Graph presentations for moments of noncentral Wishart distributions and their applications

Satoshi Kuriki · Yasuhide Numata

Abstract   We provide formulas for the moments of the real and complex noncentral Wishart distributions of general degrees. The obtained formulas for the real and complex cases are described in terms of the undirected and directed graphs, respectively. By considering degenerate cases, we give explicit formulas for the moments of bivariate chi-square distributions and 2×2 Wishart distributions by enumerating the graphs. Noting that the Laguerre polynomials can be considered to be moments of a noncentral chi-square distributions formally, we demonstrate a combinatorial interpretation of the coefficients of the Laguerre polynomials.

1 Introduction

For $t = 1, \ldots, \nu$, let $X_t = (x_{ti})_{1 \le i \le p}$ be a p-dimensional random column vector distributed independently according to the normal distribution $N_p(\mu_t, \Sigma)$ with mean vector $\mu_t = (\mu_{ti})_{1 \le i \le p}$ and covariance matrix $\Sigma = (\sigma_{ij})_{1 \le i, j \le p}$. We define the (real)

S. Kuriki (✉)
The Institute of Statistical Mathematics, ROIS, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
e-mail: kuriki@ism.ac.jp

Y. Numata
Department of Mathematical Informatics, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
e-mail: numata@stat.t.u-tokyo.ac.jp

S. Kuriki · Y. Numata
Japan Science and Technology Agency (JST), CREST,
Sanbancho, Chiyoda-ku, Tokyo 102-0075, Japan

noncentral Wishart distribution $W_p(\nu, \Sigma, \Delta)$ by the distribution of a $p \times p$ symmetric random matrix

$$W = (w_{ij})_{1 \le i,j \le p}, \quad w_{ij} = \sum_{t=1}^{\nu} x_{ti} x_{tj}, \tag{1}$$

where

$$\Delta = (\delta_{ij})_{1 \le i,j \le p}, \quad \delta_{ij} = \sum_{t=1}^{\nu} \mu_{ti} \mu_{tj}$$

is the mean square matrix. The distribution of W depends on μ_t's through Δ because its moment generating function is

$$E\left[e^{\mathrm{tr}(\Theta W)}\right] = \det(I - 2\Theta\Sigma)^{-\frac{\nu}{2}} e^{\mathrm{tr}(I - 2\Theta\Sigma)^{-1}\Theta\Delta}, \tag{2}$$

where Θ is a $p \times p$ symmetric parameter matrix (Muirhead 1982).

Note that $\Delta = 0$ if and only if $\mu_t = 0$ for all t. The Wishart distribution with $\Delta = 0$ is referred to as the central Wishart distribution $W_p(\nu, \Sigma)$. Conventionally, the triplet (ν, Σ, Ω) with $\Omega = \Sigma^{-1}\Delta$ is used for describing the noncentral Wishart distribution rather than (ν, Σ, Δ). The matrix Ω is called the noncentrality matrix. In our paper, we adopt the triplet (ν, Σ, Δ) for simplicity in describing theorems.

For $t = 1, \ldots, \nu$, let $\begin{pmatrix} X_t \\ Y_t \end{pmatrix}$ be a $2p$-dimensional random column vector distributed independently according to the normal distribution with mean vector $\begin{pmatrix} \xi_t \\ \eta_t \end{pmatrix}$ and covariance matrix $\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$, where A and B are $p \times p$ symmetric and skew-symmetric matrices, respectively. The distribution of a complex-valued random vector $Z_t = (z_{ti})_{1 \le i \le p} = X_t + \sqrt{-1}Y_t$ is referred to as the complex normal distribution $CN_p(\mu_t, \Sigma)$ with mean $\mu_t = (\mu_{ti})_{1 \le i \le p} = \xi_t + \sqrt{-1}\eta_t$ and covariance matrix $\Sigma = (\sigma_{ij})_{1 \le i,j \le p} = 2(A + \sqrt{-1}B)$. Actually, Σ is a "covariance" in the sense of

$$\sigma_{ij} = E[(z_{ti} - \mu_{ti})(\overline{z_{tj} - \mu_{tj}})].$$

Here, the overline denotes the complex conjugate. From the complex random vectors Z_t, we define the complex noncentral Wishart distribution $CW_p(\nu, \Sigma, \Delta)$ as the distribution of a $p \times p$ Hermitian random matrix

$$W = (w_{ij})_{1 \le i,j \le p}, \quad w_{ij} = \sum_{t=1}^{\nu} z_{ti} \overline{z_{tj}}, \tag{3}$$

where

$$\Delta = (\delta_{ij})_{1 \le i,j \le p}, \quad \delta_{ij} = \sum_{t=1}^{\nu} \mu_{ti} \overline{\mu_{tj}}$$

is the mean square parameter matrix. As in the real case, the distribution of W depends on μ_t's through Δ since its moment generating function is

$$E\left[e^{\text{tr}(\Theta W)}\right] = \det(I - \Theta\Sigma)^{-\nu}e^{\text{tr}(I-\Theta\Sigma)^{-1}\Theta\Delta}, \tag{4}$$

where Θ is a $p \times p$ Hermitian parameter matrix. See Goodman (1963) for the central case.

The primary purpose of this paper is to obtain expressions for the moments $E[w_{ab} w_{cd} \cdots w_{ef}]$ of the real and complex noncentral Wishart distributions in terms of graphs, where $a, b, c, d, \ldots, e, f \in \{1, \ldots, p\}$ are arbitrary indices.

Considering the cases where the mean vectors μ_t and the covariance matrix Σ take particular values, we will obtain several identities of moments of some distributions associated with the Wishart distributions. We shall see that the derivations are reduced to enumerating graphs of various types. This is the secondary purpose of our paper.

The Wishart distribution originates with a paper by Wishart (1928) around 80 years ago. Since then, it is considered to be a fundamental distribution not only in mathematical statistics but also in other fields such as random matrices theory and signal processing (e.g., Bai 1999; Maiwald and Kraus 2000). Despite this, the structure of moments of the Wishart distributions is still an active research topic. In the central case, Lu and Richards (2001), Graczyk et al. (2003, 2005) provided formulas for moments of the real and complex Wishart distributions. These studies are based on expansions of the moment generating functions of the central Wishart distributions $\det(I-2\Theta\Sigma)^{-\frac{\nu}{2}}$. The graph presentations of moments have also been discussed in these studies. Prior to these studies, it was known that terms in the expansion of $\det(I - Y)^{-\alpha}$ around $Y = 0$ have some combinatorial structures (e.g., Vere-Jones 1988), which are closely related to the problem of Wishart moment. More recently, Letac and Massam (2008) provided a method to calculate moments of the noncentral Wishart distribution by combining expansions of the moment generating function $\det(I - 2\Theta\Sigma)^{-\frac{\nu}{2}}$ and the "noncentral part" $e^{\text{tr}(I-2\Theta\Sigma)^{-1}\Theta\Delta}$ in (2).

The outline of this paper is as follows. In Sect. 2, we treat the real noncentral Wishart matrices. Formulas for the general terms of moments of the Wishart distributions are given in terms of undirected graphs. This is an extension of Takemura (1991), who treated the central case. Then, by letting the mean vectors μ_t and the covariance matrix Σ have particular kinds of structures, we obtain explicit formulas for moments of the noncentral chi-square distribution, Kibble's (1941) bivariate chi-square (gamma) distribution, and the 2×2 central Wishart distribution with $\Sigma = I$. Noting a formal correspondence between the moments of the noncentral chi-square distributions and the Laguerre polynomials, we will show that the coefficients of the Laguerre polynomials have a combinatorial interpretation.

In Sect. 3, we treat the case of the noncentral complex Wishart matrices. Major parts of discussions are parallel to the real case. One remarkable difference is that moments in the complex case are not described in terms of undirected graphs but directed graphs.

2 Moments of the real noncentral Wishart distribution

2.1 A graph presentation

In this subsection, we provide a graph presentation formula for moments of the real noncentral Wishart distributions of general degrees. Our results are generalizations of Theorem 4.3 of Takemura (1991) where the central real Wishart matrices are treated. Our basic tool is the following formula for moments of Gaussian random vectors. This is just a moment-cumulant relation in the Gaussian case. For the proof, see McCullagh (1987). In the central case $\mu = 0$, this is sometimes referred to as the Wick formula.

Lemma 1 (Moment of the real normal distribution) *Let $X = (x_i)$ be a Gaussian random vector with mean $\mu = (\mu_i)$, and covariance matrix $\Sigma = (\sigma_{ij})$. Then,*

$$E[x_1 x_2 \cdots x_n] = \sum \sigma_{i_1 i_2} \cdots \sigma_{i_{2m-1} i_{2m}} \mu_{i_{2m+1}} \cdots \mu_{i_n},$$

where the summation is taken over all partitions of n indices $\{1, 2, \ldots, n\}$ into unordered m pairs and $n - 2m$ singletons

$$(i_1, i_2), \ldots, (i_{2m-1}, i_{2m}), (i_{2m+1}), \ldots, (i_n).$$

Remark 1 Although Lemma 1 just states an expression for $E[x_1 x_2 \cdots x_n]$, it indeed gives general forms of the moments $E[x_a x_b \cdots x_c]$ by considering a degenerate case. For example, we have $E[x_1 x_2^2] = E[\tilde{x}_1 \tilde{x}_2 \tilde{x}_3]$, where

$$\begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{pmatrix} \sim N_3 \left(\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} & \sigma_{22} \\ \sigma_{21} & \sigma_{22} & \sigma_{22} \end{pmatrix} \right).$$

Throughout the paper, we will use this degeneracy argument many times.

