

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

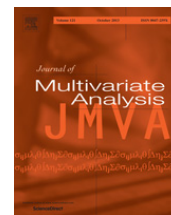
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/authorsrights>



Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

On a symbolic representation of non-central Wishart random matrices with applications



Elvira Di Nardo

Department of Mathematics, Computer Science and Economics, University of Basilicata, Viale dell'Ateneo Lucano 10,
85100 Potenza, Italy

ARTICLE INFO

Article history:

Received 12 February 2013

Available online 17 December 2013

AMS subject classifications:

primary 62H05

60B20

secondary 60E05

62E17

Keywords:

Random matrix

Moment method

Umbral calculus

Sheffer polynomial sequence

Complete homogeneous polynomial

Cyclic polynomial

Cumulant

Necklace

Permanent

Spectral polykay

Free probability

ABSTRACT

By using a symbolic method, known in the literature as the classical umbral calculus, the trace of a non-central Wishart random matrix is represented as the convolution of the traces of its central component and of a formal variable matrix. Thanks to this representation, the moments of this random matrix are proved to be a Sheffer polynomial sequence, allowing us to recover several properties. The multivariate symbolic method generalizes the employment of Sheffer representation and a closed form formula for computing joint moments and cumulants (also normalized) is given. By using this closed form formula and a combinatorial device, known in the literature as necklace, an efficient algorithm for their computations is set up. Applications are given to the computation of permanents as well as to the characterization of inherited estimators of cumulants, which turn useful in dealing with minors of non-central Wishart random matrices. An asymptotic approximation of generalized moments involving free probability is proposed.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Let $\{X_{(1)}, \dots, X_{(n)}\}$ be random row vectors independently drawn from a p -variate normal distribution with zero mean $\mathbf{0}$ and full rank covariance matrix Σ . Let $\mathbf{m}_1, \dots, \mathbf{m}_n$ be row vectors of dimension p . The *non-central Wishart distribution* is the distribution of the square random matrix of order p

$$W_p(n, \Sigma, M) = \sum_{i=1}^n (X_{(i)} - \mathbf{m}_i)^T (X_{(i)} - \mathbf{m}_i), \quad \text{with } M = \sum_{i=1}^n \mathbf{m}_i^T \mathbf{m}_i, \quad (1)$$

denoted by $W(n)$ for simplicity. The matrix $\Omega = \Sigma^{-1}M$ is called the non-centrality matrix and it is usually employed instead of M to parametrize the Wishart distribution. For $M = \mathbf{0}$ the Wishart distribution is said to be *central* and denoted by $\widehat{W}(n)$ in the following.

The most popular application of Wishart distributions is within maximum likelihood estimation in connection with the sample covariance matrix. Therefore these distributions are successfully employed in several areas: a good review is given

E-mail addresses: elvira.dinardo@unibas.it, dinardo.el@gmail.com.

URL: <http://www.unibas.it/utenti/dinardo/>.

in [28]. In parallel to its spread in the applications, also its mathematical properties have received great attention in the literature, including those involving determinants and eigenvalues. According to the so-called moment method [26], which relies on eigenvalue distribution, the i th moment of a $p \times p$ random matrix A is $E[\text{Tr}_p(A^i)]$, where Tr_p denotes the normalized trace. In particular, the computation of joint moments

$$E \left\{ \text{Tr} [W(n) H_1]^{i_1} \cdots \text{Tr} [W(n) H_m]^{i_m} \right\} \quad \text{with } H_1, \dots, H_m \in \mathbb{C}^{p \times p} \tag{2}$$

is a very general task: for example, when H_1, \dots, H_m are sparse matrices, Eq. (2) returns joint moments of entries of $W(n)$. Moments of Wishart random matrices are employed not only to characterize asymptotic properties of the sample covariance matrix [19] but also to quantify the performance of many multidimensional signal processing algorithms [25]. For example, within wireless communication theory, most literature focuses on channel capacity modeled by a memoryless linear vector of the form

