

# PROPERTIES OF THE INVERSE OF A NONCENTRAL WISHART MATRIX

GRANT HILLIER  
*University of Southampton*

RAYMOND KAN  
*University of Toronto*

The inverse of a noncentral Wishart matrix occurs in a variety of contexts in multivariate statistical work, including instrumental variable (IV) regression, but there has been very little work on its properties. In this paper, we first provide an expression for the expectation of the inverse of a noncentral Wishart matrix, and then go on to do the same for a number of scalar-valued functions of the inverse. The main result is obtained by exploiting simple but powerful group-equivariance properties of the expectation map involved. Subsequent results exploit the consequences of other invariance properties.

## 1. INTRODUCTION

Many inference problems in multivariate analysis, and in econometrics, involve various properties of the noncentral Wishart distribution. For example, all of the estimators for the parameters in structural equation models (Limited information maximum likelihood [LIML], Full information maximum likelihood [FIML], ordinary least squares [OLS], instrumental variable [IV], etc.) are functions of a noncentral Wishart matrix, and the distribution theory for them derives directly from this fact. See, for instance, Phillips (1983), and the many references therein. Recent work on partially identified models (Phillips, 1989) and weak instruments (e.g., Andrews and Stock, 2005) involves the same structure. In some applications, however, the relevant matrix is the *inverse* of a noncentral Wishart matrix, rather than the matrix itself, and this makes the problem considerably more difficult. That is the motivation for the present paper.

If the rows of the  $n \times m$  matrix  $Z$  are independent normal with covariance matrix  $\Sigma$ , and  $\mathbf{E}[Z] = M$ , then, if  $n \geq m$ , the density of  $W = Z'Z$  is given by (Muirhead, 1982, p. 442)

$$\begin{aligned} & \text{pdf}(W; \Sigma, \Lambda) \\ &= c_{m,n} |\Sigma|^{-\frac{n}{2}} \text{etr} \left( -\frac{\Sigma^{-1}(\Lambda + W)}{2} \right) |W|^{\frac{n}{2} - \frac{m+1}{2}} {}_0F_1 \left( \frac{n}{2}; \frac{\Sigma^{-1} \Lambda \Sigma^{-1} W}{4} \right), \end{aligned} \quad (1)$$

Thanks to the Editor, Co-Editor, and two referees for helpful comments on the first version of the paper. Address correspondence to Grant Hillier, Economics Division, School of Social Sciences, University of Southampton, Highfield, Southampton SO17 1BJ, UK; e-mail: ghh@soton.ac.uk.

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

with  $\Lambda = M'M$ ,<sup>1</sup>

$$c_{m,n} = \frac{1}{2^{\frac{nm}{2}} \Gamma_m\left(\frac{n}{2}\right)}, \tag{2}$$

where  $\text{etr}(A)$  denotes  $\exp(\text{tr}(A))$ ,  $\text{tr}(A)$  is the trace of  $A$ ,  $|A|$  is the determinant of  $A$ , and

$$\Gamma_m(t) = \pi^{\frac{m(m-1)}{4}} \prod_{i=1}^m \Gamma\left(t - \frac{i-1}{2}\right) \tag{3}$$

is the multivariate gamma function, with  $\Gamma(t)$  the univariate gamma function. Here and throughout, we use the usual notation for the hypergeometric functions of matrix argument (see Constantine, 1966, or Muirhead, 1982):

$${}_pF_q(a_1, \dots, a_p; c_1, \dots, c_q; A) = \sum_{j=0}^{\infty} \frac{1}{j!} \sum_{\alpha \vdash j} \frac{\prod_{i=1}^p (a_i)_{\alpha}}{\prod_{r=1}^q (c_r)_{\alpha}} C_{\alpha}(A). \tag{4}$$

Here,  $\alpha = (j_1, j_2, \dots, j_m)$  is a partition of  $j$  with at most  $m$  parts (i.e.,  $j_1 \geq j_2 \geq \dots \geq j_m \geq 0$ , and  $\sum_{s=1}^m j_s = j$ ), this being denoted by  $\alpha \vdash j$ . The numerical coefficients  $(c)_{\alpha}$  are defined by

$$(c)_{\alpha} = \prod_{s=1}^m \left(c - \frac{s-1}{2}\right)_{j_s}, \tag{5}$$

where  $(c)_r = c(c+1)\dots(c+r-1)$ . Finally,  $C_{\alpha}(\cdot)$  denotes the zonal polynomial in the elements of the indicated matrix corresponding to the partition  $\alpha \vdash j$  (see Muirhead, 1982, Chapter 7, or Macdonald, 1995). The important properties of these polynomials for our purposes are their invariance under orthogonal transformations, i.e.,  $C_{\alpha}(HAH') = C_{\alpha}(A)$ ,  $H \in \mathbb{O}(m)$ , where  $\mathbb{O}(m)$  is the group of  $m \times m$  orthogonal matrices, and the integral identity (Constantine, 1966, equation (5)):

$$\int_{\mathbb{O}(m)} C_{\alpha}(HAH'B) (dH) = \frac{C_{\alpha}(A)C_{\alpha}(B)}{C_{\alpha}(I_m)}. \tag{6}$$

Here,  $(dH)$  denotes the normalized invariant (Haar) measure on  $\mathbb{O}(m)$ . We will also use the averaged hypergeometric functions based on this integral, and defined by:

$${}_pF_q^{(m)}(a_1, \dots, a_p; c_1, \dots, c_q; A, B) = \int_{\mathbb{O}(m)} {}_pF_q(a_1, \dots, a_p; c_1, \dots, c_q; HAH'B) (dH) \tag{7}$$

(see Muirhead, 1982, p. 259).

<sup>1</sup>The noncentrality matrix is usually defined as  $\Omega = \Sigma^{-1}\Lambda$ . Here, we shall use the more symmetric notation. In addition,  $n$  in this definition need not be an integer: the expression given is a density for the positive definite symmetric matrix  $W$  for any real number  $n > m - 1$ .

We abbreviate the fact that  $W$  has this density by  $W \sim W_m(n, \Sigma, \Lambda)$ , or  $W \sim W_m(n, \Sigma)$  when  $\Lambda = 0$ . Expectations of functions of  $W$  when  $W$  has this density will be denoted by expressions like  $\mathbf{E}_{\Sigma, \Lambda}[\cdot]$ , with  $\Sigma$  omitted in the subscript when  $\Sigma = I_m$ , and  $\Lambda$  omitted in the central case  $\Lambda = 0$ . When  $\Lambda = 0$  and  $\Sigma = I_m$ , both subscripts are omitted. Analogous notation will also be used for other random variables.

In this paper, we examine some of the properties of the inverse of such a matrix  $W$ , beginning with the expectation of  $W^{-1}$ . For the problem of evaluating the expectation of  $W^{-1}$ , we may assume that  $\Sigma = I_m$ , so that the density of  $W$  depends only on  $\Lambda$ . For other functions of  $W^{-1}$  discussed later, this assumption is restrictive, but the more general case of arbitrary  $\Sigma$  is more complicated. So, later, we will discuss both of the cases  $\Sigma = I_m$  and  $\Sigma$  arbitrary. We first seek an expression for the mean,  $\mathbf{E}_{\Lambda}[W^{-1}]$ . In the central case where  $W \sim W_m(n, \Sigma)$ , the result is easily obtained (Muirhead, 1982, p. 97), and is:

$$\mathbf{E}_{\Sigma}[W^{-1}] = (n - m - 1)^{-1} \Sigma^{-1}. \tag{8}$$

In the (noncentral) scalar case ( $m = 1$ ), the result is also well known, since  $W \sim \chi_n^2(\lambda)$  (i.e., noncentral chi-square with noncentrality parameter  $\lambda$ ) and we have, for  $n > 2$ , (Krishnan, 1967)

$$\mathbf{E}[\chi_n^2(\lambda)^{-1}] = (n - 2)^{-1} {}_1F_1\left(1; \frac{n}{2}; -\frac{\lambda}{2}\right), \tag{9}$$

where  ${}_1F_1(a, b; z)$  is the confluent hypergeometric function. However, as far as we know, no explicit results are available for the noncentral case when  $m > 1$ , although some attempts on the problem have been made (Ullah, 1994; Letac and Massam, 2008). These known results are not explicit, but are expressed in terms of unresolved integrals and/or differential operators.<sup>2</sup> The expressions obtained below are, in contrast, explicit, and relatively simple.

After evaluating  $\mathbf{E}_{\Lambda}[W^{-1}]$  itself, we then go on to discuss the expectations of certain scalar functions of  $W^{-1}$ . Here, we confine attention to orthogonally invariant functions (i.e., functions  $f(W^{-1})$  invariant under  $W \rightarrow HWH'$ ,  $H \in \mathbb{O}(m)$ ). In particular, we evaluate the expectations of the elements of a particular basis for the vector space of such functions, the zonal polynomials. The expectations of functions such as  $(\text{tr}(W^{-1}))^k$  and  $\text{tr}(W^{-k})$  naturally have expansions in terms of expectations of zonal polynomials.

We denote by  $\mathbb{P}_m$  the space of  $m \times m$  symmetric matrices. The notation  $A > 0$  denotes the set of matrices  $A \in \mathbb{P}_m$  that are positive-definite symmetric. Short proofs are given in the text, and more elaborate ones are given in Appendixes A–C.

<sup>2</sup>The expression given by Ullah (1994) is deceptively simple. However, the differential operator, in fact, needs to be applied to an infinite series of zonal polynomials in the variables involved in the differential operator (equation (61)). More seriously, the action of the operator on zonal polynomials is completely unknown.

**1.1. Application: IV Regression**

In an IV regression *with known reduced form covariance matrix*, we have (after suitable standardizing transformations),<sup>3</sup> the system

$$y = X\beta + u, \tag{10}$$

$$X = Z\Pi + U, \tag{11}$$

with  $y$  ( $n \times 1$ ) independent of  $X$  ( $n \times k$ ) and  $Z$  ( $n \times p$ ) fixed (exogenous). The IV estimator for  $\beta$  is the OLS estimator

$$b = (X'X)^{-1}X'y, \tag{12}$$

and, assuming standard normality for the rows of  $(u, U)$ , we have that, conditionally, given  $X$ ,  $(b - \beta)|X \sim N(0, (X'X)^{-1})$ . Here,  $W = X'X \sim W_k(n, I_k, \Lambda)$ , with

$$\Lambda = \Pi'Z'Z\Pi. \tag{13}$$

Thus, the conditional covariance matrix of  $(b - \beta)$  is  $W^{-1}$ , and the unconditional covariance matrix its expectation.<sup>4</sup>

**1.2. A Useful Representation**

The IV regression example just discussed is useful for expository purposes. Thus, we introduce a vector  $x$  whose conditional distribution, given  $W$ , is  $N(0, W^{-1})$ . That is, the conditional covariance matrix of  $x$ , given  $W$ , is  $E_{x|W}[xx'] = W^{-1}$ , and its unconditional covariance matrix is the object of our desire. But, confining attention to the case  $\Sigma = I_m$ ,

$$E_{\Lambda}[W^{-1}] = E_{\Lambda}E_{x|W}[xx'] = E_{x,W}[xx'] = E_{\Lambda}[xx'], \tag{14}$$

where the final expectation is now with respect to the marginal density of  $x$ :

$$pdf(x; \Lambda) = \int_{W>0} pdf(x|W)pdf(W; \Lambda)(dW). \tag{15}$$

<sup>3</sup>Briefly, we start from an equation involving  $T \times (k + 1)$  variables  $(\tilde{y}, \tilde{Y})$ , say. The IV equation is  $\tilde{y} = \tilde{Y}\beta + u$ , and in the background is a reduced form for  $\tilde{Y}, \tilde{Y} = \tilde{Z}\Pi + U$ , with  $\tilde{Z} T \times p$ . Since we are assuming a known covariance matrix for the rows of  $(u, U)$ , we may assume that standardizing transformations have reduced this to an identity matrix. In addition, if the IV equation involves other variables, say  $Z_1$  (including a constant term), we assume these have been partialled out (i.e.,  $(\tilde{y}, \tilde{Y})$  are actually residuals after regression on  $Z_1$ ).