Let $X_t = (x_{ti})$ $(t = 1, \ldots, \nu)$ be independent Gaussian random vectors with mean μ_t and covariance matrix Σ. Let $W = (w_{ij})$ be a Wishart matrix made of X_t's as in (1). In the following, we give a formula for the moment $E[w_{ab} w_{cd} \cdots w_{ef}]$ with $a, b, c, d, \ldots, e, f$ arbitrary indices. By applying the degeneracy argument again, we can restrict our attention to the moment $E[w_{12} w_{34} \cdots w_{2n-1, 2n}]$ without loss of generality. For example $E[w_{11} w_{12}^2] = E[\tilde{w}_{12} \tilde{w}_{34} \tilde{w}_{56}]$, where $(\tilde{w}_{ij}) \sim W_6(\nu, (\tilde{\sigma}_{ij}), (\tilde{\delta}_{ij}))$,

$$(\tilde{\sigma}_{ij}, \tilde{\delta}_{ij}) = \begin{cases} (\sigma_{11}, \delta_{11}), & i, j \in \{1, 2, 3, 5\}, \\ (\sigma_{12}, \delta_{12}), & i \in \{1, 2, 3, 5\}, \ j \in \{4, 6\}, \\ (\sigma_{21}, \delta_{21}), & i \in \{4, 6\}, \ j \in \{1, 2, 3, 5\}, \\ (\sigma_{22}, \delta_{22}), & i, j \in \{4, 6\}. \end{cases}$$

Let $V = \{1, 2, \ldots, 2n-1, 2n\}$ be the set of indices appearing in the argument of the expectation $E[w_{12} \cdots w_{2n-1, 2n}]$. In the following, we consider an undirected graph

whose vertices are the elements of V. First consider an undirected graph $G_0 = (V, E_0)$ with the edges

$$E_0 = \{(1, 2), \ldots, (2n - 1, 2n)\}.$$

For each partition of $\{1, 2, \ldots, 2n - 1, 2n\}$ into m pairs and $2n - 2m$ singletons,

$$(i_1, i_2), \ldots, (i_{2m-1}, i_{2m}), (i_{2m+1}), \ldots, (i_{2n}), \tag{5}$$

we define a set of undirected edges

$$E = \{(i_1, i_2), \ldots, (i_{2m-1}, i_{2m})\}.$$

By adding the edges of E to G_0, we have a graph

$$G = (V, E_0 \cup E). \tag{6}$$

Each connected component of G is classified as a "cycle" (a path without terminals) and a "chain" (a path with two terminals). For the partition (5), the number of chains is $n - m$. The number of cycles of G is denoted by $\mathrm{len}(G)$. Note that $\mathrm{len}(G) \leq m$. Let $(j_1, j_2), \ldots, (j_{2n-2m-1}, j_{2n-2m})$ be pairs of two terminal vertices of $n - m$ chains of G, and let

$$\check{E} = \{(j_1, j_2), \ldots, (j_{2n-2m-1}, j_{2n-2m})\}.$$

Using these notations, the general form for the moments is given below.

Theorem 1 (Moment of the real noncentral Wishart distribution) *Let* $(w_{ij}) \sim W(\nu, (\sigma_{ij}), (\delta_{ij}))$. *Then,*

$$E[w_{12} \cdots w_{2n-1,2n}] = \sum_E \nu^{\mathrm{len}(G)} \sigma^E \delta^{\check{E}}, \tag{7}$$

where

$$\sigma^E = \prod_{(i,i') \in E} \sigma_{ii'} = \sigma_{i_1 i_2} \cdots \sigma_{i_{2m-1} i_{2m}},$$

$$\delta^{\check{E}} = \prod_{(j,j') \in \check{E}} \delta_{jj'} = \delta_{j_1 j_2} \cdots \delta_{j_{2n-2m-1} j_{2n-2m}}.$$

The summation $\sum_E$ *is taken over all partitions of* $\{1, 2, \ldots, 2n\}$ *of the form* (5).

Example 1 Consider the evaluation of the moment $E[w_{12}w_{34}w_{56}]$. Then, $V = \{1, 2, 3, 4, 5, 6\}$ and $E_0 = \{(1, 2), (3, 4), (5, 6)\}$. There are 76 partitions of V into pairs

and singletons. Figure 1 is the graph $G = (V, E_0 \cup E)$ for $E = \{(1, 6), (2, 5)\}$ ($\check{E} = \{(3, 4)\}$). Summing up 76 possibilities, we have the following:

$$
\begin{aligned}
E[w_{12}w_{34}w_{56}] = {}& v^3\sigma_{12}\sigma_{34}\sigma_{56} + v^2\sigma_{23}\sigma_{14}\sigma_{56}[6] + v\sigma_{23}\sigma_{45}\sigma_{16}[8] \\
& + v^2\sigma_{12}\sigma_{34}\delta_{56}[3] + v\sigma_{23}\sigma_{14}\delta_{56}[6] + v\sigma_{12}\sigma_{45}\delta_{36}[12] \\
& + \sigma_{23}\sigma_{45}\delta_{16}[24] + v\sigma_{12}\delta_{34}\delta_{56}[3] + \sigma_{23}\delta_{14}\delta_{56}[12] \\
& + \delta_{12}\delta_{34}\delta_{56}.
\end{aligned}
$$

Here, $[n]$ means that there are n terms of similar form.

Proof For $i = 1, \ldots, n$, let $e(i) = [(i + 1)/2]$ (the integer part of $(i + 1)/2$). Noting that $w_{ij} = \sum_{t=1}^{v} x_{ti}x_{tj}$, and from Lemma 1, we have

$$
\begin{aligned}
& E[w_{12} \cdots w_{2n-1,2n}] \\
& = \sum_{t_1=1}^{v} \cdots \sum_{t_n=1}^{v} E[x_{t_1,1}x_{t_1,2} \cdots x_{t_n,2n-1}x_{t_n,2n}] \\
& = \sum_{t_1} \cdots \sum_{t_n} E[x_{t_{e(1)},1}x_{t_{e(2)},2} \cdots x_{t_{e(2n-1)},2n-1}x_{t_{e(2n)},2n}] \\
& = \sum_{E} \sum_{t_1} \cdots \sum_{t_n} \mathrm{Cov}(x_{t_{e(i_1)},i_1}, x_{t_{e(i_2)},i_2}) \cdots \mathrm{Cov}(x_{t_{e(i_{2m-1})},i_{2m-1}}, x_{t_{e(i_{2m})},i_{2m}}) \\
& \quad \times E[x_{t_{e(i_{2m+1})},i_{2m+1}}] \cdots E[x_{t_{e(i_{2n})},i_{2n}}]. \tag{8}
\end{aligned}
$$

Since $\{i_1, \ldots, i_n\} = V$, the indices $i_1, \ldots, i_n$ can be divided into connected components of the graph G. Let $\{j_1, \ldots, j_{2k}\}$ be a set of vertices of a connected component. Then, as we already pointed out, it forms either a chain

$$
(j_1, j_2), (j_2, j_3), \ldots, (j_{2k-2}, j_{2k-1}), (j_{2k-1}, j_{2k}),
$$

or a cycle

$$
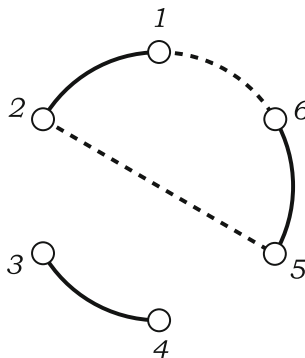(j_1, j_2), (j_2, j_3), \ldots, (j_{2k-1}, j_{2k}), (j_{2k}, j_1),
$$

Fig. 1 Graph $G = (V, E_0 \cup E)$ (E_0 *solid line*, E *dotted line*) presenting the term $v^1\sigma_{16}\sigma_{25}\delta_{34}$ ($n = 6$, $m = 2$, $\mathrm{len}(G) = 1$)

and in both cases

$$(j_1, j_2), (j_3, j_4), \ldots, (j_{2k-1}, j_{2k}) \in E_0.$$

Since the running indices $t_1, \ldots, t_n$ correspond to n edges of E_0, and $e(j_1) = e(j_2)$, $e(j_3) = e(j_4), \ldots, e(j_{2k-1}) = e(j_{2k})$, the argument of the summation $\sum_E$ in (8) is written as a product of terms of the form

$$\sum_{t_1} \cdots \sum_{t_k} E[x_{t_1, j_1}] \mathrm{Cov}(x_{t_1, j_2}, x_{t_2, j_3}) \cdots$$
$$\times \mathrm{Cov}(x_{t_{k-1}, j_{2k-2}}, x_{t_k, j_{2k-1}}) E[x_{t_k, j_{2k}}] \tag{9}$$

in the chain case, or

$$\sum_{t_1} \cdots \sum_{t_k} \mathrm{Cov}(x_{t_1, j_2}, x_{t_2, j_3}) \cdots$$
$$\times \mathrm{Cov}(x_{t_{k-1}, j_{2k-2}}, x_{t_k, j_{2k-1}}) \mathrm{Cov}(x_{t_k, j_{2k}}, x_{t_1, j_1}) \tag{10}$$

in the cycle case. Here, we used a reindexing

$$t_1 := t_{e(j_1)} = t_{e(j_2)}, \ldots, t_k := t_{e(j_{2k-1})} = t_{e(j_{2k})}.$$

Noting that $\mathrm{Cov}(x_{si}, x_{tj}) = 1_{\{s=t\}} \sigma_{ij}$ and $\sum_{t=1}^{\nu} E[x_{ti}] E[x_{tj}] = \delta_{ij}$, we see that (9) $= \sigma_{j_2 j_3} \cdots \sigma_{j_{2k-2} j_{2k-1}} \delta_{j_{2k} j_1}$ and (10) $= \nu \sigma_{j_2 j_3} \cdots \sigma_{j_{2k-2}, j_{2k-1}} \sigma_{j_{2k} j_1}$. This completes the proof. $\square$

2.2 Enumeration of undirected graphs

In this subsection, we will enumerate the graphs $G = (V, E_0 \cup E)$ defined in (6) under the condition that the number $l = \mathrm{len}(G)$ of cycles and the number m of edges of E are given. Let $f_{l,m,n}$ be the number of such graphs.

