School of Social Sciences
Economics Division
University of Southampton
Southampton SO17 1BJ, UK

Discussion Papers in
Economics and Econometrics

Generating Functions and Short Recursions, with Applications to the Moments of Quadratic Forms in Noncentral Normal Vectors

Grant Hillier, Raymond Kan & Xiaolu Wang

No. 0918

Generating Functions and Short Recursions, with Applications to the Moments of Quadratic Forms in Noncentral Normal Vectors

GRANT HILLIER*, RAYMOND KAN**, and XIAOLU WANG***

*CeMMAP and Department of Economics
University of Southampton
Highfield, Southampton SO17 1BJ
United Kingdom
email: ghh@soton.ac.uk
tel: (44) 02380-592659;
**Joseph L. Rotman School of Management
University of Toronto
105 St. George Street
Toronto, Ontario M5S 3E6
Canada
email: kan@chass.utoronto.ca
tel: (416)978-4291;
***Joseph L. Rotman School of Management
University of Toronto
105 St. George Street
Toronto, Ontario M5S 3E6
Canada
email: xiaolu.wang04@rotman.utoronto.ca
tel: (416)946-8064

Abstract

Using generating functions, the top-order zonal polynomials that occur in much distribution theory under normality can be recursively related to other symmetric functions (power-sum and elementary symmetric functions, Ruben (1962), Hillier, Kan, and Wang (2009)). Typically, in a recursion of this type the k-th object of interest, d_k say, is expressed in terms of all lower-order d_j's. In Hillier, Kan, and Wang (2009) we pointed out that, in the case of top-order zonal polynomials (and generalizations of them), a shorter (i.e., fixed length) recursion can be deduced. The present paper shows that the argument in Hillier, Kan, and Wang (2009) generalizes to a large class of objects/generating functions. The results thus obtained are then applied to various problems involving quadratic forms in noncentral normal vectors.

Keywords: Generating Functions, Invariant Polynomials, Non-central Normal Distibution, Recursions, Symmetric Functions, Zonal Polynomials.

JEL Classifications: C01, C46, C63

1. INTRODUCTION

Relations between the generating functions for different mathematical objects can yield useful recurrence relations between those objects. This has long been appreciated in the theory of symmetric functions, for instance, and in statistics can be exploited to yield recurrence relations between moments and cumulants (Smith (1995)). And, the top-order zonal polynomials that occur in much statistical distribution theory under normality can, by this device, be recursively related to other symmetric functions, in particular, the power-sum and elementary symmetric functions (Ruben (1962), Hillier, Kan, and Wang (2009)).

These results greatly facilitate the efficient computation of these functions, and hence our ability to compute moments, densities, distribution functions, etc., that are expressed in terms of the objects of interest d_k, say. However, such recursions typically express d_k in terms of all lower-order d_j's, and in Hillier, Kan, and Wang (2009) (henceforth abbreviated to HKW) we pointed out that, in the case of top-order zonal polynomials (and invariant polynomials with several matrix arguments), a shorter (i.e., fixed length) recursion can also be deduced by exploiting the relations between several generating functions. In this paper we show that the argument in HKW applies much more generally. We first show that *any* generating function may be used to define an associated function that induces a recurrence relation of exactly the same form as holds between the top-order zonal polynomials and the power-sum symmetric functions. Then, we show that, under certain often-satisfied conditions on the associated function, there is a short recursion that can considerably improve the efficiency of the recursion for computational purposes. The results are applied to various problems involving quadratic forms in noncentral normal vectors, including: moments, product moments, densities, distribution functions, and expectations of ratios of powers of quadratic forms. In all of these cases, we provide new short recursions that are extremely efficient for computation. We begin by briefly explaining the results in HKW, and a little more of the background.

The top-order zonal polynomials of a symmetric matrix A, $C_k(A)$, and the top-order invariant polynomials with several matrix arguments introduced by Davis (1979) (1981), $C_{k_1,k_2,\ldots,k_r}(A_1,\ldots,A_r)$, occur sufficiently frequently in multivariate calculations as to deserve special attention. For exam-

ple, if $z \sim N(0_n, I_n)$, the moments of the quadratic form $q = z'Az$ are given by

$$\mu_k = E[q^k] = 2^k \left(\frac{1}{2}\right)_k C_k(A), \tag{1}$$

and the product-moments of the several quadratic forms $q_i = z'A_i z$, $i = 1, \ldots, r$, are given by

$$\mu_\kappa = E[q_1^{k_1} q_2^{k_2} \cdots q_r^{k_r}] = 2^k \left(\frac{1}{2}\right)_k C_\kappa(A_1, \cdots, A_r), \tag{2}$$

where $\kappa = (k_1, \ldots, k_r)$, $k = |\kappa| = \Sigma_{i=1}^r k_i$, and $(c)_k = c(c+1) \cdots (c+k-1)$ is the usual Pochhammer symbol. These expressions follow easily from the moment generating function (MGF) of q, and the joint moment generating function of the q_i, both of which have expansions in terms of these polynomials (see below). Ruben (1962) and James (1961) essentially give (1), while Chikuse (1987) gives (2). The density function of q may also be expressed as an infinite series in the $C_k(A)$, see James (1961), Ruben (1962).

Many alternative expressions for these polynomials (or, equivalently, moments) have appeared in the literature, but for computation purposes the most efficient expressions have, until recently, been the recurrence relations due to Ruben (1962) for the $C_k(A)$, and Chikuse (1987) for the $C_\kappa(A_1, \cdots, A_r)$. These recursions involve the power-sum symmetric functions, p_k say, in the eigenvalues of A, and, in the multivariate case, generalizations of them defined in terms of a multivariate generating function. Although superior to the explicit formulae for the polynomials, these recurrence relations have length k, and hence have computation complexity of order $O(k^2)$, which means that the recursions are computationally quite inefficient. However, in HKW, we have given new recurrence relations for both cases that involve, instead, the elementary symmetric functions, and appropriate generalizations of them for the multivariate case. These have length at most equal to n, the dimension of the matrix (or matrices) involved, that does not increase with k. These new recursions therefore improve computational complexity to $O(k)$, and the fact that they involve only a fixed number of terms, whatever the degree of the polynomial, means that there is also a substantial saving on storage requirements.

The new recurrence relations in HKW were derived by exploiting properties of, and relations between, the various generating functions for the polynomials that are involved. If we normalize the top-order zonal polynomial $C_k(A)$ by writing

$$d_k(A) = \frac{\left(\frac{1}{2}\right)_k C_k(A)}{k!},$$

the (ordinary) generating function for the d_k is:

$$D(t) = |I_n - tA|^{-\frac{1}{2}} = \sum_{k=0}^{\infty} d_k t^k,$$

while the power sums p_k and elementary symmetric functions e_k have generating functions $P(t) = \text{tr}(tA(I_n - tA)^{-1})$ and $E(t) = |I_n - tA|$, respectively,[1] where $\text{tr}(\cdot)$ is the trace operator. Here, $d_0 = e_0 = 1$, and, crucially, $e_k = 0$ for $k > n$. Note that, in terms of the d_k, $\mu_k = 2^k k! d_k(A)$.

These generating functions may easily be shown to satisfy the differential equations:

$$tE'(t) = -E(t)P(t),$$

and

$$tD'(t) = \frac{1}{2}D(t)P(t).$$

The second of these immediately yields (on equating coefficients of like powers of t on both sides) the recursion in Ruben (1962):

$$d_k = \frac{1}{2k}\sum_{j=1}^{k} p_j d_{k-j}, \tag{3}$$

while combining the two leads to the alternative recursion in terms of the e_k given in HKW:

$$d_k = \sum_{j=1}^{\min[k,n]} \left(\frac{j}{2k} - 1\right) e_j d_{k-j}.$$

In HKW, we show that these relations generalize in the obvious way to the multivariate case. For brevity we refer to the recursions involving the e_j as the "short" recursions, and those involving the p_j as the "long" recursions.

Our first purpose in this paper is to show that the generating function relationships that underpin the short recursions given in HKW for the polynomials d_k and d_κ apply much more generally than we at first appreciated. These results are given in Section 2, first for univariate generating functions, then for the general multivariate case. In the remainder of the paper we apply these results to various problems connected with the properties of quadratic forms in noncentral normal vectors, and of functions of such forms. In Section 3, we show that the recurrence formulae that hold for the moments of q when $z \sim N(0_n, I_n)$, also hold in the noncentral case with $z \sim N(\mu, I_n)$, given suitable modifications of definitions of the p_k and e_k. Section 4 presents some analogous results for the product moments of several quadratic forms. Finally, in Section 5, we study the

3

expectation of a ratio of powers of two quadratic forms in noncentral normal vectors. Throughout the paper we use the notation in Wilf (2005) for coefficients in generating functions: the expression $[t^k]f(t)$ denotes the coefficient of t^k in the power series expansion of the function $f(t)$ in powers of t.

2. GENERATING FUNCTIONS AND RECURSIONS

2.1 Generating Functions with a Single Variable

Let $D(t)$ be an arbitrary generating function for the objects d_k, which themselves will in general be functions of other variables,

$$D(t) = \sum_{k=0}^{\infty} d_k t^k.$$

In applications $D(t)$ will typically be the MGF of some random variable of interest, or a more general generating function for moment-like quantities associated with one or more random variables. Note that we treat $D(t)$ as a formal power series, without using any of the function-theoretic properties of the function that may be represented by the series, and without worrying about whether such a series converges or not.

We define a second generating function $P(t)$ by the formula:

$$P(t) = t\frac{\partial \ln D(t)}{\partial t} = \frac{tD'(t)}{D(t)} = \sum_{k=1}^{\infty} p_k t^k, \tag{4}$$

so that

$$tD'(t) = D(t)P(t). \tag{5}$$

Equating coefficients on both sides of this identity yields the recursion in (3), except for the factor $\frac{1}{2}$:

$$d_k = \frac{1}{k}\sum_{j=1}^{k} p_j d_{k-j} = \frac{1}{k}\sum_{j=0}^{k-1} d_j p_{k-j}. \tag{6}$$

With the initial condition on $d_0 = D(0)$, this recurrence relation allows us to recursively obtain the d_k using the p_k. The usefulness of this type of result depends, of course, on whether the functions p_k are significantly easier to compute than are the d_k themselves. And, unless $P(t)$ is a finite order polynomial, the length of the recursion increases with k, so it may be computationally inefficient to use this recurrence relation when k is large. In the case $D(t) = |I_n - tA|^{-\frac{1}{2}}$ (the MGF

4

of $q/2 = z'Az/2$ when $z \sim N(0_n, I_n)$), $P(t) = \frac{1}{2}\text{tr}(tA(I_n - tA)^{-1})$, so that the $p_j = \frac{1}{2}\text{tr}(A^j)$ are essentially the power-sum symmetric functions, and these are indeed easily computed. However, it is now clear that this same recursion applies for any generating functions $D(t)$ and $P(t)$ related by (5).

Remark 1. *Note that, if $M(t) = \sum_{r=0}^{\infty} \frac{\mu_r}{r!} t^r$, say, is the moment generating function for a random variable with cumulant generating function $K(t) = \ln M(t) = \sum_{r=1}^{\infty} \frac{\kappa_r}{r!} t^r$, say, then $d_k = \mu_k/k!$ and $P(t) = tK'(t) = \sum_{r=1}^{\infty} \frac{\kappa_r}{(r-1)!} t^r$, so that $p_r = \kappa_r/(r-1)!$. Thus, (6) gives the well-known recursion for moments in terms of cumulants:*

$$\mu_k = \sum_{j=1}^{k} \binom{k-1}{j-1} \kappa_j \mu_{k-j} = \sum_{j=0}^{k-1} \binom{k-1}{j} \mu_j \kappa_{k-j}.$$

See, for instance, Smith (1995).

Now, suppose that $P(t)$, as defined in (4), is a rational function of t and can be written as

$$P(t) = \frac{G(t)}{E(t)},$$

where $G(t) = \sum_{k=1}^{m_1} g_k t^k$ and $E(t) = \sum_{k=0}^{m_2} e_k t^k$ are *both finite order polynomials* in t. Note that $g_0 = 0$ because $p_0 = 0$, and that g_k can be obtained by using the fact that $g_k = \sum_{i=0}^{k-1} e_i p_{k-i}$, which follows from the identity $G(t) = E(t)P(t)$. Without loss of generality, we normalize the two polynomials $G(t)$ and $E(t)$ so that $e_0 = 1$. The following result generalizes the result given in equation (24) in HKW:

LEMMA 1. *Suppose that the generating function $P(t)$ defined in (4) can be written as*

$$P(t) = \frac{G(t)}{E(t)},$$

with both $G(t)$ and $E(t)$ polynomials of finite order, say m_1 and m_2 respectively, and $e_0 = 1$. Then the d_k may be determined recursively from the relation

$$d_k = \sum_{j=1}^{\min[k,m]} \left[\frac{c_j}{k} - e_j \right] d_{k-j}, \tag{7}$$

together with the initial condition d_0, where $m = \max[m_1, m_2]$ and $c_j = je_j + g_j$.

Proof. Define

$$F(t) = E(t)D(t) = \sum_{k=0}^{\infty} f_k t^k,$$

with

$$f_k = \sum_{j=0}^{\min[k,m_2]} e_j d_{k-j}. \tag{8}$$

Differentiating $F(t)$, and making use of the relationship in (5),

$$F'(t) = E'(t)D(t) + E(t)D'(t) = E'(t)D(t) + \frac{1}{t}E(t)P(t)D(t) = \left[E'(t) + \frac{G(t)}{t}\right]D(t).$$

Thus,

$$\sum_{k=1}^{\infty} k f_k t^{k-1} = \left(\sum_{j=1}^{m}(je_j + g_j)t^{j-1}\right)\left(\sum_{i=0}^{\infty} d_i t^i\right),$$

where $m = \max[m_1, m_2]$. Equating coefficients of like powers of t on both sides and using (8) we obtain:

$$k\sum_{j=0}^{\min[k,m_2]} e_j d_{k-j} = k f_k = \sum_{j=1}^{\min[k,m]}(je_j + g_j)d_{k-j}.$$

Rearranging this and using the fact that $e_0 = 1$ gives the stated relation. ∎

The key advantage of (7) over (6) is that at most m terms are needed to compute d_k. As a result, the computation time for the d_k does not increase with k, and there is no need to store all previous values of the c_j, so the memory requirement also does not increase with k. Again, though, the usefulness of the result depends on whether or not the c_j are significantly easier to compute than the d_k themselves. As we shall see, this is certainly the case in the applications involving quadratic forms in normal variates that we discuss below.