$$\mathbf{y} = B\mathbf{x} + \mathbf{n} \tag{3}$$

with \mathbf{x} a p -dimensional input vector, \mathbf{y} a n -dimensional output vector and \mathbf{n} a n -dimensional vector representing a circularly symmetric Gaussian noise. The $n \times p$ random matrix B encodes the characteristics of the channel system having n receivers and p transmitters. Quite often performance measures of (3) rely on the product $B^\dagger B$, which is a non-central Wishart random matrix when the entries of B are Gaussian i.i.d. random variables (r.v.'s) [27]. Numerical evaluations of these indexes involve cumbersome integrals, but fortunately, much deeper insights can be obtained using the tools provided by asymptotic random matrix theory, that is to characterize the asymptotic spectrum of B as the number of columns and rows goes to infinity. Usually the Stieltjes transformation is employed as uniquely determines the distribution function of the spectrum. As an alternative, especially when the inversion formula of Stieltjes–Perron could not be applied, the evaluation of normalized moments $\text{Tr}(B^\dagger B)^i$ together with the study of their convergence provides a tool to characterize the limiting spectrum through the moment convergence theorem. Since random matrices are non-commutative objects, for dimensions greater than 8 free probability tools can be employed to recognize if two or more channels are independent. Within free probability, the classical notion of independence is replaced by the notion of freeness. Freeness allows us to recover the asymptotic spectrum of the sum of random matrices from the individual asymptotic spectra. In order to recognize the freeness property a great number of joint moments need to be computed, so that efficient algorithms are necessary.

For finite matrix theory, the computation of joint moments is still an object of in-depth analysis due to its complexity. In order to compute (2) when $i_1 = \dots = i_m = 1$, the theory of representation group together with the differential of the Laplace transform in suitable directions have been employed in [12,16]. Weighted generating functions on matchings of graphs have been used in [15]. In order to compute joint moments in terms of joint cumulants, multivariate Bell polynomials are resorted in [28] but the resulting expressions are cumbersome to deal with. Via Edgeworth expansion, joint cumulants of Wishart distributions are employed in approximating density or distribution functions of some statistics relied on quadratic forms of normal samples [14].

The aim of this paper is twofold: to give a closed form formula to compute joint moments (2), thus generalizing the result given in [16], and to introduce a symbolic method, known in the literature as the classical umbral calculus, as a natural way to deal with moments of Wishart distributions. This method allows us to manage moment sequences without making hypothesis on the existence of an underlying distribution probability. Many results rely on the representation of the trace of a non-central Wishart random matrix as the convolution of the traces of its central component and of a formal variable (or symbol) matrix, whose entries are *uncorrelated* with those of the central component. This representation allows us to express $\{E[\text{Tr}(W(n)^i)]\}$ as a Sheffer polynomial sequence [7]. Then many Sheffer properties are shared: for example the role played by complete homogeneous symmetric polynomials as well as cyclic polynomials is highlighted when referred to the central component, whose moments result to be a sequence of binomial type. When the indeterminate of a Sheffer polynomial sequence is replaced by a suitable r.v., a randomized non-central Wishart distribution is recovered. The multivariate version of this symbolic method extends the employment of Sheffer polynomial sequences to the computation of joint moments (2). Moreover, in order to take into account the cyclic property of traces, the notion of necklace is fruitfully employed in setting up an efficient symbolic procedure. The algorithm in Maple 12 is available on demand. Parallelisms with other techniques proposed for computing (2) are given forward.

The usefulness of the symbolic representation of non-central Wishart distributions is not only confined to computational issues. To widening its applicability, further applications are proposed at the end of the paper. The first concerns the computation of permanents of special classes of matrices. The second characterizes spectral polykays which are unbiased estimators of cumulants. Spectral polykays have the advantage to preserve its expansion in terms of power sum symmetric polynomials when referred to principal submatrices. Last application allows us an asymptotic approximation of generalized moments relied on some results borrowed from free probability.

In order to make the paper self-contained, a short introduction on the symbolic method and on its multivariate version is given in Sections 2 and 4 respectively. Open problems still concern an exact computation of generalized moments as well as the employment of the symbolic method in computing moments of a non-central Wishart random matrix inverse, which are also employed within signal processing [19]. The difficulty of this task relies on the circumstance that these moments are usually expressed in terms of zonal polynomials whose efficient handling is still an appealing problem from a computational point of view.

2. Moment symbolic method: the univariate case

The moment symbolic method we use relies on the classical umbral calculus introduced by Rota and Taylor in 1994 [23]. The method has been developed and refined in a series of papers starting from [9,10] reducing the overall computational apparatus to few fundamental relations. The key point consists in working with symbols instead of scalar or polynomial sequences.

