le \gamma_0 d^{r/2}$. Therefore, we obtain from (5.34) the inequality

$$\Phi_{d,p;m}(A,\Sigma)| \le \alpha_p R_m\big(\gamma_0\gamma_1 p^{1/2}(\operatorname{tr} A_+)d^{-(1-r)/2}\big), \tag{5.35}$$

where

$$R_m(t) = \sum_{k=m}^{\infty} \frac{t^k}{(k!)^{1/2}}, \qquad t > 0. \tag{5.36}$$

By applying the ratio test, we find that the series (5.36) converges for all $t$, and therefore (5.35) converges for all $d$. This establishes the bound (2.5).

To obtain a closed-form upper bound for $R_m(t)$ we begin by applying Stirling's well-known inequality for the factorial function [15, Theorem 1]; viz., for all $k \ge 1$,

$$k! \ge (2\pi)^{1/2}k^{(2k+1)/2}e^{-k}.$$

Inverting Stirling's inequality and applying the result to (5.36), we obtain

$$R_m(t) \le (2\pi)^{-1/4} \sum_{k=m}^{\infty} k^{-(2k+1)/4}e^{k/2}t^k. \tag{5.37}$$

We now apply the telescoping method to obtain an upper bound for the right-hand side of (5.37). For $j \ge m$, let

$$a_j = j^{-(2j+1)/4}e^{j/2}t^j, \tag{5.38}$$

then we obtain through straightforward algebraic manipulations

$$\frac{a_{j+1}}{a_j} = \big(1+j^{-1}\big)^{-j/2}\frac{j^{1/4}}{(j+1)^{3/4}}e^{1/2}t. \tag{5.39}$$

Since the function $j \mapsto \big(1+j^{-1}\big)^{-j/2}$, $j \ge m$, is decreasing, then

$$\big(1+j^{-1}\big)^{-j/2} \le \big(1+m^{-1}\big)^{-m/2}, \tag{5.40}$$

for all $j \ge m$, and by applying the latter inequality to (5.39), we obtain

$$\frac{a_{j+1}}{a_j} \le \big(1+m^{-1}\big)^{-m/2}e^{1/2}t\frac{j^{1/4}}{(j+1)^{3/4}} = c_m^{1/2}t\frac{j^{1/4}}{(j+1)^{3/4}},$$

where $c_m = \big(1+m^{-1}\big)^{-m}e$. It follows that, for all $k \ge m$,

$$a_k = a_m \prod_{j=m}^{k-1} \frac{a_{j+1}}{a_j} \le a_m \prod_{j=m}^{k-1} c_m^{1/2}t\frac{j^{1/4}}{(j+1)^{3/4}} = a_m c_m^{(k-m)/2}t^{k-m} \prod_{j=m}^{k-1} \frac{j^{1/4}}{(j+1)^{3/4}}. \tag{5.41}$$

Since $\prod_{j=m}^{k-1} j = (k-1)!/(m-1)!$, then

$$\prod_{j=m}^{k-1} \frac{j^{1/4}}{(j+1)^{3/4}} = \frac{[(k-1)!/(m-1)!]^{1/4}}{[k!/m!]^{3/4}} = \frac{m^{3/4}[(m-1)!]^{1/2}}{k^{3/4}[(k-1)!]^{1/2}}. \tag{5.42}$$

On applying to (5.41) the identity (5.42) and the explicit expression for $a_m$, as given in (5.38), we obtain

$$a_k \leq m^{-(2m+1)/4}\mathrm{e}^{m/2}t^m \cdot c_m^{(k-m)/2}t^{k-m}\frac{m^{3/4}[(m-1)!]^{1/2}}{k^{3/4}[(k-1)!]^{1/2}}$$

$$= (\mathrm{e}/mc_m)^{m/2}(m!)^{1/2}\frac{c_m^{k/2}t^k}{k^{3/4}[(k-1)!]^{1/2}}. \tag{5.43}$$