Remark 2. *For the case $D(t) = |I_n - tA|^{-\frac{1}{2}}$, $P(t) = \text{tr}(tA(I_n - tA)^{-1})/2$ and $P(t)$ can be written as $G(t)/E(t)$, where $G(t) = \text{tr}((tA)\text{adj}(I_n - tA))/2$ with $\text{adj}(I_n - tA)$ denotes the adjoint matrix of $I_n - tA$, and $E(t) = |I_n - tA|$. Since the elements of $\text{adj}(I_n - tA)$ are polynomials of degree $n - 1$ in t, both $G(t)$ and $E(t)$ are polynomials of degree n in t. Thus, $P(t)$ satisfies the hypotheses of the Lemma.*

2.2 Multivariate Generating Functions

We now extend the results in the previous subsection to deal with generating functions of more than one variable. Special cases of these results were given in Section 3 of HKW. For the rest of

6

the paper, we shall adopt the following notation: $t = (t_1, \ldots, t_r)$, $\kappa = (k_1, \ldots, k_r)$, the k_i being nonnegative integers, $|\kappa|$ will denote the sum of the parts of κ, i.e., $|\kappa| = \sum_{i=1}^{r} k_i$, $t^\kappa = \prod_{i=1}^{r} t_i^{k_i}$, and $\kappa! = \prod_{i=1}^{r} k_i!$.

With this notation, we can also extend Wilf's notation for the coefficients in a generating function

$$G(t) = \sum_{k=0}^{\infty} \sum_{|\kappa|=k} g_\kappa t^\kappa,$$

say, by writing

$$g_\kappa = [t^\kappa] G(t).$$

Also, generalizing the familiar relation between the coefficients in the product of two (formal) power series with those of the two constituent series, we have that, if $G(t) = P(t)E(t)$, say, where $P(t)$ and $E(t)$ are at this stage arbitrary, then, if $k = |\kappa|$,

$$g_\kappa = [t^\kappa] P(t)E(t) = \sum_{i=0}^{k} \sum_{\substack{|\nu|=i \\ \nu \leq \kappa}} e_\nu p_{\kappa-\nu},$$

where the notation $\nu \leq \kappa$ means that $\nu = (\nu_1, \ldots, \nu_r)$ is a sequence of nonnegative integers satisfying $0 \leq \nu_i \leq k_i$ for all i.

Next, for a given (ordinary) generating function

$$f(t) = \sum_{k=0}^{\infty} \sum_{|\kappa|=k} f_\kappa t^\kappa,$$

we define

$$\dot{f}(t) = \sum_{i=1}^{r} t_i \frac{\partial f(t)}{\partial t_i} = \sum_{k=1}^{\infty} k \sum_{|\kappa|=k} f_\kappa t^\kappa$$

as a generalization of $tf'(t)$ for the single variable case.

Assume given, as in the univariate case, an arbitrary multivariate generating function $D(t)$ for objects d_κ, i.e.,

$$D(t) = \sum_{k=0}^{\infty} \sum_{|\kappa|=k} d_\kappa t^\kappa.$$

Then, exactly as in the case with single variable, we define $P(t)$ by the equation

$$P(t) = \frac{\dot{D}(t)}{D(t)} = \sum_{i=1}^{r} t_i \frac{\partial \ln D(t)}{\partial t_i}, \tag{9}$$

7

so that

$$\dot{D}(t) = P(t)D(t). \tag{10}$$

Since $P(0) = 0$, we can write $P(t)$ as

$$P(t) = \sum_{k=1}^{\infty} \sum_{|\kappa|=k} p_\kappa t^\kappa$$

and rewrite (10) as

$$\sum_{k=1}^{\infty} \sum_{|\kappa|=k} k d_\kappa t^\kappa = \left(\sum_{i=1}^{\infty} \sum_{|\nu|=i} p_\nu t^\nu \right) \left(\sum_{j=0}^{\infty} \sum_{|\lambda|=j} d_\lambda t^\lambda \right).$$

Comparing the coefficients of t^κ on both sides, we obtain the multivariate version of the recurrence relation (6):

$$d_\kappa = \frac{1}{k} \sum_{i=1}^{k} \sum_{\substack{|\nu|=i \\ \nu \leq \kappa}} p_\nu d_{\kappa-\nu}, \tag{11}$$

where $k = |\kappa|$. Together with the boundary condition $d_0 = D(0)$, this result provides a (long) recursive algorithm for computing the d_κ, given the p_κ's, and is a generalization of (6) for the single variable case. However, (11) expresses d_κ as a linear combination of $\prod_{i=1}^{r}(k_i + 1) - 1$ different d_ν's, so it is extremely inefficient when any of the k_i's are large.

Before presenting the generalized version of (7), we note that a different, and potentially slightly shorter recursive algorithm for the d_κ can be obtained by using a different generalization of the expression $t f'(t)$. Instead of computing $\dot{D}(t)$, we can pick a j such that $k_j > 0$, and consider just the derivative of $D(t)$ with respect to t_j. This gives us

$$t_j \frac{\partial D(t)}{\partial t_j} = t_j \frac{\partial \ln(D(t))}{\partial t_j} D(t),$$

which implies:

$$\sum_{k=1}^{\infty} \sum_{|\kappa|=k} k_j d_\kappa t^\kappa = \left(\sum_{k=1}^{\infty} \sum_{|\kappa|=k} \frac{k_j}{k} p_\kappa t^\kappa \right) \left(\sum_{k=0}^{\infty} \sum_{|\kappa|=k} d_\kappa t^\kappa \right),$$

since the coefficient of t^κ in $\ln(D(t))$ is p_κ/k. Comparing the coefficients of t^κ on both sides, we obtain a second recursive algorithm for the d_κ:

$$d_\kappa = \frac{1}{k_j} \sum_{i=1}^{k} \sum_{\substack{|\nu|=i \\ \nu \leq \kappa}} \frac{\nu_j}{i} p_\nu d_{\kappa-\nu}, \tag{12}$$

8

which can also be considered as a multivariate generalization of (6).

Equation (12) expresses d_κ as a linear combination of $[k_j/(k_j + 1)] \prod_{i=1}^{r}(k_i + 1)$ different d_ν's with $\nu < \kappa$. While (12) works for any j with $k_j > 0$, it is best to pick the j with the smallest nonzero k_j in order to achieve the shortest recursion. When $k_j = 1$, the length of recursion in (12) is only half of that of (11). Nevertheless, there is no substantial computational advantage of using (12) over (11). This is because, while (12) requires summing fewer terms than (11), each term in the recursion entails an extra multiplication by ν_j/i.

As in the single variable case, we can obtain a shorter recurrence relation for the d_κ if $P(t)$ is a rational function of t and can be expressed as

$$P(t) = \frac{G(t)}{E(t)},\tag{13}$$

where both

$$G(t) = \sum_{k=1}^{m_1} \sum_{|\kappa|=k} g_\kappa t^\kappa$$

and

$$E(t) = \sum_{k=0}^{m_2} \sum_{|\kappa|=k} e_\kappa t^\kappa$$

are finite-order polynomials in t and $e_0 = 1$. Note that $g_0 = G(0) = 0$ because $P(0) = 0$, and that the coefficients g_κ in $G(t)$ can be obtained from

$$g_\kappa = \sum_{i=0}^{k-1} \sum_{\substack{|\nu|=i, \\ \nu \leq \kappa}} e_\nu p_{\kappa-\nu}$$

as in the single variable case. We have, in generalization of Lemma 1:

LEMMA 2. *Given an arbitrary multivariate generating function $D(t)$, defining $P(t)$ as in (9), and assuming that $P(t)$ is a rational function of t as in (13), with both $G(t)$ and $E(t)$ finite of degrees m_1 and m_2, respectively, then the d_κ can be determined recursively from the short recurrence relation:*

$$d_\kappa = \sum_{i=1}^{\min[k,m]} \sum_{\substack{|\nu|=i, \\ \nu \leq \kappa}} \left(\frac{c_\nu}{k} - e_\nu\right) d_{\kappa-\nu},\tag{14}$$

where $m = \max[m_1, m_2]$ and $c_\nu = |\nu|e_\nu + g_\nu$.

9

Proof. Defining, as in the single variable case,

$$F(t) = E(t)D(t) = \sum_{k=0}^{\infty} \sum_{|\kappa|=k} f_\kappa t^\kappa,$$

where

$$f_\kappa = \sum_{i=0}^{\min[k,m_2]} \sum_{\substack{|\nu|=i \\ \nu \leq \kappa}} e_\nu d_{\kappa-\nu}. \tag{15}$$

Then

$$\dot{F}(t) = \dot{E}(t)D(t) + E(t)\dot{D}(t) = \left[\dot{E}(t) + G(t)\right]D(t)$$

on using (10) and (13). Hence

$$\sum_{k=1}^{\infty} \sum_{|\kappa|=k} k f_\kappa t^\kappa = \left(\sum_{i=1}^{m} \sum_{|\nu|=i} (ie_\nu + g_\nu) t^\kappa \right) \left(\sum_{k=0}^{\infty} \sum_{|\kappa|=k} d_\kappa t^\kappa \right),$$

where $m = \max[m_1, m_2]$. Equating the coefficients of t^κ on both sides gives us

$$k f_\kappa = \sum_{i=1}^{\min[k,m]} \sum_{\substack{|\nu|=i, \\ \nu \leq \kappa}} (ie_\nu + g_\nu) d_{\kappa-\nu}.$$

Finally, using (15) and rearranging terms gives the stated result for the d_κ. ∎

As before — and in contrast to (11) and (12) — the short recurrence relation only ever uses at most $(m+r)!/(m!r!) - 1$ terms, and so significantly reduces the computation time and memory requirement.

In the remainder of the paper we present a variety of applications of these results to problems involving properties of quadratic forms in normal random variables. From now on we reserve the notation $D(t)$ for the multivariate generating function for the top-order invariant polynomials d_κ

$$D(t) = |I_n - A(t)|^{-\frac{1}{2}} = \sum_{k=0}^{\infty} \sum_{|\kappa|=k} d_\kappa t^\kappa, \tag{16}$$

where $A(t) = t_1 A_1 + ... + t_r A_r$, and $P(t)$ for the generalized power-sum generating function associated with it:

$$P(t) = \text{tr}(A(t)(I_n - A(t))^{-1}) = \sum_{k=1}^{\infty} \sum_{|\kappa|=k} p_\kappa t^\kappa.$$

10

Also, we reserve $E(t)$ for the determinant $|I_n - A(t)|$. In all other applications of the results given in this Section we add a tilde to D, P, and E, and their associated coefficients d_κ, p_κ, and e_κ, to indicate that these are not the basic forms. Beware, though, that this means that the same symbol will appear in different places with different meanings.

Remark 3. *$P(t)$ can be written as $G(t)/E(t)$, where $G(t) = \mathrm{tr}\,(A(t)\mathrm{adj}\,(I_n - A(t)))$. Since the elements of $\mathrm{adj}\,(I_n - A(t))$ are polynomials of degree $n - 1$, both $E(t)$ and $G(t)$ are polynomials of degree n. Therefore, $P(t)$ satisfies the hypotheses of Lemma 2.*

In the applications to follow we always have

$$\tilde{D}(t) = D(t)\exp\left(\frac{1}{2}K(t)\right)$$

for some choice of $K(t)$. In this case we obviously have the simplification:

$$\tilde{P}(t) = \frac{\dot{\tilde{D}}_t}{\tilde{D}(t)} = \frac{1}{2}P(t) + \frac{1}{2}\dot{K}(t).$$

Since we already know that $P(t)$ is a ratio of finite-order polynomials, $\tilde{P}(t)$ will have that property if $\dot{K}(t)$ does.

3. FIRST APPLICATION: MOMENTS OF QUADRATIC FORMS

Before presenting results on the moments of the q_i and other applications to follow, we note the following useful lemma:

LEMMA 3. *Any property of any statistic that depends only on quadratic forms in z, with $z \sim N(\mu, I_n)$, will be unchanged if the $N(\mu, I_n)$ density for z,*

$$\alpha(z) = (2\pi)^{-\frac{n}{2}}\exp\left(-\frac{z'z}{2}\right)\exp\left(-\frac{\mu'\mu}{2}\right)\exp(z'\mu)$$

is replaced by

$$\bar{\alpha}(z) = (2\pi)^{-\frac{n}{2}}\exp\left(-\frac{z'z}{2}\right)\exp\left(-\frac{\mu'\mu}{2}\right){}_0F_1\left(\frac{1}{2}; \frac{z'\mu\mu'z}{4}\right). \tag{17}$$

This result follows from the observation that quadratic forms of the type $z'z, z'Az, z'Bz$, etc., are invariant under $z \to -z$. If $z \sim N(\mu, I_n)$, any property of a statistic that is a function only of

such quadratic forms will, as a consequence, be unchanged if the term $\exp(z'\mu)$ in $\alpha(z)$ is replaced by its average over $\pm z$ (i.e., $[\exp(z'\mu) + \exp(-z'\mu)]/2$), which is precisely the final hypergeometric function in (17).

3.1 Formulae for the Moments

It is straightforward to obtain an explicit formula for the moments of q when $z \sim N(\mu, I_n)$, and a variety of such expressions are extant in the literature. Virtually all of these are most parsimoniously expressed in terms of the normalized top-order zonal polynomials d_k. Using (17), we can write the moment generating function of $q = z'Az$ as:

$$
\begin{aligned}
M_q(\tau) &= E[\exp(\tau z'Az)] = \sum_{k=0}^{\infty} \frac{\mu_k}{k!}\tau^k \\
&= \exp\left(-\frac{\delta}{2}\right) \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{\tau^k}{k!j!2^{2j}\left(\frac{1}{2}\right)_j} E_0\left[(z'Az)^k(z'\mu\mu'z)^j\right],
\end{aligned}
$$

where $E_0[\cdot]$ denotes expectation with respect to $z \sim N(0_n, I_n)$, and $\delta = \mu'\mu$. But, from HKW equation (48) we have

$$
E_0\left[(z'Az)^k(z'\mu\mu'z)^j\right] = 2^{j+k}j!k!d_{k,j}(A, \mu\mu').
$$

Hence,

$$
M_q(\tau) = \exp\left(-\frac{\delta}{2}\right) \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{(2\tau)^k}{2^j\left(\frac{1}{2}\right)_j} d_{k,j}(A, \mu\mu')
$$

and

$$
\mu_k = 2^k k! \exp\left(-\frac{\delta}{2}\right) \sum_{j=0}^{\infty} \frac{1}{\left(\frac{1}{2}\right)_j 2^j} d_{k,j}(A, \mu\mu'). \tag{18}
$$

It is straightforward to check (by majorization) that the series converges for all A and μ. This generalizes the result mentioned earlier for the central ($\mu = 0_n$) case:

$$
\mu_k = E_0[(z'Az)^k] = 2^k k! d_k(A).
$$

In the central case the short recursion given in HKW for evaluating the d_k is extremely efficient, and the short recursion given there for evaluating the $d_{k,j}$ can be used in conjunction with (18) to compute the moments in the noncentral case. However, we shall now show that the results in Section 2 provide an even more efficient procedure in the noncentral case.