Consider a degenerate noncentral Wishart matrix $W = (w_{ij})$ such that $\sigma_{ij} \equiv 1$ and $\delta_{ij} \equiv \delta$. This happens when every component of X_t making up W takes the same value with probability one. Accordingly, all elements of W take the same value w, say, with probability one. In this setting, (7) in Theorem 1 is reduced to a moment formula for the distribution of w, the noncentral chi-square distribution $\chi_\nu^2(\delta)$ with ν degrees of freedom and the noncentrality parameter δ. Using the coefficient $f_{l,m,n}$, the nth moment of w is given as follows.

$$E[w^n] = \sum_{m=0}^{n} \sum_{l \geq 0} \nu^l f_{l,m,n} \delta^{n-m}. \tag{11}$$

The coefficient $f_{l,m,n}$ satisfies the following recurrence formula.

Lemma 2

$$f_{l,m,n} = 2(2n - m - 1) f_{l,m-1,n-1}$$
$$+ f_{l-1,m-1,n-1} + f_{l,m,n-1}, \qquad (12)$$

with boundary conditions

$$f_{l,0,n} = \begin{cases} 1 & (l = 0), \\ 0 & (l \geq 1) \end{cases} \quad \text{for } n \geq 1, \qquad (13)$$

and

$$f_{l,1,1} = \begin{cases} 0 & (l = 0), \\ 1 & (l = 1). \end{cases} \qquad (14)$$

Proof In the following, we sometimes refer to an edge from E_0 as a "solid line" edge and an edge from E as a "dashed line" edge as shown in Fig. 1.

The connected components of G are classified as cycles and chains. In each cycle, the number of solid line edges is equal to the number of dashed line edges. In each chain, the number of solid line edges is one more than the number of dashed line edges. Thus, the number of connected chains is $n - m$, the difference between the number of solid line edges and the number of dashed line edges.

Consider a graph G' made by removing two vertices $2n - 1$ and $2n$, and all (solid line and dashed line) edges connecting to these two vertices. Note that the (solid line) edge $(2n - 1, 2n)$ is deleted.

One of the following will meet: The edge $(2n - 1, 2n)$ is contained in (i) a cycle with four or more edges, or a chain with three or more edges (the edges (a, b), (c, d), (i, j), or (k, l) in Fig. 2); (ii) a cycle with two edges $((e, f)$ in Fig. 2); (iii) a chain consisting of a edge $((g, h)$ in Fig. 2).

Case (i). In the graph G' made by removing two vertices $2n - 1$ and $2n$, and all connected edges, there are $n - 1$ solid line edges. Since one dashed line edge is removed together, there are $m - 1$ dashed line edges. The number of chains still remains $(n - 1) - (m - 1) = n - m$. To the graph G', consider adding the edge $(2n - 1, 2n)$ again. There are $n - 1 + (n - m) = 2n - m - 1$ places where $(2n - 1, 2n)$ can be inserted, and considering the direction of the inserted edge, there are $2(2n - m - 1)$ ways in which the insertion can be done. By this operation, the number of dashed lines increases by 1, whereas the number of cycles is invariant. The contribution of the number of graphs made by this operation to $f_{l,m,n}$ is

$$2(2n - m - 1) f_{l,m-1,n-1}.$$

Case (ii). Consider adding the edge $(2n - 1, 2n)$ to the graph G' again to make a cycle with the edge $(2n - 1, 2n)$ and a dashed line edge. By this operation, both the number of dashed lines and the number of cycles increases by 1. The contribution of the number of graphs made by this operation to $f_{l,m,n}$ is
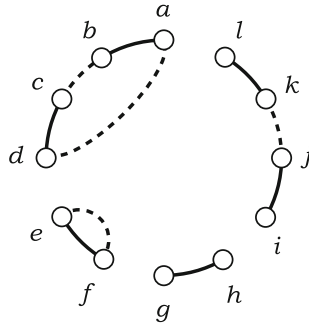
Fig. 2 A figure for the proof of Lemma 2

$$f_{l-1,m-1,n-1}.$$

Case (iii). Consider adding the edge $(2n - 1, 2n)$ to the graph G' again to make a chain consisting of one edge $(2n - 1, 2n)$. By this operation, both the number of dashed lines and the number of cycles are invariant. The contribution of the number of graphs made by this operation to $f_{l,m,n}$ is

$$f_{l,m,n-1}.$$

Summing up the three cases (i), (ii), and (iii), we obtain the recurrence formula (12). □

Theorem 2 *The generating function of $f_{l,m,n}$ with respect to the number l of cycles is given by*

$$\Phi_{m,n}(v) = \sum_{l \geq 0} v^l f_{l,m,n} = \binom{n}{m} \prod_{i=1}^{m} (v + 2(n - i)) \quad (0 \leq m \leq n, \, n \geq 1). \quad (15)$$

Here, we use a convention $\prod_{i=1}^{0} = 1$.

Proof Noting that $f_{-1,m,n} = 0$, the generating function $\Phi_{m,n}(v) = \sum_{l \geq 0} v^l f_{l,m,n}$ has to satisfy

$$\Phi_{m,n} = 2(2n - m - 1)\Phi_{m-1,n-1} + v\Phi_{m-1,n-1} + \Phi_{m,n-1}. \quad (16)$$

To solve the recurrence formula (16), consider the boundary conditions. From (13), we have

$$\Phi_{0,n}(v) = 1 \quad (n \geq 1). \quad (17)$$

Moreover, from (14), we have

$$\Phi_{1,1}(v) = v.$$

Furthermore, since $\Phi_{n,n-1} = 0$,

$$\Phi_{n,n} = 2(n-1)\Phi_{n-1,n-1} + \nu\Phi_{n-1,n-1} = (\nu + 2n - 2)\Phi_{n-1,n-1}$$

$$= \cdots = \prod_{i=1}^{n}(\nu + 2(n-i)) \quad (n \geq 1).$$ (18)

The recurrence formula (16) combined with the boundary conditions (17) and (18) determines $\Phi_{m,n}$ for all m and n.

In the following, we see that $\Phi_{m,n}(\nu)$ in (15) is actually the solution for the recurrence formula (16). The boundary conditions (17) and (18) are satisfied. We only have to make sure that (15) really satisfies (16). Indeed,

$$\Phi_{m,n} - \Phi_{m,n-1} = \binom{n}{m}\prod_{i=1}^{m}(\nu + 2n - 2i) - \binom{n-1}{m}\prod_{i=1}^{m}(\nu + 2n - 2 - 2i)$$

$$= \binom{n-1}{m-1}\frac{1}{m}\prod_{i=1}^{m-1}(\nu + 2n - 2 - 2i)$$

$$\times \{n(\nu + 2n - 2) - (n-m)(\nu + 2n - 2 - 2m)\}$$

$$= 2(2n - m - 1)\Phi_{m-1,n-1} + \nu\Phi_{m-1,n-1}.$$

$\square$

Corollary 1

$$\Phi_{m,n}(1) = \binom{n}{m}(2n-1)(2n-3)\cdots(2n-2m+1) = \binom{2n}{2m}(2m-1)!!$$

is the number of undirected graphs G, and

$$\Phi_{m,n}(0) = \binom{n}{m}(2n-2)(2n-4)\cdots(2n-2m)$$

$$= \frac{2^m\, n!(n-1)!}{m!\,(n-m)!(n-m-1)!}$$

is the number of undirected graphs G without cycles.

Remark 2 Nonnegative integers $s_n(m, l)$ defined by a generating function

$$\sum_{l=0}^{m}\nu^l s_n(m, l) = \prod_{i=1}^{m}(\nu + n - i)$$ (19)

are called the noncentral Stirling numbers of the first kind (Koutras 1982). Since

$$\sum_{\nu\geq 0}\nu^l f_{l,m,n} = \binom{n}{m}2^m\prod_{i=1}^{m}(\nu/2 + n - i) = \binom{n}{m}2^m\sum_{l\geq 0}(\nu/2)^l s_n(m, l),$$

we have

$$f_{l,m,n} = \binom{n}{m} 2^{m-l} s_n(m,l).$$

2.3 Moments of the noncentral chi-square distribution and the Laguerre polynomial

As stated in the beginning of the previous subsection, the moment of the noncentral chi-square distribution is described with the coefficient $f_{l,m,n}$.

In view of (11) and Theorem 2, the nth moment of $w \sim \chi_\nu^2(\delta)$, the noncentral chi-square distribution with ν degrees of freedom and the noncentrality parameter δ, is written as

$$
\begin{aligned}
E[w^n] &= \sum_{m=0}^{n} \sum_{l \geq 0} \nu^l f_{l,m,n} \delta^{n-m} \\
&= \sum_{m=0}^{n} \Phi_{m,n}(\nu) \delta^{n-m} \\
&= \sum_{m=0}^{n} \binom{n}{m} \prod_{i=1}^{m} (\nu + 2(n-i)) \delta^{n-m}.
\end{aligned}
\tag{20}
$$

This is a well-known expression for the moment of noncentral chi-square distribution (e.g., Johnson et al. 1995).

Remark 3 Koutras (1982) pointed out that moments of some noncentral distributions are described with the noncentral Stirling numbers of the first kind.

The moment generating function of the noncentral chi-square distribution $\chi_\nu^2(\delta)$ is

$$(1 - 2t)^{-\nu/2} e^{\delta t(1-2t)^{-1}}.$$

This can be obtained by letting $\Theta = t$, $\Sigma = 1$, $\Delta = \delta$ with (2).

