Journal of Multivariate Analysis **82**, 240–262 (2002)
doi:10.1006/jmva.2001.2016

# Moment Properties of the Multivariate Dirichlet Distributions

Rameshwar D. Gupta[1]

*University of New Brunswick, St. John, New Brunswick, Canada*

and

Donald St. P. Richards[2]

*Specola Vaticana, Vatican City*

Received June 21, 2000; published online February 5, 2002

Let $X_1, \ldots, X_n$ be real, symmetric, $m \times m$ random matrices; denote by $I_m$ the $m \times m$ identity matrix; and let $a_1, \ldots, a_n$ be fixed real numbers such that $a_j > (m-1)/2$, $j = 1, \ldots, n$. Motivated by the results of J. G. Mauldon (*Ann. Math. Statist.* **30** (1959), 509–520) for the classical Dirichlet distributions, we consider the problem of characterizing the joint distribution of $(X_1, \ldots, X_n)$ subject to the condition that $\mathbb{E}\,|I_m - \sum_{j=1}^n T_j X_j|^{-(a_1+\cdots+a_n)} = \prod_{j=1}^n |I_m - T_j|^{-a_j}$ for all $m \times m$ symmetric matrices $T_1, \ldots, T_n$ in a neighborhood of the $m \times m$ zero matrix. Assuming that the joint distribution of $(X_1, \ldots, X_n)$ is orthogonally invariant, we deduce the following results: each $X_j$ is positive-definite, almost surely; $X_1 + \cdots + X_n = I_m$, almost surely; the marginal distribution of the sum of any proper subset of $X_1, \ldots, X_n$ is a multivariate beta distribution; and the joint distribution of the determinants $(|X_1|, \ldots, |X_n|)$ is the same as the joint distribution of the determinants of a set of matrices having a multivariate Dirichlet distribution with parameter $(a_1, \ldots, a_n)$. In particular, for $n = 2$ we obtain a new characterization of the multivariate beta distribution. © 2002 Elsevier Science (USA)

AMS 1990 subject classifications: 62E15; 62H10; 60E15; 62H05; 62G30.

*Key words and phrases:* characterizations; confluent hypergeometric function; multivariate beta distribution; multivariate Dirichlet distribution; Gaussian hypergeometric function; generalized power function; Laplace transform; multivariate beta distribution; multivariate gamma function; symmetric cone; Weyl fractional derivative; Wishart distribution; zonal polynomial.

[1] Research supported in part by Natural Sciences and Engineering Research Council Grant OGP-0004850.

[2] Permanent address: Department of Statistics, University of Virginia, Charlottesville, VA 22904. Research supported in part by National Science Foundation Grant DMS-9703705.



0047-259X/02 $35.00
© 2002 Elsevier Science (USA)
All rights reserved.

## 1. INTRODUCTION

Suppose $X_1, ..., X_n$ are real, symmetric, positive-definite, $m \times m$ random matrices; $a_1, ..., a_n$ are fixed real numbers with $a_j > (m-1)/2$ for all $j = 1, ..., n$; and $a_{\bullet} \equiv a_1 + \cdots + a_n$. In this paper, we consider the problem of characterizing the joint distribution of $(X_1, ..., X_n)$ by means of the moment function

$$M(T_1, ..., T_n) := \mathbb{E} \left| I_m - \sum_{j=1}^{n} T_j X_j \right|^{-a_{\bullet}}, \tag{1.1}$$

where the variables $T_1, ..., T_n$ are real, symmetric $m \times m$ matrices. To ensure the existence of the expectation (1.1) we necessarily must assume that, with probability one, the norm of $\sum_{j=1}^{n} T_j X_j$ is less than one for all $T_1, ..., T_n$ for which the expectation (1.1) exists. Specifically, we address the problem of characterizing the class of distributions for which the expectation in (1.1) equals

$$\prod_{j=1}^{n} |I_m - T_j|^{-a_j} \tag{1.2}$$

for all $T_1, ..., T_n$ with norm less than one.

We are motivated to study this problem because of a remarkable article of J. G. Mauldon (1959). In the one-dimensional case, $m = 1$, Mauldon (1959) proved under certain regularity conditions that (1.1) and (1.2) are identical if and only if the random vector $(X_1, ..., X_n)$ follows the classical Dirichlet distribution with parameter $(a_1, ..., a_n)$. Moreover, by also treating the expectation property

$$\mathbb{E}\left(1 - \sum_{j=1}^{n} t_j X_j\right)^{-(a_1 + \cdots + a_r)} = \prod_{j=1}^{r} \left(1 - \sum_{k=1}^{n} c_{k,j} t_k\right)^{-a_j}, \tag{1.3}$$

which is assumed to hold for all sufficiently small $t_1, ..., t_n \in \mathbb{R}$ and for given constants $c_{j,k}$, Mauldon characterized a class of distributions, which he named the *basic $\beta$-distributions* and which contain the Dirichlet distributions as special cases. It is also noteworthy that Mauldon's paper, despite its seminal nature, has languished in undeserved obscurity; indeed, in a search of the Science Citation Index, we discovered that since its appearance, only two publications have cited Mauldon's article.

In addition to Mauldon's results, several authors developed applications or characterizations based on formulas similar to (1.3). In work on the distribution theory of serial correlations, Watson (1956) showed that for $m = 1$ and $a_1 = \cdots = a_n = 1$, (1.1) and (1.2) were identical in the case of the

Dirichlet distribution with parameter $(1, \ldots, 1)$. Karlin *et al.* (1986) observed later that Watson's method established the equality of (1.1) and (1.2) for all Dirichlet distributions, and they generalized Watson's formula to multivariable settings. Other characterizations and applications were developed by Volodin *et al.* (1993), Chamayou and Letac (1994), Letac and Scarsini (1998), and Gupta and Richards (2000).

In the general case, $m \geqslant 1$, further motivation for our work is provided by results of Letac and Massam (1998) and Letac *et al.* (2000). Denote by $I_m$ the $m \times m$ identity matrix, and let us write $X > 0$ whenever $X$ is a real symmetric positive-definite $m \times m$ matrix. For symmetric matrices $X_1, \ldots, X_n$, let

$$S_n := \{(X_1, \ldots, X_n) : X_j > 0 \text{ for all } 1 \leqslant j \leqslant n, \text{ and } X_1 + \cdots + X_n = I_m\}$$

denote the matrix analog of a simplex. For $a \in \mathbb{R}$, $a > (m-1)/2$, let

$$\Gamma_m(a) = \pi^{m(m-1)/4} \prod_{k=1}^{m} \Gamma(a - \tfrac{1}{2}(k-1)) \tag{1.4}$$

(cf. Muirhead (1982, p. 62)). Following Olkin and Rubin (1964) (cf. Johnson and Kotz, 1972)), the $m \times m$ random matrices $X_1, \ldots, X_n$ are said to follow a multivariate Dirichlet distribution with parameters $a_1, \ldots, a_n$ if, relative to Lebesgue measure on the simplex $S_n$, the joint probability density function of $(X_1, \ldots, X_n)$ is

$$\frac{\Gamma_m(a_\bullet)}{\Gamma_m(a_1) \cdots \Gamma_m(a_n)} \prod_{j=1}^{n} |X_j|^{a_j - (m+1)/2}, \quad (X_1, \ldots, X_n) \in S_n. \tag{1.5}$$

We will write $(X_1, \ldots, X_n) \sim D_m(a_1, \ldots, a_n)$ whenever (1.5) holds. For $n = 2$, the density function (1.5) reduces to the well known multivariate beta distribution; we will designate this by the notation $X_1 \sim \boldsymbol{\beta}_m(a_1, a_2)$.

Assuming that $(X_1, \ldots, X_n) \sim D_m(a_1, \ldots, a_n)$, Letac and Massam (1998) established the equality of (1.1) and (1.2) for the case in which $T_1, \ldots, T_n$ are scalar matrices; i.e., $T_j = t_j I_m$, $t_j \in \mathbb{R}$, $j = 1, \ldots, n$. Subsequently, Letac *et al.* (2000) derived an even more general expectation formula, which may be described as follows. For any $m \times m$ real symmetric positive-definite matrix $X$, denote by $\Delta_k(X)$ the $k$th principal minor of $X$, $k = 1, \ldots, m$. Further, for $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m) \in \mathbb{R}^m$ define the *generalized power function* by

$$\Delta_{\boldsymbol{\alpha}}(X) := |X|^{\alpha_m} \prod_{k=1}^{m-1} \Delta_k(X)^{\alpha_k - \alpha_{k+1}}. \tag{1.6}$$

We define the *multivariate gamma function* (Faraut and Korányi, 1994; Letac *et al.* 2000) on the set $\{\boldsymbol{\alpha}=(\alpha_1, ..., \alpha_m) \in \mathbb{R}^m\colon \alpha_j > (j-1)/2,\ j=1, ..., m, \}$ by

$$\Gamma_m(\boldsymbol{\alpha}) = \pi^{m(m-1)/4} \prod_{k=1}^{m} \Gamma(\alpha_k - \tfrac{1}{2}(k-1)). \tag{1.7}$$

For the case in which $\alpha_1 = \cdots = \alpha_m = a$, (1.7) reduces to (1.4), and the two notations are consistent.