3.2 Recursions for the Noncentral Case

It is easy to show that the MGF of q may also be written as

$$M_q(\tau) = |I_n - 2\tau A|^{-\frac{1}{2}} \exp\left(\frac{\mu'(I_n - 2\tau A)^{-1}\mu - \mu'\mu}{2}\right).$$

Let

$$\tilde{D}(t) = |I_n - tA|^{-\frac{1}{2}} \exp\left(\frac{\mu'(I_n - tA)^{-1}\mu - \mu'\mu}{2}\right) = \sum_{k=0}^{\infty} \tilde{d}_k(A,\mu)t^k, \tag{19}$$

so that the moments of q are given by

$$\mu_k = E[q^k] = 2^k k! \tilde{d}_k.$$

Note that the coefficients $\tilde{d}_k$ are functions of both A and μ (see (18) above), but we omit this dependence in the notation when A and μ are clear from the context.

Using the fact that when t is sufficiently small, $(I_n - tA)^{-1} = \sum_{k=0}^{\infty} A^k t^k$, and $\ln|I_n - tA| = -\sum_{k=1}^{\infty} \frac{1}{k}\mathrm{tr}(A^k)t^k$, we can write $\ln(\tilde{D}(t))$ as

$$\ln(\tilde{D}(t)) = -\frac{1}{2}\ln|I_n - tA| + \frac{1}{2}\mu'[(I_n - tA)^{-1} - I_n]\mu = \frac{1}{2}\sum_{k=1}^{\infty}\left[\mu'A^k\mu + \frac{\mathrm{tr}(A^k)}{k}\right]t^k.$$

Defining $\tilde{P}(t)$ as in (4) above, we therefore have

$$\tilde{P}(t) = \frac{1}{2}\sum_{k=1}^{\infty}\left[k\mu'A^k\mu + \mathrm{tr}(A^k)\right]t^k,$$

so that, in this case,

$$\tilde{p}_k = \frac{1}{2}\left[k\mu'A^k\mu + \mathrm{tr}(A^k)\right].$$

The recursion (6) thus applies and gives a long recurrence relation for the $\tilde{d}_k$:

$$\tilde{d}_k = \frac{1}{k}\sum_{j=1}^{k}\tilde{p}_j\tilde{d}_{k-j}. \tag{20}$$

Obviously, $\tilde{p}_j$ reduces to $p_j/2$ when $\mu = 0_n$.

To see that $\tilde{P}(t)$ is a rational polynomial with both numerator and denominator of finite degree, so that the result in Lemma 1 also applies, first note that, by definition,

$$\begin{aligned}
\tilde{P}(t) &= t\frac{\partial \ln|I_n - tA|^{-\frac{1}{2}}}{\partial t} + \frac{1}{2}t\frac{\partial}{\partial t}\mu'(I_n - tA)^{-1}\mu \\
&= \frac{1}{2}\mathrm{tr}\left(tA(I_n - tA)^{-1}\right) + \frac{1}{2}\mu'(I_n - tA)^{-1}(tA)(I_n - tA)^{-1}\mu.
\end{aligned} \tag{21}$$

13

Writing $(I_n - tA)^{-1} = \mathrm{adj}(I_n - tA)/|I_n - tA|$ we have $\tilde{P}(t) = \tilde{G}(t)/\tilde{E}(t)$, where

$$\tilde{G}(t) = \frac{1}{2}|I_n - tA|\mathrm{tr}(tA[\mathrm{adj}(I_n - tA)]) + \frac{1}{2}\mu'[\mathrm{adj}(I_n - tA)](tA)[\mathrm{adj}(I_n - tA)]\mu, \qquad (22)$$

a polynomial of degree $2n$, and

$$\tilde{E}(t) = |I_n - tA|^2, \qquad (23)$$

also of degree $2n$. In view of Lemma 1 we may state:

THEOREM 1. *The moments of a quadratic form $q = z'Az$, with $z \sim N(\mu, I_n)$, satisfy **exactly the same recurrence relations** — those given in (6) and (7) — whether μ is zero or not. In the central case $\tilde{D}(t) = D(t)$, $\tilde{P}(t) = P(t)/2$, and $\tilde{E}(t) = E(t)$, while in the noncentral case $\tilde{D}(t)$, $\tilde{P}(t)$, and $\tilde{E}(t)$ are as in (19), (21) and (23), respectively.*

Note again that, in applying the result of Lemma 1, the $\tilde{g}_k$ may be computed indirectly from the $\tilde{p}_k$ and $\tilde{e}_k$ by using the identity $\tilde{G}(t) = \tilde{P}(t)\tilde{E}(t)$ (rather than directly from the expansion of $\tilde{G}(t)$ in (22)). Also, since $\tilde{E}(t) = E(t)E(t)$,

$$\tilde{e}_k = \sum_{j=0}^{\min[k,n]} e_j e_{k-j},$$

where the e_k are the elementary symmetric functions of the eigenvalues of $-A$.

A second expression for the moments of q that also leads to a simple recursion may be obtained as follows. Let

$$\phi(t) = \mu'(I_n - tA)^{-1}\mu - \mu'\mu = \sum_{i=1}^{\infty} \eta_i t^i,$$

where $\eta_i = \mu' A^i \mu$, and define functions $a_{r,l}$ by the equation

$$a_{r,l} = [t^r]D(t)\phi(t)^l. \qquad (24)$$

Note that $a_{r,0} = d_r(A)$, and that the lowest-order term in $\phi(t)^l$ is t^l, so that $a_{r,l} = 0$ for $l > r$. We then have, from the generating function for $d_{r,k}(A, \mu\mu')$ in (16),

LEMMA 4. *With the functions $a_{r,l}$ as defined by (24),*

$$d_{r,k}(A, \mu\mu') = \frac{\left(\frac{1}{2}\right)_k}{k!} \sum_{l=0}^{k} \binom{k}{l} \delta^{k-l} a_{r,l},$$

where $\delta = \eta_0 = \mu'\mu$.

14

The proof of Lemma 4 is given in Appendix. Using this result in (19) and simplifying we obtain the formula:

$$\mu_k = 2^k k! [t^k] D(t) \exp\left(\frac{\phi(t)}{2}\right) = 2^k k! \sum_{l=0}^{k} \frac{a_{k,l}}{l! 2^l}.$$

This is evidently simpler than (18) in that the sum is finite. In addition, though, the $a_{k,l}$ themselves satisfy a very simple recurrence relation. To see this, simply note that

$$D(t)\phi(t)^l = \left[D(t)\phi(t)^{l-1}\right]\phi(t) = \left(\sum_{j=0}^{\infty} a_{j,l-1} t^j\right)\left(\sum_{i=1}^{\infty} \eta_i t^i\right).$$

Equating coefficients of like powers of t on both sides, and taking account of the fact that $a_{r,l} = 0$ for $r < l$, gives the following recursion for the $a_{r,l}$:

LEMMA 5. *For $l \geq 1$, the functions $a_{r,l}$ defined by (24) satisfy the recursion:*

$$a_{r,l} = \sum_{j=1}^{r-l+1} \eta_j a_{r-j,l-1}, \tag{25}$$

where $\eta_j = \mu' A^j \mu$ and we have the boundary conditions $a_{r,0} = d_r(A)$.

The functions $a_{r,l}$ will also be useful later in Section 5 where some low-order cases are given explicitly.

3.3 Special Cases: Repeated Eigenvalues and the Partially Central Case

The recurrence relations for moments given so far hold for any values of the eigenvalues of A, and any value of μ. However, some further improvement is possible if either some eigenvalues of A occur with multiplicity greater than one, and/or the noncentrality present is of dimension lower than n. To see this, first let $A = H\Lambda H'$, where $\Lambda = \text{Diag}(\lambda_1, \cdots, \lambda_n)$ is a diagonal matrix containing the eigenvalues of A, and $H = [h_1, \ldots, h_n]$ is an orthogonal matrix of the corresponding eigenvectors. Using this decomposition, we can write

$$q = z' A z = z' H \Lambda H' z = \tilde{z}' \Lambda \tilde{z},$$

where $\tilde{z} = H'z \sim N(H'\mu, I_n) = N(\tilde{\mu}, I_n)$, say.

Now, suppose that the eigenvalues λ_i are not distinct, but that the $s \leq n$ distinct eigenvalues λ_i occur with multiplicities n_i, where $n = \Sigma_{i=1}^s n_i$. This setup occurs naturally in the context of much-studied statistics of the form

$$q = \lambda_1 q_1 + \cdots + \lambda_s q_s,$$

where the q_i are independent noncentral $\chi^2_{n_i}(\delta_i)$ random variables (see Ruben (1962) and Press (1966)). Letting $\tilde{z}_i \sim N(\tilde{\mu}_i, I_{n_i})$ denote the sub-vector of $\tilde{z}$ associated with λ_i, we set $\delta_i = \tilde{\mu}_i' \tilde{\mu}_i$ for $i = 1, \ldots, s$. We wish to consider the case where, in addition to the possibility of repeated eigenvalues, some of the noncentrality parameters δ_i may also vanish. Without loss of generality, we assume that $\delta_i \neq 0$ for $i = 1, \ldots, r$, and $\delta_i = 0$ for $i = r+1, \ldots, s$.

With these assumptions and notation the MGF of $q/2$ in (19) becomes:

$$\tilde{D}(t) = \left[\prod_{i=1}^s (1 - t\lambda_i)^{-\frac{n_i}{2}} \right] \exp\left(\frac{1}{2} \sum_{i=1}^r \frac{t\delta_i\lambda_i}{1 - t\lambda_i} \right).$$

Thus, defining $\tilde{P}(t)$ as in (4) again, we have from (21)

$$\tilde{P}(t) = \frac{1}{2} \left[\sum_{i=1}^s \frac{tn_i\lambda_i}{1 - t\lambda_i} + \sum_{i=1}^r \frac{t\delta_i\lambda_i}{(1 - t\lambda_i)^2} \right] = \frac{1}{2} \sum_{k=1}^\infty \left(\sum_{i=1}^s n_i\lambda_i^k + k \sum_{i=1}^r \delta_i\lambda_i^k \right) t^k.$$

Hence, in this case,

$$\tilde{p}_k = \frac{1}{2} \left(\sum_{i=1}^s n_i\lambda_i^k + k \sum_{i=1}^r \delta_i\lambda_i^k \right), \tag{26}$$

and the recursion (6) applies with these $\tilde{p}_k$. However, as before, $\tilde{P}(t)$ is a rational polynomial with both denominator polynomial

$$\tilde{E}(t) = \left(\prod_{i=1}^r (1 - t\lambda_i)^2 \right) \left(\prod_{i=r+1}^s (1 - t\lambda_i) \right) = \sum_{i=0}^{r+s} \tilde{e}_i t^i, \tag{27}$$

say, and numerator polynomial

$$\tilde{G}(t) = \frac{1}{2} \left[\left(\prod_{i=1}^r (1 - t\lambda_i) \right) \left(\sum_{i=1}^s tn_i\lambda_i \prod_{\substack{j=1\\j\neq i}}^s (1 - t\lambda_j) \right) + \left(\prod_{i=r+1}^s (1 - t\lambda_i) \right) \left(\sum_{i=1}^r t\delta_i\lambda_i \prod_{\substack{j=1\\j\neq i}}^r (1 - t\lambda_j)^2 \right) \right]$$

of finite degree. Applying Lemma 1 we have:

THEOREM 2. *Suppose A has s distinct eigenvalues λ_1, ..., λ_s, with multiplicities $n_1, \ldots, n_s$ ($n_1 + \cdots + n_s = n$), respectively, and there are $r \leq s$ non-vanishing noncentrality parameters δ_i (as defined above). Then $\tilde{d}_k$ may be computed from the short recursion given in (7), with the $\tilde{p}_k$ as given in (26), and the $\tilde{e}_k$ as defined by (27). The recursion has length at most $r + s$ (rather than $2n$ as in the case where $r = s = n$).*

Remark 4. *To illustrate the improvement afforded by this new recursive algorithm, we consider an example with $A = I_n$, so that $q = z'z \sim \chi_n^2(\delta)$ is a noncentral chi-square variate. Using (20), we obtain the following recurrence relation for the $\mu_k \equiv E[q^k]$:*

$$\mu_k = \frac{1}{2k} \sum_{i=1}^{k} (k - i + 1)_i 2^i (n + i\delta)\mu_{k-i} \quad \text{for } k > 0.$$

However, applying Theorem 2, we obtain the following much more efficient recurrence relation:

$$\mu_k = (4k + \delta + n - 4)\mu_{k-1} - 2(k - 1)(2k + n - 4)\mu_{k-2} \quad \text{for } k > 1,$$

with the boundary conditions $\mu_0 = 1$ and $\mu_1 = n + \delta$.

Remark 5. *For the special case of $\mu = 0_n$, i.e., $z \sim N(0_n, I_n)$, we have $\tilde{d}_k = d_k(A)$, where $d_k(A)$ is the normalized top-order zonal polynomial. For this case, Theorem 2 gives a more efficient short recurrence relation for top-order zonal polynomials when some of the eigenvalues of A are repeated (see HKW, Section 2.3).*

4.   SECOND APPLICATION: PRODUCT MOMENTS OF SEVERAL QUADRATIC FORMS

Let A_1 to A_r be r $n \times n$ real symmetric matrices, and let $q_i = z'A_i z$, $i = 1, \ldots, r$, denote the variates of interest, with $z \sim N(\mu, I_n)$. Explicit expressions of μ_κ have, at least for low-order cases, long been available in the statistics literature. However, most of the existing work expresses μ_κ as a sum of various products of the traces of $|\kappa|$ matrices related to A_i's and are extremely inefficient for computational purposes. Kan (2008) provides a review of this literature, and a discussion of why current methods are impractical for computing μ_κ even for moderately large $|\kappa|$. See also Mathai and Provost (1992) for an excellent review of quadratic forms in random variables.