Therefore, by (5.37), (5.38), and (5.43),

$$R_m(t) \leq (2\pi)^{-1/4}(\mathrm{e}/mc_m)^{m/2}(m!)^{1/2}\sum_{k=m}^{\infty}\frac{c_m^{k/2}t^k}{k^{3/4}[(k-1)!]^{1/2}}$$

$$\leq (2\pi)^{-1/4}(\mathrm{e}/mc_m)^{m/2}(m!)^{1/2}\left(\sum_{k=m}^{\infty}\frac{1}{k^{3/2}}\right)^{1/2}\left(\sum_{k=m}^{\infty}\frac{c_m^k t^{2k}}{(k-1)!}\right)^{1/2}, \tag{5.44}$$

where the latter inequality follows by the Cauchy–Schwarz inequality.

To obtain an upper bound for the first sum in (5.44), we observe from graphical considerations using Riemann sums that, for $m \geq 2$,

$$\sum_{k=m}^{\infty}\frac{1}{k^{3/2}} \leq \int_m^{\infty}\frac{\mathrm{d}t}{(t-1)^{3/2}} = \frac{2}{(m-1)^{1/2}} \leq \left(\frac{8}{m}\right)^{1/2}. \tag{5.45}$$

As for the second sum in (5.44), we apply the inequality $\frac{1}{(k+m-1)!} \leq \frac{1}{k!(m-1)!}$ to obtain

$$\sum_{k=m}^{\infty}\frac{t^{2k}}{(k-1)!} = \sum_{k=0}^{\infty}\frac{t^{2(k+m)}}{(k+m-1)!} \leq \frac{t^{2m}}{(m-1)!}\sum_{k=0}^{\infty}\frac{t^{2k}}{k!} = \frac{t^{2m}}{(m-1)!}\exp(t^2). \tag{5.46}$$

Applying (5.45) and (5.46) to (5.44), we obtain

$$R_m(t) \leq (2\pi)^{-1/4}(\mathrm{e}/mc_m)^{m/2}(m!)^{1/2}\left(\frac{8}{m}\right)^{1/4}\left(\frac{c_m^m t^{2m}}{(m-1)!}\exp(c_m t^2)\right)^{1/2}$$

$$= (4\mathrm{e}/\pi)^{1/4}(\mathrm{e}/m)^{(m/2)-(1/4)}t^m\exp(c_m t^2/2), \tag{5.47}$$

and by inserting this bound at (5.35), we obtain (2.6).

Finally, it follows from (2.5) that $\Phi_{d,p;m}(A,\Sigma) = O\left(d^{-(1-r)m/2}\right)$ as $d \to \infty$. ∎

**Comments on Remark 2.3.** Denote by $b_m$ the bound in (5.47). By direct calculation, we obtain

$$\frac{b_{m+1}}{b_m} = t\left(\frac{m}{m+1}\right)^{(m/2)-(1/4)}\left(\frac{\mathrm{e}}{m+1}\right)^{1/2}\exp\left((c_{m+1} - c_m)t^2/2\right). \tag{5.48}$$

By (5.40), $c_{m+1} - c_m < 0$; hence the latter three terms in (5.48) each are strictly less than 1, so we obtain $b_{m+1}/b_m < t$. Therefore, if $t \leq 1$, then $b_{m+1} < b_m$, i.e., the sequence $\{b_m, m \geq 2\}$, is strictly decreasing; and by applying this result to (2.6), we obtain the statement in the first part

of Remark 2.3. We have also determined from simple numerical calculations that $b_m$ remains strictly decreasing for $t \leq 1.5$.

For larger values of $t$, the sequence $b_m$ is unimodal. This can be proved by treating $m$ (temporarily) as a continuous variable, calculating the logarithmic derivative of (5.48) with respect to $m$, and verifying that there exists an $m_0$ such that the derivative is positive for all $m \leq m_0$ and negative for $m > m_0$. By graphing the bound (5.47) as a function of $m$, we have observed that it appears to be unimodal for all $t \geq 1.6$.