Suppose $T$ is a positive-definite symmetric $m \times m$ matrix; $t_1, ..., t_m \in \mathbb{R}$ are sufficiently small; $a_1, ..., a_n > (m-1)/2$; and also $s_1, ..., s_n \in \mathbb{R}^m$ are vectors whose components are sufficiently large. If $(X_1, ..., X_n)$ follows the multivariate Dirichlet distribution (1.5) then Letac *et al.* (2000) proved that

$$\begin{aligned}\mathbb{E}\, &\Delta_{s_1}(X_1) \cdots \Delta_{s_n}(X_n)\, \Delta_{s_1+\cdots+s_n+a_\bullet}((T+t_1X_1+\cdots+t_nX_n)^{-1}) \\ &= \frac{\Gamma_m(a_\bullet)}{\Gamma_m(s_1+\cdots+s_n+a_\bullet)} \prod_{j=1}^{n} \frac{\Gamma_m(s_j+a_j)}{\Gamma_m(a_j)}\, \Delta_{s_j+a_j}((T+t_jI_m)^{-1}).\end{aligned} \tag{1.8}$$

Letac *et al.* (2000) deduced the maximal range of values of $s_1, ..., s_n$ and $t_1, ..., t_n$ for which (1.8) is valid, and established analogs of (1.8) on spaces of symmetric cones, of which the space of positive-definite symmetric matrices is only one example. Moreover, Letac *et al.* (2000) derived (1.8) for the "singular" multivariate Dirichlet distributions.

In this paper we shall consider two problems. The first problem we address is the calculation of the expectation (1.1) for arbitrary symmetric matrices $T_1, ..., T_n$ when $(X_1, ..., X_n)$ follows the multivariate Dirichlet distribution (1.5). For $n=2$, we show that (1.1) and (1.2) are identical. For $n \geqslant 3$, we also find sufficient conditions on $T_1, ..., T_n$ for (1.1) and (1.2) to be equal. Further, we describe the difficulties intrinsic to evaluation of (1.1) for $n \geqslant 3$ and arbitrary $T_1, ..., T_n$.

The second problem we consider is the characterization of the joint distribution of $(X_1, ..., X_n)$ through the hypothesis that (1.1) and (1.2) are identical. Under certain invariance assumptions, we show for $n=2$ that the equality of (1.1) and (1.2) characterizes the Dirichlet distribution $D_m(a_1, a_2)$; since this result is valid for all $m \geqslant 1$ we have then a partial extension of Mauldon's characterization of the Dirichlet distributions.

For $n \geqslant 3$, under similar invariance assumptions, we show that if (1.1) and (1.2) are identical then the matrices $X_1, ..., X_n$ satisfy the following conditions: (i) For all $j=1, ..., n$, $X_j > 0$, and $X_1+\cdots+X_n = I_m$ almost surely; (ii) the marginal distribution of the sum, $X = X_{k_1}+\cdots+X_{k_r}$, of any subset of $X_1, ..., X_n$ has a multivariate beta distribution $\boldsymbol{\beta}_m(a_{k_1}+\cdots+a_{k_r}; a_\bullet - a_{k_1} - \cdots - a_{k_r})$; (iii) the joint distribution of the determinants

$(|X_1|, ..., |X_n|)$ is the same as the joint distribution of the determinants of a set of random matrices having a multivariate Dirichlet distribution; and (iv) the principal submatrices of each $X_j$ all satisfy analogs of (i)–(iii). Noting that the Dirichlet distributions $D_m(a_1, ..., a_n)$ in (1.5) satisfy all of (i)–(iv), we then have a partial characterization of those distributions.

## 2. EVALUATIONS OF THE MOMENT FUNCTION $M(T_1, ..., T_n)$

Throughout this section, we assume that $(X_1, ..., X_n) \sim D_m(a_1, ..., a_n)$. As we noted earlier, Letac and Massam (1998) evaluated the function $M(T_1, ..., T_n)$ in (1.1) for the case in which $T_1, ..., T_n$ are scalar matrices; and their result was later extended by Letac *et al.* (2000) to the formula (1.8). We will begin this section by evaluating (1.1) for a more general class of matrices. First, we note from Letac *et al.* (2000) a direct method of evaluating (1.1) for the case in which $T_1, ..., T_n$ are scalar; as we show below, this direct method also yields the evaluation of (1.1) for a broader class of matrices $T_1, ..., T_n$.

Now suppose that $T_j = t_j I_m$ for $t_j \in \mathbb{R}$, $j = 1, ..., n$. It is well known that

$$\prod_{j=1}^{n} (1-t_j)^{-ma_j} = \prod_{j=1}^{n} |(1-t_j)\, I_m|^{-a_j} = \prod_{j=1}^{n} \mathbb{E} \exp(t_j \operatorname{tr} V_j), \tag{2.1}$$

where $V_1, ..., V_n$ are mutually independent positive-definite symmetric $m \times m$ random matrices, and each $V_j$ has a Wishart distribution with probability density function

$$\frac{1}{\Gamma_m(a_j)} |V_j|^{a_j-(m+1)/2} \exp(-\operatorname{tr} V_j), \qquad V_j > 0. \tag{2.2}$$

It is well known that the moment-generating function of $V_j$ is

$$\mathbb{E} \exp(\operatorname{tr} TV_j) = |I_m - T|^{-a_j}, \tag{2.3}$$

for $\|T\| < 1$, where $\|T\|$ denotes the maximum of the absolute values of all eigenvalues of $T$.

From (2.1) and the mutual independence of $V_1, ..., V_n$, we obtain

$$\prod_{j=1}^{n} (1-t_j)^{-ma_j} = \mathbb{E} \exp\left( \operatorname{tr} \sum_{j=1}^{n} t_j V_j \right). \tag{2.4}$$

By Gupta and Richards (1987), Proposition 7.3(i), we find that the random matrices $V_1, ..., V_n$ satisfy the stochastic representation

$$(V_1, ..., V_n) \stackrel{\mathrm{d}}{=} V^{1/2}(X_1, ..., X_n)\, V^{1/2}, \tag{2.5}$$

where $V$ and $(X_1, ..., X_n)$ are mutually independent; $(X_1, ..., X_n) \sim D_m(a_1, ..., a_n)$; and $V \stackrel{\mathrm{d}}{=} V_1 + \cdots + V_n$. Thus, $V$ has the density function

$$\frac{1}{\Gamma_m(a_\bullet)} |V|^{a_\bullet - (m+1)/2} \exp(-\operatorname{tr} V), \qquad V > 0. \tag{2.6}$$

Therefore by (2.5),

$$\mathbb{E} \exp\left(\operatorname{tr} \sum_{j=1}^{n} t_j V_j\right) = \mathbb{E} \exp\left(\operatorname{tr} \sum_{j=1}^{n} t_j X_j V\right) = \mathbb{E} \left| I_m - \sum_{j=1}^{n} t_j X_j \right|^{-a_\bullet}, \tag{2.7}$$

where the second equality follows from (2.3) and (2.6). Comparing (2.4) and (2.7) we obtain

$$\mathbb{E} \left| I_m - \sum_{j=1}^{n} t_j X_j \right|^{-a_\bullet} = \prod_{j=1}^{n} (1 - t_j)^{-m a_j}, \tag{2.8}$$

which establishes the equality of (1.1) and (1.2) for the case in which $T_1, ..., T_n$ are scalar matrices.

More general than the foregoing is the following result.