A straightforward generalization of the result in Section 3.1 yields an expression for the product moments μ_κ in terms of the normalized Davis polynomials d_κ. And, the results in Section 2.2 give both long and short recursions for their computation that are extremely efficient.

4.1 Explicit Formulae

Direct expansion of the MGF

$$M_{q_1,\dots,q_r}(\tau) = E\left[\exp(\tau_1 q_1 + \cdots + \tau_r q_r)\right],$$

together with Lemma 3, gives, on using equation (48) from HKW:

$$M_{q_1,\dots,q_r}(\tau) = \exp\left(-\frac{\mu'\mu}{2}\right) \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{2^k}{2^j \left(\frac{1}{2}\right)_j} \sum_{|\kappa|=k} d_{\kappa,j}(A_1,\dots,A_r,\mu\mu')\tau^\kappa,$$

so that

$$\mu_\kappa = E\left[q_1^{k_1} q_2^{k_2} \cdots q_r^{k_r}\right] = 2^k \kappa! \exp\left(-\frac{\mu'\mu}{2}\right) \sum_{j=0}^{\infty} \frac{1}{\left(\frac{1}{2}\right)_j 2^j} d_{\kappa,j}(A_1,\dots,A_r,\mu\mu'),$$

again generalizing the result for the central case:

$$\mu_\kappa = 2^k \kappa! d_\kappa(A_1,\dots,A_r).$$

The recursions presented in HKW for the invariant polynomials d_κ provide one way of computing these moments. In addition, Kan (2008) presents a much more efficient method for computing μ_κ. Proposition 4 of Kan (2008) shows that

$$\mu_\kappa = \frac{1}{k!} \sum_{0 \leq \nu \leq \kappa} (-1)^{|\nu|} \binom{\kappa}{\nu} E[(z'B_\nu z)^k], \tag{28}$$

where $B_\nu = \sum_{i=1}^{r} \left(\frac{k_i}{2} - \nu_i\right) A_i$, $k = |\kappa|$, and

$$\binom{\kappa}{\nu} = \frac{\kappa!}{\nu!(\kappa-\nu)!}.$$

As noted in Kan (2008), half of the terms on the right hand side of (28) are repeated, so one can compute μ_κ by computing the k-th moments of $\lfloor \prod_{i=1}^{r}(k_i+1)/2 \rfloor$ different quadratic forms in z, where $\lfloor x \rfloor$ stands for the integral part of x. Kan (2008) suggests using the recurrence relation (20) to compute $E[(z'B_\nu z)^k]$. With the new recursive algorithm given above, we can now significantly

improve the computation speed of $E[(z'B_\nu z)^k]$, especially when k is large. Although, with this procedure, (28) is quite efficient, there are circumstances where we still prefer to use a recurrence relation to compute μ_κ. This is particularly so if we need to compute not just a single μ_κ but require all μ_ν with $0 \leq \nu \leq \kappa$. In addition, when n or r is small, the new recursive algorithm based on the result in Section 2.2 is very short, and it significantly dominates (28) in terms of computation speed.

4.2   Recursions Long and Short

The joint moment generating function of $(q_1, \ldots, q_r)$ may also be written as:

$$M_{q_1,\ldots,q_r}(\tau) = |I_n - 2A(\tau)|^{-\frac{1}{2}} \exp\left(\frac{\mu'(I_n - 2A(\tau))^{-1}\mu}{2} - \frac{\mu'\mu}{2}\right) = \sum_{k=0}^{\infty} \sum_{|\kappa|=k} \frac{\mu_\kappa}{\kappa!}\tau^\kappa,$$

where $A(\tau) = \tau_1 A_1 + \cdots + \tau_r A_r$ (see, for example, Lemma 5 of Magnus (1986)).

Let

$$\tilde{D}(t) = |I_n - A(t)|^{-\frac{1}{2}} \exp\left(\frac{\mu'(I_n - A(t))^{-1}\mu}{2} - \frac{\mu'\mu}{2}\right) = \sum_{k=0}^{\infty} \sum_{|\kappa|=k} \tilde{d}_\kappa t^\kappa, \tag{29}$$

where $A(t) = t_1 A_1 + \cdots + t_r A_r$. The product moments themselves are given by:

$$\mu_\kappa = E[q_1^{k_1} q_2^{k_2} \cdots q_r^{k_r}] = 2^{|\kappa|}\kappa!\tilde{d}_\kappa.$$

Defining $\tilde{P}(t)$ as in (9), we have

$$\tilde{P}(t) = \frac{1}{2}\sum_{k=1}^{\infty}[\text{tr}(A(t)^k) + k\mu'A(t)^k\mu] = \sum_{k=1}^{\infty} \sum_{|\kappa|=k} \tilde{p}_\kappa t^\kappa, \tag{30}$$

where, for $k = |\kappa|$,

$$\tilde{p}_\kappa = \frac{1}{2}[t^\kappa][\text{tr}(A(t)^k) + k\mu'A(t)^k\mu].$$

We can also write $\tilde{P}(t)$ in the form

$$\tilde{P}(t) = \frac{1}{2}\text{tr}(A(t)(I_n - A(t))^{-1}) + \frac{1}{2}\mu'(I_n - A(t))^{-1}A(t)(I_n - A(t))^{-1}\mu,$$

so it is clear that this satisfies the hypotheses of Lemma 2, with $m = 2n$. Defining

$$\tilde{E}(t) = |I_n - A(t)|^2 = \sum_{k=0}^{2n} \sum_{|\kappa|=k} \tilde{e}_\kappa t^\kappa, \tag{31}$$

19

we see that both $\tilde{E}(t)$ and

$$\tilde{G}(t) = \tilde{E}(t)\tilde{P}(t) = \sum_{k=1}^{2n} \sum_{|\kappa|=k} \tilde{g}_\kappa t^\kappa \tag{32}$$

are polynomials of degree $2n$ in t.

In view of the results in Section 2.2, we can use (11), (12) and (14) to obtain the following three recurrence relations for the functions $\tilde{d}_\kappa$ defined by (29):

THEOREM 3. *Using $\tilde{p}_\kappa$, $\tilde{e}_\kappa$ and $\tilde{g}_\kappa$ as defined by (30), (31) and (32), the $\tilde{d}_\kappa$ in (29) can be recursively obtained from one of the following recurrence relations:*

$$\tilde{d}_\kappa = \frac{1}{k} \sum_{i=1}^{k} \sum_{\substack{|\nu|=i, \\ \nu \leq \kappa}} \tilde{p}_\nu \tilde{d}_{\kappa-\nu}, \tag{33}$$

$$\tilde{d}_\kappa = \frac{1}{k_j} \sum_{i=1}^{k} \sum_{\substack{|\nu|=i, \\ \nu \leq \kappa}} \frac{\nu_j}{i} \tilde{p}_\nu \tilde{d}_{\kappa-\nu} \quad \text{when } k_j > 0, \tag{34}$$

$$\tilde{d}_\kappa = \sum_{i=1}^{\min[k,2n]} \sum_{\substack{|\nu|=i, \\ \nu \leq \kappa}} \left[\frac{c_\nu}{k} - \tilde{e}_\nu \right] \tilde{d}_{\kappa-\nu}, \tag{35}$$

where $c_\nu = |\nu|\tilde{e}_\nu + \tilde{g}_\nu$, and we have the boundary condition $\tilde{d}_0 = 1$.

In contrast to the recurrence relations (33) and (34) which are in terms of the p_κ, our new recurrence relation (35) only involves the $\tilde{e}_\kappa$ and $\tilde{g}_\kappa$, and these vanish for $|\kappa| > 2n$. Regardless of the value of κ, (35) suggests that $\tilde{d}_\kappa$ can be expressed as a linear combination of at most $(2n+r)!/[(2n)!r!] - 1$ other $\tilde{d}_\nu$'s with $\nu < \kappa$. Therefore, (35) can provide a significant improvement over (33) and (34) when k_i's are large.

In order to use the above recursive algorithms to compute $\tilde{d}_\kappa$, we need to first obtain the coefficients $\tilde{p}_\kappa$ and $\tilde{e}_\kappa$. When n is very small, we can use a symbolic mathematics program to compute $\tilde{p}_\kappa$ and $\tilde{e}_\kappa$. However, this is extremely time consuming even when n is only moderately large. Therefore, it is crucial that we have efficient numerical algorithms for computing the $\tilde{p}_\kappa$ and $\tilde{e}_\kappa$. HKW provide an efficient method for computing the coefficients of t^k in the expansion of $\text{tr}(A(t)^k)$, which then allows us to easily obtain the $\tilde{e}_\kappa$. In addition, their algorithm can be extended in a straightforward manner to compute the coefficients of t^κ in the expansion of $\mu' A(t)^k \mu$. Therefore, both the $\tilde{e}_\kappa$ and $\tilde{p}_\kappa$ can be efficiently computed by the methods described in HKW.[2]

20

5. FINAL APPLICATION: RATIOS OF POWERS OF QUADRATIC FORMS

In this section we give results for the more complicated problem of evaluating expectations of the form

$$\mu_s^r = E\left[\frac{(z'Az)^r}{(z'Bz)^s}\right], \tag{36}$$

where A is a symmetric $n \times n$ matrix, B is a positive definite $n \times n$ matrix, $z \sim N(\mu, I_n)$, r is a nonnegative integer and s is a positive real number. We shall assume throughout that the largest eigenvalue of A is positive (i.e., that A is not negative definite). If A is negative definite the results to follow can be applied to $(-1)^r \mu_s^r$, rather than μ_s^r itself. It is easy to show that the expectation in (36) exists if and only if $\frac{n}{2} + r > s$, and we shall assume that this condition is satisfied throughout this section.

Many estimators in statistics take the form of ratio of quadratic form in normal random variables. As a result, the problem of computing expectations of the form given in (36) has attracted the attention of many researchers. Most of the results in this literature are based on a lemma due to Sawa (1972), and present formulae that take the form of a one-dimensional integral that must be evaluated numerically. For the development of this type of formula, see the excellent papers of Magnus (1986) and Meng (2005) and the references therein. We shall briefly describe the result here before presenting our own results.

Let $B = P\Lambda P'$, where Λ is a diagonal matrix of the eigenvalues of B, and P is an orthogonal matrix of the corresponding eigenvectors. By combining the results of Theorem 6 of Magnus (1986) and Lemma 1 of Meng (2005), we obtain

$$\mu_s^r = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} |\Delta| \exp\left(\frac{\mu' P[(I_n + 2t\Lambda)^{-1} - I_n]P'\mu}{2}\right) E[(w'Rw)^r] \mathrm{d}t, \tag{37}$$

where $\Delta = (I_n + 2t\Lambda)^{-\frac{1}{2}}$, $R = \Delta P'AP\Delta$, and $w \sim N(\Delta P'\mu, I_n)$. Currently, this is the only practical method that can be used for numerical evaluation of μ_s^r. However, there are two problems associated with the use of this formula. The first is in the computation of $E[(w'Rw)^r]$, which we have discussed in Section 3.1 above. As we have seen, both explicit formulae for this term, and efficient recursions for evaluating it, are available. However, because R is a function of t, this expectation must be evaluated many times. The second problem is that it is difficult to control the

21

accuracy of the numerical integration: there is no general result in the literature that allows us to analyze and control the errors in the numerical integration of (37). For these reasons, we seek here a more efficient method for evaluating the μ_s^r based on the results in Section 2. Before doing so, we briefly describe the exact formulae that are available.

5.1 Explicit Formulae

Smith (1989) provides a very different expression for the μ_s^r. He shows that μ_s^r can be expressed as a doubly infinite series involving the top-order invariant polynomials d_κ. In our notation, his expression is:

$$\mu_s^r = \frac{2^{r-s}\beta^s r! \Gamma\left(\frac{n}{2}+r-s\right)e^{-\frac{\delta}{2}}}{\Gamma\left(\frac{n}{2}+r\right)} \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{(s)_j \left(\frac{n}{2}+r-s\right)_k}{2^k \left(\frac{1}{2}\right)_k \left(\frac{n}{2}+r\right)_{j+k}} d_{r,j,k}\left(A, I_n - \beta B, \mu\mu'\right), \qquad (38)$$

reducing to

$$\mu_s^r = \frac{2^{r-s}\beta^s r! \Gamma\left(\frac{n}{2}+r-s\right)}{\Gamma\left(\frac{n}{2}+r\right)} \sum_{j=0}^{\infty} \frac{(s)_j}{\left(\frac{n}{2}+r\right)_j} d_{r,j}\left(A, I_n - \beta B\right), \qquad (39)$$

when $\mu = 0_n$. Here, $\delta = \mu'\mu$, and β is a constant that satisfies $0 < \beta < 2/b_{\max}$, with $b_{\max}$ the largest eigenvalue of B.[3] When $B = I_n$ we may choose $\beta = 1$, so that the sum on j in (38) vanishes, and we have the simpler result

$$\mu_s^r = \frac{2^{r-s} r! \Gamma\left(\frac{n}{2}+r-s\right)e^{-\frac{\delta}{2}}}{\Gamma\left(\frac{n}{2}+r\right)} \sum_{k=0}^{\infty} \frac{\left(\frac{n}{2}+r-s\right)_k}{2^k \left(\frac{1}{2}\right)_k \left(\frac{n}{2}+r\right)_k} d_{r,k}\left(A, \mu\mu'\right). \qquad (40)$$

Finally, when both $B = I_n$ and $\mu = 0_n$ we have

$$\mu_s^r = \frac{2^{r-s} r! \Gamma(\frac{n}{2}+r-s)}{\Gamma(\frac{n}{2}+r)} d_r\left(A\right),$$

a multiple of the corresponding moment of q dealt with in Section 3.1 above. Smith (1993) makes an attempt to use these formulae to compute the moments for the case $r = 1$, and with either $\mu = 0_n$ or $B = I_n$, but there has been great difficulty in using this formula for computation in the general case.