Let \mathbb{C} be the complex field whose elements are called scalars. An umbral calculus consists in a generating set $\mathcal{A} = \{\alpha, \beta, \dots\}$, called *alphabet*, whose elements are named *umbræ*, a polynomial ring $\mathbb{C}[\mathcal{A}]$ and a linear functional $E: \mathbb{C}[\mathcal{A}] \rightarrow \mathbb{C}$ called *evaluation*. The linear functional E is such that $E[1] = 1$ and

$$E[\alpha^i \beta^j \dots \gamma^k] = E[\alpha^i] E[\beta^j] \dots E[\gamma^k], \quad (\text{uncorrelation property}) \quad (4)$$

for any set of distinct umbræ in \mathcal{A} and for i, j, k nonnegative integers. In particular, we set $E[\alpha^i] = a_i$ for $i = 0, 1, \dots$ with $a_0 = 1$ and $\alpha \in \mathcal{A}$. The scalar a_i is the i th moment of α and we say that the sequence $\{a_i\}$ is *umbrally represented* by α . Conversely, a sequence of scalars $\{a_i\}$ with $a_0 = 1$ is represented by an umbra α if $E[\alpha^i] = a_i$ for $i = 0, 1, \dots$. Indeed it is always possible to insert in the alphabet new symbols such that each scalar sequence $\{a_i\}$ corresponds to an element in \mathcal{A} , which is not necessarily unique. Indeed, two umbræ α and γ may represent the same sequence of moments when $E[\alpha^i] = E[\gamma^i]$ for all nonnegative integers i . In this case, α and γ are said to be similar, in symbols $\alpha \equiv \gamma$. For umbral polynomials $p, q \in \mathbb{C}[\mathcal{A}]$ a nimbler equivalence is the umbral equivalence: $p \simeq q$ if and only if $E[p] = E[q]$. Umbral polynomials with disjoint supports¹ are uncorrelated, that is $E[pq] = E[p]E[q]$.

The alphabet \mathcal{A} contains special symbols representing special sequences. For example the *augmentation umbra* ε is such that $E[\varepsilon^i] = 0$ for all positive integers i and the *unity umbra* u is such that $E[u^i] = 1$ for all positive integers i . Further special umbræ will be introduced when necessary. The formal power series in $\mathbb{C}[\mathcal{A}][[z]]$

$$u + \sum_{i \geq 1} \alpha^i \frac{z^i}{i!} \quad (5)$$

is the *generating function* (g.f.) of an umbra α and denoted by $e^{\alpha z}$. Moreover, for any exponential formal power series in $\mathbb{C}[[z]]$

$$f(z) = 1 + \sum_{i \geq 1} a_i \frac{z^i}{i!} \quad (6)$$

there exists an umbra α such that $E[e^{\alpha z}] = f(z)$, by extending coefficient-wise the action of E , cf. [4]. The formal power series $f(z)$ in (6) is the g.f. of α , usually denoted by $f(\alpha, z)$ to avoid misunderstandings.

For distinct (and so uncorrelated) umbræ α and γ , we have $f(\alpha + \gamma, z) = f(\alpha, z)f(\gamma, z)$. More in general $f(\alpha' + \dots + \alpha'', z) = [f(\alpha, z)]^n$ if α', \dots, α'' are n uncorrelated umbræ similar to an umbra α . The summation $\alpha' + \dots + \alpha''$ is denoted by the auxiliary symbol $n \cdot \alpha$ which is named as the *dot product* of the nonnegative integer n and the umbra α . Its moments are [4]

$$E[(n \cdot \alpha)^i] = \sum_{\lambda \vdash i} (n)_{l(\lambda)} d_\lambda a_\lambda \quad (7)$$

where $\lambda = (1^{r_1}, 2^{r_2}, \dots)$ is a partition² of the integer i , $a_\lambda = a_1^{r_1} a_2^{r_2} \dots$ and

$$d_\lambda = \frac{i!}{(1!)^{r_1} r_1! (2!)^{r_2} r_2! \dots}$$

Let us recall that if $\alpha, \gamma \in \mathcal{A}$ then

$$n \cdot (\alpha + \gamma) \equiv n \cdot \alpha + n \cdot \gamma \quad \text{and} \quad (n + m) \cdot \alpha \equiv n \cdot \alpha + m \cdot \alpha. \quad (8)$$