2.1. Proposition. *Suppose* $(X_1, ..., X_n) \sim D_m(a_1, ..., a_n)$ *and* $T_1, ..., T_n$ *are* $m \times m$ *symmetric matrices for which* $\|T_j\| < 1$, $j = 1, ..., n$. *If either*

(i) $n = 2$, *or*

(ii) $n \geqslant 3$ *and the matrices* $T_1, ..., T_n$ *are such that, for some* $r$, $1 \leqslant r \leqslant n$, $T_j - T_r$ *is a scalar matrix for all* $j = 1, ..., n$; *then*

$$\mathbb{E} \left| I_m - \sum_{j=1}^{n} T_j X_j \right|^{-a_\bullet} = \prod_{j=1}^{n} |I_m - T_j|^{-a_j}. \tag{2.9}$$

*Proof.* (i) For $n = 2$ the joint distribution of $(X_1, X_2)$ is concentrated on the matrix simplex $S_2$, so that $X_2 = I_m - X_1$. Then

$$\begin{aligned}
\mathbb{E}\, |I_m - T_1 X_1 - T_2 X_2|^{-(a_1+a_2)} &= \mathbb{E}\, |I_m - T_1 X_1 - T_2 (I_m - X_1)|^{-(a_1+a_2)} \\
&= \mathbb{E}\, |(I_m - T_2) - (T_1 - T_2)\, X_1|^{-(a_1+a_2)} \\
&= |I_m - T_2|^{-(a_1+a_2)}\, \mathbb{E}\, |I_m - T_0 X_1|^{-(a_1+a_2)},
\end{aligned} \tag{2.10}$$

where $T_0 = (I_m - T_2)^{-1/2}\, (T_1 - T_2)(I_m - T_2)^{-1/2}$.

Note that $X_1 \sim \boldsymbol{\beta}_m(a_1, a_2)$, a multivariate beta distribution with probability density function

$$\frac{\Gamma_m(a_1+a_2)}{\Gamma_m(a_1)\,\Gamma_m(a_2)}\,|X_1|^{a_1-(m+1)/2}\,|I_m - X_1|^{a_2-(m+1)/2}, \tag{2.11}$$

$0 < X_1 < I_m$. By applying a well known Euler integral for the Gauss hypergeometric function, ${}_2F_1$, of matrix argument (cf. Muirhead, 1982, p. 264) we deduce from (2.10) and (2.11)

$$\begin{aligned}\mathbb{E}\,|I_m - T_0 X_1|^{-(a_1+a_2)} &= \frac{\Gamma_m(a_1+a_2)}{\Gamma_m(a_1)\,\Gamma_m(a_2)} \int_{0<X_1<I_m} |X_1|^{a_1-(m+1)/2} \\ &\quad \times |I_m - X_1|^{a_2-(m+1)/2}\,|I_m - T_0 X_1|^{-(a_1+a_2)}\,dX_1 \\ &= {}_2F_1(a_1, a_1+a_2; a_1+a_2; T_0).\end{aligned}$$

This last equality requires that $\|T_0\| < 1$.

By a well known result for the Gauss hypergeometric functions of matrix argument (cf. Herz, 1955),

$${}_2F_1(a_1, a_1+a_2; a_1+a_2; T_0) = |I_m - T_0|^{-a_1}; \tag{2.12}$$

therefore the right-hand side of (2.10) reduces to

$$\begin{aligned}&|I_m - T_2|^{-(a_1+a_2)}\,|I_m - T_0|^{-a_1} \\ &\quad = |I_m - T_2|^{-(a_1+a_2)}\,|I_m - (I_m - T_2)^{-1/2}\,(T_1 - T_2)(I_m - T_2)^{-1/2}|^{-a_1} \\ &\quad = |I_m - T_2|^{-(a_1+a_2)}\,|I_m - (I_m - T_2)^{-1}\,(T_1 - T_2)|^{-a_1} \\ &\quad = |I_m - T_2|^{-a_2}\,|(I_m - T_2)(I_m - (I_m - T_2)^{-1}\,(T_1 - T_2))|^{-a_1} \\ &\quad = |I_m - T_1|^{-a_1}\,|I_m - T_2|^{-a_2}.\end{aligned}$$

This establishes the result (i). Note also that, by applying analytic continuation along the lines of Herz (1955), the final result is seen to be valid for all symmetric $T_1$ and $T_2$ satisfying $\|T_j\| < 1$, $j = 1, 2$.

(ii) We shall assume without loss of generality that $r = n$, so that $T_j - T_n$ is scalar for all $j = 1, \ldots, n-1$. Let $V_1, \ldots, V_n$ be independent random matrices with density functions as in (2.2). By the moment-generating function (2.3) and the stochastic representation (2.5), we have

$$\begin{aligned}\prod_{j=1}^{n} |I_m - T_j|^{-a_j} &= \mathbb{E}\exp\left(\operatorname{tr} \sum_{j=1}^{n} T_j V_j\right) \\ &= \mathbb{E}\exp\left(\operatorname{tr} \sum_{j=1}^{n} T_j V^{1/2} X_j V^{1/2}\right).\end{aligned}$$

Since $X_n = I_m - (X_1 + \cdots + X_{n-1})$,

$$\sum_{j=1}^{n} T_j V^{1/2} X_j V^{1/2} = \sum_{j=1}^{n-1} T_j V^{1/2} X_j V^{1/2} + T_n V^{1/2} \left( I_m - \sum_{j=1}^{n-1} X_j \right) V^{1/2}$$
$$= \sum_{j=1}^{n-1} (T_j - T_n)\, V^{1/2} X_j V^{1/2} + T_n V.$$

By hypothesis, $T_j - T_n = t_j I_m$, a scalar matrix, for all $j = 1, \ldots, n-1$; therefore,

$$\operatorname{tr} \sum_{j=1}^{n-1} (T_j - T_n)\, V^{1/2} X_j V^{1/2} + \operatorname{tr} T_n V = \operatorname{tr} \sum_{j=1}^{n-1} t_j V^{1/2} X_j V^{1/2} + \operatorname{tr} T_n V$$
$$= \operatorname{tr} \sum_{j=1}^{n-1} t_j X_j V + \operatorname{tr} T_n V$$
$$= \operatorname{tr} \sum_{j=1}^{n} T_j X_j V,$$

where the last equality holds by virtue of $X_1 + \cdots + X_n = I_m$. Therefore,

$$\prod_{j=1}^{n} |I_m - T_j|^{-a_j} = \mathbb{E} \exp\left( \operatorname{tr} \sum_{j=1}^{n} T_j X_j V \right)$$
$$= \mathbb{E} \left| I_m - \sum_{j=1}^{n} T_j X_j \right|^{-a_\bullet}.$$

Now the proof of (2.9) is complete. ▮

In general, it appears to be difficult to obtain simple expressions for the expectation (2.9) without restrictions on $T_1, \ldots, T_n$ for $n \geqslant 3$. Nevertheless, in our next result we derive a single matrix integral representation for the expectation $\mathbb{E}\, |I_m - \sum_{j=1}^{n} t_j X_j|^{-a}$, where $t_1, \ldots, t_n \in \mathbb{R}$ and $a \in \mathbb{R}$ with $a < a_\bullet$. As a limiting case of this result we obtain a new proof of (2.8).

In preparation for this result, we introduce some notation. For a scalar-valued function $f$ defined on the space $\{S > 0\}$, the Weyl fractional integral operator is defined as

$$\mathscr{W}_T^a f = \frac{1}{\Gamma_m(a)} \int_{S>0} |S|^{a-(m+1)/2} f(S+T)\, dS, \tag{2.13}$$

$T > 0$, where $a > (m-1)/2$. As a well known example, it is easy to calculate from (2.13) that if $f_0(S) = \exp(-\mathrm{tr}\, SV)$, $S > 0$, where $V > 0$ is fixed, then

$$\mathscr{W}_T^a f_0 = |V|^{-a} \exp(-\mathrm{tr}\, TV), \tag{2.14}$$

$T > 0$, $a > (m-1)/2$.

All the properties of the operator $\mathscr{W}_T^a$ we shall need are given in detail by Gupta and Richards (1987, 1995); in particular, $\mathscr{W}_T^a$ satisfies the semigroup property

$$\mathscr{W}_T^{a+b} = \mathscr{W}_T^a \mathscr{W}_T^b,$$

for $a, b > (m-1)/2$. By means of the semigroup property, we can extend the range of $a \in \mathbb{C}$ for which $\mathscr{W}_T^a$ is well defined into the left half-plane through analytic continuation in $a$. As a consequence, $\mathscr{W}_T^0$ may be identified with the identity operator in that

$$\lim_{a \to 0+} \mathscr{W}_T^a f = f,$$

where the limit is in the pointwise sense. Moreover, by means of the semigroup property, (2.14) is well defined for all $a \in \mathbb{C}$.