The IV estimators are of the form  $b = (\tilde{Y}'P\tilde{Y})^{-1}\tilde{Y}'P\tilde{y}$ , where  $P$  is an idempotent of rank  $n \geq k$ . It therefore has the form  $V(V'V)^{-1}V'$  for some choice of instruments  $V$  ( $n \times k$ ). Both  $n$  and  $V$  depend on the choice of estimator. Now, define  $X = (V'V)^{-\frac{1}{2}}V'\tilde{Y}, y = (V'V)^{-\frac{1}{2}}V'\tilde{y}$  ( $n \times k$  and  $n \times 1$ , respectively), and  $Z = (V'V)^{-\frac{1}{2}}V'\tilde{Z}$ . The estimator is then, as defined in the text,  $b = (X'X)^{-1}X'y$ . Further details can be found in Hillier, Kinal, and Srivastava (1984).

<sup>4</sup>The result we obtain below for this expectation, although reasonably complex, is considerably simpler than that given in Hillier et al. (1984) for the second moment matrix of the IV estimator. That is because we are here assuming that the covariance matrix of  $(u, U)$  is known (and so can be taken to be the identity). That assumption was not made in the earlier paper, and both the mean and covariance matrix of the conditional distribution of  $b$ , given  $X$ , there depend on  $X$ . This complicates matters considerably.

Using a standard result on Laplace transforms of zonal polynomials (see Constantine, 1963, equation (1), for example), the marginal density of  $x$  is easily seen to be

$$pdf(x; \Lambda) = \frac{\Gamma_m\left(\frac{n+1}{2}\right) \text{etr}\left(-\frac{\Lambda}{2}\right)}{\pi^{\frac{m}{2}} \Gamma_m\left(\frac{n}{2}\right)} \times (1+x'x)^{-\frac{n+1}{2}} {}_1F_1\left(\frac{n+1}{2}; \frac{n}{2}; \frac{1}{2}\Lambda(I_m+xx')^{-1}\right). \tag{16}$$

In the IV regression example, this equation (with  $m = k$  and  $\Lambda = \Pi'Z'Z\Pi$ ) is simply the unconditional density of the recentered IV estimator  $b - \beta$ .

We note, for latter use, the fact that, when evaluating the expectation of any function of  $x$  that is invariant under  $x \rightarrow Hx$ ,  $H \in \mathbb{O}(m)$ , the density (16) may be replaced by its average over  $\mathbb{O}(m)$  under the transformation  $x \rightarrow Hx$ ,  $H \in \mathbb{O}(m)$ . Using (6), this is:

$$pdf^{(m)}(x; \Lambda) = \frac{\Gamma_m\left(\frac{n+1}{2}\right) \text{etr}\left(-\frac{\Lambda}{2}\right)}{\pi^{\frac{m}{2}} \Gamma_m\left(\frac{n}{2}\right)} \times (1+x'x)^{-\frac{n+1}{2}} {}_1F_1^{(m)}\left(\frac{n+1}{2}; \frac{n}{2}; \frac{\Lambda}{2}, (I_m+xx')^{-1}\right). \tag{17}$$

An exactly analogous argument applies, of course, to the density of  $W$  itself (see equation (37) below).

**Remark 1.** This argument can easily be generalized to the case of an  $m \times r$  matrix  $X$  whose columns are independent  $N(0, W^{-1})$  vectors. See Section 4.2.2.

## 2. PRELIMINARIES

In the representation in terms of  $x$ , note that since the conditional mean  $\mathbf{E}_{x|W}[x]$  is zero, so is the unconditional mean. The required covariance matrix  $\mathbf{E}_\Lambda[W^{-1}] = \mathbf{E}_\Lambda[xx']$  will be a matrix-valued function of  $\Lambda$ , say  $\Phi(\Lambda)$ . The expectation operator  $\mathbf{E}_\Lambda[W^{-1}] = \mathbf{E}_\Lambda[xx']$  thus defines a mapping from the space of symmetric matrices,  $\mathbb{P}_m$ , to itself. This map has the following simple but important property.

LEMMA 1. *The map  $\Lambda \rightarrow \Phi(\Lambda) : \mathbb{P}_m \rightarrow \mathbb{P}_m$  defined by  $\Phi(\Lambda) = \mathbf{E}_\Lambda[W^{-1}] = \mathbf{E}_\Lambda[xx']$  is equivariant under the conjugate action of the orthogonal group  $\mathbb{O}(m)$  on  $\Lambda$ . That is, for all  $H \in \mathbb{O}(m)$ ,*

$$\Phi(H\Lambda H') = H\Phi(\Lambda)H'. \tag{18}$$

**Proof.** On making the transformation  $\Lambda \rightarrow H\Lambda H'$ , in the integral defining  $\Phi$ , we may then transform  $x \rightarrow \tilde{x} = H'x$ , leaving the density of  $\tilde{x}$  as was. Thus,  $\Phi(H\Lambda H') = \mathbf{E}_{H\Lambda H'}[xx'] = \mathbf{E}_\Lambda[H\tilde{x}\tilde{x}'H'] = H\Phi(\Lambda)H'$ . □

Lemma 1 has several valuable implications for our problem, gathered in the following corollary.

COROLLARY 1.

(i) If  $\Lambda = LDL'$ ,  $L \in \mathbb{O}(m)$ ,  $D = \text{diag}(\omega_1, \dots, \omega_m)$ , is the spectral decomposition of  $\Lambda$ , then

$$\Phi(\Lambda) = \Phi(LDL') = L\Phi(D)L'. \tag{19}$$

(ii) If  $\Lambda = \alpha I_m$ , then  $\Phi(\Lambda) = \bar{\psi}(\alpha)I_m$ , for some scalar function  $\bar{\psi}(\alpha)$  of  $\alpha$ .  
 (iii) For  $D$  diagonal,  $\Phi(D)$  is also diagonal.

**Proof.** (i) is just the equivariance property. (ii) Equivariance implies that  $\Phi(\alpha I_m) = \Phi(\alpha H I_m H') = H\Phi(\alpha I_m)H'$ , for all  $H \in \mathbb{O}(m)$ , which implies that  $\Phi(\alpha I_m)$  is proportional to an identity matrix. (iii) If  $\mathbb{T}$  is the subgroup of  $\mathbb{O}(m)$  consisting of the  $2^m$  diagonal matrices  $T = \text{diag}\{\pm 1, \pm 1, \dots, \pm 1\}$  (i.e., the group of sign changes), it is easy to see that  $A \in \mathbb{P}_m$  is invariant under the action of  $\mathbb{T}$  if and only if  $A$  is diagonal. But, by equivariance, for every  $T$  in this subgroup,  $\Phi(D) = \Phi(TDT') = T\Phi(D)T'$ , which implies that  $\Phi(D)$  is diagonal.  $\square$

Translated into conclusions about the problem of interest, these results say: (i) only  $\text{pdf}(x; D)$ , with  $D$  the diagonal matrix of eigenvalues of  $\Lambda$ , needs to be considered:  $\Phi(\Lambda) = \mathbf{E}_\Lambda[xx']$  can be recovered from  $\mathbf{E}_D[xx']$ . (ii) The matrices  $\Lambda$  and  $\Phi(\Lambda)$  have common eigenvectors. (iii) When  $\Lambda$  is spherical (proportional to the identity matrix), so is  $\Phi(\Lambda)$ , and (iv) variables with density  $\text{pdf}(x; D)$ ,  $D$  diagonal, are uncorrelated (although manifestly not independent). Thus, only the variances of the  $x_i$  are needed to determine  $\Phi(D)$ . When  $D = \alpha I_m$ ,  $\text{tr}(\Phi(D)) = \mathbf{E}_\Lambda[x'x] = m\bar{\psi}(\alpha)$ , so, in this case, only the expectation  $\mathbf{E}_\Lambda[x'x]$  is needed.

The fact that  $\Phi(D)$  is diagonal, coupled with its equivariance, means that the  $m$  diagonal elements of  $\Phi(D)$ , together with the eigenvectors of  $\Lambda$  in  $L$ , completely determine  $\Phi(\Lambda)$ . However, even more can be said about these diagonal elements, as follows. Let  $P_r, r = 2, \dots, m$ , denote the permutation (transposition) matrix that interchanges the first and  $r$ th elements of a vector in  $\mathbb{R}^m$ . Note that the  $P_r$  are elements of  $\mathbb{O}(m)$ . We have the following lemma.

LEMMA 2. Let  $\psi_i(D), i = 1, \dots, m$ , denote the diagonal elements of  $\Psi = \Phi(D)$ . Then, for each  $r = 2, \dots, m$ ,  $\psi_r(D) = \psi_1(P_r D P_r')$ . That is, each of the  $\psi_r(D)$  can be obtained from  $\psi_1(D)$  by simply exchanging the roles of  $\omega_1$  and  $\omega_r, r = 2, \dots, m$ .

**Proof.** By Lemma 1,  $\Phi(P_r D P_r') = P_r \Phi(D) P_r'$ , for each  $r = 2, \dots, m$ . The (1, 1) element of the matrix on the left is  $\psi_1(P_r D P_r')$ . On the right, the (1, 1) element is  $\psi_r(D)$ , establishing the result.<sup>5</sup>  $\square$

<sup>5</sup>In fact, for every permutation matrix  $P_\sigma$ , with  $\sigma \in S_m$  (the symmetric group on  $m$  objects), we have  $\Phi(P_\sigma D P_\sigma') = P_\sigma \Phi(D) P_\sigma'$ . This imposes more structure on  $\Phi(D)$ , but the result in Lemma 2 is sufficient for our purposes.

This result means that we need only determine one of the elements  $\psi_i(D)$  of  $\Psi = \Phi(D)$ , say  $\psi_1(D)$ , since this determines all the remaining elements. We will see the properties described in Lemma 2 directly in Section 3 (Theorem 1).

**2.1. A Useful Lemma**

We will make use of the following result.