**Proof of Theorem 2.4.** Since $A$, $\Sigma$ are positive definite, then $C_\kappa(A) > 0$ and $C_\kappa(\Sigma) > 0$ for all partitions $\kappa$. By Proposition 5.7,

$$\Phi_{d,p;m}(A, \Sigma) = \sum_{k=m}^{\infty} \frac{1}{k!} \sum_{|\kappa|=k} \frac{C_\kappa(A)C_\kappa(\Sigma)}{C_\kappa(I_d)} \geq \sum_{k=m}^{\infty} \frac{1}{k!} \sum_{|\kappa|=k} C_\kappa(A)(1+|\kappa|)^{-N_d} \sigma_{(1)}^{\kappa_1} \cdots \sigma_{(p)}^{\kappa_p}$$

$$= \sum_{k=m}^{\infty} \frac{(1+k)^{-N_d}}{k!} \sum_{|\kappa|=k} C_\kappa(A) \sigma_{(1)}^{\kappa_1} \cdots \sigma_{(p)}^{\kappa_p}.$$

Since $\sigma_{(1)} \geq \cdots \geq \sigma_{(d)} > 0$, then $\sigma_{(1)}^{\kappa_1} \cdots \sigma_{(p)}^{\kappa_p} \geq \sigma_{(p)}^{|\kappa|}$ for any partition $\kappa$ of length $p$. Therefore,

$$\sum_{|\kappa|=k} C_\kappa(A) \sigma_{(1)}^{\kappa_1} \cdots \sigma_{(p)}^{\kappa_p} \geq \sigma_{(p)}^k \sum_{|\kappa|=k} C_\kappa(A) = \sigma_{(p)}^k (\operatorname{tr} A)^k,$$

so we obtain

$$\Phi_{d,p;m}(A, \Sigma) \geq \sum_{k=m}^{\infty} \frac{(1+k)^{-N_d}(\sigma_{(p)} \operatorname{tr} A)^k}{k!}.$$

We now apply the telescoping method to bound the latter series. Let $\tau \equiv \sigma_{(p)} \operatorname{tr}(A)$ and, for $j \geq m$, define

$$a_j = \frac{(1+j)^{-N_d} \tau^j}{j!}.$$

Then

$$\frac{a_{j+1}}{a_j} = \left(\frac{1+j}{2+j}\right)^{N_d} \frac{\tau}{1+j}.$$

It is straightforward that, for $j \geq m$, $\frac{1+j}{2+j} \geq \frac{1+m}{2+m}$, and therefore

$$\frac{a_{j+1}}{a_j} \geq \left(\frac{1+m}{2+m}\right)^{N_d} \frac{\tau}{1+j}.$$

Hence for all $k \geq m$,

$$a_k = a_m \prod_{j=m}^{k-1} \frac{a_{j+1}}{a_j} \geq a_m \prod_{j=m}^{k-1} \left[ \left(\frac{1+m}{2+m}\right)^{N_d} \frac{\tau}{1+j} \right] = a_m \left(\frac{1+m}{2+m}\right)^{(k-m)N_d} \frac{m! \tau^{k-m}}{k!}.$$

Substituting for $a_m$ and then simplifying, we obtain

$$a_k \geq (2+m)^{mN_d}(1+m)^{-(1+m)N_d} \left(\frac{1+m}{2+m}\right)^{kN_d} \frac{\tau^k}{k!},$$

and therefore

$$\Phi_{d,p;m}(A, \Sigma) \geq \sum_{k=m}^{\infty} a_k \geq (2+m)^{mN_d}(1+m)^{-(1+m)N_d} \sum_{k=m}^{\infty} \left(\frac{1+m}{2+m}\right)^{kN_d} \frac{\tau^k}{k!}.$$

The proof now is complete. ∎

### 5.4 The proofs for Section 3

**Proof of Theorem 3.1.** Since $B'B$ is positive semidefinite, then all its eigenvalues are non-negative; hence, by (5.9), $C_\kappa\left(\frac{1}{4}B'B\right) \geq 0$. Also, it follows from (2.1) that $(d/2)_\kappa > 0$. Therefore, it follows from (3.1) that $\Psi_{d,p;m}(B) \geq 0$.