We also let

$$\hat{\phi}(T) = \int_{S>0} \exp(-\mathrm{tr}\, ST)\, \phi(S)\, dS, \qquad T > 0,$$

denote the Laplace transform on the space $\{S > 0\}$.

With these conventions in place, we have the following result.

2.2. THEOREM. *Suppose* $(X_1, ..., X_n) \sim D_m(a_1, ..., a_n)$, $t_1, ..., t_n \in \mathbb{R}$, $a \in \mathbb{R}$ *with* $a < a_{\bullet}$, *and* $R$ *is an* $m \times m$, *positive-definite symmetric matrix. Then*

$$\mathbb{E}\left|R + \sum_{j=1}^n t_j X_j\right|^{-a} = \frac{\Gamma_m(\sum_{j=1}^n a_j)}{\Gamma_m(a)\, \Gamma_m(a_{\bullet} - a)} \times \int_{S>0} |S|^{a_{\bullet} - a - (m+1)/2} \prod_{j=1}^n |S + R + t_j I_m|^{-a_j}\, dS. \tag{2.15}$$

*Proof.* By a well known formula (the Laplace transform of the Wishart distribution), we have

$$\mathbb{E}\left|R + \sum_{j=1}^n t_j X_j\right|^{-a} = \mathbb{E}\frac{1}{\Gamma_m(a)} \int_{S>0} |S|^{a-(m+1)/2} \exp\left(-\mathrm{tr}\, S\left(R + \sum_{j=1}^n t_j X_j\right)\right) dS.$$

Applying Fubini's theorem to interchange the expectation and integral, and substituting $X_n = I_m - \sum_{j=1}^{n-1} X_j$, we deduce

$$
\begin{aligned}
\mathbb{E}\left|R + \sum_{j=1}^{n} t_j X_j\right|^{-a} &= \frac{1}{\Gamma_m(a)} \int_{S>0} |S|^{a-(m+1)/2} \\
&\quad \times \exp(-\operatorname{tr} SR)\, \mathbb{E} \exp\left(-\operatorname{tr} S \sum_{j=1}^{n} t_j X_j\right) dS \\
&= \frac{1}{\Gamma_m(a)} \int_{S>0} |S|^{a-(m+1)/2} \exp(-\operatorname{tr} S(R + t_n I_m)) \\
&\quad \times \mathbb{E} \exp\left(\operatorname{tr} \sum_{j=1}^{n-1} (t_n - t_j)\, S X_j\right) dS. \qquad (2.16)
\end{aligned}
$$

We introduce the transformation

$$
X_j = \begin{cases} Y_j, & j = 1, \ldots, n-2 \\ \left(I_m - \sum_{j=1}^{n-2} Y_j\right)^{1/2} Y_{n-1} \left(I_m - \sum_{j=1}^{n-2} Y_j\right)^{1/2}, & j = n-1. \end{cases} \qquad (2.17)
$$

On calculating the Jacobian of the transformation (2.17), we deduce that $(Y_1, \ldots, Y_{n-2})$ and $Y_{n-1}$ are mutually independent, and also that $(Y_1, \ldots, Y_{n-2}, I_m - \sum_{j=1}^{n-2} Y_j) \sim D_m(a_1, \ldots, a_{n-2}, a_{n-1} + a_n)$ and $Y_{n-1} \sim \boldsymbol{\beta}_m(a_{n-1}, a_n)$. Thus

$$
\begin{aligned}
&\mathbb{E} \exp\left(\operatorname{tr} \sum_{j=1}^{n-1} (t_n - t_j)\, S X_j\right) \\
&\quad = \mathbb{E}_{Y_1, \ldots, Y_{n-2}} \prod_{j=1}^{n-2} \exp(\operatorname{tr}(t_n - t_j)\, S Y_j) \\
&\quad \times \mathbb{E}_{Y_{n-1}} \exp\left(\operatorname{tr}(t_n - t_{n-1})\, S \left(I_m - \sum_{j=1}^{n-2} Y_j\right)^{1/2} Y_{n-1} \left(I_m - \sum_{j=1}^{n-2} Y_j\right)^{1/2}\right). \qquad (2.18)
\end{aligned}
$$

Recall the confluent, or ${}_1F_1$, hypergeometric function with matrix argument (cf. Muirhead, 1982, p. 264), which may be defined by the integral formula

$$
\begin{aligned}
&{}_1F_1(a_{n-1}; a_{n-1} + a_n; S) \\
&\quad = \frac{\Gamma_m(a_{n-1} + a_n)}{\Gamma_m(a_{n-1})\, \Gamma_m(a_n)} \int_{0<Y<I_m} |Y|^{a_{n-1}-(m+1)/2} |I_m - Y|^{a_n-(m+1)/2} \exp(\operatorname{tr} SY)\, dY, \qquad (2.19)
\end{aligned}
$$

valid for any $m\times m$ symmetric matrix $S$, $a_{n-1}>(m-1)/2$ and $a_n>(m-1)/2$.

Since $Y_{n-1}\sim \boldsymbol{\beta}_m(a_{n-1}, a_n)$, it follows from (2.19) that, conditional on $Y_1, \ldots, Y_{n-2}$,

$$\begin{aligned}
&\mathbb{E}_{Y_{n-1}} \exp\left(\operatorname{tr}(t_n-t_{n-1})\, S\left(I_m-\sum_{j=1}^{n-2} Y_j\right)^{1/2} Y_{n-1}\left(I_m-\sum_{j=1}^{n-2} Y_j\right)^{1/2}\right)\\
&\quad={}_1F_1(a_{n-1}; a_{n-1}+a_n; (t_n-t_{n-1})\left(I_m-\sum_{j=1}^{n-2} Y_j\right)^{1/2} S\left(I_m-\sum_{j=1}^{n-2} Y_j\right)^{1/2})\\
&\quad={}_1F_1(a_{n-1}; a_{n-1}+a_n; (t_n-t_{n-1})\, S\left(I_m-\sum_{j=1}^{n-2} Y_j\right)). \qquad (2.20)
\end{aligned}$$

Substituting (2.20) into (2.18) and inserting the result into (2.16), we obtain an expression which we write in integral form,

$$\begin{aligned}
\mathbb{E}\left|R+\sum_{j=1}^{n} t_jX_j\right|^{-a} &=\frac{\Gamma_m(\sum_{j=1}^{n} a_j)}{\Gamma_m(a)\,\Gamma_m(a_{n-1}+a_n)}\int_{S>0} |S|^{a-(m+1)/2} \exp(-\operatorname{tr} S(R+t_nI_m))\\
&\quad\times\int_{0<\sum_{j=1}^{n-2}}\cdots\int_{Y_j<I_m}\left|I_m-\sum_{j=1}^{n-2} Y_j\right|^{a_{n-1}+a_n-(m+1)/2}\\
&\quad\times{}_1F_1\left(a_{n-1}; a_{n-1}+a_n; (t_n-t_{n-1})\, S\left(I_m-\sum_{j=1}^{n-2} Y_j\right)\right)\\
&\quad\times\prod_{j=1}^{n-2}\frac{1}{\Gamma_m(a_j)}\,|Y_j|^{a_j-(m+1)/2}\exp(\operatorname{tr}(t_n-t_j)\, SY_j)\, dY_j\, dS.
\end{aligned}\qquad (2.21)$$

Next we substitute $Y_j=S^{-1/2}W_jS^{-1/2}$, $j=1, \ldots, n-2$. Then the right-hand side of (2.21) becomes

$$\begin{aligned}
&\frac{\Gamma_m(\sum_{j=1}^{n} a_j)}{\Gamma_m(a)\,\Gamma_m(a_{n-1}+a_n)}\int_{S>0} |S|^{a-a_\bullet} \exp(-\operatorname{tr} S(R+t_nI_m))\\
&\quad\times\int_{0<\sum_{j=1}^{n-2}}\cdots\int_{W_j<S}\left|S-\sum_{j=1}^{n-2} W_j\right|^{a_{n-1}+a_n-(m+1)/2}\\
&\quad\times{}_1F_1\left(a_{n-1}; a_{n-1}+a_n; (t_n-t_{n-1})\left(S-\sum_{j=1}^{n-2} W_j\right)\right)\\
&\quad\times\prod_{j=1}^{n-2}\frac{1}{\Gamma_m(a_j)}\,|W_j|^{a_j-(m+1)/2}\exp(\operatorname{tr}(t_n-t_j)\, W_j)\, dW_j\, dS. \qquad (2.22)
\end{aligned}$$