LEMMA 3. *Let  $\rho$  be a partition of an integer  $r > 0$ ,  $C_\rho(\cdot)$  the zonal polynomial associated with  $\rho$ , and  $n > m - 1$  a positive real number. We have*

$$\int_{\mathbb{R}^m} (1 + x'x)^{-\frac{n+1}{2}} C_\rho((I_m + xx')^{-1})(dx) = \frac{\pi^{\frac{m}{2}} \Gamma(\frac{n-m+1}{2})}{\Gamma(\frac{n+1}{2})} \frac{(\frac{n}{2})_\rho}{(\frac{n+1}{2})_\rho} C_\rho(I_m), \tag{20}$$

and

$$\int_0^\infty q^{\frac{m}{2}-1} (1+q)^{-\frac{n+1}{2}} C_\rho((I_m + qe_1e_1')^{-1}) dq = B\left(\frac{m}{2}, \frac{n-m+1}{2}\right) \frac{(\frac{n}{2})_\rho}{(\frac{n+1}{2})_\rho} C_\rho(I_m), \tag{21}$$

where  $B(a, c) = \Gamma(a)\Gamma(c) / \Gamma(a + c)$  is the beta coefficient.

The proof is in Appendix A. Using Lemma 3, the following corollary, which will be useful later, is easily established.

COROLLARY 2. *For  $x$  with density (16) or (17), and any  $s < \frac{n-m+1}{2}$ ,*

$$\begin{aligned} \mathbf{E}_\Lambda[(1 + x'x)^s] &= \frac{\Gamma(\frac{n+1}{2}) \Gamma(\frac{n-m+1}{2} - s)}{\Gamma(\frac{n-m+1}{2}) \Gamma(\frac{n+1}{2} - s)} \\ &\times \text{etr}\left(-\frac{\Lambda}{2}\right) {}_2F_2\left(\frac{n+1}{2}, \frac{n}{2} - s; \frac{n}{2}, \frac{n+1}{2} - s; \frac{\Lambda}{2}\right). \end{aligned} \tag{22}$$

**Proof.** Multiply (17) by  $(1 + x'x)^s$  and integrate using Lemma 3 with  $n$  replaced by  $n - 2s$ . This produces (22). Since  $x'x$  is invariant under transformations  $x \rightarrow Hx, H \in \mathbb{O}(m)$ , (22) also holds when  $x$  has density (16). □

**3. RESULTS FOR  $E[W^{-1}]$**

**3.1. Special Cases**

In the central case ( $\Lambda = 0$ ), we have already noted the result in equation (8). In the spherical case with  $\Lambda = \alpha I_m, \Sigma = I_m$ , part (ii) of Corollary 1 says that we only need to find  $\tilde{\psi}(\alpha)$ , and this can be deduced from the fact that  $\text{tr}(\Phi) = \mathbf{E}_\Lambda[x'x] = m\tilde{\psi}(\alpha)$ . Thus, we need simply to evaluate  $\mathbf{E}_\Lambda[x'x]$  for this case. The result is easily obtained as an application of Corollary 2.

PROPOSITION 1. When  $\Lambda = \alpha I_m$ , we have  $\Phi(\alpha I_m) = \bar{\psi}(\alpha)I_m$ , with

$$\bar{\psi}(\alpha) = \frac{1}{m} \left[ \left( \frac{n-1}{n-m-1} \right) e^{-\frac{m\alpha}{2}} {}_2F_2 \left( \frac{n+1}{2}, \frac{n-2}{2}; \frac{n}{2}, \frac{n-1}{2}; \frac{\alpha}{2} I_m \right) - 1 \right]. \tag{23}$$

**Proof.** Using Corollary 2, we have, when  $\Lambda = \alpha I_m$ ,

$$\mathbf{E}_\Lambda[1 + x'x] = \left( \frac{n-1}{n-m-1} \right) e^{-\frac{m\alpha}{2}} {}_2F_2 \left( \frac{n+1}{2}, \frac{n-2}{2}; \frac{n}{2}, \frac{n-1}{2}; \frac{\alpha}{2} I_m \right). \tag{24}$$

This gives the result stated. □

When  $\Lambda$  is of rank  $r < m$ , it can easily be shown that diagonal elements of  $\Phi(D)$  corresponding to zero eigenvalues of  $\Lambda$  are all equal to  $(n - m - 1)^{-1}$ . The remaining elements will follow from the main result, the proof of which applies whatever the rank of  $\Lambda$ .

### 3.2. Main Result

According to Lemma 2, we need only consider the case  $\Lambda = D$ , and also need only evaluate  $\psi_1(D) = \text{Var}[x_1]$ , the remaining terms being obtained by transpositions  $(1, r)$  of the diagonal elements of  $D$ . We prove the following theorem.

THEOREM 1. If  $W \sim W_m(n, I_m, \Lambda)$ , and  $\Lambda = LDL'$ ,  $L \in \mathbb{O}(m)$ ,  $D = \text{diag}(\omega_1, \dots, \omega_m)$ , is the spectral decomposition of  $\Lambda$ , then  $\mathbf{E}_\Lambda[W^{-1}] = \Phi(\Lambda) = L\Psi L'$ , with  $\Psi = \text{diag}(\psi_i(D))$ . When  $n > m + 1$ , the elements on the diagonal of  $\Psi = \Phi(D)$  are given by

$$\psi_i(D) = \frac{\text{etr}\left(-\frac{D_i}{2}\right)}{n-m-1} \sum_{k=0}^{\infty} \frac{\left(-\frac{\omega_i}{2}\right)^k}{\left(\frac{n}{2}\right)_k} {}_2F_2 \left( \frac{n}{2}, \frac{n-1}{2} + k; \frac{n-1}{2}, \frac{n}{2} + k; \frac{D_i}{2} \right), \tag{25}$$

$i = 1, \dots, m$ , where  $D_i$  is  $D$  with its  $i$ th row and column removed. The  $\psi_i(D)$  are the eigenvalues of  $\Phi(\Lambda)$ , and  $\Phi(\Lambda) = L\Psi L'$  is the spectral decomposition of  $\Phi(\Lambda)$ .

The proof can be found in Appendix B. It is straightforward to check that this result agrees with the known result for the case  $m = 1$ , i.e., the case  $\mathbf{E}[(x_n^2(\omega))^{-1}]$ . As expected (because  $\Phi(\Lambda)$  is the (unconditional) covariance matrix of  $x$ , and  $W^{-1} > 0$  almost surely), we have the following corollary.

COROLLARY 3. The matrix  $\Phi(\Lambda) = \mathbf{E}_\Lambda[W^{-1}]$  is positive definite whatever the rank of  $\Lambda$ . The  $\psi_i(D)$  corresponding to zero eigenvalues of  $\Lambda$  are  $(n - m - 1)^{-1}$ .

**Remark 2.** When  $W \sim W_m(n, \Sigma, \Lambda)$ ,  $\tilde{W} = \Sigma^{-\frac{1}{2}} W \Sigma^{-\frac{1}{2}} \sim W_m(n, I_m, \Sigma^{-\frac{1}{2}} \Lambda \Sigma^{-\frac{1}{2}})$ . Thus, for the case  $W \sim W_m(n, \Sigma, \Lambda)$ , we have

$$\mathbf{E}_{\Sigma, \Lambda}[W^{-1}] = \Sigma^{-\frac{1}{2}} \mathbf{E}_{\Lambda_1}[\tilde{W}^{-1}] \Sigma^{-\frac{1}{2}}, \tag{26}$$



where  $\Lambda_1 = \Sigma^{-\frac{1}{2}} \Lambda \Sigma^{-\frac{1}{2}}$ , with  $E_{\Lambda_1}[\tilde{W}^{-1}]$  as given in [Theorem 1](#) with  $\Lambda$  replaced by  $\Lambda_1$ .

**Remark 3.** Note that, for  $n$  large, the term  $\mu_k(D_1)$  as defined in the proof of [Theorem 1](#) is approximately equal to  $\frac{\binom{n-m+1}{2}}{k} / \binom{n}{2}_k$ , so that

$$\text{Var}[x_1] \simeq \frac{1}{n-m-1} \sum_{k=0}^{\infty} \frac{\left(-\frac{\omega_1}{2}\right)^k}{\binom{n}{2}_k} = \frac{1}{n-m-1} {}_1F_1\left(1; \frac{n}{2}; -\frac{\omega_1}{2}\right). \tag{27}$$

Thus, for  $n$  large

$$\Psi \simeq \frac{1}{n-m-1} \text{diag} \left\{ {}_1F_1\left(1, \frac{n}{2}; -\frac{\omega_i}{2}\right) \right\}, \tag{28}$$

with diagonal elements approximately equal to

$$\psi_i(D) \simeq \frac{n-2}{n-m-1} \mathbf{E}[\chi_n^2(\omega_i)^{-1}], \quad i = 1, \dots, m. \tag{29}$$

**Remark 4.** [Theorem 1](#) evidently enables the evaluation of the expectation of any linear function of the elements of  $W^{-1}$ . For instance, for any fixed matrix  $A$ ,

$$E_{\Lambda} [\text{tr}(AW^{-1})] = \text{tr}(AE_{\Lambda}[W^{-1}]) = \text{tr}(AL\Psi L'). \tag{30}$$

### 3.3. Estimation of the Precision Matrix $\Sigma^{-1}$

In the central case, equation (8) means that  $(n-m-1)W^{-1}$  is an unbiased estimator of the precision matrix  $\Sigma^{-1}$ . In the noncentral case, this is no longer the case. Instead, as noted above,

$$E_{\Sigma, \Lambda} [(n-m-1)W^{-1}] = (n-m-1)\Sigma^{-\frac{1}{2}} \Phi(\Lambda_1)\Sigma^{-\frac{1}{2}} \tag{31}$$

$$= \Sigma^{-1} - \Sigma^{-\frac{1}{2}} [I_m - (n-m-1)\Phi(\Lambda_1)] \Sigma^{-\frac{1}{2}}. \tag{32}$$

To the extent that  $(n-m-1)\Phi(\Lambda_1)$  differs from an identity matrix, the estimator is biased by the second factor in this expression. In the case  $\Lambda_1 = \alpha I_m$ , for instance,  $I_m - (n-m-1)\Phi(\Lambda_1) = (1 - (n-m-1)\bar{\psi}(\alpha))I_m$  and

$$E_{\Sigma, \Lambda} [(n-m-1)W^{-1}] = (n-m-1)\bar{\psi}(\alpha)\Sigma^{-1}, \tag{33}$$

with  $\bar{\psi}(\alpha)$  as in equation (23). For  $n$  large,  $(n-m-1)\bar{\psi}(\alpha)$  is near 1, but the estimator is, in general, biased for moderate  $n$  even in this simple case.