By [14, Section 5.6], $\|B'B\| \leq \|B\|^2$. Denoting $\frac{1}{4}B'B$ by $\Lambda$, it follows from the assumption on $B$ that $\|\Lambda\| = \frac{1}{4}\|B'B\| \leq \frac{1}{4}\|B\|^2 \leq \gamma_0 d^{r/2}$. Since $\Lambda$ is positive semidefinite, then by Proposition 5.4,

$$C_\kappa(\Lambda) \leq \left(\prod_{i=1}^{p} \frac{(d^{1/2}/2)_{p\kappa_i}}{(d/2)_{p\kappa_i}}\right)^{1/p} C_\kappa(I_p)\|\Lambda\|^{|\kappa|}.$$

Applying this bound to (3.1), we find that

$$\Psi_{d,p;m}(B) \leq \sum_{k=m}^{\infty} \frac{\|\Lambda\|^k}{k!} \sum_{|\kappa|=k} \frac{C_\kappa(I_p)}{(d/2)_\kappa} \left(\prod_{i=1}^{p} \frac{(d^{1/2}/2)_{p\kappa_i}}{(d/2)_{p\kappa_i}}\right)^{1/p}.$$

Next we apply the lower bound in (5.15) and the bound in (5.16) to obtain

$$\Psi_{d,p;m}(B) \leq \sum_{k=m}^{\infty} \frac{\|\Lambda\|^k}{k!} \sum_{|\kappa|=k} C_\kappa(I_p) \cdot (2p)^{|\kappa|} d^{-|\kappa|} \cdot \alpha_p \gamma_1^{|\kappa|} d^{-|\kappa|/2} p^{|\kappa|/2} (|\kappa|!)^{1/2}$$

$$= \alpha_p \sum_{k=m}^{\infty} \frac{\left(2\gamma_1 p^{3/2} d^{-3/2}\|\Lambda\|\right)^k}{(k!)^{1/2}} \sum_{|\kappa|=k} C_\kappa(I_p).$$

By (5.3) and the bound $\|\Lambda\| \leq \gamma_0 d^{r/2}$, we deduce that

$$\Psi_{d,p;m}(B) \leq \alpha_p \sum_{k=m}^{\infty} \frac{\left(2\gamma_1 p^{3/2} d^{-3/2} \cdot \gamma_0 d^{r/2}\right)^k}{(k!)^{1/2}} p^k = \alpha_p \sum_{k=m}^{\infty} \frac{\left(2\gamma_0\gamma_1 p^{5/2} d^{-(3-r)/2}\right)^k}{(k!)^{1/2}}.$$

Noting that the latter series is identical with $R_m\left(2\gamma_0\gamma_1 p^{5/2} d^{-(3-r)/2}\right)$, where $R_m(t)$ is given in (5.36), we apply the inequality (5.47) to obtain (3.3).

Finally, it follows from (3.2) that $\Psi_{d,p;m}(B) = O\left(d^{-(3-r)m/2}\right)$ as $d \to \infty$. ∎

**Proof of Theorem 3.2.** Since $B$ is of full rank, then $B'B$ is positive definite. Letting $\beta_{(1)} \geq \cdots \geq \beta_{(p)}$ denote the ordered eigenvalues of $\frac{1}{4}B'B$, then by Proposition 5.7,

$$\Psi_{d,p;m}(B) = \sum_{k=m}^{\infty} \frac{1}{k!} \sum_{|\kappa|=k} \frac{C_\kappa\left(\frac{1}{4}B'B\right)}{(d/2)_\kappa} \geq \sum_{k=m}^{\infty} \frac{1}{k!} \sum_{|\kappa|=k} \frac{(1+|\kappa|)^{-N_p} C_\kappa(I_p)}{(d/2)_\kappa} \beta_{(1)}^{\kappa_1} \cdots \beta_{(p)}^{\kappa_p}$$

$$\geq \sum_{k=m}^{\infty} \frac{(1+k)^{-N_p} \beta_{(p)}^k}{k!} \sum_{|\kappa|=k} \frac{C_\kappa(I_p)}{(d/2)_\kappa} \beta_{(1)}^{\kappa_1} \cdots \beta_{(p)}^{\kappa_p},$$

since the ordering of the $\beta_{(j)}$ implies that $\beta_{(1)}^{\kappa_1} \cdots \beta_{(p)}^{\kappa_p} \geq \beta_{(p)}^{|\kappa|}$.