Define the functions $\phi_1, ..., \phi_{n-1}$ on the space $\{S>0\}$ by

$$\phi_j(S) = \frac{1}{\Gamma_m(a_j)} |S|^{a_j-(m+1)/2} \exp(\operatorname{tr}(t_n - t_j) S) \tag{2.23}$$

for $j = 1, ..., n-2$, and

$$\phi_{n-1}(S) = \frac{1}{\Gamma_m(a_{n-1}+a_n)} |S|^{a_{n-1}+a_n-(m+1)/2} {}_1F_1(a_{n-1}; a_{n-1}+a_n; (t_n - t_{n-1}) S). \tag{2.24}$$

Denote by

$$(f_1 * f_2)(S) = \int_{0<W<S} f_1(W) f_2(S-W) \, dW$$

the convolution of two functions $f_1$ and $f_2$ on the space of positive-definite matrices. In analogy with the classical convolution on the real line, it is simple to deduce that the convolution of a collection of functions $f_1, ..., f_{n-1}$ is given by

$$(f_1 * \cdots * f_{n-1})(S) = \int_{0<\sum_{j=1}^{n-2}} \cdots \int_{W_j<S} f_{n-1}\left(S - \sum_{j=1}^{n-2} W_j\right) \prod_{j=1}^{n-2} f_j(W_j) \, dW_j.$$

Therefore, by (2.22), we obtain

$$\begin{aligned} \mathbb{E}\left|R + \sum_{j=1}^{n} t_j X_j\right|^{-a} &= \frac{\Gamma_m(\sum_{j=1}^{n} a_j)}{\Gamma_m(a)} \int_{S>0} |S|^{a-a_\bullet} \\ &\quad \times \exp\left(-\operatorname{tr} S(R + t_n I_m)\right) (\phi_1 * \cdots * \phi_{n-1})(S) \, dS, \end{aligned} \tag{2.25}$$

where $\phi_1, ..., \phi_{n-1}$ are defined in (2.23) and (2.24).

By (2.14),

$$\mathscr{W}_T^{-a+a_\bullet} \exp(-\operatorname{tr} ST) = |S|^{a-a_\bullet} \exp(-\operatorname{tr} ST). \tag{2.26}$$

Substituting (2.26) into the right-hand side of (2.25), we obtain

$$\begin{aligned} \mathbb{E}\left|R + \sum_{j=1}^{n} t_j X_j\right|^{-a} &= \frac{\Gamma_m(\sum_{j=1}^{n} a_j)}{\Gamma_m(a)} \mathscr{W}_T^{-a+a_\bullet} \\ &\quad \times \int_{S>0} \exp(-\operatorname{tr} ST)(\phi_1 * \cdots * \phi_{n-1})(S) \, dS \Big|_{T=R+t_n I_m} \\ &= \frac{\Gamma_m(\sum_{j=1}^{n} a_j)}{\Gamma_m(a)} \mathscr{W}_T^{-a+a_\bullet} \prod_{j=1}^{n-1} \hat{\phi}_j(T) \Big|_{T=R+t_n I_m}, \end{aligned} \tag{2.27}$$

where the second equality follows from the convolution theorem for the Laplace transform.

In order to simplify (2.27), we need to calculate the Laplace transforms $\hat{\phi}_j$, $j=1, ..., n-1$. In the case of the functions $\phi_1, ..., \phi_{n-2}$, we have

$$\hat{\phi}_j(T) = |T + (t_j - t_n)\, I_m|^{-a_j}, \tag{2.28}$$

$j = 1, ..., n-2$. In the case of $\phi_{n-1}$, the calculation of the Laplace transform can be obtained by means of a basic connection between the confluent (${}_1F_1$) and the Gaussian hypergeometric (${}_2F_1$) functions of matrix argument. Indeed,

$$\begin{aligned}\hat{\phi}_{n-1}(T) &= \frac{1}{\Gamma_m(a_{n-1}+a_n)} \int_{S>0} \exp(-\mathrm{tr}\, ST)\, |S|^{a_{n-1}+a_n-(m+1)/2} \\ &\quad \times {}_1F_1(a_{n-1};\, a_{n-1}+a_n;\, (t_n - t_{n-1})\, S)\, dS \\ &= |T|^{-(a_{n-1}+a_n)}\, {}_2F_1(a_{n-1},\, a_{n-1}+a_n;\, a_{n-1}+a_n;\, (t_n - t_{n-1})\, T^{-1}),\end{aligned}$$

where the last equality follows from Herz (1955, p. 485, Eq. (2.1)). By applying (2.12) we obtain

$$\begin{aligned}\hat{\phi}_{n-1}(T) &= |T|^{-(a_{n-1}+a_n)}\, |I - (t_n - t_{n-1})\, T^{-1}|^{-a_{n-1}} \\ &= |T|^{-a_n}\, |T + (t_{n-1} - t_n)\, I_m|^{-a_{n-1}}.\end{aligned} \tag{2.29}$$

Substituting (2.28) and (2.29) into (2.27) we obtain

$$\begin{aligned}\mathbb{E}\left|R + \sum_{j=1}^{n} t_j X_j\right|^{-a} &= \frac{\Gamma_m(\sum_{j=1}^n a_j)}{\Gamma_m(a)}\, \mathscr{W}_T^{-a+a.}\, |T|^{-a_n} \prod_{j=1}^{n-1} |T + (t_j - t_n)\, I_m|^{-a_j} \Bigg|_{T = R + t_n I_m} \\ &= \frac{\Gamma_m(\sum_{j=1}^n a_j)}{\Gamma_m(a)}\, \mathscr{W}_T^{-a+a.} \prod_{j=1}^{n} |T + (t_j - t_n)\, I_m|^{-a_j} \Bigg|_{T = R + t_n I_m} \\ &= \frac{\Gamma_m(\sum_{j=1}^n a_j)}{\Gamma_m(a)\, \Gamma_m(a. - a)} \int_{S>0} |S|^{a. - a - (m+1)/2} \prod_{j=1}^{n} |S + R + t_j I_m|^{-a_j}\, dS.\end{aligned} \tag{2.30}$$

The proof of (2.15) is now complete. ▮

2.3. *Remark.* (1) Some special cases can be obtained directly from (2.15). For example, if $t_1 = \cdots = t_n = 1$ then the left-hand side of (2.15) reduces to $|R + I_m|^{-a}$. The right-hand side can be calculated directly by making the transformation $S \to (R + I_m)^{-1/2}\, S (R + I_m)^{-1/2}$, and observing that the resulting integral is the normalizing constant for an inverted multivariate beta matrix distribution.

(2) Theorem 2.2 can be extended to cases in which the determinant on the left-hand side of (2.15) is replaced by certain products of powers of principal minors. In this setting, the proof remains similar to that given above. The details of the proof requires the hypergeometric functions and the fractional derivative operators of Gindikin (1964).

(3) The formula (2.15) is valid for $a < a_{\bullet}$. In addition, we may take limits as $a \to a_{\bullet}$. Since $\mathscr{W}_T^0$ is the identity operator then, by (2.30), we obtain

$$\mathbb{E}\left|R+\sum_{j=1}^{n} t_j X_j\right|^{-a_{\bullet}} = \prod_{j=1}^{n} |T+(t_j-t_n)\, I_m|^{-a_j}\Big|_{T=R+t_n I_m}$$
$$= \prod_{j=1}^{n} |R+t_j I_m|^{-a_j}.$$

This constitutes a new proof of equality of (1.1) and (1.2) for the case in which the matrices $T_1, ..., T_n$ are scalar.