## 4. EXPECTATION OF THE ZONAL POLYNOMIALS $C_{\kappa}(W^{-1})$

The zonal polynomials  $C_{\kappa}(A)$ ,  $\kappa \vdash k$ , form a basis for the space of orthogonally invariant homogeneous polynomials of degree  $k$  in the elements of a symmetric matrix  $A$  (see, for instance, Macdonald, 1995). That is, any scalar function of  $W^{-1}$  that is homogeneous and invariant under  $W \rightarrow HWH'$ ,  $H \in \mathbb{O}(m)$  can, in principle,

be expressed as a linear combination of the corresponding zonal polynomials. Scalar functions of  $W^{-1}$  of potential interest in this class include the following three cases:<sup>6</sup>

powers of the determinant,

$$\delta_k(\Sigma, \Lambda) = \mathbf{E}_{\Sigma, \Lambda} [|W^{-1}|^k], \quad k = 1, 2, \dots, \tag{34}$$

powers of the trace,

$$\tau_k(\Sigma, \Lambda) = \mathbf{E}_{\Sigma, \Lambda} [(\text{tr}(W^{-1}))^k], \quad k = 1, 2, \dots, \tag{35}$$

and the power sums

$$\pi_k(\Sigma, \Lambda) = \mathbf{E}_{\Sigma, \Lambda} [\text{tr}(W^{-k})], \quad k = 1, 2, \dots \tag{36}$$

Other examples include  $f_\nu(W) = \prod_{i=1}^s \text{tr}(W^{-i})^{\nu_i}$ , with  $\nu_1 + 2\nu_2 + \dots + s\nu_s = k$ , which can be written as linear combination of the  $C_\kappa(W^{-1})$ , for  $\kappa \vdash k$ . The cases  $(\text{tr}(W^{-1}))^k = \sum_{\kappa \vdash k} C_\kappa(W^{-1})$  and  $\text{tr}(W^{-k})$  (i.e.,  $\tau_k$  and  $\pi_k$ ) are both in this class, with  $s = 1$  and  $\nu_1 = k$  in the first case, and  $s = k, \nu_i = 0$ , for  $i = 1, \dots, k - 1$ , and  $\nu_k = 1$  in the second case. Hence, if the expectations of the  $C_\kappa(W^{-1})$ , for all  $\kappa \vdash k$ , are known, so are the expectations of all other functions of  $W^{-1}$  in this class. This fact is the motivation for this section. The computation of the functions appearing in the results in this section is discussed in Section 5.

For orthogonally invariant functions  $f(W)$  (including the zonal polynomials), the following results hold.

LEMMA 4. *Let  $f(W)$  satisfy  $f(HWH') = f(W)$ , for all  $H \in \mathbb{O}(m)$ . Then,*

1. *In the case  $\Sigma = I_m$ ,  $\mathbf{E}_\Lambda[f(W)]$  is invariant under  $\Lambda \rightarrow H\Lambda H'$ ,  $H \in \mathbb{O}(m)$ , and, when  $\Lambda = 0$ ,  $\mathbf{E}_\Sigma[f(W)]$  is invariant under  $\Sigma \rightarrow H\Sigma H'$ ,  $H \in \mathbb{O}(m)$ .*
2. *In the case of general  $\Sigma \in \mathbb{P}_m$ , and  $\Lambda \neq 0$ ,  $\mathbf{E}_{\Sigma, \Lambda}[f(W)]$  is invariant under the simultaneous transformations  $\Lambda \rightarrow H\Lambda H'$ ,  $\Sigma \rightarrow H\Sigma H'$ ,  $H \in \mathbb{O}(m)$ .*

The proofs are trivial and omitted. Evidently, in the general case, the expectations  $\mathbf{E}_{\Sigma, \Lambda}[f(W)]$  will involve a class of functions of  $(\Sigma, \Lambda)$  more general than symmetric polynomials in either argument. We will see this directly below. However, when  $\Sigma = I_m$ , the expectations of functions in the class referred to in the lemma depend only on the eigenvalues of  $\Lambda$ , i.e.,  $D$ , and may be computed from the averaged density of  $W$ ,

$$pdf^{(m)}(W; \Lambda) = \int_{\mathbb{O}(m)} pdf(W; H\Lambda H')(dH)$$

<sup>6</sup>Of course, for  $k = 1$ ,

$$\tau_1(\Sigma, \Lambda) = \pi_1(\Sigma, \Lambda) = \mathbf{E}_{\Sigma, \Lambda} [\text{tr}(W^{-1})],$$

and the result follows immediately from Corollary 2 (with  $s = 1$ ) when  $\Sigma = I_m$ , since  $\mathbf{E}_\Lambda[\text{tr}(W^{-1})] = \text{tr}(\Phi(\Lambda)) = \mathbf{E}_\Lambda[x'x]$ .

$$= c_{m,n} \text{etr} \left( -\frac{\Lambda + W}{2} \right) |W|^{\frac{n}{2} - \frac{m+1}{2}} {}_0F_1^{(m)} \left( \frac{n}{2}; \frac{\Lambda}{2}, \frac{W}{2} \right). \tag{37}$$

We begin with the simplest case, the top-order zonal polynomials, with  $\Sigma = I_m$ .

### 4.1. Expectations of Top-Order Zonal Polynomials

The case of top-order zonal polynomials can be dealt with by exploiting the conditioning argument explained in Section 1.2.<sup>7</sup> When  $x|W \sim N(0, W^{-1})$ , it is straightforward to see (using the moment generating function of  $q = x'x$ ) that the conditional moments of  $q$  are given by

$$\mathbf{E}_{x|W}[q^k|W] = 2^k \left( \frac{1}{2} \right)_k C_k(W^{-1}), \tag{38}$$

and therefore,

$$\mathbf{E}_\Lambda[C_k(W^{-1})] = \frac{1}{2^k \left( \frac{1}{2} \right)_k} \mathbf{E}_\Lambda[q^k], \tag{39}$$

with the latter expectation evaluated in the density given in equation (17).

We may evaluate the expectations  $\mathbf{E}_\Lambda[(1 + q)^k]$  first, using Corollary 2, then use the (binomial) expansion

$$q^k = \sum_{l=0}^k \binom{k}{l} (-1)^l (1 + q)^{k-l} \tag{40}$$

to recover  $\mathbf{E}_\Lambda[q^k]$ . Defining  $\varphi_k(\Lambda) := \mathbf{E}_\Lambda[(1 + q)^k]$ ,  $\varphi_0 = 1$ , we will then have

$$\mathbf{E}_\Lambda[q^k] = \sum_{l=0}^k \binom{k}{l} (-1)^l \varphi_{k-l}(\Lambda).$$

Thus, directly from Corollary 2, we have the following theorem.

**THEOREM 2.** For  $W \sim W_m(n, I_m, \Lambda)$  and  $k < \frac{n-m+1}{2}$ , the expectation of the top-order zonal polynomial  $C_k(W^{-1})$  is given by

$$\mathbf{E}_\Lambda[C_k(W^{-1})] = \frac{1}{2^k \left( \frac{1}{2} \right)_k} \sum_{l=0}^k \binom{k}{l} (-1)^l \varphi_{k-l}(\Lambda), \tag{41}$$

<sup>7</sup>The conditioning argument used here can evidently be applied to any function of  $x$ , say  $g(x)$ , which, when  $x \sim N(0, \Sigma)$ , has known expectation  $G(\Sigma)$ . In particular, for any  $g(\cdot)$  that is orthogonally invariant ( $g(Hx) = g(x) \forall H \in \mathbb{O}(m)$ ),  $g$  depends on  $x$  only through  $q = x'x, g(x) = \bar{g}(q)$  (since  $q = x'x$  is a maximal invariant under this action). In such cases, we will then have

$$\mathbf{E}_\Lambda[G(W^{-1})] = \mathbf{E}_\Lambda[\bar{g}(q)].$$

with the expectation on the right evaluated in the density in equation (17). The usefulness of this fact depends, of course, on whether the expectation of the function  $G(W^{-1})$  is itself of interest.

where, for  $r < \frac{n-m+1}{2}$ ,

$$\varphi_r(\Lambda) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n-m+1}{2}\right)} \frac{\Gamma\left(\frac{n-m+1}{2} - r\right)}{\Gamma\left(\frac{n+1}{2} - r\right)} \times \text{etr}\left(-\frac{\Lambda}{2}\right) {}_2F_2\left(\frac{n}{2} - r, \frac{n+1}{2}; \frac{n+1}{2} - r, \frac{n}{2}; \frac{\Lambda}{2}\right). \tag{42}$$

### 4.2. General Zonal Polynomials

4.2.1. *The Central Case.* Before tackling the noncentral case, we record here the results for the central case, for ease of reference later. The following result for  $C_\kappa(W)$  is standard:

$$\mathbf{E}_\Sigma[C_\kappa(W)] = 2^k \left(\frac{n}{2}\right)_\kappa C_\kappa(\Sigma). \tag{43}$$

For  $C_\kappa(W^{-1})$ , the result is essentially given by Constantine (1966, equation 10) and (in different notation) Khatri (1966, Lemma 4), which we repeat here for completeness.<sup>8</sup>

LEMMA 5. For  $Z$ , a complex symmetric matrix with  $\text{Re}[Z] > 0$ ,  $T$  complex symmetric, and  $t > k_1 + \frac{m-1}{2}$ ,

$$\int_{W>0} \text{etr}(-ZW) |W|^{t-\frac{m+1}{2}} C_\kappa(TW^{-1})(dW) = \Gamma_m(t, -\kappa) |Z|^{-t} C_\kappa(TZ), \tag{44}$$

where, for  $\kappa = (k_1, k_2, \dots, k_m) \vdash k$ ,

$$\Gamma_m(t, -\kappa) = \pi^{\frac{m(m-1)}{4}} \prod_{i=1}^m \Gamma\left(t - k_i - \frac{m-i}{2}\right). \tag{45}$$

Hence, for  $\kappa \vdash k$ , with  $k < \frac{n-m+1}{2}$ , and  $W \sim W_m(n, \Sigma)$ , we have (taking  $T = I_m$ ),

$$\mathbf{E}_\Sigma[C_\kappa(W^{-1})] = \frac{\Gamma_m\left(\frac{n}{2}, -\kappa\right)}{2^k \Gamma_m\left(\frac{n}{2}\right)} C_\kappa(\Sigma^{-1}). \tag{46}$$

4.2.2. *The Noncentral Case.* For the general case, the conditioning device used in the previous subsection for the top-order polynomials is one possible route. For, from Kushner and Meisner (1984, equation (2.12)), we have an integral expression for a zonal polynomial  $C_\kappa(W^{-1})$ :

$$C_\kappa(W^{-1}) = \frac{1}{2^k \left(\frac{m}{2}\right)_\kappa} C_\kappa(I_m) \mathbf{E}_X|W [ \Xi_\kappa(X'X) ], \tag{47}$$

<sup>8</sup>The formulas (43) and (46) reflect the fact that both  $C_\kappa(W)$  and  $C_\kappa(W^{-1})$  are eigenfunctions of the family of  $W_m(n, \Sigma)$  distributions ( $n \geq m$ ), i.e., that  $\mathbf{E}_W[f(W)] = \lambda_{f,n} f(\Sigma)$ , with  $\lambda_{f,n}$  a constant (eigenvalue; cf. Kushner et al., 1981, Theorem 3.6).

where  $X$  is  $m \times r$  with independent columns each distributed as  $N(0, W^{-1})$ . Here,  $r \leq m$  is the number of nonzero parts of  $\kappa$ , and the function  $\Xi_\kappa(\cdot)$ , for  $\kappa = (k_1, k_2, \dots, k_m) \vdash k$ , is defined, for a symmetric matrix  $A$ , by

$$\Xi_\kappa(A) = |A_1|^{k_1 - k_2} |A_2|^{k_2 - k_3} \dots |A_{m-1}|^{k_{m-1} - k_m} |A_m|^{k_m}, \tag{48}$$

where  $A_i$  is the upper left  $i \times i$  principal submatrix of  $A$  (so  $A_1 = a_{11}$  and  $A_m = A$ ). The expectation can, in principle, therefore be evaluated, as before, in the unconditional density of  $X$  by interchanging the expectation operations. In the case  $r = 1$  (i.e.,  $C_\kappa(\cdot)$  is the top-order polynomial), this yields the results given in the previous subsection, since  $\Xi_{(k)}(X'X) = x_1' x_1$ , where  $x_1$  is the first column of  $X$ . In general, however, this approach produces an integral that seems difficult to evaluate. We therefore adopt a different, but closely related, direct approach.