Recently, HKW have given an efficient recursive algorithm for computing the top-order invariant polynomials. In principle, their algorithm can be used to compute the $d_{r,j,k}(A, I_n - \beta B, \mu\mu')$, and the moments approximated by truncating the double series in (38) at some suitable point. However, this process is extremely inefficient. As a result, HKW focus only on the simpler special case of $\mu = 0_n$,

when only a singly infinite series of top-order invariant polynomials with two matrix arguments is involved. In addition, for the case $\mu = 0_n$, HKW give an upper bound on the approximation error when truncating the infinite series at $j = M$. For the general case of $\mu \neq 0_n$, it remains a significant challenge to bound this truncation error.

5.2 New Formulae for the μ_s^r

To address these difficulties, in this section we provide two new formulae that greatly simplify the evaluation of the μ_s^r for the general case when $\mu \neq 0_n$. Unlike Smith's formula, both of our two new expressions involve only a singly infinite series, and the coefficients are easily obtained using a fast recursive algorithm based on Lemma 2. In addition, we also provide error control, so we can compute the expectation up to any desired level of accuracy.

The results we develop are based on the following formal representation for μ_s^r — which is also the basis of the formula (37) (see Sawa (1972) or Cressie, Davis, Folks, and Policello (1981)):

$$\mu_s^r = \frac{r![t^r]}{\Gamma(s)} \int_0^\infty x^{s-1} M_{q_1,q_2}(t, -x) \mathrm{d}x,$$

where $M_{q_1,q_2}(t_1, t_2) = E[\exp(t_1 z' A z + t_2 z' B z)]$ is the joint moment generating function of $q_1 = z' A z$ and $q_2 = z' B z$ when $z \sim N(\mu, I_n)$, i.e.,

$$M_{q_1,q_2}(t_1, t_2) = |I_n - 2t_1 A - 2t_2 B|^{-\frac{1}{2}} \exp\left(\frac{\mu'(I_n - 2t_1 A - 2t_2 B)^{-1}\mu}{2} - \frac{\delta}{2}\right),$$

where $\delta = \mu'\mu$. Our starting point is thus the following integral expression for μ_s^r:

$$\mu_s^r = \frac{r![t^r]}{\Gamma(s)} \int_0^\infty x^{s-1} |I_n - 2tA + 2xB|^{-\frac{1}{2}} \exp\left(\frac{\mu'(I_n - 2tA + 2xB)^{-1}\mu}{2} - \frac{\delta}{2}\right) \mathrm{d}x.$$

We discuss the existence of the integral as necessary below. For convenience later we transform to $y = x/\beta$ in the integral, with β a positive constant to be chosen. This leads to the following expression for μ_s^r:

$$\mu_s^r = \frac{\beta^s r![t^r]}{\Gamma(s)} \int_0^\infty y^{s-1} |I_n - 2tA + 2y\beta B|^{-\frac{1}{2}} \exp\left(\frac{\mu'(I_n - 2tA + 2y\beta B)^{-1}\mu}{2} - \frac{\delta}{2}\right) \mathrm{d}y. \tag{41}$$

We derive two different expressions for μ_s^r from (41), each suited to different circumstances. These arise from slightly different ways of rewriting the matrix $I_n - 2tA + 2y\beta B$.

To obtain the first result of this type, observe that we can write

$$I_n - 2tA + 2y\beta B = (1 + 2y)\left(I_n - \frac{2t}{1 + 2y}A - \frac{2y}{1 + 2y}\tilde{B}\right), \tag{42}$$

where $\tilde{B} = I_n - \beta B$. Define, for fixed μ, functions $h_{r,j}(A_1, A_2)$ by the generating function[4]

$$
\begin{aligned}
H_{A_1, A_2}(t_1, t_2) & = |I - t_1 A_1 - t_2 A_2|^{-\frac{1}{2}} \exp\left(\frac{(1 - t_2)\mu'(I - t_1 A_1 - t_2 A_2)^{-1}\mu}{2} - \frac{\mu'\mu}{2} \right) \\
& = \sum_{r=0}^{\infty} \sum_{j=0}^{\infty} h_{r,j}(A_1, A_2) t_1^r t_2^j,
\end{aligned} \tag{43}
$$

In fact, the $h_{r,j}(A_1, A_2)$ may be expressed explicitly in terms of the $d_{r,j,l}(A_1, A_2, \mu\mu')$, as described in the following lemma, which is proved in Appendix:

LEMMA 6. *The functions $h_{r,j}(A_1, A_2)$ defined by (43) can be expressed in terms of the invariant polynomials d_κ as follows:*

$$
h_{r,j}(A_1, A_2) = \exp\left(-\frac{\delta}{2} \right) \sum_{m=0}^{\infty} \frac{1}{2^m m!} \sum_{l=0}^{j} \frac{(-1)^l}{2^l l!} \frac{(l+m)!}{\left(\frac{1}{2}\right)_{l+m}} d_{r,j-l,l+m}(A_1, A_2, \mu\mu').
$$

This result is useful for proving the convergence of the infinite series involved in the result to follow, but not directly for computation purposes.

Transforming now to $b = 2y/(1 + 2y)$ in (42), the integrand in (41) has the form

$$
2^{-s} b^{s-1} (1-b)^{\frac{n}{2} - s - 1} \sum_{r=0}^{\infty} \sum_{j=0}^{\infty} (2t(1-b))^r b^j h_{r,j}(A, \tilde{B}),
$$

so that the coefficient of t^r is

$$
2^{r-s} b^{s-1} (1-b)^{\frac{n}{2} + r - s - 1} \sum_{j=0}^{\infty} b^j h_{r,j}(A, \tilde{B}).
$$

We show in the Appendix that term-by-term integration can be justified when $\frac{n}{2} + r > s$ and $0 < \beta < 2/b_{\max}$. We may therefore state:

THEOREM 4. *For (r, s) satisfying $\frac{n}{2} + r > s$, and any choice of β satisfying $0 < \beta < 2/b_{\max}$, where $b_{\max}$ is the largest eigenvalue of B, we have the following expression for the moments μ_s^r:*

$$
\mu_s^r = \frac{2^{r-s} \beta^s r! \Gamma\left(\frac{n}{2} + r - s\right)}{\Gamma\left(\frac{n}{2} + r\right)} \sum_{j=0}^{\infty} \frac{(s)_j}{\left(\frac{n}{2} + r\right)_j} h_{r,j}(A, I_n - \beta B). \tag{44}
$$

Note that when $\mu = 0_n$, $h_{r,j}(A, I_n - \beta B) = d_{r,j}(A, I_n - \beta B)$, and (44) specializes to (39).

To obtain the second expression of this type, note that we can alternatively write

$$I_n - 2tA + 2y\beta B = \beta(1 + 2y)B^{\frac{1}{2}}\left(I_n - \frac{2t}{1 + 2y}\hat{A} - \frac{1}{1 + 2y}\hat{B}\right)B^{\frac{1}{2}},$$

where $\hat{A} = B^{-\frac{1}{2}}AB^{-\frac{1}{2}}/\beta$ and $\hat{B} = I_n - (\beta B)^{-1}$. Changing the variable of integration in (41) to $b = 1/(1 + 2y)$, $0 < b < 1$, we then have

$$
\begin{aligned}
\mu_s^r &= \frac{\beta^s r! e^{-\frac{\delta}{2}}[t^r]}{2^s \Gamma(s)|\beta B|^{\frac{1}{2}}} \int_0^1 b^{\frac{n}{2}-s-1}(1-b)^{s-1}|I_n - 2tb\hat{A} - b\hat{B}|^{-\frac{1}{2}} \\
&\quad \times \exp\left(\frac{b\hat{\mu}'(I_n - 2tb\hat{A} - b\hat{B})^{-1}\hat{\mu}}{2}\right) db,
\end{aligned}
\tag{45}
$$

where $\hat{\mu} = (\beta B)^{-\frac{1}{2}}\mu$.

Now, for fixed $\hat{\mu}$, define functions $\tilde{h}_{r,j}(A_1, A_2)$ by the generating function

$$
\begin{aligned}
\tilde{H}_{A_1,A_2}(t_1, t_2) &= |I_n - t_1 A_1 - t_2 A_2|^{-\frac{1}{2}}\exp\left(\frac{t_2\hat{\mu}'(I_n - t_1 A_1 - t_2 A_2)^{-1}\hat{\mu}}{2}\right) \\
&= \sum_{r=0}^{\infty}\sum_{j=0}^{\infty}\tilde{h}_{r,j}(A_1, A_2)t_1^r t_2^j.
\end{aligned}
\tag{46}
$$

These functions $\tilde{h}_{r,j}(A_1, A_2)$ can again be expressed directly in terms of the $d_{r,j,l}(A_1, A_2, \hat{\mu}\hat{\mu}')$, as described in the following lemma:

LEMMA 7. *The functions $\tilde{h}_{r,j}(A_1, A_2)$ defined by (46) are given by:*

$$\tilde{h}_{r,j}(A_1, A_2) = \sum_{l=0}^{j}\frac{1}{2^l\left(\frac{1}{2}\right)_l}d_{r,j-l,l}(A_1, A_2, \hat{\mu}\hat{\mu}').$$

Again, this result is useful analytically, but not directly for computational work.

Now, the integrand in (45) has the form

$$b^{\frac{n}{2}-s-1}(1-b)^{s-1}\sum_{r=0}^{\infty}\sum_{j=0}^{\infty}(2tb)^r b^j \tilde{h}_{r,j}(\hat{A}, \hat{B}),$$

so that the coefficient of t^r is

$$2^r b^{\frac{n}{2}+r-s-1}(1-b)^{s-1}\sum_{j=0}^{\infty}b^j \tilde{h}_{r,j}(\hat{A}, \hat{B}).$$

We show in the Appendix that term-by-term integration is justified if $\frac{n}{2} + r > s$ and $\beta > 1/(2b_{\min})$, where $b_{\min}$ is the smallest eigenvalue of B. We may therefore state:

25

THEOREM 5. *For (r, s) satisfying $\frac{n}{2} + r > s$, and any choice of β satisfying $\beta > 1/(2b_{\min})$, where $b_{\min}$ is the smallest eigenvalue of B, we have the following expression for the moments μ_s^r :*

$$\mu_s^r = \frac{2^{r-s}\beta^s r! e^{-\frac{\delta}{2}} \Gamma\left(\frac{n}{2} + r - s\right)}{|\beta B|^{\frac{1}{2}} \Gamma\left(\frac{n}{2} + r\right)} \sum_{j=0}^{\infty} \frac{\left(\frac{n}{2} + r - s\right)_j}{\left(\frac{n}{2} + r\right)_j} \tilde{h}_{r,j}(\hat{A}, \hat{B}), \tag{47}$$

where the functions $\tilde{h}_{r,j}$ are defined by (46), $\hat{A} = B^{-\frac{1}{2}} A B^{-\frac{1}{2}}/\beta$, and $\hat{B} = I_n - (\beta B)^{-1}$.

Note that when $\mu = 0_n$, $\tilde{h}_{r,j}(\hat{A}, \hat{B}) = d_{r,j}(\hat{A}, \hat{B})$, so (47) gives us the following new expression for μ_s^r in the case $\mu = 0_n$:

$$\mu_s^r = \frac{2^{r-s}\beta^s r! \Gamma\left(\frac{n}{2} + r - s\right)}{|\beta B|^{\frac{1}{2}} \Gamma\left(\frac{n}{2} + r\right)} \sum_{j=0}^{\infty} \frac{\left(\frac{n}{2} + r - s\right)_j}{\left(\frac{n}{2} + r\right)_j} d_{r,j}(\hat{A}, \hat{B}).$$

The expressions for μ_s^r given in Theorem 4 and 5 seem superficially similar to Smith's expression (38) given above. However, there are two important simplifications that make (44) and (47) much more efficient for computation purposes than is (38). The first is simply that these new expressions involve only a singly infinite series, rather than the doubly infinite series present in (38). The second, and more important, aspect of the results is that both of the generating functions $H_{A_1,A_2}(t_1, t_2)$ and $\tilde{H}_{A_1,A_2}(t_1, t_2)$ satisfy the hypotheses of Lemma 2, so that short recursions are available for both the the $h_{r,j}$ and the $\tilde{h}_{r,j}$. We describe these in more detail below, but first give some additional results for the special case in which $B = I_n$, when the above results simplify considerably.

5.3 The Special Case: $B = I_n$

The moments μ_s^r simplify considerably when $B = I_n$. Clearly, like the moments μ_k of $q = z'Az$, they depend only upon the matrices A and $\mu\mu'$, and in fact, like the μ_k, they can be concisely expressed in terms of the functions $a_{r,l}$ introduced at the end of Section 3.2. To see this, simply insert the result in Lemma 4 into (40) to obtain:

$$
\begin{aligned}
\mu_s^r &= \frac{2^{r-s}r! \Gamma\left(\frac{n}{2} + r - s\right) e^{-\frac{\delta}{2}}}{\Gamma\left(\frac{n}{2} + r\right)} \sum_{k=0}^{\infty} \frac{\left(\frac{n}{2} + r - s\right)_k}{2^k k! \left(\frac{n}{2} + r\right)_k} \sum_{l=0}^{k} \binom{k}{l} \delta^{k-l} a_{r,l} \\
&= \frac{2^{r-s}r! \Gamma\left(\frac{n}{2} + r - s\right) e^{-\frac{\delta}{2}}}{\Gamma\left(\frac{n}{2} + r\right)} \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \frac{\left(\frac{n}{2} + r - s\right)_{k+l} \delta^k}{2^{k+l} k! l! \left(\frac{n}{2} + r\right)_{k+l}} a_{r,l} \\
&= \frac{2^{r-s}r! \Gamma\left(\frac{n}{2} + r - s\right) e^{-\frac{\delta}{2}}}{\Gamma\left(\frac{n}{2} + r\right)} \sum_{l=0}^{r} \frac{\left(\frac{n}{2} + r - s\right)_l}{2^l l! \left(\frac{n}{2} + r\right)_l} {}_1F_1\left(\frac{n}{2} + r - s + l, \frac{n}{2} + r + l; \frac{\delta}{2}\right) a_{r,l}
\end{aligned}
$$

26

$$= 2^{r-s} r! \sum_{l=0}^{r} \frac{\Gamma\left(\frac{n}{2}+r-s+l\right)}{2^l l! \Gamma\left(\frac{n}{2}+r+l\right)} {}_1F_1\left(s, \frac{n}{2}+r+l; -\frac{\delta}{2}\right) a_{r,l},$$

where the third line follows from the fact that $a_{r,l} = 0$ for $l > r$, and the last step follows from the Kummer formula for the confluent hypergeometric function: $e^{-z} {}_1F_1(a, c; z) = {}_1F_1(c - a, c; -z)$. We may therefore state:

THEOREM 6. *When $B = I_n$,*

$$\mu_s^r = 2^{r-s} r! \sum_{l=0}^{r} \frac{\Gamma\left(\frac{n}{2}+r-s+l\right)}{2^l l! \Gamma\left(\frac{n}{2}+r+l\right)} {}_1F_1\left(s, \frac{n}{2}+r+l; -\frac{\delta}{2}\right) a_{r,l}, \tag{48}$$

where the functions $a_{r,l}$ are defined by (24), and satisfy the recursion (25).