(4) For the case in which $n=2$ and $R \to 0$, we obtain from (2.15) the result

$$\mathbb{E}\,|t_1 X_1 + t_2 X_2|^{-a} = \frac{\Gamma_m(a_1+a_2)}{\Gamma_m(a)\,\Gamma_m(a_1+a_2-a)} \times \int_{S>0} |S|^{a_1+a_2-a-(m+1)/2}\, |S+t_1 I_m|^{-a_1}\, |S+t_2 I_m|^{-a_2}\, dS. \tag{2.31}$$

Since the left hand side of (2.31) is homogeneous in $(t_1, t_2)$ we assume, without loss of generality, that $t_2 = 1$; now we can reduce (2.15) to a Gaussian hypergeometric function of matrix argument, as was shown by Letac *et al.* (2000). To do this, we make the transformation $S = \Lambda^{-1} - I_m$; then the right-hand side of (2.31) becomes

$$\frac{\Gamma_m(a_1+a_2)}{\Gamma_m(a)\,\Gamma_m(a_1+a_2-a)} \int_{0<\Lambda<I_m} |\Lambda|^{a-(m+1)/2} \times |I_m - \Lambda|^{a_1+a_2-a-(m+1)/2}\, |I_m-(1-t_1)\,\Lambda|^{-a}\, d\Lambda,$$

By Herz (1955, p. 489, Eq. (2.12)), this latter integral can be expressed in terms of the Gaussian hypergeometric function; then we obtain under the assumption $|1-t_1| < 1$ the result

$$\mathbb{E}\,|t_1 X_1 + X_2|^{-a} = {}_2F_1(a, a_1; a_1+a_2; (1-t_1)\, I_m).$$

If $|1-t_1| > 1$ then a similar result can be obtained by a symmetry argument. This formula was derived earlier by Letac *et al.* (2000, Eq. (5.11)).

## 3. CHARACTERIZATION RESULTS FOR MULTIVARIATE DIRICHLET DISTRIBUTIONS

In this section we establish the characterization results described at the end of the Introduction. As before, suppose $X_1, \ldots, X_n$ are $m \times m$ symmetric random matrices; $a_1, \ldots, a_n$ are fixed real numbers, $a_j > (m-1)/2$, $j = 1, \ldots, n$; and

$$\mathbb{E}\left| I_m - \sum_{j=1}^{n} T_j X_j \right|^{-a_\bullet} = \prod_{j=1}^{n} |I_m - T_j|^{-a_j} \tag{3.1}$$

for all symmetric $m \times m$ matrices $T_1, \ldots, T_n$ in a sufficiently small neighborhood of the zero matrix. Our purpose now is to characterize the distribution of $X_1, \ldots, X_n$ by means of (3.1).

Observe that the right-hand side of (3.1) is well defined only for $\|T_j\| < 1$, $j = 1, \ldots, n$. Similarly, we must also have $\|\sum_{j=1}^{n} T_j X_j\| < 1$, almost surely, for all $T_1, \ldots, T_n$ such that $\|T_j\| < 1$, $j = 1, \ldots, n$. Therefore (3.1) induces the assumption $\|\sum_{j=1}^{n} X_j\| \leqslant 1$, almost surely.

Denote by $O(m)$ the group of $m \times m$ orthogonal matrices. We observe that, for any $H \in O(m)$, the right-hand side of (3.1) is invariant under the transformation $T_j \to H T_j H'$, $j = 1, \ldots, n$. Therefore, to the extent that (3.1) characterizes the joint distribution of $X_1, \ldots, X_n$, it can do so only up to a similar invariance assumption on the joint distribution. Hence, we assume that the joint distribution of $X_1, \ldots, X_n$ is invariant under the transformation

$$(X_1, \ldots, X_n) \to H(X_1, \ldots, X_n)\, H', \tag{3.2}$$

for all $H \in O(m)$.

In the course of proving the main result, we shall need some basic results from the theory of zonal polynomials; cf. Muirhead (1982, p. 227 ff.). For completeness and ease of reference, we list them separately. Thus, a *partition* $\kappa = (k_1, \ldots, k_m)$ is an $m$-tuple of nonnegative integers such that $k_1 \geqslant \cdots \geqslant k_m \geqslant 0$. We denote by $|\kappa| := k_1 + \cdots + k_m$ the *length* of the partition $\kappa$.

For any symmetric $m \times m$ matrix $T$ the *zonal polynomial* $C_\kappa(T)$, corresponding to the partition $\kappa$, is a polynomial homogeneous of degree $|\kappa|$ and *orthogonally invariant*; i.e.,

$$C_\kappa(HTH') = C_\kappa(T) \tag{3.3}$$

for all $H \in O(m)$. The set of all zonal polynomials forms a basis for the vector space of all orthogonally invariant polynomials. Furthermore, by Muirhead (1982, p. 259), we have the series expansion

$$\exp(\operatorname{tr} T) = \sum_{\kappa} \frac{C_\kappa(T)}{|\kappa|!}, \tag{3.4}$$

where the sum is over the set of all partitions of all nonnegative integers; the series (3.4) converges absolutely for all symmetric $m \times m$ matrices $T$. If we denote by $dH$ the invariant Haar probability measure on the group $O(m)$ then, by Muirhead (1982, p. 243),

$$\int_{O(m)} C_\kappa(HT_1H'T_2)\, dH = \frac{C_\kappa(T_1)\, C_\kappa(T_2)}{C_\kappa(I_m)} \tag{3.5}$$

for any symmetric matrices $T_1, T_2$. It is well known that the zonal polynomials are characterized uniquely by the conditions (3.4) and (3.5).

For any $a \in \mathbb{R}$, the *partitional rising factorial* (or *generalized Pochhammer symbol*) corresponding to the partition $\kappa$ is

$$(a)_\kappa = \prod_{j=1}^{m} (a - \tfrac{1}{2}(j-1))_{k_j}, \tag{3.6}$$

where $(a)_k = a(a+1)\cdots(a+k-1)$ is the classical rising factorial. Similarly to the classical rising factorial, the partitional rising factorial in (3.6) arises in the binomial theorem for symmetric matrices,

$$|I_m - T|^{-a} = \sum_{\kappa} \frac{(a)_\kappa}{|\kappa|!} C_\kappa(T), \tag{3.7}$$

valid for $\|T\| < 1$; cf. Muirhead (1982, p. 259).

In the sequel we shall need a result involving a matrix of partial derivatives. For any $m \times m$ symmetric matrix $T = (t_{ij})$, let

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

be a symmetric matrix of partial derivatives, where $\delta_{ij}$ denotes Kronecker's delta. It is well known, and not difficult to verify, that if $P$ is any polynomial in $T$ then

$$P\left(\frac{\partial}{\partial T}\right) \exp(\operatorname{tr} VT) = P(V) \exp(\operatorname{tr} VT) \tag{3.8}$$

for all symmetric $m \times m$ matrices $V$. Of particular interest to us is the case in which $P(T) = |T|$, the determinant of $T$; in the case $m = 2$, it is known that the differential operator $|\partial/\partial T|$ is equivalent (after a one–one transformation) to the wave operator, $(\partial/\partial x)^2 - (\partial/\partial y)^2 - (\partial/\partial z)^2$ in the variables $x, y, z$.

The genesis of the following lemma goes back to Gårding (1947). Throughout, we denote by $(a)_{(1^m)}$ the partitional rising factorial in (3.6) corresponding to the partition $(1, ..., 1)$.

LEMMA 3.1. *Let $V$ and $X$ be symmetric $m \times m$ matrices, and $l$ be a nonnegative integer. Then*

$$\left|\frac{\partial}{\partial T}\right|^l |V - TX|^{-a} = \frac{\Gamma_m(a+l)}{\Gamma_m(a)} |V - TX|^{-(a+l)} |X|^l. \tag{3.9}$$

*Proof.* For $a > (m-1)/2$ and $V - TX > 0$, it follows by direct differentiation that

$$\left|\frac{\partial}{\partial T}\right|^l |V - TX|^{-a} = \frac{1}{\Gamma_m(a)} \left|\frac{\partial}{\partial T}\right|^l \int_{S>0} \exp(-\operatorname{tr} S(V - TX)) \, |S|^{a-(m+1)/2} \, dS. \tag{3.10}$$

By (3.8) we obtain

$$\left|\frac{\partial}{\partial T}\right|^l \exp(-\operatorname{tr} S(V - TX)) = |X|^l \, |S|^l \exp(-\operatorname{tr} S(V - TX)). \tag{3.11}$$

After interchanging integral and differential operator in (3.10) and applying (3.11) we see that the right-hand side of (3.10) reduces to

$$\frac{1}{\Gamma_m(a)} |X|^l \int_{S>0} \exp(-\operatorname{tr} S(V - TX)) \, |S|^{a+l-(m+1)/2} \, dS,$$

which is well known to equal the right-hand side of (3.9). This proves (3.9) for $a > (m-1)/2$ and $V - TX > 0$, and then the extension to all $V$, $X$, and $a$ follows by analytic continuation. ▌

Now we can state and prove our main result, a partial extension of Mauldon's characterization of the classical Dirichlet distributions.