### 4.3. Direct Approach, Noncentral Case

Consider first the case  $\Sigma = I_m$ . Direct computation using the averaged density for  $W$  in equation (37) means that we can write

$$\mathbf{E}_\Lambda[C_\kappa(W^{-1})] = \text{etr}\left(-\frac{\Lambda}{2}\right) \sum_{j=0}^\infty \frac{1}{j!} \sum_{\alpha \vdash j} c_{\kappa, \alpha} C_\alpha\left(\frac{\Lambda}{2}\right), \tag{49}$$

with numerical coefficients

$$c_{\kappa, \alpha} = \frac{1}{2^j \binom{n}{\alpha} C_\alpha(I_m)} \mathbf{E}[C_\kappa(W^{-1})C_\alpha(W)]. \tag{50}$$

Recall that the absence of subscripts on  $\mathbf{E}[\cdot]$  means that  $W \sim W_m(n, I_m)$ . The problem thus reduces to the evaluation of the expectations of the products  $C_\kappa(W^{-1})C_\alpha(W)$  when  $W \sim W_m(n, I_m)$ . We will give an explicit expression for these coefficients.

The key to the results given below is the following simple lemma, expressing the zonal polynomial  $C_\kappa(W^{-1})$  in terms of a power of the determinant of  $W$ , multiplied by a zonal polynomial with argument  $W$  itself.

LEMMA 6. For  $\kappa = (k_1, \dots, k_m) \vdash k$  and  $r$  any integer  $\geq k_1$ , the zonal polynomial  $C_\kappa(W^{-1})$  satisfies

$$C_\kappa(W^{-1}) = [C_\kappa(I_m)/C_{\kappa_r}(I_m)] |W|^{-r} C_{\kappa_r}(W), \tag{51}$$

where  $\kappa_r$  is given by  $\kappa_r = (r - k_m, r - k_{m-1}, \dots, r - k_1)$ , a partition of  $rm - k$ .

Constantine (1966, p. 217) mentions this relationship without proof (there the partition  $\kappa_r$  is denoted by  $\kappa^*$ ). Takemura (1984) states the result in Lemma 2, p. 54,

and gives a detailed proof. Another statement and short proof may be found in Macdonald's (2013) notes.<sup>9</sup>

Lemma 6 enables a direct evaluation of the  $c_{\kappa, \alpha}$ , but we also require the following result on the so-called linearization of a product of zonal polynomials (Constantine, 1966; Kushner, 1988; Macdonald, 1995). This says that, for  $\alpha \vdash j$ ,  $\lambda \vdash l$ , and certain coefficients  $g_{\alpha, \lambda}^\delta$ , we have

$$C_\alpha(W)C_\lambda(W) = \sum_{\delta \vdash j+l} g_{\alpha, \lambda}^\delta C_\delta(W). \tag{52}$$

The existence of such coefficients follows from the fact that the left-hand side is a homogeneous invariant polynomial of degree  $j + l$ , and the zonal polynomials span the space of such functions. The coefficients  $g_{\alpha, \lambda}^\delta$  may not all be nonzero, as  $\delta$  varies over the partitions of  $j + l$ ; they can be related to the transition matrices between the various bases of symmetric functions, and are discussed further in Section 5 below.

We also use the following easy consequence of the result in equation (43) and the linearization just given.

LEMMA 7. *If  $W \sim W_m(n, \Sigma)$ ,  $\alpha \vdash j$ , and  $\lambda \vdash l$ ,*

$$E_\Sigma[C_\alpha(W)C_\lambda(W)] = 2^{j+l} \sum_{\delta \vdash j+l} g_{\alpha, \lambda}^\delta \left(\frac{n}{2}\right)_\delta C_\delta(\Sigma). \tag{53}$$

Combining the previous two lemmas, we have the following lemma.<sup>10</sup>

LEMMA 8. *For  $\alpha \vdash j$  and  $\kappa = (k_1, \dots, k_m) \vdash k$ , and  $r$  an integer satisfying  $k_1 \leq r < \frac{n-m+1}{2}$ ,*

$$E_\Sigma[C_\kappa(W^{-1})C_\alpha(W)] = \left[ \frac{2^{j-k} \Gamma_m\left(\frac{n}{2} - r\right) C_\kappa(I_m)}{\Gamma_m\left(\frac{n}{2}\right) C_{\kappa_r}(I_m)} \right] \times |\Sigma|^{-r} \sum_{\delta \vdash rm-k+j} g_{\alpha, \kappa_r}^\delta \left(\frac{n}{2} - r\right)_\delta C_\delta(\Sigma), \tag{54}$$

where  $\kappa_r = (r - k_m, r - k_{m-1}, \dots, r - k_1) \vdash rm - k$ .

<sup>9</sup>When  $m = 2$ , the problem simplifies because, in that case, we can take  $r = k$  in Lemma 6, and observe that  $\kappa_k = \kappa$ , so that  $C_\kappa(W^{-1}) = |W|^{-k} C_\kappa(W)$ , generalizing the fact that  $(\text{tr}(W^{-1}))^k = |W|^{-k} (\text{tr}(W))^k$  when  $m = 2$ .

<sup>10</sup>Note that Lemma 8 implies that, unlike its two components, the function  $f_{\kappa, \alpha}(W) = C_\kappa(W^{-1})C_\alpha(W)$  is **not** an eigenfunction of the  $W_m(n, \Sigma)$  distribution, i.e., (regrettably)  $E_\Sigma[C_\kappa(W^{-1})C_\alpha(W)]$  is **not** a multiple of  $C_\kappa(\Sigma^{-1})C_\alpha(\Sigma)$ . The latter is, instead,

$$C_\alpha(\Sigma)C_\kappa(\Sigma^{-1}) = \frac{C_\kappa(I_m)}{C_{\kappa_r}(I_m)} |\Sigma|^{-r} \sum_{\delta \vdash rm+j-k} g_{\alpha, \kappa_r}^\delta C_\delta(\Sigma).$$

**Remark 5.** In the central case, the term  $C_\alpha(W)$  is missing, and the above result becomes

$$\mathbf{E}_\Sigma [C_\kappa(W^{-1})] = \frac{\Gamma_m\left(\frac{n}{2} - r, \kappa_r\right)}{2^k \Gamma_m\left(\frac{n}{2}\right)} C_\kappa(\Sigma^{-1}).$$

This agrees with the result in equation (46), because, for any  $k_1 \leq r < \frac{n-m+1}{2}$ , it is easy to see that  $\Gamma_m\left(\frac{n}{2} - r, \kappa_r\right) = \Gamma_m\left(\frac{n}{2}, -\kappa\right)$ .

Applying Lemma 8 with  $\Sigma = I_m$ , we have

$$c_{\kappa,\alpha} = \frac{\Gamma_m\left(\frac{n}{2} - r\right) C_\kappa(I_m)}{2^k \binom{\frac{n}{2}}{\alpha} \Gamma_m\left(\frac{n}{2}\right) C_\alpha(I_m) C_{\kappa_r}(I_m)} \sum_{\delta \vdash rm-k+j} g_{\alpha,\kappa_r}^\delta \binom{\frac{n}{2} - r}{\delta} C_\delta(I_m), \tag{55}$$

which produces the following theorem.

**THEOREM 3.** For  $\kappa = (k_1, \dots, k_m) \vdash k$  with  $k_1 < \frac{n-m+1}{2}$  and  $W \sim W_m(n, I_m, \Lambda)$ , we have

$$\mathbf{E}_\Lambda [C_\kappa(W^{-1})] = \text{etr}\left(-\frac{\Lambda}{2}\right) \sum_{j=0}^\infty \frac{1}{j!} \sum_{\alpha \vdash j} c_{\kappa,\alpha} C_\alpha\left(\frac{\Lambda}{2}\right), \tag{56}$$

with coefficients  $c_{\kappa,\alpha}$  defined as in equation (55). Here,  $\kappa_r = (r - k_m, r - k_{m-1}, \dots, r - k_1) \vdash rm - k$ , with  $k_1 \leq r < \frac{n-m+1}{2}$ , and the coefficients  $g_{\alpha,\kappa_r}^\delta$  are those appearing in the linearization (52).

#### 4.4. General $\Sigma$

When  $\Sigma$  is not proportional to  $I_m$ , the above argument using the averaged density (37) is not available, and an alternative approach is needed. Instead, we have

$$\mathbf{E}_{\Sigma,\Lambda} [C_\kappa(W^{-1})] = \text{etr}\left(-\frac{\Sigma^{-1}\Lambda}{2}\right) \sum_{j=0}^\infty \frac{1}{2^{2j} j!} \sum_{\alpha \vdash j} \frac{1}{\binom{\frac{n}{2}}{\alpha}} \mathbf{E}_\Sigma [C_\kappa(W^{-1}) C_\alpha(\Sigma^{-1}\Lambda\Sigma^{-1}W)]. \tag{57}$$

The expectation term on the right-hand side is, on using Lemma 6 and Davis (1979, equation (2.6)),

$$\begin{aligned} & \frac{c_{m,n} C_\kappa(I_m)}{c_{m,n-2r} C_{\kappa_r}(I_m)} |\Sigma|^{-r} \mathbf{E}_{W \sim W_m(n-2r, \Sigma)} [C_{\kappa_r}(W) C_\alpha(\Sigma^{-1}\Lambda\Sigma^{-1}W)] \\ &= \left[ \frac{2^{j-k} \Gamma_m\left(\frac{n}{2} - r\right) C_\kappa(I_m)}{\Gamma_m\left(\frac{n}{2}\right) C_{\kappa_r}(I_m)} \right] |\Sigma|^{-r} \sum_{\phi \in \kappa_r, \alpha} \binom{\frac{n}{2} - r}{\phi} \theta_{\phi}^{\kappa_r, \alpha} C_{\phi}^{\kappa_r, \alpha}(\Sigma, \Sigma^{-1}\Lambda), \tag{58} \end{aligned}$$

where  $\theta_{\phi}^{\kappa_r, \alpha} = C_{\phi}^{\kappa_r, \alpha}(I_m, I_m) / C_{\phi}(I_m)$ . The somewhat elaborate notation needed here is explained in Davis (1979). The polynomials  $C_{\phi}^{\lambda, \alpha}(A, B)$  are Davis's (1980, 1981) invariant polynomials with two matrix arguments, invariant under the simultaneous transformations  $A \rightarrow HAH', B \rightarrow HBH', H \in \mathbb{O}(m)$ . As anticipated in Lemma 4,

therefore,  $\mathbf{E}_{\Sigma, \Lambda}[C_{\kappa}(W^{-1})]$  is invariant under the *simultaneous* transformations  $\Sigma \rightarrow H\Sigma H', \Lambda \rightarrow H\Lambda H', H \in \mathbb{O}(m)$ . Using this result, we obtain the most general expression for the expectation of  $C_{\kappa}(W^{-1})$ .<sup>11</sup>