The Laguerre polynomials, the orthogonal polynomial systems on $(0, \infty)$ with respect to the gamma weight functions, are defined as

$$L_n^{(\nu)}(x) = 2^n \frac{d^n}{d\sigma^n} f^{(\nu)}(x; \sigma) \Big|_{\sigma=1} \Big/ f^{(\nu)}(x; 1) \quad (\nu > 0),$$

where

$$f^{(\nu)}(x; \sigma) = x^{\nu/2-1} \sigma^{-\nu/2} e^{-x/(2\sigma)}$$

(e.g., Morris 1982). From this definition, we immediately get the generating function of the Laguerre polynomial as

$$\sum_{n=0}^{\infty}(-1)^n\frac{t^n}{n!}L_n^{(\nu)}(x) = (1-2t)^{-\nu/2}e^{-(x/2)((1-2t)^{-1}-1)},$$

which has formal coincidence with the moment generating function of the chi-square distribution with ν degrees of freedom and the noncentrality parameter $-x$. Therefore, we have an expression for the Laguerre polynomials

$$L_n^{(\nu)}(x) = (-1)^n\sum_{m=0}^{n}\sum_{\nu\geq 0}\nu^l f_{l,m,n}(-x)^{n-m}$$

$$= (-1)^n\sum_{m=0}^{n}\binom{n}{m}\prod_{i=1}^{m}(\nu+2(n-i))(-x)^{n-m}.$$

This gives a combinatorial interpretation for the coefficients of the Laguerre polynomials.

This type of combinatorial interpretation for Hermite polynomials is widely known. Applying a degenerate multivariate distribution $N((\mu_i),(\sigma_{ij}))$ with $\mu_i\equiv\mu,\sigma_{ij}\equiv\sigma^2$ to Lemma 1, we see that the nth moment of the normal distribution $X\sim N(\mu,\sigma^2)$ is

$$E[X^n] = \sum_{m=0}^{[n/2]}a_{m,n}\sigma^{2m}\mu^{n-2m},$$

where

$$a_{m,n} = \binom{n}{2m}(2m-1)!! = \frac{n!}{(n-2m)!\,2^m\,m!}$$

is the number of partitions of an n-member set into (unordered) m pairs and $n-2m$ singletons. The Hermite polynomials are defined by

$$H_n(x) = (-1)^n\frac{d^n}{dx^n}e^{-x^2/2}\Big/e^{-x^2/2}.$$

The generating function of $H_n(x)$ is

$$\sum_{n=0}^{\infty}\frac{t^n}{n!}H_n(x) = e^{-(x-t)^2/2}\Big/e^{-x^2/2} = e^{xt-t^2/2},$$

which coincides with the moment generating function $e^{\mu t+\sigma^2 t^2/2}$ of the normal distribution $N(\mu,\sigma^2)$ with μ and σ^2 replaced by x and -1, respectively. Therefore, the

nth Hermite polynomial is the n moment of the normal distribution $N(x, -1)$ formally (McCullagh 1987; Kuriki and Takemura 1996; Withers and Nadarajah 2006), and hence

$$H_n(x) = \sum_{m=0}^{[n/2]} a_{m,n}(-1)^m x^{n-2m}.$$

2.4 Moments of bivariate chi-square distribution

In this subsection, we give an explicit expression for the moment of bivariate chi-square distributions as the second application of the graph presentations for the moments of the noncentral Wishart distribution.

There are several proposals for defining bivariate chi-square (gamma) distributions. The distribution we will discuss here is that given by Kibble's (1941). Kibble's bivariate chi-square distribution is defined as the distribution of the diagonal elements (w_{11}, w_{22}) of a 2×2 central Wishart distribution $(w_{ij}) \sim W_2(v, \Sigma)$ with

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

Substituting $\Theta = \text{diag}(t_{11}, t_{22})$ and $\Delta = 0$ into (2), we have the moment generating function

$$E[e^{t_{11}w_{11}+t_{22}w_{22}}] = \det \begin{pmatrix} 1 - 2t_{11} & -2\rho t_{11} \\ -2\rho t_{22} & 1 - 2t_{22} \end{pmatrix}^{-v/2}$$
$$= (1 - 2t_{11} - 2t_{22} + 4t_{11}t_{22} - 4\rho^2 t_{11}t_{22})^{-v/2}. \qquad (21)$$

The next theorem gives an expression for the moments of general degrees $E[w_{11}^b w_{22}^c]$. We evaluate this as the moment

$$E[\tilde{w}_{12} \cdots \tilde{w}_{2b-1,2b} \tilde{w}_{2b+1,2b+2} \tilde{w}_{2(b+c)-1,2(b+c)}],$$

where $(\tilde{w}_{ij}) \sim W_{2(b+c)}(v, (\sigma_{ij}))$ with

$$\sigma_{ij} = \begin{cases} 1 & (i, j \leq 2b \text{ or } i, j \geq 2b+1), \\ \rho & (\text{otherwise}). \end{cases}$$

Theorem 3 (Moment of the bivariate chi-square distribution) *Let b and c be nonnegative integers. Then,*

$$E[w_{11}^b w_{22}^c] = \sum_{a=0}^{\min(b,c)} \rho^{2a} \frac{2^a \, b! \, c!}{(b-a)! \, (c-a)! \, a!} \prod_{i=1}^{a}(v + 2(a - i))$$
$$\times \prod_{i=1}^{b-a}(v + 2(b - i)) \prod_{i=1}^{c-a}(v + 2(c - i)).$$

Remark 4 Nadarajah and Kotz (2006) derived an expression for $E[w_{11}^b w_{22}^c]$ including the Jacobi polynomials. Their derivation is a use of a series of identities of special functions, which are totally different from our combinatorial proof given below.

Proof Assume that $b \leq c$ without loss of generality. Let $G_0 = (V_1 \cup V_1, E_1 \cup E_2)$ be a union of two undirected graphs (V_i, E_i) $(i = 1, 2)$, where

$$V_1 = \{1, 2, \ldots, 2b-1, 2b\}, \quad E_1 = \{(1, 2), \ldots, (2b-1, 2b)\}$$

and

$$V_2 = \{2b+1, 2b+2, \ldots, 2(b+c)-1, 2(b+c)\},$$

$$E_2 = \{(2b+1, 2b+2), \ldots, (2(b+c)-1, 2(b+c))\}.$$

By forming $b + c$ pairs from $2(b + c)$ vertices of $V = V_1 \cup V_2$, we add $b + c$ edges joining two vertices of each pair to the graph G_0 to make G. Let $h_{l,a,b,c}$ be the number of resulting graphs G such that the number of cycles is l and the number of edges joining an element of V_1 and an element of V_2 (the number of pairs consisting of an element of V_1 and an element of V_2) is $2a$. Then, the moment that we want to evaluate is represented as

$$\sum_{l,a} v^l \rho^{2a} h_{l,a,b,c}.$$

We divide the process of adding $b + c$ edges into three steps below (Fig. 3).

Step (i). Choose $2c - 2a$ from the $2c$ vertices of V_2, and form $c - a$ pairs from the $2c - 2a$ vertices. Add $c - a$ edges joining two vertices of each pair to the graph G_0. The number of resulting graphs having l' cycles is

$$f_{l',c-a,c}.$$

Note that $2a$ vertices not chosen are the terminal vertices of a chains.

Step (ii). Form b pairs from the $2b$ vertices of V_1, and add b edges joining two vertices of each pair. The number of resulting graphs having l'' cycles is

$$f_{l'',b,b}.$$

Step (iii). Choose a edges from the b edges added in step (ii), and make "cuts" at the middle of each edge, and have the a chains generated in step (i) fit in at the a cut points. Note that this operation does not alter the number of cycles. Since the a chains have directions, the number of ways to perform this operation is
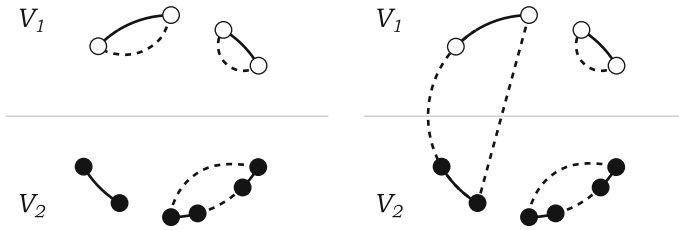
$$2^a \times \frac{b!}{(b-a)!}.$$

Fig. 3 Figures for the proof of Theorem 3 (*Left* steps (i), (ii), *right* step (iii); *open circles* vertex of V_1, *filled circles* vertex of V_2)

Summing up (i), (ii), and (iii), we get

$$h_{l,a,b,c} = \frac{2^a\, b!}{(b-a)!} \sum_{l'+l''=l} f_{l',c-a,c}\, f_{l'',b,b}.$$

Therefore,

$$\sum_{l,a} v^l \rho^{2a} h_{l,a,b,c} = \sum_a \rho^{2a} \frac{2^a\, b!}{(b-a)!} \sum_{l'} v^{l'} f_{l',c-a,c} \sum_{l''} v^{l''} f_{l'',b,b}$$

$$= \sum_a \rho^{2a} \frac{2^a\, b!}{(b-a)!} \binom{c}{c-a} \prod_{i=1}^{c-a}(v + 2(c-i)) \prod_{i=1}^{b}(v + 2(b-i)).$$

□

2.5 Moments of a 2 × 2 real Wishart distribution

As the third example, we give an explicit expression for the moments of a 2×2 central Wishart distribution with the parameter $\Sigma = I$. Let $(w_{ij}) \sim W_2(v, I)$. Substituting $\Theta = \begin{pmatrix} t_{11} & t_{12}/2 \\ t_{12}/2 & t_{22} \end{pmatrix}$ and $\Delta = 0$ into (2), we have the moment generating function

$$E[e^{t_{12}w_{12}+t_{11}w_{11}+t_{22}w_{22}}] = \det \begin{pmatrix} 1 - 2t_{11} & -t_{12} \\ -t_{12} & 1 - 2t_{22} \end{pmatrix}^{-v/2}$$

$$= (1 - 2t_{11} - 2t_{22} + 4t_{11}t_{22} - t_{12}^2)^{-v/2}. \qquad (22)$$

We first show that

$$E[w_{12}^a w_{11}^b w_{22}^c] = 0 \quad \text{for } a \text{ odd}.$$

Let X_t be Gaussian random vectors making up the Wishart matrix. Since $E[X_t] = 0$, the distribution of X_t is invariant under the change of the sign of the first coordinate.