Remark 6. *For $r = 0$, (48) is the inverse moment of a noncentral chi-squared distribution, and its expression was first provided by Bock, Judge, and Yancey (1984) when s is a positive integer. For $r = 1$, Smith (1993) uses a different approach to obtain the same expression as ours. For the case that $r = s$, Ghazal (1994) presents the results for $r = 1$ to 4. Our results are more general in that s can be an arbitrary positive real number and r can be any nonnegative integer.*

Although it is straightforward to evaluate the $a_{r,l}$ numerically, we present the explicit expressions of $a_{r,l}$ for $r = 1$ to 4 here for easy reference. Setting $\tau_i = \text{tr}(A^i)$ and $\eta_i = \mu' A^i \mu$, we have:

Table 1: The $a_{r,l}$ for $1 \le r \le 4$

l \ r	1	2	3	4
0	$\frac{\tau_1}{2}$	$\frac{\tau_1^2}{8} + \frac{\tau_2}{4}$	$\frac{\tau_1^3}{48} + \frac{\tau_1\tau_2}{8} + \frac{\tau_3}{6}$	$\frac{\tau_1^4}{384} + \frac{\tau_1^2\tau_2}{32} + \frac{\tau_2^2}{32} + \frac{\tau_1\tau_3}{12} + \frac{\tau_4}{8}$
1	η_1	$\frac{\tau_1\eta_1}{2} + \eta_2$	$\frac{\tau_1^2\eta_1}{8} + \frac{\tau_2\eta_1}{4} + \frac{\tau_1\eta_2}{2} + \eta_3$	$\frac{\tau_1^3\eta_1}{48} + \frac{\tau_1\tau_2\eta_1}{8} + \frac{\tau_1^2\eta_2}{8} + \frac{\tau_3\eta_1}{6} + \frac{\tau_2\eta_2}{4} + \frac{\tau_1\eta_3}{2} + \eta_4$
2	.	η_1^2	$\frac{\tau_1\eta_1^2}{2} + 2\eta_1\eta_2$	$\frac{\tau_1^2\eta_1^2}{8} + \frac{\tau_2\eta_1^2}{4} + \tau_1\eta_1\eta_2 + \eta_2^2 + 2\eta_1\eta_3$
3	.	.	η_1^3	$\frac{\tau_1\eta_1^3}{2} + 3\eta_1^2\eta_2$
4	.	.	.	η_1^4

5.4 Long and Short Recursions for the $h_{r,j}$ and $\tilde{h}_{r,j}$

We now show that the results in Section 2.2 provide both long and short recursions for the functions $h_{r,j}$ and $\tilde{h}_{r,j}$, and give the details for implementing these. Consider the generating function for the $h_{r,j}$, $H(t_1, t_2)$ as defined in (43), first. Defining $\tilde{P}(t)$ as in (9) (with $D(t)$ replaced by $H(t_1, t_2)$), we find that the associated function $\tilde{P}(t)$ is given by

$$
\begin{aligned}
\tilde{P}(t_1, t_2) &= t_1 \frac{\partial \ln H(t_1, t_2)}{\partial t_1} + t_2 \frac{\partial \ln H(t_1, t_2)}{\partial t_2} \\
&= \frac{1}{2} P(t) + \frac{1}{2}(1 - t_2)\mu'(I_n - t_1 A_1 - t_2 A_2)^{-2}\mu - \frac{1}{2}\mu'(I_n - t_1 A_1 - t_2 A_2)^{-1}\mu,
\end{aligned}
$$

where

$$
P(t) = \mathrm{tr}\left((t_1 A_1 + t_2 A_2)(I_n - t_1 A_1 - t_2 A_2)^{-1}\right) = \sum_{\substack{j=0 \\ j+k>0}}^{\infty} \sum_{k=0}^{\infty} p_{j,k} t_1^j t_2^k.
$$

This clearly satisfies the hypotheses of Lemma 2, with

$$
\tilde{E}(t) = |I_n - t_1 A_1 - t_2 A_2|^2 = \sum_{i=0}^{2n} \sum_{j=0}^{2n-i} \tilde{e}_{i,j} t_1^i t_2^j, \tag{49}
$$

and $\tilde{G}(t)$ both of degree $2n$.

Now, for fixed μ, define functions $\eta_{j,k}$ of matrices A_1, A_2 by

$$
\mu'(I_n - t_1 A_1 - t_2 A_2)^{-1}\mu = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \eta_{j,k} t_1^j t_2^k.
$$

Then,

$$\tilde{p}_{j,k} = \frac{1}{2}p_{j,k} + \frac{1}{2}(j+k)(\eta_{j,k} - \eta_{j,k-1}), \tag{50}$$

The corresponding function $\tilde{P}(t)$ for the case $D(t) = \tilde{H}(t_1, t_2)$, as defined in (46), is:

$$
\begin{aligned}
\tilde{P}(t_1, t_2) &= t_1 \frac{\partial \ln \tilde{H}(t_1, t_2)}{\partial t_1} + t_2 \frac{\partial \ln \tilde{H}(t_1, t_2)}{\partial t_2} \\
&= \frac{1}{2}P(t) + \frac{1}{2}t_2\hat{\mu}'(I_n - t_1 A_1 - t_2 A_2)^{-2}\hat{\mu},
\end{aligned}
$$

and again the hypotheses of Lemma 2 are satisfied, with $\tilde{E}(t) = |I_n - t_1 A_1 - t_2 A_2|^2$ again. In this case we find that

$$\tilde{p}_{j,k} = \frac{1}{2}p_{j,k} + \frac{1}{2}(j+k)\hat{\eta}_{j,k-1}, \tag{51}$$

where $\hat{\eta}_{j,k}$ are functions defined by

$$\hat{\mu}'(I_n - t_1 A_1 - t_2 A_2)^{-1}\hat{\mu} = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\hat{\eta}_{j,k}t_1^j t_2^k.$$

We may therefore state:

THEOREM 7. *(i) Using the boundary condition of $h_{0,0} = 1$, the functions $h_{r,j}$ defined by the generating function (43) may be generated by the long recursion given in (11), which has the form:*

$$h_{r,j} = \frac{1}{r+j}\sum_{\substack{k_1=0 \\ k_1+k_2>0}}^{r}\sum_{k_2=0}^{j}\tilde{p}_{k_1,k_2}h_{r-k_1,j-k_2}, \tag{52}$$

where $\tilde{p}_{k_1,k_2}$ are as in (50). Or, they may be more efficiently generated using the short recursion given in Lemma 2, which has the form:

$$h_{r,j} = \sum_{\substack{k_1=0 \\ 0<k_1+k_2\le 2n}}^{r}\sum_{k_2=0}^{j}\left[\frac{c_{k_1,k_2}}{r+j} - \tilde{e}_{k_1,k_2}\right]h_{r-k_1,j-k_2}, \tag{53}$$

with

$$c_{k_1,k_2} = (k_1 + k_2)\tilde{e}_{k_1,k_2} + \tilde{g}_{k_1,k_2},$$

and with the $\tilde{e}_{k_1,k_2}$ as in (49) and the $\tilde{g}_{k_1,k_2}$ determined by

$$\tilde{g}_{k_1,k_2} = \sum_{\substack{v_1=0 \\ v_1+v_2>0}}^{k_1}\sum_{v_2=0}^{k_2}\tilde{p}_{v_1,v_2}\tilde{e}_{k_1-v_1,k_2-v_2}.$$

(ii) Using the boundary condition of $\tilde{h}_{0,0} = 1$, the functions $\tilde{h}_{r,j}$ defined by the generating function (46) may be generated by exactly the same long and short recursions, as given in (52) and (53), respectively, except that in this case the $\tilde{p}_{k_1,k_2}$ are as in (51).

5.5   Truncation Errors

When using (44) or (47) to evaluate μ_s^r, we must in practice truncate the infinite series at $j = M$ for some value of M. In order to control the accuracy of the computation, we need to obtain an upper bound on the truncation error. For the presentation of our error bounds, we introduce the following notation. Suppose A is a symmetric matrix. We define $A^+ = A$ when A is positive semidefinite or when r is even, and $A^+ = P\Lambda^+ P'$ otherwise, where Λ^+ is a diagonal matrix of the absolute eigenvalues of A, and P is a matrix of the corresponding eigenvectors of A. With this notation, we now present an upper bound on the truncation error for (44).

THEOREM 8. *For any choice of β satisfying $0 < \beta \leq 1/b_{\max}$, where $b_{\max}$ is the largest eigenvalue of B, an upper bound on the approximation error of μ_s^r when truncating the infinite series in (44) at $j = M$ is given by*

$$
\left| \mu_s^r - \frac{2^{r-s}\beta^s r! \Gamma\left(\frac{n}{2} + r - s\right)}{\Gamma\left(\frac{n}{2} + r\right)} \sum_{j=0}^{M} \frac{(s)_j}{\left(\frac{n}{2} + r\right)_j} h_{r,j}(A, I_n - \beta B) \right|
$$
$$
\leq \frac{2^{r-s}\beta^s r! \Gamma\left(\frac{n}{2} + r - s\right)}{\Gamma\left(\frac{n}{2} + r\right)} \frac{(s)_{M+1}}{\left(\frac{n}{2} + r\right)_{M+1}} \left[\frac{e^{\frac{\bar{\delta}-\delta}{2}}\tilde{d}_r(\bar{A}, \bar{\mu})}{|\beta B|^{\frac{1}{2}}} - \sum_{j=0}^{M} \hat{h}_{r,j}(A^+, I_n - \beta B) \right],
$$

where $\bar{\mu} = \sqrt{2}(\beta B)^{-\frac{1}{2}}\mu$, $\bar{\delta} = \bar{\mu}'\bar{\mu}$, $\bar{A} = B^{-\frac{1}{2}}A^+ B^{-\frac{1}{2}}/\beta$, $\tilde{d}_r$ is defined as in (19), and the generating function of $\hat{h}_{j,k}(A_1, A_2)$ is given by

$$
\hat{H}_{A_1,A_2}(t_1, t_2) = |I_n - t_1 A_1 - t_2 A_2|^{-\frac{1}{2}} \exp\left(\frac{(1 + t_2)\mu'(I_n - t_1 A_1 - t_2 A_2)^{-1}\mu}{2} - \frac{\delta}{2} \right).
$$

Our bound on the truncation error for (47) is simpler, and it is given in the following Theorem.

THEOREM 9. *For any choice of β satisfying $\beta \geq 1/b_{\min}$ where $b_{\min}$ is the smallest eigenvalue of B, an upper bound on the approximation error of μ_s^r when truncating the infinite series in (47)*

at $j = M$ is given by

$$\left| \mu_s^r - \frac{2^{r-s}\beta^s r! e^{-\frac{\delta}{2}} \Gamma\left(\frac{n}{2} + r - s\right)}{|\beta B|^{\frac{1}{2}} \Gamma\left(\frac{n}{2} + r\right)} \sum_{j=0}^{M} \frac{\left(\frac{n}{2} + r - s\right)_j}{\left(\frac{n}{2} + r\right)_j} \tilde{h}_{r,j}\left(\hat{A}, \hat{B}\right) \right|$$

$$\leq \frac{2^{r-s}\beta^s r! e^{-\frac{\delta}{2}} \Gamma\left(\frac{n}{2} + r - s + M + 1\right)}{|\beta B|^{\frac{1}{2}} \Gamma\left(\frac{n}{2} + r + M + 1\right)} \left[|\beta B|^{\frac{1}{2}} e^{\frac{\delta}{2}} \tilde{d}_r(A^+, \mu) - \sum_{j=0}^{M} \tilde{h}_{r,j}(\bar{A}, \hat{B}) \right],$$

where $\tilde{d}_r$ is defined as in (19) and $\bar{A}$ is defined in Theorem 8.

With the results in Theorem 8 and 9, we can now approximate μ_s^r to any desired level of accuracy. However, it is not an easy matter to decide which one of these two algorithms to use for a given problem. The relative speed of convergence of (44) and (47) depends on A, B, μ, r, s, as well as the choice of β. For given values of A, B, μ, neither algorithm dominates the other for all values of (r, s). Our experience seems to indicate that (44) is more efficient when s is small whereas (47) is more efficient when s is large. Further analysis is required to better understand this issue, but we leave this topic for future research.

5.6 An Example

For illustrative purpose, we consider an example with $n = 20$, A a Toeplitz matrix with (i, j)th element given by $(|i - j| - 1)/n^2$, B a diagonal matrix with i-th diagonal element $b_{ii} = i/n^2$, and μ is set to be a vector of $\mu_i = i/n$ for $i = 1, \ldots, n$. Using the choice of $\beta = 1/b_{\max}$ for (44) and $\beta = 1/b_{\min}$ for the case of (47), Table 2 reports the value of μ_s^r for various combinations of r and s, with approximation errors less than 10^{-5}. The table also reports the number of required terms (M) to achieve the desired level of accuracy in parentheses, with the first number being the required terms when using (44) and the second number being the required terms when using (47).

From Table 2, we can observe that for a fixed r, the number of required terms (M) increases with s when using (44). In contrast, M decreases with s when using (47). For our particular example, the expression based on (44) generally converges much faster than (47) except when $s = 10$. Nevertheless, both expressions are very fast and efficient. Using either (44) or (47), it takes less than three seconds for our Matlab programs to generate Table 2 using a PC with an Intel E6600 CPU.