**THEOREM 4.** For  $\kappa \vdash k$  and any  $r$  satisfying  $k_1 \leq r < \frac{n-m+1}{2}$ , and  $W \sim W_m(n, \Sigma, \Lambda)$ ,

$$\begin{aligned} \mathbf{E}_{\Sigma, \Lambda}[C_{\kappa}(W^{-1})] &= \frac{\Gamma_m\left(\frac{n}{2} - r\right) C_{\kappa}(I_m)}{2^k \Gamma_m\left(\frac{n}{2}\right) C_{\kappa_r}(I_m)} \text{etr}\left(-\frac{\Sigma^{-1}\Lambda}{2}\right) |\Sigma|^{-r} \sum_{j=0}^{\infty} \frac{1}{2^j j!} \\ &\times \sum_{\alpha \vdash j} \frac{1}{\left(\frac{n}{2}\right)_{\alpha}} \sum_{\phi \in \kappa_r, \alpha} \left(\frac{n}{2} - r\right)_{\phi} \theta_{\phi}^{\kappa_r, \alpha} C_{\phi}^{\kappa_r, \alpha}(\Sigma, \Sigma^{-1}\Lambda). \end{aligned} \tag{59}$$

**4.5. Applications: The Functions  $\delta_k(\Sigma, \Lambda)$ ,  $\tau_k(\Sigma, \Lambda)$ , and  $\pi_k(\Sigma, \Lambda)$**

First observe that, in the case  $m = 1$ , the three functions coincide, being just the expected value of a negative power of a noncentral  $\chi_n^2(\lambda)$  variate. Hence, for  $m = 1$ , we have, for all  $k < \frac{n}{2}$ ,

$$\tau_k(\lambda) = \pi_k(\lambda) = \delta_k(\lambda) = \frac{\Gamma\left(\frac{n}{2} - k\right)}{2^k \Gamma\left(\frac{n}{2}\right)} {}_1F_1\left(k; \frac{n}{2}; -\frac{\lambda}{2}\right). \tag{60}$$

Note, in particular, that (unlike the expression for positive powers) this involves an infinite series in  $\lambda$ . It follows that, for general  $m$ , we cannot expect to obtain finite expansions for the expectations of these functions. Indeed, the simplest example is  $\delta_{\kappa}(\Sigma, \Lambda) = \mathbf{E}_{\Sigma, \Lambda}[|W|^{-k}]$ , for which the result is given in Muirhead (1982, Theorem 10.3.7), with  $r$  replaced by  $-k$ , and is a generalization of the result just given for  $m = 1$ : for  $k < \frac{n-m+1}{2}$ ,

$$\delta_k(\Sigma, \Lambda) = \frac{\Gamma_m\left(\frac{n}{2} - k\right)}{2^{km} \Gamma_m\left(\frac{n}{2}\right)} |\Sigma|^{-k} {}_1F_1\left(k; \frac{n}{2}; -\frac{\Sigma^{-1}\Lambda}{2}\right). \tag{61}$$

In contrast to the case  $\mathbf{E}_{\Sigma, \Lambda}[|W|^k]$ , which is a polynomial of degree  $km$ , the series expansion for  $\mathbf{E}_{\Sigma, \Lambda}[|W|^{-k}]$  does not terminate. We do not need to consider this case further.

**4.5.1.  $\tau_k$  and  $\pi_k$  in the Central Case.** Since  $(\text{tr}(W^{-1}))^k = \sum_{\kappa \vdash k} C_{\kappa}(W^{-1})$ , we can set  $r = k$  in Lemma 6 and, using Lemma 5, obtain at once:

$$\mathbf{E}_{\Sigma}[(\text{tr}(W^{-1}))^k] = \sum_{\kappa \vdash k} \frac{\Gamma_m\left(\frac{n}{2}, -\kappa\right)}{2^k \Gamma_m\left(\frac{n}{2}\right)} C_{\kappa}(\Sigma^{-1}). \tag{62}$$

<sup>11</sup>The analogous (infinite series) result for  $\mathbf{E}_{\Sigma, \Lambda}[C_{\kappa}(W)]$  is given as Corollary 4.1 in Díaz-García and Gutiérrez-Jáimez (2001). However, we are able to show that it is possible to obtain an expression for the expectation of  $C_{\kappa}(W)$  as a *finite* sum of invariant polynomials in the matrices  $(\Sigma, \Lambda)$ . The derivation is beyond the scope of the present paper.



The corresponding result for the power sums  $\pi_k$  is less straightforward. However, being a homogeneous and invariant function of  $W$ , the power sum  $\text{tr}(W^{-k})$  has an expansion in terms of zonal polynomials of the form:

$$\text{tr}(W^{-k}) = \sum_{\kappa \vdash k} d_{k,\kappa} C_\kappa(W^{-1}), \tag{63}$$

with coefficients  $d_{k,\kappa}$  that, we show in Appendix C, are given by

$$d_{k,\kappa} = \frac{1}{(k_1)_{k-k_1}} \prod_{i=2}^m \left(-\frac{i-1}{2}\right)_{k_i}, \tag{64}$$

for  $\kappa = (k_1, k_2, \dots, k_m) \vdash k$ , and  $(c)_0 = 1$  by convention.<sup>12</sup>

Given this expansion, the result in Lemma 5 can again be invoked to yield, for the central case,

$$\mathbf{E}_\Sigma[\text{tr}(W^{-k})] = \frac{1}{2^k} \sum_{\kappa \vdash k} \frac{d_{k,\kappa} \Gamma_m\left(\frac{n}{2}, -\kappa\right)}{\Gamma_m\left(\frac{n}{2}\right)} C_\kappa(\Sigma^{-1}). \tag{65}$$

4.5.2.  $\tau_k$  and  $\pi_k$  in the Noncentral Case. In the noncentral case, for  $\tau_k$ , we use the expansion  $(\text{tr}(W^{-1}))^k = \sum_{\kappa \vdash k} C_\kappa(W^{-1})$ , and then set  $r = k$  in Lemma 6 to obtain the following immediate consequence.

COROLLARY 4. For  $k < \frac{n-m+1}{2}$ , and  $\Sigma = I_m$ ,

$$\mathbf{E}_\Lambda [(\text{tr}(W^{-1}))^k] = \text{etr}\left(-\frac{\Lambda}{2}\right) \sum_{j=0}^\infty \frac{1}{j!} \sum_{\alpha \vdash j} b_{k,\alpha} C_\alpha\left(\frac{\Lambda}{2}\right), \tag{66}$$

with coefficients

$$b_{k,\alpha} = \left[ \frac{\Gamma_m\left(\frac{n}{2} - k\right)}{2^k \Gamma_m\left(\frac{n}{2}\right) \left(\frac{n}{2}\right)_\alpha C_\alpha(I_m)} \right] \sum_{\kappa \vdash k} \frac{C_\kappa(I_m)}{C_{\kappa_k}(I_m)} \sum_{\delta \vdash km+j-k} g_{\alpha,\kappa_k}^\delta \left(\frac{n}{2} - k\right)_\delta C_\delta(I_m). \tag{67}$$

Finally, for  $\pi_k(\Sigma, \Lambda) = \mathbf{E}_{\Sigma, \Lambda}[\text{tr}(W^{-k})]$ , we may use equation (63), and, combining this with Lemma 6, we have the expansion

$$\text{tr}(W^{-k}) = \sum_{\kappa \vdash k} d_{k,\kappa} C_\kappa(W^{-1}) = |W|^{-k} \sum_{\kappa \vdash k} d_{k,\kappa} \frac{C_\kappa(I_m)}{C_{\kappa_k}(I_m)} C_{\kappa_k}(W). \tag{68}$$

The coefficients  $d_{k,\kappa}$  are those given earlier in equation (64). Multiplying the above expansion by  $C_\alpha(W)$  and taking the expectation using Lemma 8 produces the following proposition.

<sup>12</sup>The coefficients  $d_{k,\kappa}$  are the elements in the first row of the transition matrix  $\mathbb{D}_k$ , say, that maps the zonal polynomials into the power-sum symmetric functions. The matrix  $\mathbb{D}_k$  is not known explicitly, and equation (64) does not seem to be widely known in the literature. See Section 5 and Appendix C for further discussion.

PROPOSITION 2. For  $\alpha \vdash j$  and  $k < \frac{n-m+1}{2}$ ,

$$\mathbf{E}_\Sigma [\text{tr}(W^{-k})C_\alpha(W)] = |\Sigma|^{-k} \sum_{\delta \vdash km-k+j} f_{k,\alpha}^\delta \left(\frac{n}{2} - k\right)_\delta C_\delta(\Sigma), \tag{69}$$

with

$$f_{k,\alpha}^\delta = \frac{2^{j-k} \Gamma_m\left(\frac{n}{2} - k\right)}{\Gamma_m\left(\frac{n}{2}\right)} \sum_{\kappa \vdash k} d_{k,\kappa} g_{\alpha,\kappa}^\delta \frac{C_\kappa(I_m)}{C_{\kappa_k}(I_m)}. \tag{70}$$

We therefore have, for the case  $\Sigma = I_m$ , the following theorem.

THEOREM 5. For  $k < \frac{n-m+1}{2}$ , and  $W \sim W_m(n, I_m, \Lambda)$ ,

$$\mathbf{E}_\Lambda[\text{tr}(W^{-k})] = \text{etr}\left(-\frac{\Lambda}{2}\right) \sum_{j=0}^\infty \frac{1}{j!} \sum_{\alpha \vdash j} v_{k,\alpha} C_\alpha\left(\frac{\Lambda}{2}\right), \tag{71}$$

with

$$v_{k,\alpha} = \frac{1}{2^j \left(\frac{n}{2}\right)_\alpha C_\alpha(I_m)} \sum_{\delta \vdash km-k+j} f_{k,\alpha}^\delta \left(\frac{n}{2} - k\right)_\delta C_\delta(I_m), \tag{72}$$

where  $f_{k,\alpha}^\delta$  is defined in Proposition 2.

In the case where  $\Sigma$  is not proportional to  $I_m$ , the density-averaging device is again unavailable, and we need to rely on the Davis two-matrix-argument polynomials again. The result (really a corollary of Theorem 4) is the following theorem.

THEOREM 6. For  $k < \frac{n-m+1}{2}$ , and  $W \sim W_m(n, \Sigma, \Lambda)$ ,

$$\begin{aligned} \mathbf{E}_{\Sigma, \Lambda}[\text{tr}(W^{-k})] &= \frac{\Gamma_m\left(\frac{n}{2} - k\right)}{2^k \Gamma_m\left(\frac{n}{2}\right)} \text{etr}\left(-\frac{\Sigma^{-1}\Lambda}{2}\right) |\Sigma|^{-k} \sum_{j=0}^\infty \frac{1}{2^j j!} \sum_{\alpha \vdash j} \sum_{\kappa \vdash k} \frac{d_{k,\kappa}}{\left(\frac{n}{2}\right)_\alpha} \frac{C_\kappa(I_m)}{C_{\kappa_k}(I_m)} \\ &\quad \times \sum_{\phi \in \kappa_k, \alpha} \left(\frac{n}{2} - k\right)_\phi \theta_\phi^{\kappa_k, \alpha} C_\phi^{\kappa_k, \alpha}(\Sigma, \Sigma^{-1}\Lambda). \end{aligned} \tag{73}$$

### 5. TRANSITION MATRICES AND COMPUTATION ISSUES

In order to compute  $\mathbf{E}[W^{-1}]$  in Theorem 1, and expectations of the scalar functions of  $W^{-1}$  discussed in Section 4, we need to compute zonal polynomials and hypergeometric functions with matrix argument. The fastest algorithms for computing the Jack polynomials (of which zonal polynomials are a special case) and hypergeometric functions with matrix arguments are those available in Koev and Edelman (2006) and Chan et al. (2019).