On the other hand, this change causes $(w_{12}, w_{11}, w_{22}) \mapsto (-w_{12}, w_{11}, w_{22})$. This implies that

$$E[w_{12}^a w_{11}^b w_{22}^c] = E[(-w_{12})^a w_{11}^b w_{22}^c]$$
$$= (-1)^a E[w_{12}^a w_{11}^b w_{22}^c].$$

Unless a is even, the left- and right-hand sides become 0.

In the following, we will derive the moment $E[w_{12}^{2a} w_{11}^b w_{22}^c]$, where a, b, and c are nonnegative integers. Although some methods to calculate this quantity are already known (e.g., Remark 5), we demonstrate that our combinatoric approach does get the same results.

Let

$$V_1^0 = \{1, 3, \ldots, 4a - 1\}, \quad V_1^+ = \{4a + 1, 4a + 2, \ldots, 4a + 2b\},$$

and

$$V_2^0 = \{2, 4, \ldots, 4a\}, \quad V_2^+ = \{4a + 2b + 1, 4a + 2b + 2, \ldots, 4a + 2b + 2c\}.$$

Then,

$$E[w_{12}^{2a} w_{11}^b w_{22}^c] = E[\tilde{w}_{12} \tilde{w}_{34} \cdots \tilde{w}_{4a+2b+2c-1,4a+2b+2c}],$$

where $(\tilde{w}_{ij}) \sim W_{4a+2b+2c}(\nu, (\sigma_{ij}))$ with

$$\sigma_{ij} = \begin{cases} 1 & (i, j \in V_1^+, \text{ or } i, j \in V_2^+, \text{ or } i \in V_1^0, j \in V_2^0, \text{ or } i \in V_2^0, j \in V_1^0), \\ 0 & (\text{otherwise}). \end{cases}$$

Theorem 4 (Moment of the 2×2 real Wishart distribution) *Let $a, b,$ and c be nonnegative integers. Then,*

$$E[w_{12}^{2a} w_{11}^b w_{22}^c] = (2a - 1)!! \prod_{i=1}^{a}(\nu + 2(a - i))$$

$$\times \prod_{i=1}^{b}(\nu + 2(a + b - i)) \prod_{i=1}^{c}(\nu + 2(a + c - i)).$$

Proof Let $V_1 = V_1^0 \cup V_1^+$ and $V_2 = V_2^0 \cup V_2^+$. First, define an undirected graph $G_0 = (V, E_0)$ with $4a + 2b + 2c$ vertices $V = V_1 \cup V_2$ and $2a + b + c$ edges

$$E_0 = \{(1, 2), (3, 4), \ldots, (4a + 2b + 2c - 1, 4a + 2b + 2c)\}.$$

Then, consider the addition of additional $2a + b + c$ edges to the graph G_0 to make G such that no edges joining V_1 and V_2 are added. Let $h_{l,a,b,c}$ be the number of resulting
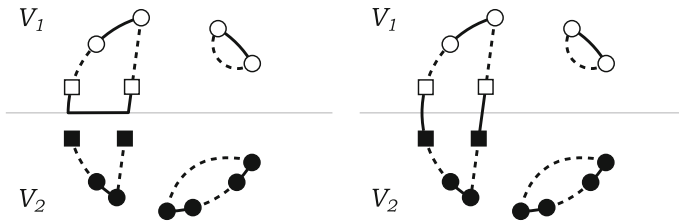
Fig. 4 Figures for the proof of Theorem 4 (*left* steps (i), (ii), *right* step (iii); *open square* vertex of V_1^0, *open circles* vertex of V_1^+, *filled square* vertex of V_2^0, *filled circles* vertex of V_2^+)

graphs G such that G has l cycles. Then,

$$E[w_{12}^{2a} w_{11}^b w_{22}^c] = \sum_{l \geq 0} v^l h_{l,a,b,c}.$$

We divide the process of adding $2a + b + c$ edges to G_0 into three steps (i), (ii), and (iii) below (Fig. 4).

Step (i). Let $0 \leq a' \leq \min(a, c)$. Choose $2c - 2a'$ elements from the $2c$ vertices of V_2^+, and form $c - a'$ pairs from them. Add $c - a'$ edges defined by the $c - a'$ pairs to the graph G_0. According to this operation, a' chains are newly generated. The number of graphs with l' cycles is

$$f_{l',c-a',c}.$$

Step (ii). Form a pairs from the $2a$ vertices of V_2^0. The number of ways in which this pairing can be done is

$$(2a - 1)!! = \frac{(2a)!}{2^a a!}.$$

Choose a' pairs from the a pairs, and assign each pair to each of the a' chains generated in step (i). Connect a vertex of the pair to one terminal vertex of the chain using a new edge, and connect the other vertex of the pair to the other terminal vertex of the chain using another (new) edge. (Add $2a'$ edges in total.) The number of the correspondences is

$$2a(2a - 2) \cdots (2a - 2a') = \frac{2^{a'} a!}{(a - a')!}.$$

For the remaining $a - a'$ pairs, connect two vertices of each pair using a new edge. (Add $a - a'$ edges in total.)

The number of edges added in steps (i) and (ii) is $(c - a') + 2a' + (a - a') = a + c$. The number of cycles is l'.

In steps (i) and (ii), summing the number of ways for $0 \leq a' \leq \min(a, c)$ yields

$$e_{l',c,a} = \sum_{a'=0}^{\min(a,c)} f_{l',c-a',c} \times (2a-1)!! \times \frac{2^{a'} a!}{(a-a')!}.$$

The coefficient $e_{l',c,a}$ can be combinatorially interpreted as follows: let $G_2 = (V_2, E_2)$ with $E_2 = \{(4a+2b+1, 4a+2b+2), \ldots, (4a+2b+2c-1, 4a+2b+2c)\}$ be an undirected graph, and add $a+c$ edges by forming $a+c$ pairs from the $2a+2c$ vertices V_2. Then, a chains are newly generated. $e_{l',c,a}$ is the number of resulting graphs having l' cycles, and the terminal vertices of the a chains are elements of V_2^0.

Step (iii). For all $2a+2b$ vertices of V_1, form $a+b$ pairs and connect the vertices of each pair with a new edge. (Add $a+b$ edges in total.) According to step (ii), the $2a$ vertices of V_1 have already been divided into a pairs, and the vertices of each pair have been connected with an edge. In step (iii), the number of graphs adding new l'' cycles is $f_{l'',a+b,a+b}$.

Summarizing (i), (ii), and (iii), we get

$$h_{l,a,b,c} = \sum_{l'+l''=l} e_{l',c,a} f_{l'',a+b,a+b},$$

and hence,

$$\sum_{l \geq 0} v^l h_{l,a,b,c} = \sum_{l' \geq 0} v^{l'} e_{l',c,a} \sum_{l'' \geq 0} v^{l''} f_{l'',a+b,a+b}. \tag{23}$$

For a nonnegative integer n, write

$$\binom{a}{n} = \frac{a(a-1)\cdots(a-n+1)}{n!}.$$

Then, the generating function of the coefficient $e_{l',c,a}$ with respect to the number of cycles l' is

$$\sum_{l' \geq 0} v^{l'} e_{l',c,a} = (2a-1)!! \sum_{a'=0}^{\min(a,c)} \sum_{l' \geq 0} v^{l'} f_{l',c-a',c} \frac{2^{a'} a!}{(a-a')!}$$

$$= (2a-1)!! \sum_{a'=0}^{\min(a,c)} \binom{c}{c-a'} \prod_{i=1}^{c-a'} (v+2(c-i)) \frac{2^{a'} a!}{(a-a')!}$$

$$= (2a-1)!! \sum_{a'=0}^{\min(a,c)} \binom{c}{c-a'} 2^{c-a'} \binom{v/2+c-1}{c-a'} (c-a')! \frac{2^{a'} a!}{(a-a')!}$$

$$= (2a-1)!! \, 2^c c! \sum_{a'=0}^{\min(a,c)} \binom{a}{a'} \binom{v/2+c-1}{c-a'}$$

$$= (2a - 1)!! \, 2^c c! \binom{\nu/2 + a + c - 1}{c}$$

$$= (2a - 1)!! \prod_{i=1}^{c} (\nu + 2(a + c - i)).$$

The fifth equality above is known as a convolution identity for two binomial coefficients. Substituting this into (23) completes the proof. □

Remark 5 Theorem 4 can also be shown by the following geometric consideration. Let $X_t = \begin{pmatrix} x_{t1} \\ x_{t2} \end{pmatrix}$ be Gaussian random vectors making up the 2×2 Wishart matrix $(w_{ij}) \sim W_2(\nu, I)$. Write $X_{(i)} = (x_{1i}, \ldots, x_{\nu i})'$ $(i = 1, 2)$, where $'$ denotes the matrix transposition. Since the ν-dimensional distribution of $X_{(i)}$ is invariant under the orthogonal transformation preserving the norm $\|X_{(i)}\| = \sqrt{X'_{(i)} X_{(i)}}$, four quantities $\|X_{(i)}\|$, $X_{(i)}/\|X_{(i)}\|$ for $i = 1, 2$ are independently distributed. Thus,

$$E[w_{12}^{2a} w_{11}^b w_{22}^c] = E\left[(X'_{(1)} X_{(2)})^{2a} \|X_{(1)}\|^{2b} \|X_{(2)}\|^{2c} \right]$$
$$= E\left[r^{2a} \right] E\left[\|X_{(1)}\|^{2(a+b)} \right] E\left[\|X_{(2)}\|^{2(a+c)} \right],$$

where $r = X'_{(1)} X_{(2)} / (\|X_{(1)}\| \|X_{(2)}\|)$. This calculation can be completed by noting that $\|X_{(i)}\|^2 \sim \chi_\nu^2$ and $r^2 \sim B\left(\frac{1}{2}, \frac{\nu-1}{2}\right)$, the beta distribution.