By the upper bound in (5.15), we have

$$\frac{1}{(d/2)_\kappa} \geq 2^{|\kappa|}(d+|\kappa|)^{-|\kappa|};$$

therefore

$$\Psi_{d,p;m}(B) \geq \sum_{k=m}^{\infty} \frac{(1+k)^{-N_p}\beta_{(p)}^k}{k!} \sum_{|\kappa|=k} 2^{|\kappa|}(d+|\kappa|)^{-|\kappa|}C_\kappa(I_p)$$

$$= \sum_{k=m}^{\infty} \frac{(1+k)^{-N_p}2^k(d+k)^{-k}\beta_{(p)}^k}{k!} \sum_{|\kappa|=k} C_\kappa(I_p)$$

$$= \sum_{k=m}^{\infty} \frac{(1+k)^{-N_p}(d+k)^{-k}}{k!}(2p\beta_{(p)})^k.$$

Since each term in the latter series is positive, then we obtain the single-term lower bound,

$$\Psi_{d,p;m}(B) \geq \frac{(1+m)^{-N_p}(d+m)^{-m}}{m!}(2p\beta_{(p)})^m,$$

and this bound also converges to 0 as $d \to \infty$ or $m \to \infty$.

We apply now the telescoping method to obtain a sharper lower bound for $\Psi_{d,p;m}(B)$. Let $\tau \equiv 2p\beta_{(p)}$ and, for $j \geq m$, define $a_j = (1+j)^{-N_p}(d+j)^{-j}\tau^j/j!$. Then

$$\frac{a_{j+1}}{a_j} = \left(\frac{1+j}{2+j}\right)^{N_p}\left(\frac{d+j}{d+1+j}\right)^{j+1} \frac{\tau}{(j+1)(d+1+j)}$$

$$\geq \left(\frac{1+m}{2+m}\right)^{N_p} e^{-1} \frac{\tau}{(j+1)(d+1+j)}$$

since $\frac{1+j}{2+j} \geq \frac{1+m}{2+m}$, $\left(\frac{d+j}{d+1+j}\right)^{j+1} \geq e^{-1}$ for all $j \geq m$ and all $d \geq 1$. Therefore, for all $k \geq m$,

$$a_k = a_m \prod_{j=m}^{k-1} \frac{a_{j+1}}{a_j} \geq a_m \prod_{j=m}^{k-1}\left[\left(\frac{1+m}{2+m}\right)^{N_p}\frac{e^{-1}\tau}{(j+1)(d+1+j)}\right]$$

$$= a_m \frac{\mu^{k-m}}{(m+1)_{k-m}(d+1+m)_{k-m}},$$

where $\mu = \left(\frac{1+m}{2+m}\right)^{N_p}e^{-1}\tau$. Therefore,

$$\Psi_{d,p;m}(B) \geq \sum_{k=m}^{\infty} a_k \geq a_m \sum_{k=m}^{\infty} \frac{\mu^{k-m}}{(m+1)_{k-m}(d+1+m)_{k-m}}$$

$$= a_m \sum_{k=0}^{\infty} \frac{\mu^k}{(m+1)_k(d+1+m)_k}$$

$$= a_m{}_1F_2(1; m+1, d+1+m; \mu).$$

This establishes (3.4). ∎

## 6   Concluding remarks

In this article, we presented an approach, to calculating the normalizing constants of the matrix Bingham and matrix Fisher distributions, that is alternative to numerical simulation methods. Although Monte Carlo computations may be adequate for approximating the normalizing constants $\Phi_{d,p}$ and $\Psi_{d,p}$, such numerical simulations can only yield estimates of the errors in the computed values of the normalizing constants and cannot provide provable bounds on those errors.