3.2. THEOREM. *Suppose $X_1, ..., X_n$ satisfies* (3.1) *and* (3.2). *Then*

(i) $X_n + \cdots + X_n = I_m$, *almost surely;*

(ii) *The marginal distribution of $X_{k_1} + \cdots + X_{k_r}$, the sum of any proper subset of $X_1, ..., X_n$, is a multivariate beta distribution $\boldsymbol{\beta}(a_{k_1} + \cdots + a_{k_r}; a_{\bullet} - a_{k_1} - \cdots - a_{k_r})$;*

(iii) *The joint distribution of the determinants* $(|X_1|, \ldots, |X_n|)$ *is the same as the joint distribution of the determinants of a set of random matrices having a multivariate Dirichlet distribution* $D_m(a_1, \ldots, a_n)$*; and*

(iv) *For any* $k = 1, \ldots, m$*, the principal* $k \times k$ *submatrices of* $X_1, \ldots, X_n$ *satisfy analogs of* (i)–(iii).

*Proof.* (i) Denote $X_1 + \cdots + X_n$ by $X$. Substituting $T_1 = \cdots = T_n = T$ in (3.1), we obtain

$$\mathbb{E}\,|I_m - TX|^{-a_\bullet} = |I_m - T|^{-a_\bullet} \tag{3.12}$$

for all symmetric $T$ in a sufficiently small neighborhood of the zero matrix. Applying (3.7) to expand both sides of (3.12) in a zonal polynomial series and interchanging expectation and summation, we obtain

$$\begin{aligned}\sum_\kappa \frac{(a_\bullet)_\kappa}{|\kappa|!} C_\kappa(T) &= |I_m - T|^{-a_\bullet} \\ &= \mathbb{E}\,|I_m - TX|^{-a_\bullet} \\ &= \mathbb{E} \sum_\kappa \frac{(a_\bullet)_\kappa}{|\kappa|!} C_\kappa(TX) \\ &= \sum_\kappa \frac{(a_\bullet)_\kappa}{|\kappa|!} \mathbb{E} C_\kappa(TX). \end{aligned} \tag{3.13}$$

By (3.3) the left-hand side of (3.13) is invariant under the transformation $T \to HTH'$, $H \in O(m)$; therefore the right-hand side of (3.13) is also invariant under the same transformation. Thus, we replace $T$ on both sides of (3.13) by $HTH'$, and then average both sides with respect to the invariant Haar probability measure, $dH$, on the group $O(m)$. This produces the result

$$\sum_\kappa \frac{(a_\bullet)_\kappa}{|\kappa|!} C_\kappa(T) = \sum_\kappa \frac{(a_\bullet)_\kappa}{|\kappa|!} \mathbb{E} \int_{O(m)} C_\kappa(HTH'X)\, dH. \tag{3.14}$$

By (3.5), we find that (3.14) reduces to

$$\sum_\kappa \frac{(a_\bullet)_\kappa}{|\kappa|!} C_\kappa(T) = \sum_\kappa \frac{(a_\bullet)_\kappa}{|\kappa|!\, C_\kappa(I_m)} C_\kappa(T)\, \mathbb{E} C_\kappa(X). \tag{3.15}$$

Since the set of all zonal polynomials forms a basis for the vector space of orthogonally invariant polynomials then we may compare coefficients of $C_\kappa(T)$ on both sides of (3.15); hence we obtain

$$\mathbb{E} C_\kappa(X) = C_\kappa(I_m) \tag{3.16}$$

for all partitions $\kappa$.

By (3.4), we have

$$\mathbb{E}\exp(\operatorname{tr} TX) = \mathbb{E}\sum_{\kappa} \frac{C_\kappa(TX)}{|\kappa|!} = \sum_{\kappa} \frac{1}{|\kappa|!}\,\mathbb{E}C_\kappa(TX). \tag{3.17}$$

Since the distribution of $X$ is invariant under $O(m)$ then we can replace $T$ by $HTH'$ and average over $O(m)$ with respect to the Haar probability measure. Applying (3.5) and (3.16) to (3.17), we find that the moment-generating function of $X$ satisfies

$$\begin{aligned}\mathbb{E}\exp(\operatorname{tr} TX) &= \sum_{\kappa} \frac{1}{|\kappa|!\,C_\kappa(I_m)}\,C_\kappa(T)\,\mathbb{E}C_\kappa(X)\\ &= \sum_{\kappa} \frac{1}{|\kappa|!\,C_\kappa(I_m)}\,C_\kappa(T)\,C_\kappa(I_m)\\ &= \sum_{\kappa} \frac{C_\kappa(T)}{|\kappa|!}\\ &= \exp(\operatorname{tr} T),\end{aligned} \tag{3.18}$$

where the last equality follows from (3.4), Therefore $X = I_m$, almost surely.

(ii) In (3.1), we substitute $T_{k_1} = \cdots = T_{k_r} = T$ and set all other $T_j$ equal to 0, the zero matrix. Denoting $X_{k_1} + \cdots + X_{k_r}$ by $X$, it follows that (3.1) reduces to

$$\mathbb{E}\,|I_m - TX|^{-a_\bullet} = |I_m - T|^{-a}, \tag{3.19}$$

where $a = a_{k_1} + \cdots + a_{k_r}$. Similar to (3.13), we expand both sides of (3.19) in a series of zonal polynomials; then we obtain

$$\sum_{\kappa} \frac{(a_\bullet)_\kappa}{|\kappa|!}\,\mathbb{E}C_\kappa(TX) = \sum_{\kappa} \frac{(a)_\kappa}{|\kappa|!}\,C_\kappa(T).$$

Applying an invariance argument as in (3.14), averaging over the orthogonal group using (3.5), and comparing coefficients of $C_\kappa(T)$, we obtain

$$\mathbb{E}C_\kappa(X) = \frac{(a)_\kappa}{(a_\bullet)_\kappa}\,C_\kappa(I_m).$$

Proceeding as in (3.17) and (3.18), we deduce that the moment-generating function of $X$ is

$$
\begin{aligned}
\mathbb{E}\exp(\operatorname{tr} TX) &= \sum_{\kappa} \frac{1}{|\kappa|!\, C_\kappa(I_m)} C_\kappa(T)\, \mathbb{E}C_\kappa(X) \\
&= \sum_{\kappa} \frac{(a)_\kappa}{(a_\bullet)_\kappa} \frac{C_\kappa(T)}{|\kappa|!} \\
&= {}_1F_1(a; a_\bullet; T),
\end{aligned}
$$

where the last equality follows from the zonal polynomial series expansions for the hypergeometric functions of matrix argument (cf. Muirhead, 1982, p. 258). By the integral representation in (2.19) for the confluent hypergeometric function of matrix argument, it follows that

$$
\begin{aligned}
\mathbb{E}\exp(\operatorname{tr} TX) = {} & \frac{\Gamma_m(a_\bullet)}{\Gamma_m(a_\bullet - a)\, \Gamma_m(a)} \int_{0<Y<I_m} \exp(\operatorname{tr} TY) \\
& \times |Y|^{a-(m+1)/2}\, |I_m - Y|^{a_\bullet - a - (m+1)/2}\, dY.
\end{aligned} \tag{3.20}
$$

Once we observe that the right-hand side of (3.20) is the moment-generating function of the multivariate beta distribution, $\boldsymbol{\beta}(a; a_\bullet - a)$, we deduce that $X$ has the stated distribution.

(iii) Our strategy here is to apply integer powers of the partial differential operators $|\partial/\partial T_1|, \ldots, |\partial/\partial T_n|$ to both sides of (3.1) and then evaluate the results at $T_1 = \cdots = T_n = 0$. If $l_1, \ldots, l_n$ are arbitrary nonnegative integers, it follows from Lemma 3.1 that

$$
\begin{aligned}
\prod_{j=1}^{n} \left|\frac{\partial}{\partial T_j}\right|^{l_j} \cdot \prod_{j=1}^{n} |I_m - T_j|^{-a_j} &= \prod_{j=1}^{n} \left|\frac{\partial}{\partial T_j}\right|^{l_j} |I_m - T_j|^{-a_j} \\
&= \prod_{j=1}^{n} \frac{\Gamma_m(a_j + l_j)}{\Gamma_m(a_j)} |I_m - T_j|^{-(a_j + l_j)}.
\end{aligned} \tag{3.21}
$$