3 Moments of the noncentral complex Wishart distribution

3.1 Preliminaries on the complex normal distribution

In this section, we will deal with the complex noncentral Wishart matrices. Major parts of the discussion are parallel to the real case. One remarkable difference is that the moments in the complex case are described in terms of directed graphs, whereas those in the real cases are described in terms of undirected graphs.

We begin by summarizing some preliminaries on the complex normal distribution and the complex Wishart distribution. Let $\overline{Z} = (\overline{z_i})$ be a complex conjugate of $Z = (z_i)$. The following lemma is a complex version of Lemma 1.

Lemma 3 (Moment of the complex normal distribution) *Let* $Z = (z_i) \sim CN(\mu, \Sigma)$, *and let* $\bar{i} = n + i$, $i = 1, \ldots, n'$. *Then,*

$$E[z_1 \cdots z_n \overline{z_{\bar{1}}} \cdots \overline{z_{\bar{n'}}}] = \sum \sigma_{i_1 j_1} \cdots \sigma_{i_m j_m} \mu_{i_{m+1}} \cdots \mu_{i_n} \overline{\mu_{j_{m+1}}} \cdots \overline{\mu_{j_{n'}}},$$

where the summation is over possible pairing $\{(i_1, j_1), (i_2, j_2), \ldots (i_m, j_m)\}$ *such that*

$$\{i_1, \ldots, i_m\} \subset \{1, 2, \ldots, n\} \text{ and } \{j_1, \ldots, j_m\} \subset \{\bar{1}, \ldots, \bar{n'}\}$$

(i.e., matching). The other indices are

$$\{i_{m+1}, \ldots, i_n\} = \{1, \ldots, n\} \setminus \{i_1, \ldots, i_k\},$$
$$\{j_{m+1}, \ldots, j_{n'}\} = \{\overline{1}, \ldots, \overline{n'}\} \setminus \{j_1, \ldots, j_k\}.$$

Proof Write $Z = X + \sqrt{-1}Y$, $\mu = \xi + \sqrt{-1}\eta$, $\Sigma = 2(A + \sqrt{-1}B)$. Let $\theta = (\theta_i)$, $\varphi = (\varphi_i)$ be parameter column vectors. Because of

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N \left(\begin{pmatrix} \xi \\ \eta \end{pmatrix}, \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \right),$$

and

$$\begin{pmatrix} Z \\ \overline{Z} \end{pmatrix} = J \begin{pmatrix} X \\ Y \end{pmatrix}, \qquad J = \begin{pmatrix} I & \sqrt{-1}I \\ I & -\sqrt{-1}I \end{pmatrix},$$

the moment generating function of $(Z, \overline{Z})$ is obtained as

$$
\begin{aligned}
E[\exp\{\theta' Z + \varphi' \overline{Z}\}] &= \exp \left\{ (\theta' \ \varphi') \, J \begin{pmatrix} \xi \\ \eta \end{pmatrix} + \frac{1}{2} (\theta' \ \varphi') \, J \begin{pmatrix} A & -B \\ B & A \end{pmatrix} J' \begin{pmatrix} \theta \\ \varphi \end{pmatrix} \right\} \\
&= \exp \left\{ (\theta' \ \varphi') \begin{pmatrix} \mu \\ \overline{\mu} \end{pmatrix} + \frac{1}{2} (\theta' \ \varphi') \begin{pmatrix} 0 & \Sigma \\ \Sigma & 0 \end{pmatrix} \begin{pmatrix} \theta \\ \varphi \end{pmatrix} \right\} \\
&= \exp \left\{ \theta' \mu + \varphi' \overline{\mu} + \frac{1}{2}(\theta' \Sigma \varphi + \varphi' \overline{\Sigma} \theta) \right\} \\
&= \exp\{\theta' \mu + \varphi' \overline{\mu} + \theta' \Sigma \varphi\}.
\end{aligned}
$$

From this, we have the joint cumulants of $(Z, \overline{Z})$ as

$$
\mathrm{Cum}(z_1, \ldots, z_n, \overline{z_{\overline{1}}}, \ldots, \overline{z_{\overline{n'}}}) = \begin{cases} \mu_1 & (n = 1, \ n' = 0), \\ \overline{\mu_{\overline{1}}} & (n = 0, \ n' = 1), \\ \sigma_{1\overline{1}} & (n = 1, \ n' = 1), \\ 0 & (\text{otherwise}). \end{cases}
$$

Lemma 3 is the moment-cumulant relation for this particular cumulants.   □

3.2 A graph presentation

Let $Z_t = (z_{ti})$ ($t = 1, \ldots, \nu$) be independent complex Gaussian random vectors with mean μ_t and covariance matrix Σ. A complex Wishart matrix $W = (w_{ij})$ is constructed from Z_t as given in (3). In this subsection, we give a formula for the moment $E[w_{ab} w_{cd} \cdots w_{ef}]$ with $a, b, c, d, \ldots, e, f$ arbitrary indices. By considering the degenerate case again, without loss of generality, we only have to treat the moment $E[w_{1\overline{1}} w_{2\overline{2}} \cdots w_{n\overline{n}}]$ with $\overline{i} = n + i$, $i = 1, \ldots, n$.

Let $V = \{1, 2, \ldots, n\}$ be the set of indices appearing in the expectation that we want to evaluate. In the following, we consider a directed graph whose vertices are the elements of V. Choose a subset V_1 of V such that its cardinality is $|V_1| = m$, and consider an injection $\pi : V_1 \to V$. The map π defines a set of directed pairs

$$E = \{(i, \pi(i)) \mid i \in V_1\}.$$

We regard the pair (V, E) as a directed graph G, where V is the set of vertices, and E is the set of directed edges. Note that E and the pair (V_1, π) have one-to-one correspondence.

As in the undirected case, for a given G, every connected component is classified as a "cycle" (a directed path without terminals) and a "chain" (a directed path with a starting terminal and an ending terminal). For the map π, the number of chains is $n - m$, where $n = |V|$, $m = |V_1|$. The number of cycles of G is denoted by $\mathrm{len}(G)$. Note that $\mathrm{len}(G) \leq m$. Let $(j_1, k_1), \ldots, (j_{n-m}, k_{n-m}) \in V \times V$ be directed pairs of ending and starting terminal vertices of $n - m$ chains of G, and let

$$\check{E} = \{(j_1, k_1), \ldots, (j_{n-m}, k_{n-m})\}.$$

Using the notations above, we give the general form for the moments as follows.

Theorem 5 (Moment of the complex noncentral Wishart distribution) *Let $(w_{ij}) \sim CW(\nu, (\sigma_{ij}), (\delta_{ij}))$, and let $\bar{i} = i + n$, $i = 1, \ldots, n$. Then,*

$$E[w_{1\bar{1}} \cdots w_{n\bar{n}}] = \sum_E \nu^{\mathrm{len}(G)} \sigma^E \delta^{\check{E}}, \tag{24}$$

where

$$\sigma^E = \prod_{(i,i') \in E} \sigma_{i\overline{i'}} = \sigma_{i_1\overline{\pi(i_1)}} \cdots \sigma_{i_m\overline{\pi(i_m)}},$$

$$\delta^{\check{E}} = \prod_{(j,j') \in \check{E}} \delta_{j\overline{j'}} = \delta_{j_1\overline{j_1'}} \cdots \delta_{j_{n-m}\overline{j_{n-m}'}}.$$

The summation $\sum_E$ is taken over all possibilities of $V_1 = \{i_1, \ldots, i_m\} \subset \{1, \ldots, n\}$, and injections $\pi : V_1 \to V$.

Example 2 Consider the evaluation of the moment $E[w_{1\bar{1}}w_{2\bar{2}}w_{3\bar{3}}]$. Then, $V = \{1, 2, 3\}$. There are 34 injections from subsets $V_1 \subset V$ to V. Figure 5 is the graph $G = (V, E)$ for $E = \{(1, 1), (2, 3)\}$ ($\check{E} = \{(3, 2)\}$). Summing up 34 possibilities, we have the following:

$$
\begin{aligned}
E[w_{1\bar{1}}w_{2\bar{2}}w_{3\bar{3}}] = {}& \nu^3 \sigma_{1\bar{1}}\sigma_{2\bar{2}}\sigma_{3\bar{3}} + \nu^2 \sigma_{1\bar{2}}\sigma_{2\bar{1}}\sigma_{3\bar{3}}[3] + \nu\sigma_{1\bar{2}}\sigma_{2\bar{3}}\sigma_{3\bar{1}}[2] \\
& + \nu^2 \sigma_{1\bar{1}}\sigma_{2\bar{2}}\delta_{3\bar{3}}[3] + \nu\sigma_{1\bar{2}}\sigma_{2\bar{1}}\delta_{3\bar{3}}[3] + \nu\sigma_{1\bar{1}}\sigma_{2\bar{3}}\delta_{3\bar{2}}[6] + \sigma_{1\bar{2}}\sigma_{2\bar{3}}\delta_{3\bar{1}}[6] \\
& + \nu\sigma_{1\bar{1}}\delta_{2\bar{2}}\delta_{3\bar{3}}[3] + \sigma_{1\bar{2}}\delta_{2\bar{1}}\delta_{3\bar{3}}[6] \\
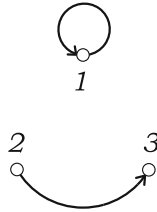& + \delta_{1\bar{1}}\delta_{2\bar{2}}\delta_{3\bar{3}}.
\end{aligned}
$$

Fig. 5 Directed graph $G = (V, E)$ presenting the term $\nu^1 \sigma_{1\bar{1}} \sigma_{2\bar{3}} \delta_{3\bar{2}}$ ($n = 3, m = 2, \text{len}(G) = 1$)

Here, $[n]$ means that there are n terms of similar form.