The approach given in the present article provides, under stated hypotheses, guaranteed upper bounds for the remainder terms, and therefore is able to calculate the normalizing constants to any desired level of accuracy. Further, the methods developed here provides explicit rates of convergence as $d \to \infty$ or $m \to \infty$, and also deduces in some instances lower bounds on the accuracy of each remainder term.

We note that the methods used in this article are applicable to a general class of distributions, on the Stiefel and Grassmann manifolds, as described by Chikuse [4, pp. 35–36]. Further, the approach given here applies to the complex matrix Bingham and complex matrix Langevin distributions; this can be done by making the necessary changes from the real zonal polynomials to the complex zonal polynomials and from the generalized hypergeometric functions of real symmetric matrix argument to the generalized hypergeometric functions of Hermitian matrix argument [16]. In particular, we note that the complex analog of (5.1), the crucial integral representation for the zonal polynomials, is provided in [11, Theorem 4.8].

Finally, we remark that all the results in this article extend, *mutatis mutandis*, to arbitrary symmetric cones. A basic principle observed from [6, 11] is that methods that are successful for the cone of real positive definite matrices will remain successful for any symmetric cone. Indeed, the results that we have established can be extended to derive rates of convergence of the spherical polynomial series expansions of all generalized hypergeometric functions defined on the symmetric cones.

## Acknowledgements

# References

[1] Bagyan A., Richards D., Complete asymptotic expansions and the high-dimensional Bingham distributions, *TEST* **33** (2024), 540–563, arXiv:2303.04575.

[2] Bingham C., An antipodally symmetric distribution on the sphere, *Ann. Statist.* **2** (1974), 1201–1225.

[3] Bingham C., Chang T., Richards D., Approximating the matrix Fisher and Bingham distributions: applications to spherical regression and Procrustes analysis, *J. Multivariate Anal.* **41** (1992), 314–337.

[4] Chikuse Y., Statistics on special manifolds, *Lect. Notes Stat.*, Vol. 174, Springer, New York, 2003.

[5] de Waal D.J., On the normalizing constant for the Bingham–von Mises–Fisher matrix distribution, *South African Statist. J.* **13** (1979), 103–112.

[6] Ding H., Gross K.I., Richards D.St.P., Ramanujan's master theorem for symmetric cones, *Pacific J. Math.* **175** (1996), 447–490.

[7] Dryden I.L., Statistical analysis on high-dimensional spheres and shape spaces, *Ann. Statist.* **33** (2005), 1643–1665, arXiv:math/0508279.

[8] Edelman A., Kostlan E., Shub M., How many eigenvalues of a random matrix are real?, *J. Amer. Math. Soc.* **7** (1994), 247–267.

[9] Faraut J., Korányi A., Analysis on symmetric cones, *Oxford Math. Monogr.*, The Clarendon Press, Oxford University Press, New York, 1994.

[10] Granados-Garcia G., Fiecas M., Babak S., Fortin N.J., Ombao H., Brain waves analysis via a non-parametric Bayesian mixture of autoregressive kernels, *Comput. Statist. Data Anal.* **174** (2022), 107409, 181 pages, arXiv:2102.11971.

[11] Gross K.I., Richards D.St.P., Special functions of matrix argument. I. Algebraic induction, zonal polynomials, and hypergeometric functions, *Trans. Amer. Math. Soc.* **301** (1987), 781–811.

[12] Hoff P.D., Simulation of the matrix Bingham-von Mises–Fisher distribution, with applications to multivariate and relational data, *J. Comput. Graph. Statist.* **18** (2009), 438–456, arXiv:0712.4166.

[13] Holbrook A., Vandenberg-Rodes A., Shahbaba B., Bayesian inference on matrix manifolds for linear dimensionality reduction, arXiv:1606.04478.