For the results in Section 4, we need to compute the transition matrices between the different bases for the space of homogeneous symmetric functions in the eigenvalues of  $W$ . If  $w_1, \dots, w_m$  are the eigenvalues of  $W$ , the three bases that we are interested in are (1) the monomial symmetric functions, (2) the power-sum symmetric functions, and (3) the zonal polynomials. For  $\kappa = (k_1, \dots, k_m) \vdash k$ , the monomial symmetric function is defined as

$$m_\kappa(W) = \sum_{1 \leq i_1 < i_2 < \dots < i_m \leq m} w_{i_1}^{k_1} w_{i_2}^{k_2} \dots w_{i_m}^{k_m}. \tag{74}$$

The power-sum symmetric functions are defined as

$$p_\kappa(W) = p_{k_1} p_{k_2} \dots p_{k_m}, \tag{75}$$

where  $p_i = \sum_{j=1}^m w_j^i = \text{tr}(W^i)$ . For a given value of  $k$ , we arrange these symmetric functions in reverse lexicographical order and define the corresponding basis vectors as

$$m^{(k)} = \begin{bmatrix} m^{(k)}(W) \\ m^{(k-1,1)}(W) \\ \vdots \\ m^{(1^k)}(W) \end{bmatrix}, \quad p^{(k)} = \begin{bmatrix} p^{(k)}(W) \\ p^{(k-1,1)}(W) \\ \vdots \\ p^{(1^k)}(W) \end{bmatrix}, \quad C^{(k)} = \begin{bmatrix} C^{(k)}(W) \\ C^{(k-1,1)}(W) \\ \vdots \\ C^{(1^k)}(W) \end{bmatrix}. \tag{76}$$

Our particular objective is to compute the transition matrix  $\mathbb{D}_k$  between  $C^{(k)}$  and  $p^{(k)}$ , defined by

$$p^{(k)} = \mathbb{D}_k C^{(k)}. \tag{77}$$

The matrix  $\mathbb{D}_k$  (and other analogous transition matrices) is universal, i.e., its elements do not depend on the argument matrix  $W$ , so, once computed, it is known. Its dimension, though—being the number of partitions of  $k$ —grows rapidly with  $k$ . For  $k \leq 5$ ,  $\mathbb{D}_k$  is available from James (1961), and it is obviously desirable to be able to compute  $\mathbb{D}_k$  for arbitrary  $k$ . This task has hitherto been regarded as prohibitively time-consuming. For example, Gutiérrez, Rodríguez, and Saez (2000) spent about 8 days to compute the transition matrix that transforms  $C^{(k)}$  into  $m^{(k)}$ , for  $k = 20$ , using double precision. However, we have recently written programs that are vastly more efficient when compared with existing alternatives.<sup>13</sup> Programs to implement the calculation of  $\mathbb{D}_k$ , as well as other results in the paper, are available from the authors upon request. We shall briefly outline how these are obtained.

We first compute the inverse of  $\mathbb{D}_k$ . We do this in two steps, first computing the transition matrix  $\mathbb{C}_k$  that expands  $C^{(k)}$  in terms of  $m^{(k)}$ , then the transition matrix

<sup>13</sup>For example, on an Intel i7-4790K CPU, our SageMath program can compute  $\mathbb{D}_{20}$  exactly (i.e., symbolically) in less than 10 seconds. Using double precision in Matlab, the same algorithm can compute  $\mathbb{D}_{20}$  in less than 1.1 seconds. The details of our algorithm will be reported elsewhere.

$\mathbb{T}_k$  that expands  $m^{(k)}$  in terms of  $p^{(k)}$ ,

$$C^{(k)} = \mathbb{C}_k m^{(k)}, \tag{78}$$

$$m^{(k)} = \mathbb{T}_k p^{(k)}. \tag{79}$$

We can then compute  $\mathbb{D}_k^{-1} = \mathbb{C}_k \mathbb{T}_k$ . The most efficient method for computing  $\mathbb{C}_k$  was developed by James (1968), and implemented by Dumitriu, Edelman, and Shuman (2007) and Jiu and Koutschan (2020). The matrix  $\mathbb{T}_k$  can be computed using a recursive algorithm suggested by Merca (2015).

After obtaining  $\mathbb{D}_k^{-1}$ , we can take its inverse to obtain  $\mathbb{D}_k$ , but a more efficient method is to rely on the orthogonality property of the coefficients of zonal polynomials to obtain  $\mathbb{D}_k$ . Let  $d_{\kappa,\lambda}$  be a typical element of  $\mathbb{D}_k$ , where  $\kappa \vdash k$  and  $\lambda \vdash k$ . We can use James (1961, equation (15)) to show that

$$d_{\kappa,\lambda} = \frac{d^{\lambda,\kappa} d^{k,(1^k)}}{d^{k,\kappa} d^{\lambda,(1^k)}}, \tag{80}$$

where  $d^{\lambda,\kappa}$  is a typical element of  $\mathbb{D}_k^{-1}$ .

For the most general results obtained in the paper, the ability to compute the expressions obtained depends on the availability of transition matrices for the Davis invariant polynomials. These are matrices  $\mathbb{D}_{k,l}$ , say, that map a vector of polynomials  $C_\phi^{\kappa,\lambda}(A,B)$  with  $\kappa \vdash k, \lambda \vdash l$ , and  $\phi \vdash k+l$ , suitably ordered, into a vector of *distinct* products of traces like

$$\left(\text{tr}(A^{a_1} B^{b_1} A^{c_1} \dots)\right)^{r_1} \left(\text{tr}(A^{a_2} B^{b_2} A^{c_2} \dots)\right)^{r_2} \dots \tag{81}$$

of total degree  $k$  in  $A, l$  in  $B$ . Davis (1979) gives these matrices up to degree  $k+l=5$ , and explains their construction in Davis (1981). We have so far not been able to extend these beyond that degree.

Finally, we need to compute the linearization coefficients  $g_{\alpha,\lambda}^\delta$  as defined in (52). These coefficients have been extensively studied (see, e.g., Kushner, 1988), and there are combinatorial formulas available for the linearization coefficients for products of the more general class of Jack polynomials (Stanley, 1989; Macdonald, 1995), of which the zonal polynomials are a special case. However, when  $\mathbb{D}_j^{-1}$ ,  $\mathbb{D}_l^{-1}$ , and  $\mathbb{D}_{j+l}$  are available, we can simply use Theorem 3.1 of Richards (1982) to compute  $g_{\alpha,\lambda}^\delta$  with  $\alpha \vdash j, \lambda \vdash l$ , and  $\delta \vdash j+l$  as

$$g_{\alpha,\lambda}^\delta = \sum_{\mu \vdash j+l} d_{\mu\delta} \sum_{\nu \vdash j} \sum_{\substack{\rho \vdash l \\ \nu + \rho = \mu}} d^{\alpha\nu} d^{\lambda\rho}. \tag{82}$$

### 6. CONCLUDING COMMENT

The key to the main result of this paper is the equivariance property in Lemma 1, and it seems likely that this type of argument may well be useful in other contexts.

It appears to be novel in the econometric literature. Although very technical, the results obtained for the expectations of zonal polynomials provide the basis for solving a large class of moment problems that are potentially of interest. These also seem to be new. Finally, it may be noted that all of the results presented here can easily be generalized to the complex noncentral Wishart case. This distribution has important applications in physics and electronics.

## APPENDICES

### A. Proof of Lemma 3

The functions in the integrand on the left are invariant under  $x \rightarrow Hx$ ,  $H \in \mathbb{O}(m)$ . We may therefore use the density (17) in evaluating them. Integrating over  $x \in \mathbb{R}^m$ , the coefficients of the polynomials  $C_\rho(\Lambda)$  in the expansion of the hypergeometric function with two matrix arguments are the numerical coefficients multiplied by the expressions

$$\frac{1}{C_\rho(I_m)} \int_{\mathbb{R}^m} (1+x'x)^{-\frac{n+1}{2}} C_\rho((I_m+xx')^{-1})(dx). \tag{83}$$

Since the integral must evaluate to unity, we can equate coefficients of the polynomials  $C_\rho(\Lambda)$  in this expression and in the expansion of  $\text{etr}(\Lambda/2)$  to obtain equation (20). The second expression arises by transforming  $x \rightarrow (q, h)$  in the first integral with  $q = x'x$  and  $h = x(x'x)^{-\frac{1}{2}}$ . The integrand is constant on the unit sphere  $h'h = 1$ , so integrating over  $h'h = 1$  produces the result.<sup>14</sup>

### B. Proof of Theorem 1

In this appendix, we use the following notation: for any  $n \times k$  matrix  $A$  with  $\text{rank}(A) = k$  (so  $n \geq k$ ), we define  $M_A = I_T - A(A'A)^{-1}A'$ . In addition, we denote by  $e_1$  the unit vector ( $e_1 = (1, 0, \dots, 0)'$ ), whatever its dimension. According to Lemma 2, we need only consider the case  $\Lambda = D$ , i.e.,  $W \sim W_m(n, I_m, D)$ . The conditional variance of  $x_1$  given  $W$  is the (1, 1) element of  $W^{-1}$ :

$$\text{Var}[x_1|W] = e_1' W^{-1} e_1 = \frac{|W_{22}|}{|W|}, \tag{84}$$

where  $W_{22}$  is a submatrix of  $W$  without its first row and column. We need the expectation of this with respect to the distribution of  $W$ . Writing  $W = Z'Z$ , we can assume that

$$Z \sim N\left(\begin{pmatrix} D^{\frac{1}{2}} \\ 0 \end{pmatrix}, I_n \otimes I_m\right).$$

We also write  $Z = (z_1, Z_2)$ , with  $z_1 n \times 1$  and  $Z_2 n \times (m - 1)$ . In this notation,

$$\text{Var}[x_1|Z] = (z_1' M_{Z_2} z_1)^{-1}. \tag{85}$$

<sup>14</sup>An alternative proof of (21) can be based on the generalized binomial expansion for zonal polynomials (Constantine, 1966; Bingham, 1974; Muirhead, 1982, p. 267).