Proof Note that $w_{ij} = \sum_{t=1}^{\nu} z_{ti} \overline{z_{tj}}$. In view of Lemma 3, we have

$$E[w_{1\bar{1}} \cdots w_{n\bar{n}}]$$

$$= \sum_{t_1=1}^{\nu} \cdots \sum_{t_n=1}^{\nu} E[z_{t_1,1} \overline{z_{t_1,\bar{1}}} \cdots z_{t_n,n} \overline{z_{t_n,\bar{n}}}]$$

$$= \sum_{E} \sum_{t_1} \cdots \sum_{t_n} \text{Cum}(z_{t_{i_1},i_1}, \overline{z_{t_{\pi(i_1)},\pi(i_1)}}) \cdots \text{Cum}(z_{t_{i_m},i_m}, \overline{z_{t_{\pi(i_m)},\pi(i_m)}})$$

$$\times E[z_{t_{i_{m+1}},i_{m+1}}] \cdots E[z_{t_{i_n},i_n}] E[\overline{z_{t_{i'_{m+1}},i'_{m+1}}}] \cdots E[\overline{z_{t_{i'_n},i'_n}}], \tag{25}$$

where $V_1 = \{i_1, \ldots, i_m\}$,

$$\{i_{m+1}, \ldots, i_n\} = V \setminus V_1, \quad \{i'_{m+1}, \ldots, i'_n\} = V \setminus \pi(V_1).$$

Since $\{i_1, \ldots, i_n\} = V$, the indices $i_1, \ldots, i_n$ can be divided into connected components of the graph G. A connected component having vertices $j_1, \ldots, j_k$ forms either a directed chain

$$(j_1, j_2), (j_2, j_3), \ldots, (j_{k-2}, j_{k-1}), (j_{k-1}, j_k)$$

or a directed cycle

$$(j_1, j_2), (j_2, j_3), \ldots, (j_{k-1}, j_k), (j_k, j_1),$$

where $\pi(j_i) = j_{i+1}$ (and $\pi(j_k) = j_1$ in the cycle case). Since the running n indices $t_1, \ldots, t_n$ correspond to n vertices of V, the argument of the summation $\sum_E$ in (25) is a product of terms of the form

$$\sum_{t_{j_1}} \sum_{t_{j_2}} \cdots \sum_{t_{j_k}} E[\overline{z_{t_{j_1},\bar{j_1}}}] \text{Cum}(z_{t_{j_1},j_1}, \overline{z_{t_{j_2},\bar{j_2}}}) \text{Cum}(z_{t_{j_2},j_2}, \overline{z_{t_{j_3},\bar{j_3}}}) \cdots$$

$$\times \text{Cum}(z_{t_{j_{k-1}},j_{k-1}}, \overline{z_{t_{j_k},\bar{j_k}}}) E[z_{t_{j_k},j_k}] \tag{26}$$

in the chain case, or

$$\sum_{t_{j_1}} \sum_{t_{j_2}} \cdots \sum_{t_{j_k}} \text{Cum}(z_{t_{j_1},j_1}, \overline{z_{t_{j_2},\overline{j_2}}}) \text{Cum}(z_{t_{j_2},j_2}, \overline{z_{t_{j_3},\overline{j_3}}}) \cdots$$
$$\times \text{Cum}(z_{t_{j_{k-1}},j_{k-1}}, \overline{z_{t_{j_k},\overline{j_k}}}) \text{Cum}(z_{t_{j_k},j_k}, \overline{z_{t_{j_1},\overline{j_1}}}) \tag{27}$$

in the cycle case. Noting that

$$\text{Cum}(z_{si}, \overline{z_{t\overline{j}}}) = \text{Cov}(z_{si}, z_{t\overline{j}}) = 1_{\{s=t\}}\sigma_{i\overline{j}}$$

and $\sum_{t=1}^{\nu} E[z_{ti}]E[\overline{z_{t\overline{j}}}] = \delta_{i\overline{j}}$, we see that (26) $= \sigma_{j_1\overline{j_2}}\sigma_{j_2\overline{j_3}}\cdots\sigma_{j_{k-1}\overline{j_k}}\delta_{j_k\overline{j_1}}$ and (27) $= \nu\sigma_{j_1\overline{j_2}}\sigma_{j_2\overline{j_3}}\cdots\sigma_{j_{k-1}\overline{j_k}}\sigma_{j_k\overline{j_1}}$. This completes the proof. $\qquad\square$

3.3 Enumeration of directed graphs

In this subsection, we evaluate the number of directed graphs appearing in the expression (24) for the moments of the complex Wishart distribution.

Let $V = \{1, \ldots, n\}$. Let V_1 be a subset of V such that $|V_1| = m$. Let π be an injection $V_1 \to V$. As explained in the previous subsection, we can define a directed graph $G = (V, E)$ with $E = \{(i, \pi(i)) \mid i \in V_1\}$. The connected components of G are either a directed cycle or a directed chain (the length may be 0). Note that the number of chains is $n - m$. Let $g_{l,m,n}$ be the number of such graphs having l cycles. The coefficient $g_{l,m,n}$ satisfies the following recurrence formula.

Lemma 4

$$g_{l,m,n} = g_{l-1,m-1,n-1} + g_{l,m,n-1} + (2n - m - 1)g_{l,m-1,n-1} \tag{28}$$

with boundary conditions

$$g_{l,0,n} = \begin{cases} 1 & (l = 0), \\ 0 & (l \geq 1) \end{cases} \quad \text{for } n \geq 1, \tag{29}$$

and

$$g_{l,1,1} = \begin{cases} 0 & (l = 0), \\ 1 & (l = 1). \end{cases} \tag{30}$$

Proof We consider removing the vertex n and the adjacent edges from the graph G. There are three types of status about adjacent edges.

Case (i). $\pi(n) = n$. In this case, the vertex n is contained in a cycle with length 1. Removing the vertex n and the edge $(n, \pi(n))$ yields a graph whose values of l, m and n are one less than those of G. This corresponds to the first term in the right-hand side of (28).

Case (ii). Neither $\pi(n)$ nor $\pi^{-1}(n)$ exists. In this case, the vertex n is an isolated vertex. Removing the vertex n yields a graph whose l and m are invariant and n are one less than that of G. This corresponds to the second term in the right-hand side of (28).

Case (iii). Otherwise. In this case, the vertex n is contained in a cycle with length more than or equal to 2, or contained in a chain with length more than or equal to 1. Remove the vertex n and one edge adjacent to the vertex n. This manipulation yields a graph whose l is invariant, and whose m and n are one less than those of G. Conversely, when we rebuild the graph G from this smaller graph, there are $(n-1) + \{(n-1) - (m-1)\} = 2n - m - 1$ places where the vertex n and one edge can be inserted. This corresponds to the third term in the right-hand side of (28). □

Theorem 6 *The generating function of the coefficient $g_{l,m,n}$ with respect to the cycle number l,*

$$\Psi_{m,n}(\nu) = \sum_{l \geq 0} \nu^l g_{l,m,n},$$

is given by

$$\Psi_{m,n}(\nu) = \binom{n}{m} \prod_{i=1}^{m} (\nu + n - i) \quad (0 \leq m \leq n, \ n \geq 1). \tag{31}$$

Here, we use a convention $\prod_{i=1}^{0} = 1$.

Proof Because $g_{-1,m,n} = 0$, we see

$$\Psi_{m,n} = \nu \Psi_{m-1,n-1} + \Psi_{m,n-1} + (2n - m - 1) \Psi_{m-1,n-1}. \tag{32}$$

The boundary conditions $\Psi_{0,n} = 1 \ (n \geq 1)$ due to (29) and $\Psi_{1,1} = \nu$ due to (30) satisfy (31). In addition, since $\Psi_{n,n-1} = 0$,

$$\Psi_{n,n} = (\nu + n - 1)\Psi_{n-1,n-1} = \cdots = \prod_{i=1}^{n} (\nu + n - i) \quad (n \geq 1).$$

Hence, it is sufficient to ensure that (31) satisfies (32). Indeed, we have

$$\Psi_{m,n} - \Psi_{m,n-1} = \binom{n}{m} \prod_{i=1}^{m} (\nu + n - i) - \binom{n-1}{m} \prod_{i=1}^{m} (\nu + n - 1 - i)$$

$$= \binom{n-1}{m-1} \frac{1}{m} \prod_{i=1}^{m-1} (\nu + n - 1 - i)$$

$$\times \{n(\nu + n - 1) - (n - m)(\nu + n - 1 - m)\}$$

$$= (2n - m - 1)\Psi_{m-1,n-1} + \nu \Psi_{m-1,n-1}.$$

□

Remark 6 Comparing (31) with (19) in Remark 2, we have

$$g_{l,m,n} = \binom{n}{m} s_n(m, l),$$

where $s_n(m, l)$ is the noncentral Stirling numbers of the first kind. In particular, $g_{l,n,n} = s_n(n, l)$ is the Stirling numbers of the first kind. Let $V = \{1, \dots, n\}$ and let $\pi : V \to V$ (bijection). It is well-known that the Stirling number of the first kind $s_n(n, l)$ is the number of directed graphs (V, E), $E = \{(i, \pi(i)) \mid i \in V\}$ having l cycles (Stanley (2000)).