31

Table 2: Expectation of Ratio of Quadratic Forms in Noncentral Normal Vectors

The table presents $E[(z'Az)^r/(z'Bz)^s]$ for various values of r and s, where $z \sim N(\mu, I_n)$, $n = 20$, A is a Toeplitz matrix with its (i,j)th element as $a_{ij} = (|i - j| - 1)/n^2$, B is a diagonal matrix with its ith diagonal element as $b_{ii} = i/n^2$ and $\mu_i = i/n$. The approximation error is set to be less than 10^{-5} and the number of terms required to achieve this level of accuracy is reported in the parentheses, with the first number being the required terms when using (44) and the second number being the required terms when using (47).

	$s = 1$	$s = 2$	$s = 3$	$s = 4$	$s = 5$	$s = 10$
$r = 0$	1.42721	2.36909	4.67693	11.30111	34.72798	n/a
	(63, 487)	(91, 464)	(128, 444)	(176, 426)	(236, 409)	
$r = 1$	1.40950	1.91118	2.96700	5.36157	11.50669	7638.94030
	(69, 551)	(98, 525)	(135, 501)	(181, 479)	(239, 459)	(726, 383)
$r = 2$	4.19497	5.18942	7.28829	11.80941	22.53012	27925.79115
	(74, 580)	(102, 552)	(137, 527)	(179, 503)	(232, 482)	(660, 400)
$r = 3$	13.34410	14.79819	18.34967	25.75133	41.50710	8655.50979
	(86, 695)	(118, 662)	(156, 631)	(202, 602)	(256, 576)	(678, 470)
$r = 4$	59.03048	60.36432	68.43545	86.92433	125.28018	10856.79180
	(89, 703)	(118, 670)	(152, 640)	(192, 611)	(240, 584)	(606, 478)
$r = 5$	295.93344	279.52112	290.15474	333.89538	430.35843	14607.30704
	(108, 854)	(143, 815)	(183, 779)	(229, 744)	(282, 711)	(668, 575)
$r = 10$	6425021.47108	4505458.62224	3383790.18983	2734240.84284	2389287.33517	5009200.42040
	(151, 1147)	(185, 1103)	(220, 1061)	(258, 1020)	(300, 980)	(579, 807)

6.   CONCLUDING REMARKS

We have shown in this paper that, given a generating function for some objects of interest (moments, the coefficients in a series expansion, etc.), an associated generating function may be defined that induces a recurrence relation between the original objects of interest and a set of associated objects. This generalizes some known results relating moments and cumulants, and also results relating top-order zonal polynomials and power-sum symmetric functions. We then showed that, when the associated generating function is a ratio of two generating functions of finite order, more efficient recurrence relations of fixed bounded length can be deduced.

These general results have been applied here to a number of problems involving quadratic forms in noncentral normal vectors, including the following much-studied problems: the moments of a single quadratic form, product-moments for several quadratic forms, and the moments of a ratio

of powers of two quadratic forms. In addition to their intrinsic interest, these examples show that the methodology is certainly useful for a number of different distribution-theoretic problems in statistics.

Many other distribution problems share many of the features present in the examples treated here. For example, the density and the distribution of a quadratic form of noncentral normal random variables have various series expansions (see James (1961), Ruben (1962), Shah and Khatri (1961, 1963), and Kotz, Johnson, and Boyd (1967)), and the coefficients in these series expansions can be easily shown to have a short recursive relation. In addition, the density of a ratio of a linear to a quadratic form is important in econometrics (generalizations of the cases discussed in Sawa (1972)), and can be expressed as multiple infinite series involving the Davis-Chikuse invariant polynomials. It seems highly likely that our methodology will prove useful there, and we are confident many other new applications of the results will follow.

Acknowledgement

Notes

[1]Note here that, for convenience, we define $E(t)$ in such a way that the e_k are the elementary symmetric functions of the eigenvalues of $-A$, rather than of A.

[2]A set of Matlab programs for implementing the three recursive algorithms is available upon request.

[3]This condition is needed to ensure that the expansion of $[1 - v'(I_n - \beta B)v]^{-s}$ as a power series in $v'(I_n - \beta B)v$ (from which (38) is derived) actually converges uniformly over the region of integration. This is so if and only if $|1 - \beta b_{\max}| < 1$, the condition stated.

[4]We include the term $\exp(-\mu'\mu/2)$ in the generating function to ensure that $h_{0,0}(A_1, A_2) = 1$.

REFERENCES

Bock, M. E., G. G. Judge, & T. A. Yancey (1984) A simple form for the inverse moments of non-central χ^2 and F random variables and certain confluent hypergeometric functions. *Journal of Econometrics* 25, 217–234.

Chikuse, Y. (1987) Methods for constructing top order invariant polynomials. *Econometric Theory* 3, 195–207.

Cressie, N., A. S. Davis, J. L. Folks, & G. E. Policello II (1981) The moment-generating function and negative integer moments. *American Statistican* 35, 148–150.

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* A31, 465–485.

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* A33, 297–313.

Ghazal, G. A. (1994) Moments of the ratio of two dependent quadratic forms. *Statistics and Probability Letters* 20, 313–319.

Hillier, G. H., R. Kan, & X. Wang (2009) Computationally efficient recursions for top-order invariant polynomials with applications. *Econometric Theory* 25, 211–242.

James, A. T. (1961) Zonal polynomials of the real positive definite symmetric matrices. *Annals of Mathematics* 74, 456–469.

Kan, R. (2008) From moments of sum to moments of product. *Journal of Multivariate Analysis* 99, 542–554.

Kotz, S., N. L. Johnson, & D. W. Boyd (1967) Series representations of distributions of quadratic forms in normal random variables II: Non-central case. *Annals of Mathematical Statistics* 38, 838–848.

Magnus, J. (1986) The exact moments of a ratio of quadratic forms in normal variables. *Annales d'Économie et de Statistique* 4, 95–109.

Mathai, A. M. & S. B. Provost (1992) *Quadratic Forms in Random Variables: Thoery and Applications*, Marcel Dekker.

Meng, X. L. (2005) From unit root to Stein's estimator to Fisher's k statistics: If you have a moment, I can tell you more. *Statistical Science* 20, 141–162.

Press, S. J. (1966) Linear combinations of non-central chi-square variates. *Annals of Mathematical Statistics* 37, 480–487.

Ruben, H. (1962) Probability content of regions under spherical normal distributions, IV: The distribution of homogeneous and non-homogeneous quadratic functions of normal variables. *Annals of Mathematical Statistics* 33, 542–570.

Sawa, T. (1972) Finite sample properties of the k-class estimators. *Econometrica* 40, 653–680.

Shah, B. K.& C. G. Khatri (1961) Distribution of a definite quadratic forms in non-central normal variates. *Annals of Mathematical Statistics* 32, 883–887.

Shah, B. K.& C. G. Khatri (1963) Correction to "Distribution of a definite quadratic forms in non-central normal variates." *Annals of Mathematical Statistics* 34, 673.

Smith, M. (1989) On the expectation of a ratio of quadratic forms in normal random variables. *Journal of Multivariate Analysis* 31, 244–257.

Smith, M. (1993) Expectations of ratios of quadratic forms in normal variables: Evaluating some top-order invariant polynomials. *Australian Journal of Statistics* 35, 271–282.

Smith, P. J. (1995) A recursive formulation of the old problem of obtaining moments from cumulants and vice versa. *American Statistican* 49, 217–218.

Wilf, H. S. (2005) *Generatingfunctionology*, third edition, A. K. Peters.

Proof of LEMMA 4. We have that

$$
\begin{aligned}
d_{r,k}(A, \mu\mu') &= [t_1^r t_2^k]|I_n - t_1 A - t_2 \mu\mu'|^{-\frac{1}{2}} \\
&= [t_1^r t_2^k]|I_n - t_1 A|^{-\frac{1}{2}}(1 - t_2 \mu'(I_n - t_1 A)^{-1}\mu)^{-\frac{1}{2}} \\
&= \frac{(\frac{1}{2})_k}{k!}[t_1^r]|I_n - t_1 A|^{-\frac{1}{2}}(\mu'(I_n - t_1 A)^{-1}\mu)^k \\
&= \frac{(\frac{1}{2})_k}{k!}[t_1^r]|I_n - t_1 A|^{-\frac{1}{2}}(\delta + \phi(t_1))^k \\
&= \frac{(\frac{1}{2})_k}{k!}\sum_{l=0}^{k}\binom{k}{l}\delta^{k-l}[t_1^r]|I_n - t_1 A|^{-\frac{1}{2}}\phi(t_1)^l \\
&= \frac{(\frac{1}{2})_k}{k!}\sum_{l=0}^{k}\binom{k}{l}\delta^{k-l}a_{r,l},
\end{aligned}
$$

by the definition of the $a_{r,l}$ in (24). ∎

Proof of LEMMA 6. Since

$$
\begin{aligned}
H_{A_1,A_2}(t_1,t_2) &= \exp\left(-\frac{\delta}{2}\right)|I_n - t_1 A_1 - t_2 A_2|^{-\frac{1}{2}}\sum_{m=0}^{\infty}\frac{(1-t_2)^m}{2^m m!}(\mu'(I_n - t_1 A_1 - t_2 A_2)^{-1}\mu)^m \\
&= \exp\left(-\frac{\delta}{2}\right)|I_n - t_1 A_1 - t_2 A_2|^{-\frac{1}{2}}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-t_2)^l}{2^{m+l}l!m!}(\mu'(I_n - t_1 A_1 - t_2 A_2)^{-1}\mu)^{l+m},
\end{aligned}
$$

we have

$$
\begin{aligned}
h_{r,j}(A_1, A_2) &= [t_1^r t_2^j]H_{A_1,A_2}(t_1,t_2) \\
&= \exp\left(-\frac{\delta}{2}\right)\sum_{m=0}^{\infty}\frac{1}{2^m m!} \\
&\quad \times \sum_{l=0}^{j}\frac{(-1)^l}{2^l l!}[t_1^r t_2^{j-l}]|I_n - t_1 A_1 - t_2 A_2|^{-\frac{1}{2}}(\mu'(I_n - t_1 A_1 - t_2 A_2)^{-1}\mu)^{l+m} \\
&= \exp\left(-\frac{\delta}{2}\right)\sum_{m=0}^{\infty}\frac{1}{2^m m!}\sum_{l=0}^{j}\frac{(-1)^l}{2^l l!}\frac{(l+m)!}{(\frac{1}{2})_{l+m}}d_{r,j-l,l+m}(A_1, A_2, \mu\mu'),
\end{aligned}
$$

where the last equality follows because

$$
\begin{aligned}
d_{i,j,k}(A_1, A_2, \mu\mu') &= [t_1^i t_2^j t_3^k]|I_n - t_1 A_1 - t_2 A_2 - t_3 \mu\mu'|^{-\frac{1}{2}} \\
&= [t_1^i t_2^j t_3^k]|I_n - t_1 A_1 - t_2 A_2|^{-\frac{1}{2}}(1 - t_3 \mu'(I_n - t_1 A_1 - t_2 A_2)^{-1}\mu)^{-\frac{1}{2}} \\
&= \frac{(\frac{1}{2})_k}{k!}[t_1^i t_2^j]|I_n - t_1 A_1 - t_2 A_2|^{-\frac{1}{2}}(\mu'(I_n - t_1 A_1 - t_2 A_2)^{-1}\mu)^k. \qquad (54)
\end{aligned}
$$

■

Proof of LEMMA 7. By definition, we have

$$
\begin{aligned}
\tilde{h}_{r,j}(A_1, A_2) &= [t_1^r t_2^j]|I_n - t_1 A_1 - t_2 A_2|^{-\frac{1}{2}} \exp\left(\frac{1}{2} t_2 \hat{\mu}'(I_n - t_1 A_1 - t_2 A_2)^{-1}\hat{\mu}\right) \\
&= \sum_{l=0}^{j} \frac{1}{l! 2^l} [t_1^r t_2^{j-l}]|I_n - t_1 A_1 - t_2 A_2|^{-\frac{1}{2}} (\hat{\mu}'(I_n - t_1 A_1 - t_2 A_2)^{-1}\hat{\mu})^l \\
&= \sum_{l=0}^{j} \frac{1}{2^l \left(\frac{1}{2}\right)_l} d_{r,j-l,l}(A_1, A_2, \hat{\mu}\hat{\mu}'),
\end{aligned}
$$

where the last equality is obtained by using (54). ■

Proof of Convergence in (44). We shall show that, if β is chosen so that $0 < \beta < 2/b_{\max}$, the series in (44) converges, justifying the term-by-term integration. We shall make use here of the following lemma:

LEMMA 8. *If a_i is the largest absolute eigenvalue of A_i for each i, then*

$$
|d_\kappa(A_1, \ldots, A_r)| \leq \frac{1}{\kappa!} \left(\frac{n}{2}\right)_k \prod_{i=1}^{r} a_i^{k_i}. \tag{55}
$$

Proof. From equation (81) in HKW,

$$
|d_\kappa(A_1, \ldots, A_r)| = \frac{1}{\kappa!} \left(\frac{n}{2}\right)_k \left|\int_{v'v=1} \left(\prod_{i=1}^{r}(v'A_i v)^{k_i}\right)(\mathrm{d}v)\right| \leq \frac{1}{\kappa!} \left(\frac{n}{2}\right)_k \int_{v'v=1} \prod_{i=1}^{r}|v'A_i v|^{k_i}(\mathrm{d}v).
$$

But, it is well-known that

$$
\sup_{v'v=1} |v'A_i v| = a_i.
$$

This immediately yields the stated inequality. ■

Let a and $\tilde{b}$ be the largest absolute eigenvalue of A and $\tilde{B}$, respectively. Using Lemma 6 and (55), we can bound $|h_{r,j}(A, \tilde{B})|$ by

$$
\begin{aligned}
\left|h_{r,j}(A, \tilde{B})\right| &\leq e^{-\frac{\delta}{2}} \sum_{m=0}^{\infty} \frac{1}{2^m m!} \sum_{l=0}^{j} \frac{(l+m)!}{2^l l! \left(\frac{1}{2}\right)_{l+m}} \left|d_{r,j-l,l+m}(A, \tilde{B}, \mu\mu')\right| \\
&\leq e^{-\frac{\delta}{2}} \sum_{m=0}^{\infty} \frac{1}{2^m m!} \sum_{l=0}^{j} \frac{(l+m)!}{2^l l! \left(\frac{1}{2}\right)_{l+m}} \frac{\left(\frac{n}{2}\right)_{r+j+m} a^r \tilde{b}^{j-l} \delta^{l+m}}{r!(j-l)!(l+m)!} \\
&= \frac{e^{-\frac{\delta}{2}} a^r \left(\frac{n}{2}\right)_r}{r!} \sum_{m=0}^{\infty} \sum_{l=0}^{j} \frac{\left(\frac{n}{2}+r\right)_{j+m} \left(\frac{\delta}{2}\right)^{l+m} \tilde{b}^{j-l}}{m! l!(j-l)! \left(\frac{1}{2}\right)_{l+m}}.
\end{aligned}
$$