Here,  $z_1$  is independent of  $Z_2$ , and  $z_1 \sim N(\sqrt{\omega_1}e_1, I_n)$ , so that, conditionally,

$$z_1' M_{Z_2} z_1 | Z_2 \sim \chi_{n-m+1}^2(\omega_1 e_1' M_{Z_2} e_1). \tag{86}$$

Therefore, as in equation (9) in the text (with  $n$  replaced by  $n - m + 1$ ),

$$\mathbf{E} \left[ (z_1' M_{Z_2} z_1)^{-1} | Z_2 \right] = \frac{1}{n-m-1} {}_1F_1 \left( 1; \frac{n-m+1}{2}; -\frac{\omega_1 e_1' M_{Z_2} e_1}{2} \right). \tag{87}$$

When  $\omega_1 > 0$ , the unconditional variance of  $x_1$  is the expectation of this with respect to the distribution of  $Z_2$  :

$$\text{Var}[x_1] = \frac{1}{n-m-1} \sum_{k=0}^{\infty} \frac{\left(-\frac{\omega_1}{2}\right)^k}{\binom{n-m+1}{k}} \mu_k(D_1), \tag{88}$$

where

$$\mu_k(D_1) = \mathbf{E}_{Z_2} [(e_1' M_{Z_2} e_1)^k]. \tag{89}$$

Now, let

$$Z_2 = \begin{pmatrix} \tilde{z}'_{21} \\ Z_{22} \end{pmatrix}, \tag{90}$$

with  $Z_{22} \sim N\left(\begin{pmatrix} D_1^{\frac{1}{2}} \\ 0 \end{pmatrix}, I_{n-1} \otimes I_{m-1}\right)$ , and  $\tilde{z}_{21} \sim N(0, I_{m-1})$  independent of  $Z_{22}$ . Then, using the properties of partitioned determinants, we have

$$e_1' M_{Z_2} e_1 = \frac{|(e_1, Z_2)'(e_1, Z_2)|}{|Z_2' Z_2|} = \frac{|Z_{22}' Z_{22}|}{|Z_{22}' Z_{22} + \tilde{z}_{21} \tilde{z}'_{21}|} = \left(1 + \tilde{z}'_{21} R^{-1} \tilde{z}_{21}\right)^{-1}, \tag{91}$$

where  $R = Z_{22}' Z_{22} \sim W_{m-1}(n-1, I_{m-1}, D_1)$  is independent of  $\tilde{z}_{21}$ . Writing  $z = R^{-\frac{1}{2}} \tilde{z}_{21}$ , we have  $z|R \sim N(0, R^{-1})$ . That is, the joint distribution properties of  $(z, R)$  are exactly as above for  $(x, W)$ , but with  $(n, m)$  replaced by  $(n - 1, m - 1)$  and  $D$  replaced by  $D_1$ . Since  $e_1' M_{Z_2} e_1 = (1 + z'z)^{-1}$ , we need to evaluate the expectations

$$\mu_k(D_1) = \mathbf{E}_z [(1 + z'z)^{-k}] \tag{92}$$

when  $z$  has density  $pdf^{(m-1)}(z; D_1)$ . These clearly exist for all  $k \geq 0$ , and we can use **Corollary 2**, with  $(n, m)$  replaced by  $(n - 1, m - 1)$  and  $s$  replaced by  $-k$  to obtain

$$\mu_k(D_1) = \frac{\binom{n-m+1}{2}^k}{\binom{n}{2}^k} \text{etr}\left(-\frac{D_1}{2}\right) {}_2F_2\left(k + \frac{n-1}{2}, \frac{n}{2}; k + \frac{n}{2}, \frac{n-1}{2}; \frac{D_1}{2}\right). \tag{93}$$

Substituting this into equation (88) gives  $\psi_1(D)$ , and **Lemma 2** then implies the result stated. Note that, because the hypergeometric function is symmetric in the elements of its argument, the order of the terms in  $D_i$  is irrelevant.

**C. Other Proofs**

**Proof of (64).** Consider the alternative normalization of  $C_\kappa$ , denoted by  $Z_\kappa$  by James (1961). The expansion of  $Z_\kappa$  in terms of power-sum symmetric polynomials is given by

$$Z_\kappa = \sum_{\lambda \vdash k} \theta_{\kappa, \lambda} p_\lambda. \tag{94}$$

Writing equation (80) in terms of the  $\theta_{\kappa, \lambda}$ , we can write  $d_{k, \kappa}$  as

$$d_{k, \kappa} = \frac{\theta_{\kappa, k} \theta_{k, (1^k)}}{\theta_{k, k} \theta_{\kappa, (1^k)}} = \frac{\theta_{\kappa, k}}{\theta_{k, k}}, \tag{95}$$

because  $\theta_{\kappa, (1^k)} = 1$ , for  $k \vdash \kappa$  (see Macdonald, 1995, equation 10.29, p. 382). An explicit expression for  $\theta_{\kappa, k}$  is given in Macdonald (1995, Example 6.10.1(b), p. 383). Setting  $\alpha = 2$  in that expression and simplifying, we obtain

$$\theta_{\kappa, k} = 2^{k-1} (k_1 - 1)! \prod_{i=2}^m \left( -\frac{i-1}{2} \right)_{k_i}. \tag{96}$$

For the special case of  $\kappa = (k)$ , we have

$$\theta_{k, k} = 2^{k-1} (k - 1)!. \tag{97}$$

It follows that

$$d_{k, \kappa} = \frac{1}{(k_1)_{k-k_1}} \prod_{i=2}^m \left( -\frac{i-1}{2} \right)_{k_i}. \tag{98}$$

□

**Proof of Lemma 8.** For any integer  $k_1 \leq r < \frac{n-m+1}{2}$ ,

$$\mathbf{E}_\Sigma \left[ C_\kappa(W^{-1}) C_\alpha(W) \right] = \frac{c_{m, n} C_\kappa(I_m)}{c_{m, n-2r} C_{\kappa_r}(I_m)} |\Sigma|^{-r} \mathbf{E}_{W \sim W_m(n-2r, \Sigma)} \left[ C_\alpha(W) C_{\kappa_r}(W) \right]. \tag{99}$$

Thus, replacing  $n$  with  $n - 2r$  and  $\lambda$  with  $\kappa_r$  in Lemma 7, we have

$$\begin{aligned} \mathbf{E}_\Sigma \left[ C_\kappa(W^{-1}) C_\alpha(W) \right] &= 2^{j+rm-k} \frac{c_{m, n} C_\kappa(I_m)}{c_{m, n-2r} C_{\kappa_r}(I_m)} |\Sigma|^{-r} \sum_{\delta \vdash rm+j-k} g_{\alpha, \kappa_r}^\delta \left( \frac{n}{2} - r \right)_\delta C_\delta(\Sigma) \\ &= \frac{2^{j-k} \Gamma_m \left( \frac{n}{2} - r \right) C_\kappa(I_m)}{\Gamma_m \left( \frac{n}{2} \right) C_{\kappa_r}(I_m)} |\Sigma|^{-r} \sum_{\delta \vdash rm+j-k} g_{\alpha, \kappa_r}^\delta \left( \frac{n}{2} - r \right)_\delta C_\delta(\Sigma). \end{aligned} \tag{100}$$

□

**REFERENCES**

Andrews, D. W. K. & J. H. Stock (2005) Inference with weak instruments. In R. Blundell, W. K. Newey, and T. Persson (eds.), *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, vol. III. Cambridge University Press. 122–173.

- Bingham, C. (1974) An identity involving partitioned generalized binomial coefficients. *Journal of Multivariate Analysis* 4, 210–223.
- Chan, C., V. Drensky, A. Edelman, R. Kan & P. Koev (2019) On computing Schur functions and series thereof. *Journal of Algebraic Combinatorics* 50, 127–141.
- Constantine, A. G. (1963) Some non-central distribution problems in multivariate analysis. *Annals of Mathematical Statistics* 34, 1270–1285.
- Constantine, A. G. (1966) The distribution of Hotelling's generalized  $T_0^2$ . *Annals of Mathematical Statistics* 37, 215–225.
- Davis, A. W. (1979) Invariant polynomials with two matrix arguments extending the zonal polynomials: Applications to multivariate distribution theory. *Annals of the Institute of Statistical Mathematics* 31(Part A), 465–485.
- Davis, A. W. (1980) Invariant polynomials with two matrix arguments extending the zonal polynomials. In P. R. Krishnaiah (ed.), *Multivariate Analysis-V*. North-Holland Publishing Company.
- Davis, A. W. (1981) On the construction of a class of invariant polynomials in several matrices, extending the zonal polynomials. *Annals of the Institute of Statistical Mathematics* 33(Part A), 297–313.
- Díaz-García, J. & R. Gutiérrez-Jáimez (2001) The expected value of zonal polynomials. *Test* 10, 133–145.
- Dumitriu, I., A. Edelman & G. Shuman (2007) MOPS: Multivariate orthogonal polynomials (symbolically). *Journal of Symbolic Computation* 42, 587–620.
- Gutiérrez, R., J. Rodríguez & J. Saez (2000) Approximation of hypergeometric functions with matricial argument through their development in series of zonal polynomials. *Electronic Transactions on Numerical Analysis* 11, 121–130.
- Hillier, G. H., T. Kinal & V. K. Srivastava (1984) On the moments of ordinary least squares and instrumental variables estimators in a general structural equation. *Econometrica* 52, 185–202.
- James, A. T. (1961) The distribution of non-central means with known covariance. *Annals of Mathematical Statistics* 32, 874–882.
- James, A. T. (1968) Calculation of zonal polynomial coefficients by use of the Laplace–Beltrami operator. *Annals of Mathematical Statistics* 39, 1711–1718.
- Jiu, L. & C. Koutschan (2020) Calculation and properties of zonal polynomials. *Mathematics in Computer Science* 14, 623–640.
- Khatrı, C. G. (1966) On certain distribution problems based on positive definite quadratic functions of normal vectors. *Annals of Mathematical Statistics* 37, 468–479.
- Koev, P. & A. Edelman (2006) An efficient evaluation of the hypergeometric function of a matrix argument. *Mathematics of Computation* 75, 833–846.
- Krishnan, M. (1967) The moments of a doubly non-central t-distribution. *Journal of the American Statistical Association* 62, 278–287.
- Kushner, H. B. (1988) The linearization of the product of two zonal polynomials. *SIAM Journal of Mathematical Analysis* 19, 687–717.
- Kushner, H. B., A. Lebow & M. Meisner (1981) Eigenfunctions of expected value operators in the Wishart distribution, II. *Journal of Multivariate Analysis* 11, 418–433.
- Kushner, H. B. & M. Meisner (1984) Formula for zonal polynomials. *Journal of Multivariate Analysis* 14, 336–347.
- Letac, G. & H. Massam (2008) The non-central Wishart as an exponential family, and its moments. *Journal of Multivariate Analysis* 99, 1393–1417.
- Macdonald, I. G. (1995) *Symmetric Functions and Hall Polynomials, 2nd Edition*. Oxford University Press.
- Macdonald, I. G. (2013) Hypergeometric functions, I. *Preprint*. arXiv:1309.4568v1 [math.CA].
- Merca, M. (2015) Augmented monomials in terms of power sums. *SpringerPlus* 4, 724.
- Muirhead, R. J. (1982) *Aspects of Multivariate Statistical Theory*. Wiley.
- Phillips, P. C. B. (1983) Exact small sample theory in the simultaneous equations model. In Z. Griliches and M. D. Intriligator (eds.), *Handbook of Econometrics*, vol. I: North-Holland Publishing Company.



- Phillips, P. C. B. (1989) Partially identified econometric models. *Econometric Theory* 5, 181–240.
- Richards, D. S. P. (1982) Differential operators associated with zonal polynomials. I. *Annals of the Institute of Statistical Mathematics* 34, 111–117.
- Stanley, R. (1989) Some combinatorial properties of Jack symmetric functions. *Advances in Mathematics* 77, 76–115.
- Takemura, A. (1984) *A Statistical Approach to Zonal Polynomials*. Lecture Notes Monograph Series, vol. 4. Institute of Mathematical Statistics.
- Ullah, A. (1994) On the inverse moments of the non central Wishart matrix. *Parisankhyan Samikkha* 1, 49–50.