Corollary 2

$$\Psi_{m,n}(1) = \binom{n}{m} n(n-1) \cdots (n-m+1) = \binom{n}{m}^2 m!$$

is the number of directed graphs G, and

$$\Psi_{m,n}(0) = \binom{n}{m}(n-1)(n-2) \cdots (n-m)$$

$$= \frac{n!(n-1)!}{m!(n-m)!(n-m-1)!}$$

is the number of directed graphs G without cycles.

3.4 Degenerate cases

3.4.1 The noncentral chi-square distribution

As in the real case, we can obtain several identities for moments by assuming that the parameters Σ and Δ have particular kinds of structures. First, we consider the case where $\Sigma = (\sigma_{ij})$, $\sigma_{ij} \equiv 2$, and $\Delta = (\delta_{ij})$, $\delta_{ij} \equiv \delta$. Then, every element of W has the same value w, say, with probability one, and the distribution of w is the noncentral chi-square distribution $\chi^2_{2\nu}(\delta)$ with 2ν degrees of freedom and the noncentrality parameter δ. The nth moment of $w \sim \chi^2_{2\nu}(\delta)$ is

$$\begin{aligned}
E[w^n] &= E[w_{1\bar{1}} \cdots w_{n\bar{n}}] \\
&= \sum_{m=0}^{n} \sum_{l \geq 0} \nu^l g_{l,m,n} 2^m \delta^{n-m} = \sum_{m=0}^{n} \Psi_{m,n}(\nu) 2^m \delta^{n-m} \\
&= \sum_{m=0}^{n} \binom{n}{m} \prod_{i=1}^{m} (\nu + n - i) 2^m \delta^{n-m} = \sum_{m=0}^{n} \binom{n}{m} \prod_{i=1}^{m} (2\nu + 2(n-i)) \delta^{n-m}.
\end{aligned}$$

This coincides with the formula (20) obtained from the real Wishart distribution.

3.4.2 Moments of a bivariate chi-square distribution associated with the complex Wishart distribution

Let $(w_{ij}) \sim CW_2(\nu, \Sigma)$ be a 2×2 central complex Wishart matrix. We consider a particular structure of the parameter:

$$\Sigma = 2 \begin{pmatrix} 1 & \overline{\rho} \\ \rho & 1 \end{pmatrix}, \quad \rho \in \mathbb{C}.$$

The diagonal elements (w_{11}, w_{22}) are distributed according to a sort of bivariate chi-square distribution, since the marginal distribution of w_{11} and w_{22} are the chi-square distribution $\chi^2_{2\nu}$ and they are correlated. At a glance, (w_{11}, w_{22}) has a different distribution from Kibble's distribution since it has a different origin. However, from (4), the moment generating function is

$$E[e^{t_{11}w_{11}+t_{22}w_{22}}] = \det \begin{pmatrix} 1 - 2t_{11} & -2\overline{\rho}t_{11} \\ -2\rho t_{22} & 1 - 2t_{22} \end{pmatrix}^{-\nu}$$
$$= (1 - 2t_{11} - 2t_{22} + 4t_{11}t_{22} - 4\rho\overline{\rho}t_{11}t_{22})^{-\nu},$$

which is equal to the moment generating function (21) of Kibble's bivariate chi-square distribution with ν and ρ replaced by 2ν and $\sqrt{\rho\overline{\rho}}$, respectively.

3.4.3 Moments of a 2×2 complex Wishart distribution

Consider a 2×2 complex Wishart matrix $(w_{ij}) \sim CW_2(\nu, I)$. We first show that

$$E[w_{12}^a w_{21}^{a'} w_{11}^b w_{22}^c] = 0 \quad \text{if } a \neq a'.$$

Let Z_t be complex Gaussian random variables from which the Wishart matrix W is constructed. Since $E[Z_t] = 0$, its distribution is invariant when the first element of Z_t is multiplied by $\sqrt{-1}$. On the other hand, this manipulation causes

$$\begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} \mapsto \begin{pmatrix} w_{11} & \sqrt{-1}w_{12} \\ -\sqrt{-1}w_{21} & w_{22} \end{pmatrix}.$$

Therefore,

$$E[w_{12}^a w_{21}^{a'} w_{11}^b w_{22}^c] = E[(\sqrt{-1}w_{12})^a (-\sqrt{-1}w_{21})^{a'} w_{11}^b w_{22}^c]$$
$$= \sqrt{-1}^{a-a'} E[w_{12}^a w_{21}^{a'} w_{11}^b w_{22}^c].$$

The left- and right-hand sides become 0 unless $a = a'$.

Substituting $\Theta = (t_{ij})'$ and $\Delta = 0$ into (4), we obtain the moment generating function

$$E[e^{t_{12}w_{12}+t_{21}w_{21}+t_{11}w_{11}+t_{22}w_{22}}] = \det \begin{pmatrix} 1-t_{11} & -t_{12} \\ -t_{21} & 1-t_{22} \end{pmatrix}^{-\nu}$$

$$= (1 - t_{11} - t_{22} + t_{11}t_{22} - t_{12}t_{21})^{-\nu}. \quad (33)$$

Let

$$\phi_\nu(u, v, w) = (1 - u - v + uv - w)^{-\nu}, \quad \phi_\nu^{(k)}(u, v, w) = \left(\frac{\partial}{\partial w}\right)^k \phi_\nu(u, v, w).$$

Then,

$$(33) = \phi_\nu(t_{11}, t_{22}, t_{12}t_{21}) = \sum_{k \geq 0} \frac{(t_{12}t_{21})^k}{k!} \phi_\nu^{(k)}(t_{11}, t_{22}, 0).$$

On the other hand, the moment generating function of the 2×2 real Wishart distribution is rewritten as

$$(22) = \phi_{\nu/2}(2t_{11}, 2t_{22}, t_{12}^2) = \sum_{k \geq 0} \frac{t_{12}^{2k}}{k!} \phi_{\nu/2}^{(k)}(2t_{11}, 2t_{22}, 0).$$

Comparing the two functions, we can immediately obtain the moments of the 2×2 complex Wishart distribution from the results for the real case in Theorem 4.

Theorem 7 (Moment of the 2×2 complex Wishart distribution) *Let a, b, and c be nonnegative integers. Then,*

$$E[w_{11}^b(w_{12}w_{21})^a w_{22}^c] = a! \prod_{i=1}^a (\nu + a - i) \prod_{i=1}^b (\nu + a + b - i) \prod_{i=1}^c (\nu + a + c - i).$$

References

Bai, Z. D. (1999). Methodologies in spectral analysis of large dimensional random matrices: A review. *Statistica Sinica, 9*, 611–677.

Goodman, N. R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution: An introduction. *The Annals of Mathematical Statistics, 34*, 152–177.

Graczyk, P., Letac, G., Massam, H. (2003). The complex Wishart distribution and the symmetric groups. *The Annals of Statistics, 31*, 287–309.

Graczyk, P., Letac, G., Massam, H. (2005). The hyperoctahedral group, symmetric group representations and the moments of the real Wishart distribution. *Journal of Theoretical Probability, 18*, 1–42.

Johnson, N. L., Kotz, S., Balakrishnan, N. (1995). *Continuous univariate distributions* (Vol. 2, 2nd ed.). New York: Wiley-Interscience.

Kibble, W. F. (1941). A two-variate gamma type distribution. *Sankhya, 5A*, 137–150.

Koutras, M. (1982). Noncentral Stirling numbers and some applications. *Discrete Mathematics, 42*, 73–89.

Kuriki, S., Takemura, A. (1996). Asymptotic expansion of null distribution of likelihood ratio statistic in multiparameter exponential family to an arbitrary order. In S. Watanabe, M. Fukushima, Yu. V. Prohorov, A. N. Shiryaev (Eds.), *Probability theory and mathematical statistics: Proceedings of the seventh Japan–Russia symposium* (pp. 244–255). Singapore: World Scientific.

Letac, G., Massam, H. (2008). The noncentral Wishart as an exponential family, and its moments. *Journal of Multivariate Analysis, 99*, 1393–1417.

Lu, I.-L., Richards, D. St. P. (2001). MacMahon's master theorem, representation theory, and moments of Wishart distributions. *Advances in Applied Mathematics, 27*, 531–547.

Maiwald, D., Kraus, D. (2000). Calculation of moments of complex Wishart and complex inverse Wishart distributed matrices. *IEE Proceedings Radar, Sonar & Navigation, 147*, 162–168.

McCullagh, P. (1987). *Tensor methods in statistics*. London: Chapman & Hall.

Morris, C. N. (1982). Natural exponential families with quadratic variance functions. *The Annals of Statistics, 10*, 65–80.

Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. New York: Wiley.

Nadarajah, S., Kotz, S. (2006). Product moments of Kibble's bivariate gamma distribution. *Circuits, Systems, and Signal Processing, 25*, 567–570.

Stanley, R.P. (2000). *Enumerative combinatorics* (Vol. 1, 2nd ed.) Cambridge: Cambridge University Press.

Takemura, A. (1991). *Foundations of multivariate statistical inference* (in Japanese). Tokyo: Kyoritsu Shuppan.

Vere-Jones, D. (1988). A generalization of permanents and determinants. *Linear Algebra and its Applications, 111*, 119–124.

Wishart, J. (1928). The generalised product moment distribution in samples from a normal multivariate population. *Biometrika, 20A*, 32–52.

Withers, C., Nadarajah, S. (2006). Simple representations for Hermite polynomials. *Electronics Letters, 42*, 1368–1369.