[14] Horn R.A., Johnson C.R., Matrix analysis, 2nd ed., Cambridge University Press, Cambridge, 2012.

[15] Impens C., Stirling's series made easy, Amer. Math. Monthly 110 (2003), 730–735.

[16] James A.T., Distributions of matrix variates and latent roots derived from normal samples, Ann. Math. Statist. 35 (1964), 475–501.

[17] James A.T., Calculation of zonal polynomial coefficients by use of the Laplace–Beltrami operator, Ann. Math. Statist. 39 (1968), 1711–1718.

[18] James A.T., The variance information manifold and the functions on it, in Multivariate Analysis III, Academic Press, New York, 1973, 157–169.

[19] Jauch M., Hoff P.D., Dunson D.B., Monte Carlo simulation on the Stiefel manifold via polar expansion, J. Comput. Graph. Statist. 30 (2021), 622–631, arXiv:1906.07684.

[20] Jiu L., Koutschan C., Calculation and properties of zonal polynomials, Math. Comput. Sci. 14 (2020), 623–640, arXiv:2001.11599.

[21] Jupp P.E., Mardia K.V., Maximum likelihood estimators for the matrix von Mises–Fisher and Bingham distributions, Ann. Statist. 7 (1979), 599–606.

[22] Jupp P.E., Mardia K.V., A unified view of the theory of directional statistics, 1975–1988, Int. Stat. Rev. 57 (1989), 261–294.

[23] Khatri C.G., Mardia K.V., The von Mises–Fisher matrix distribution in orientation statistics, J. Roy. Statist. Soc. Ser. B 39 (1977), 95–106.

[24] Kume A., Preston S.P., Wood A.T.A., Saddlepoint approximations for the normalizing constant of Fisher–Bingham distributions on products of spheres and Stiefel manifolds, Biometrika 100 (2013), 971–984.

[25] Kume A., Wood A.T.A., Saddlepoint approximations for the Bingham and Fisher–Bingham normalising constants, Biometrika 92 (2005), 465–476.

[26] Kushner H.B., On the expansion of $C_\rho^*(V + I)$ as a sum of zonal polynomials, J. Multivariate Anal. 17 (1985), 84–98.

[27] Ledermann W., Introduction to group characters, Cambridge University Press, Cambridge, 1977.

[28] Mantoux C., Couvy-Duchesne B., Cacciamani F., Epelbaum S., Durrleman S., Allassonnière S., Understanding the variability in graph data sets through statistical modeling on the Stiefel manifold, Entropy 23 (2021), 490, 34 pages.

[29] Mardia K.V., Jupp P.E., Directional statistics, Wiley Ser. Probab. Stat., John Wiley & Sons, Chichester, 1999.

[30] Muirhead R.J., Aspects of multivariate statistical theory, Wiley Ser. Probab. Math. Stat., John Wiley & Sons, New York, 1982, .

[31] Oualkacha K., Rivest L.P., On the estimation of an average rigid body motion, Biometrika 99 (2012), 585–598.

[32] Parkhurst A.M., James A.T., Zonal polynomials of order 1 through 12, in Selected Tables in Mathematical Statistics, Vol. II, American Mathematical Society, Providence, RI, 1974, 199–388.

[33] Prentice M.J., Antipodally symmetric distributions for orientation statistics, J. Statist. Plann. Inference 6 (1982), 205–214.

[34] Richards D.St.P., Applications of invariant differential operators to multivariate distribution theory, SIAM J. Appl. Math. 45 (1985), 280–288.

[35] Richards D.St.P., Functions of matrix argument, in NIST Handbook of Mathematical Functions, Vol. 35, U.S. Department of Commerce, National Institute of Standards and Technology, Washington, DC, 2010, 767–774.

[36] Sudlow C., Gallacher J., Allen N., Beral V., Burton P., Danesh J., Downey P., Elliott P., Green J., Landray M., UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age, PLoS Med. 12 (2015), e1001779, 11 pages.