Next, by another application of Lemma 3.1 we have

$$
\begin{aligned}
\prod_{j=1}^{n} \left|\frac{\partial}{\partial T_j}\right|^{l_j} \cdot \left|I_m - \sum_{j=1}^{n} T_j X_j\right|^{-a_\bullet} &= \prod_{j=1}^{n-1} \left|\frac{\partial}{\partial T_j}\right|^{l_j} \left|\frac{\partial}{\partial T_m}\right|^{l_m} \left|I_m - \sum_{j=1}^{n-1} T_j X_j - T_n X_n\right|^{-a_\bullet} \\
&= \prod_{j=1}^{n-1} \left|\frac{\partial}{\partial T_j}\right|^{l_j} \frac{\Gamma_m(a_\bullet + l_n)}{\Gamma_m(a_\bullet)} \left|I_m - \sum_{j=1}^{n} T_j X_j\right|^{-(a_\bullet + l_n)} |X_n|^{l_n}.
\end{aligned}
$$

Repeating this process, we finally obtain

$$\prod_{j=1}^{n} \left|\frac{\partial}{\partial T_j}\right|^{l_j} \cdot \left|I_m - \sum_{j=1}^{n} T_j X_j\right|^{-a.}$$
$$= \frac{\Gamma_m(a. + l_1 + \cdots + l_n)}{\Gamma_m(a.)} \left|I_m - \sum_{j=1}^{n} T_j X_j\right|^{-(a.+l_1+\cdots+l_n)} \prod_{j=1}^{n} |X_j|^{l_j}. \tag{3.22}$$

Thus, when the operator $|\partial/\partial T_1|^{l_1} \cdots |\partial/\partial T_n|^{l_n}$ is applied to both sides of (3.1) and the outcome is evaluated at $T_1 = \cdots = T_n = 0$, it follows from (3.21) and (3.22) that

$$\mathbb{E} \prod_{j=1}^{n} |X_j|^{l_j} = \frac{\Gamma_m(a.)}{\Gamma_m(a. + l_1 + \cdots + l_n)} \prod_{j=1}^{n} \frac{\Gamma_m(a_j + l_j)}{\Gamma_m(a_j)}. \tag{3.23}$$

Since the random determinants $|X_1|, ..., |X_n|$ are bounded, their distribution is uniquely determined by their joint moments. It is straightforward to verify that the right-hand side of (3.23) is the joint moments of the determinants of a set of random matrices having the multivariate Dirichlet distribution $D_m(a_1, ..., a_n)$, and this completes the proof of (iii).

(iv) If $1 \leqslant r \leqslant n$ then we can verify that the $r \times r$ principal submatrices of $X_1, ..., X_n$ satisfy natural analogs of (i)–(iii) by writing each $T_j$ in (3.1) in block-decomposition form,

$$T_j = \begin{pmatrix} \tilde{T}_j & 0 \\ 0 & 0 \end{pmatrix},$$

where each $\tilde{T}_j$ is an $r \times r$ symmetric matrix. Then (3.1) reduces to a similar condition for the $r \times r$ principal submatrices of $X_1, ..., X_n$ and, by proceeding as we did before, we obtain analogs of (i)–(iii) for those submatrices. ∎

3.3. *Remark.* (i) As we observed earlier, it is a consequence of the previous result that for the case in which $n = 2$, the condition (3.1) characterizes the Dirichlet distribution $D_m(a_1, a_2)$; since this result is valid for all $m \geqslant 1$ we have then a partial extension of Mauldon's characterization of the classical Dirichlet distributions.

(ii) We can also represent the distribution of the determinants $(|X_1|, ..., |X_n|)$ in terms of the components of classical Dirichlet random vectors. To do this we apply to (3.23) the product formula (1.4) for the multivariate gamma function, thereby obtaining the result

$$\mathbb{E} \prod_{j=1}^{n} |X_j|^{l_j} = \prod_{k=1}^{m} \frac{\Gamma(a. - \frac{1}{2}(k-1))}{\Gamma(a. - \frac{1}{2}(k-1) + l_1 + \cdots + l_n)} \prod_{j=1}^{n} \frac{\Gamma(a_j - \frac{1}{2}(k-1) + l_j)}{\Gamma(a_j - \frac{1}{2}(k-1))}. \tag{3.24}$$

Now define $\mu_{j,k} = a_j - \frac{1}{2}(k-1)$ and $\nu_k = \frac{1}{2}(n-1)(k-1)$ for $j = 1, ..., n$, $k = 1, ..., m$. Further, for $k = 1, ..., m$ let $(U_{1,k}, ..., U_{n,k}, U_{n+1,k})$ be mutually independent Dirichlet random vectors, with $(U_{1,k}, ..., U_{n,k}, U_{n+1,k}) \sim D_1(\mu_{1,k}, ..., \mu_{n,k}, \nu_k)$. It is straightforward from (3.24) to verify that

$$\mathbb{E} \prod_{j=1}^{n} |X_j|^{l_j} = \prod_{k=1}^{m} \prod_{j=1}^{n} \mathbb{E} U_{j,k}^{l_j} = \mathbb{E} \prod_{j=1}^{n} \left( \prod_{k=1}^{m} U_{j,k} \right)^{l_j},$$

from which we conclude that

$$(|X_1|, ..., |X_n|) \stackrel{\mathrm{d}}{=} \left( \prod_{k=1}^{m} U_{1,k}, ..., \prod_{k=1}^{m} U_{n,k} \right).$$

In particular, for $j = 1, ..., n$, we have $|X_j| \stackrel{\mathrm{d}}{=} \prod_{k=1}^{m} U_{j,k}$.

(iii) Based on the results of Theorem 3.2, we find it natural to conjecture that the multivariate Dirichlet distributions are the only orthogonally invariant distributions which satisfy the condition (3.1). Further, we conjecture that there are many non-invariant distributions which satisfy (3.1).

## ACKNOWLEDGMENTS

The major portion of the research appearing in this article was carried out in summer, 1999, when one of the authors (D.R.) was a guest at the Vatican Observatory Summer School on Astronomy, held at Specola Vaticana (Vatican Observatory), Castel Gandolfo, Italy. D.R. expresses his deepest gratitude to Dr. George Coyne, S.J., Director of Specola Vaticana, and his colleagues for their gracious hospitality and collegiality during his stay.

## REFERENCES

J.-F. Chamayou and G. Letac, A transient random walk on stochastic matrices with Dirichlet distributions, *Ann. Probab.* **22** (1994), 424–430.

J. Faraut and A. Korányi, "Analysis on Symmetric Cones," Oxford Univ. Press, New York, 1994.

L. Gårding, The solution of Cauchy's problem for two totally hyperbolic linear differential equations by means of Riesz integrals, *Ann. of Math. (2)* **48** (1947), 785–826.

S. G. Gindikin, Analysis in homogeneous domains, *Uspekhi Mat. Nauk* **19** (1964), 3–92. [English translation: *Russian Math. Surveys* **19**, 1–90]

R. D. Gupta and D. St. P. Richards, Multivariate Liouville distributions, *J. Multivariate Anal.* **23** (1987), 233–256.

R. D. Gupta and D. St. P. Richards, Multivariate Liouville distributions, IV, *J. Multivariate Anal.* **54** (1995), 1–17.

R. D. Gupta and D. St. P. Richards, The history of the Dirichlet and Liouville distributions, *Internat. Statist. Rev.* (2001), to appear.

N. L. Johnson and S. Kotz, "Distributions in Statistics: Continuous Multivariate Distributions," Wiley, New York, 1972.

S. Karlin, C. A. Micchelli, and Y. Rinott, Multivariate splines: A probabilistic perspective, *J. Multivariate Anal.* **20** (1986), 69–90.

G. Letac and M. Scarsini, Random nested tetrahedra, *Adv. in Appl. Probab.* **30** (1998), 619–627.

G. Letac and H. Massam, A formula on multivariate Dirichlet distributions, *Statist. Probab. Lett.* **38** (1998), 247–253.

G. Letac, H. Massam, and D. Richards, An expectation formula for the multivariate Dirichlet distribution, *J. Multivariate Anal.* **77** (2001), 117–137.

J. G. Mauldon, A generalization of the beta-distribution, *Ann. Math. Statist.* **30** (1959), 509–520.

R. J. Muirhead, "Aspects of Multivariate Statistical Theory," Wiley, New York, 1982.

I. Olkin and H. Rubin, Multivariate beta distributions and independence properties of the Wishart distribution, *Ann. Math. Statist.* **35** (1964), 261–269.

N. A. Volodin, S. Kotz, and N. L. Johnson, Use of moments in distribution theory: A multivariate case, *J. Multivariate Anal.* **46** (1993), 112–119.

G. S. Watson, On the joint distribution of the circular serial correlation coefficients, *Biometrika* **43** (1956), 161–168.

*Printed in The Netherlands*