38

Under the condition $0 < \beta < 2/b_{\max}$, we have $0 \le \tilde{b} < 1$. Together with the condition $\frac{n}{2} + r > s$, we can bound the absolute value of the terms in the infinite series in (44) by

$$
\begin{aligned}
\sum_{j=0}^{\infty} \frac{(s)_j}{\left(\frac{n}{2} + r\right)_j} |h_{r,j}(A, \tilde{B})| &\le \sum_{j=0}^{\infty} |h_{r,j}(A, \tilde{B})| \\
&= \frac{e^{-\frac{\delta}{2}} a^r \left(\frac{n}{2}\right)_r}{r!} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \sum_{l=0}^{j} \frac{\left(\frac{n}{2} + r\right)_{j+m} \left(\frac{\delta}{2}\right)^{l+m} \tilde{b}^{j-l}}{m! l! (j-l)! \left(\frac{1}{2}\right)_{l+m}} \\
&= \frac{e^{-\frac{\delta}{2}} a^r \left(\frac{n}{2}\right)_r}{r!} \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} \frac{\left(\frac{n}{2} + r\right)_{l+m} \left(\frac{\delta}{2}\right)^{l+m}}{m! l! \left(\frac{1}{2}\right)_{l+m}} \sum_{j=0}^{\infty} \frac{\left(\frac{n}{2} + r + l + m\right)_j \tilde{b}^j}{j!} \\
&= \frac{e^{-\frac{\delta}{2}} a^r \left(\frac{n}{2}\right)_r}{(1-\tilde{b})^{\frac{n}{2}+r} r!} \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} \frac{\left(\frac{n}{2} + r\right)_{l+m} \left(\frac{\delta}{2(1-\tilde{b})}\right)^{l+m}}{m! l! \left(\frac{1}{2}\right)_{l+m}} \\
&= \frac{e^{-\frac{\delta}{2}} a^r \left(\frac{n}{2}\right)_r}{(1-\tilde{b})^{\frac{n}{2}+r} r!} \sum_{k=0}^{\infty} \frac{\left(\frac{n}{2} + r\right)_k \left(\frac{\delta}{2(1-\tilde{b})}\right)^k 2^k}{\left(\frac{1}{2}\right)_k k!} \\
&= \frac{e^{-\frac{\delta}{2}} a^r \left(\frac{n}{2}\right)_r}{(1-\tilde{b})^{\frac{n}{2}+r} r!} {}_1F_1\left(\frac{n}{2} + r; \frac{1}{2}; \frac{\delta}{1-\tilde{b}}\right),
\end{aligned}
$$

where the fourth line is obtained by using the identity

$$
\sum_{j=0}^{\infty} \frac{\left(\frac{n}{2} + r + l + m\right)_j \tilde{b}^j}{j!} = (1-\tilde{b})^{-\frac{n}{2}-r-l-m}
$$

when $0 \le \tilde{b} < 1$, and the fifth line is obtained by using the identify $\sum_{l=0}^{k} \frac{k!}{l!(k-l)!} = 2^k$ and setting $k = l + m$. Since the ${}_1F_1$ converges uniformly for all values of its argument, this confirms the claim that term-by-term integration in (41) is justified. ∎

Proof of Convergence in (47). In this case any choice of β such that $\beta > 1/(2b_{\min})$ will ensure that $\hat{b}$, the largest absolute eigenvalue of $\hat{B}$, is less than one. Using Lemma 7 and (55), and assuming that β is so chosen, we have that

$$
\left|\tilde{h}_{r,j}(\hat{A}, \hat{B})\right| \le \frac{\hat{a}^r}{r!} \left(\frac{n}{2}\right)_{r+j} \sum_{l=0}^{j} \frac{\hat{b}^{j-l} \left(\frac{\hat{\delta}}{2}\right)^l}{l! (j-l)! \left(\frac{1}{2}\right)_l},
$$

where $\hat{a}$ is the largest absolute eigenvalue of $\hat{A}$ and $\hat{\delta} = \hat{\mu}' \hat{\mu}$. And the infinite series in (47) is

dominated termwise by

$$\sum_{j=0}^{\infty} \frac{\left(\frac{n}{2} + r - s\right)_j}{\left(\frac{n}{2} + r\right)_j} |\tilde{h}_{r,j}(\hat{A}, \hat{B})|$$

$$\leq \frac{\hat{a}^r}{r!} \left(\frac{n}{2}\right)_r \sum_{j=0}^{\infty} \left(\frac{n}{2} + r - s\right)_j \sum_{l=0}^{j} \frac{\hat{b}^{j-l} \left(\frac{\hat{\delta}}{2}\right)^l}{l!(j-l)! \left(\frac{1}{2}\right)_l}$$

$$= \frac{\hat{a}^r}{r!} \left(\frac{n}{2}\right)_r \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} \frac{\left(\frac{n}{2} + r - s\right)_{j+l} \hat{b}^j}{j! l! \left(\frac{1}{2}\right)_l} \left(\frac{\hat{\delta}}{2}\right)^l$$

$$= \frac{\hat{a}^r}{r!} \left(\frac{n}{2}\right)_r (1 - \hat{b})^{-\left(\frac{n}{2} + r - s\right)} \sum_{l=0}^{\infty} \frac{\left(\frac{n}{2} + r - s\right)_l}{l! \left(\frac{1}{2}\right)_l} \left(\frac{\hat{\delta}}{2(1 - \hat{b})}\right)^l$$

$$= \frac{\hat{a}^r}{r!} \left(\frac{n}{2}\right)_r (1 - \hat{b})^{-\left(\frac{n}{2} + r - s\right)} {}_1F_1 \left(\frac{n}{2} + r - s; \frac{1}{2}; \frac{\hat{\delta}}{2(1 - \hat{b})}\right).$$

Since the ${}_1F_1$ converges uniformly for all values of its argument, this confirms the claim that term-by-term integration is justified. ∎

Proof of THEOREM 8. Following the proof of Lemma 6, we can show that

$$\hat{h}_{r,j}(A_1, A_2) = e^{-\frac{\delta}{2}} \sum_{m=0}^{\infty} \sum_{l=0}^{j} \binom{l+m}{l} \frac{d_{r,j-l,l+m}(A_1, A_2, \mu\mu')}{2^{l+m} \left(\frac{1}{2}\right)_{l+m}}.$$

Under the assumption $0 < \beta \leq 1/b_{\max}$, $I_n - \beta B$ is positive semidefinite. Therefore, $d_{r,j,k}(A, I_n - \beta B, \mu\mu')$ is nonnegative when A is positive semidefinite or r is even. When A is not positive semidefinite and r is odd, we have $|z'Az| = |z'P\Lambda P'z| \leq z'P\Lambda^+ P'z = z'A^+ z$. Using the fact that $z'(I_n - \beta B)z \geq 0$ and $z'\mu\mu'z \geq 0$, we have $|(z'Az)^r (z'(I_n - \beta B)z)^j (z'\mu\mu'z)^k| \leq (z'A^+ z)^r (z'(I_n - \beta B)z)^j (z'\mu\mu'z)^k$, which implies

$$\left|d_{r,j,k}(A, I_n - \beta B, \mu\mu')\right| \leq d_{r,j,k}(A^+, I_n - \beta B, \mu\mu').$$

Using Lemma 6, we have

$$|h_{r,j}(A, I_n - \beta B)| \leq e^{-\frac{\delta}{2}} \sum_{m=0}^{\infty} \sum_{l=0}^{j} \binom{l+m}{l} \frac{|d_{r,j-l,l+m}(A, I_n - \beta B, \mu\mu')|}{2^{l+m} \left(\frac{1}{2}\right)_{l+m}} \leq \hat{h}_{r,j}(A^+, I_n - \beta B).$$

Using this result and the fact that $\frac{n}{2} + r > s$, we obtain

$$
\left| \sum_{j=M+1}^{\infty} \frac{(s)_j}{\left(\frac{n}{2}+r\right)_j} h_{r,j}(A, I_n - \beta B) \right| \leq \sum_{j=M+1}^{\infty} \frac{(s)_j}{\left(\frac{n}{2}+r\right)_j} |h_{r,j}(A, I_n - \beta B)|
$$

$$
\leq \sum_{j=M+1}^{\infty} \frac{(s)_j}{\left(\frac{n}{2}+r\right)_j} \hat{h}_{r,j}(A^+, I_n - \beta B)
$$

$$
\leq \frac{(s)_{M+1}}{\left(\frac{n}{2}+r\right)_{M+1}} \left[\frac{e^{\frac{\bar{\delta}-\delta}{2}} \tilde{d}_r(\bar{A}, \bar{\mu})}{|\beta B|^{\frac{1}{2}}} - \sum_{j=0}^{M} \hat{h}_{r,j}(A^+, I_n - \beta B) \right].
$$

The last inequality holds because

$$
\sum_{j=0}^{\infty} \hat{h}_{r,j}(A^+, I_n - \beta B)
$$

$$
= [t_1^r] |I_n - t_1 A^+ - (I_n - \beta B)|^{-\frac{1}{2}} \exp\left(\mu' \left[I_n - t_1 A^+ - (I_n - \beta B) \right]^{-1} \mu - \frac{\delta}{2} \right)
$$

$$
= [t_1^r] |\beta B|^{-\frac{1}{2}} e^{-\frac{\delta}{2}} |I_n - t_1 \bar{A}|^{-\frac{1}{2}} \exp\left(\frac{\bar{\mu}'(I_n - t_1 \bar{A})^{-1} \bar{\mu}}{2} \right)
$$

$$
= \frac{e^{\frac{\bar{\delta}-\delta}{2}}}{|\beta B|^{\frac{1}{2}}} \tilde{d}_r(\bar{A}, \bar{\mu}).
$$

∎

Proof of Theorem 9. For any choice of β satisfying $\beta \geq 1/b_{\min}$, $\hat{B} = I_n - (\beta B)^{-1}$ is positive semidefinite. When A is positive semidefinite or r is even, $d_{r,j,k}(\hat{A}, \hat{B}, \hat{\mu}\hat{\mu}')$ is nonnegative. When A is not positive semidefinite and r is odd, we have $|z'\hat{A}z| = \left| z'B^{-\frac{1}{2}} P\Lambda P'B^{-\frac{1}{2}} z/\beta \right| \leq z'B^{-\frac{1}{2}} P\Lambda^+ P'^{-\frac{1}{2}} z/\beta = z'\bar{A}z$. Using the fact that $z'\hat{B}z \geq 0$ and $z'\hat{\mu}\hat{\mu}'z \geq 0$, we have

$$
\left| d_{r,j,k}\left(\hat{A}, \hat{B}, \hat{\mu}\hat{\mu}' \right) \right| \leq d_{r,j,k}\left(\bar{A}, \hat{B}, \hat{\mu}\hat{\mu}' \right).
$$

Applying Lemma 7, we have

$$
\left| \tilde{h}_{r,j}\left(\hat{A}, \hat{B} \right) \right| \leq \sum_{l=0}^{j} \frac{|d_{r,j-l,l}\left(\hat{A}, \hat{B}, \hat{\mu}\hat{\mu}' \right)|}{2^l \left(\frac{1}{2}\right)_l} \leq \sum_{l=0}^{j} \frac{d_{r,j-l,l}\left(\bar{A}, \hat{B}, \hat{\mu}\hat{\mu}' \right)}{2^l \left(\frac{1}{2}\right)_l} = \tilde{h}_{r,j}\left(\bar{A}, \hat{B} \right).
$$

Using this result and the fact that $\frac{n}{2} + r > s$, we obtain

$$
\left| \sum_{j=M+1}^{\infty} \frac{\left(\frac{n}{2} + r - s\right)_j}{\left(\frac{n}{2} + r\right)_j} \tilde{h}_{r,j}\left(\hat{A}, \hat{B}\right) \right| \leq \sum_{j=M+1}^{\infty} \frac{\left(\frac{n}{2} + r - s\right)_j}{\left(\frac{n}{2} + r\right)_j} \left| \tilde{h}_{r,j}\left(\hat{A}, \hat{B}\right) \right|
$$

$$
\leq \sum_{j=M+1}^{\infty} \frac{\left(\frac{n}{2} + r - s\right)_j}{\left(\frac{n}{2} + r\right)_j} \tilde{h}_{r,j}\left(\bar{A}, \hat{B}\right)
$$

$$
\leq \frac{\left(\frac{n}{2} + r - s\right)_{M+1}}{\left(\frac{n}{2} + r\right)_{M+1}} \sum_{j=M+1}^{\infty} \tilde{h}_{r,j}\left(\bar{A}, \hat{B}\right)
$$

$$
= \frac{\left(\frac{n}{2} + r - s\right)_{M+1}}{\left(\frac{n}{2} + r\right)_{M+1}} \left[|\beta B|^{\frac{1}{2}} e^{\frac{\delta}{2}} \tilde{d}_r(A^+, \mu) - \sum_{j=0}^{M} \tilde{h}_{r,j}(\bar{A}, \hat{B}) \right].
$$

The last equality holds because

$$
\sum_{j=0}^{\infty} \tilde{h}_{r,j}\left(\bar{A}, \hat{B}\right) = [t_1^r]\left| I_n - t_1\bar{A} - \hat{B} \right|^{-\frac{1}{2}} \exp\left(\frac{\hat{\mu}' \left(I_n - t_1\bar{A} - \hat{B} \right)^{-1} \hat{\mu}}{2} \right)
$$

$$
= [t_1^r]|\beta B|^{\frac{1}{2}} \left| I_n - t_1 A^+ \right|^{-\frac{1}{2}} \exp\left(\frac{\mu' \left(I_n - t_1 A^+ \right)^{-1} \mu}{2} \right)
$$

$$
= |\beta B|^{\frac{1}{2}} e^{\frac{\delta}{2}} \tilde{d}_r(A^+, \mu).
$$

∎