The integer n in (7) can be replaced by -1 . The auxiliary umbra $-1 \cdot \alpha$ is the *inverse umbra* of α with g.f. $f(-1 \cdot \alpha, z) = f(\alpha, z)^{-1}$. Also the composition of g.f.'s can be symbolically represented by using the dot-product. The first step is to replace the integer n with an umbra $\gamma \in \mathcal{A}$, so that

$$f(\gamma \cdot \alpha, z) = f(\gamma, \log[f(\alpha, z)]), \quad (9)$$

which is not yet the composition of $f(\alpha, z)$ and $f(\gamma, z)$ but plays a fundamental role in the symbolic method. Indeed, if in (9) the umbra γ is replaced by the *singleton umbra* χ with g.f. $f(\chi, z) = 1 + z$, then $f(\chi \cdot \alpha, z) = 1 + \log[f(\alpha, z)]$ and the

¹ The support of an umbral polynomial $p \in \mathbb{C}[\mathcal{A}]$ is the set of all umbræ which occur.

² Recall that a partition of an integer i is a sequence $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_l)$, where λ_j are weakly decreasing integers and $\sum_{j=1}^l \lambda_j = i$. The integers λ_j are named *parts* of λ . The *length* of λ is the number of its parts and will be denoted by $l(\lambda)$. A different notation is $\lambda = (1^{r_1}, 2^{r_2}, \dots)$, where r_j is the number of parts of λ equal to j and $r_1 + r_2 + \dots = l(\lambda)$. We use the classical notation $\lambda \vdash i$ with the meaning “ λ is a partition of i ”.

moments of the auxiliary umbra $\varkappa_\alpha \equiv \chi \cdot \alpha$ result to be the formal cumulants³ of α . The umbra \varkappa_α is the α -cumulant umbra, cf. [10]. A special cumulant umbra, employed very often in the rest of the paper, is $\varkappa_{\bar{u}} \equiv \chi \cdot \bar{u}$, with \bar{u} the boolean unity [8]. The boolean unity represents the sequence $\{i!\}$ and its g.f. is the ordinary formal power series

$$f(\bar{u}, z) = 1 + \sum_{i \geq 1} z^i = \frac{1}{1-z}. \tag{10}$$

From (9) and (10), we have $f(\chi \cdot \bar{u}, z) = 1 + \log(1-z)^{-1} = 1 + \sum_{i \geq 1} \frac{z^i}{i}$ so that $E[(\chi \cdot \bar{u})^i] = (i-1)!$ for all positive integers i .

If in $\gamma \cdot \alpha$ we replace the umbra γ with the Bell umbra⁴ β , the dot-product $\beta \cdot \alpha$ denotes the α -partition umbra. One of its main properties is

$$\beta \cdot (\alpha \dot{+} \gamma) \equiv \beta \cdot \alpha + \beta \cdot \gamma \tag{11}$$

where the symbol $\alpha \dot{+} \gamma$, denotes the disjoint sum of α and γ and umbrally represents the sequence $\{a_n + g_n\}$ with $\{a_n\}$ and $\{g_n\}$ moments of α and γ respectively. More in general, the symbol $\dot{+}_{k=1}^n \alpha_k$ denotes an auxiliary umbra whose i th moment is the summation of the i th moments of $\alpha_1, \dots, \alpha_n$ respectively, that is

$$\mathcal{S}_i = E\left[\left(\dot{+}_{k=1}^n \alpha_k\right)^i\right] = a_{i,1} + \dots + a_{i,n}, \tag{12}$$

with $a_{i,j} = E[\alpha_j^i]$ for all nonnegative integers i and $j = 1, 2, \dots, n$. Since $\mathcal{S}_i \simeq \alpha_1^i + \dots + \alpha_n^i$, the symbol \mathcal{S}_i is used in analogy with the i th power sum symmetric polynomial. A special auxiliary umbra employed in the following is $-1 \cdot \beta \cdot \alpha$ with g.f. $f(-1 \cdot \beta \cdot \alpha, z) = \exp[1 - f(\alpha, z)]$.

Nesting dot-products, one of which is the α -partition umbra, we get the composition umbra. Indeed the symbol $\gamma \cdot (\beta \cdot \alpha)$, where β is the Bell umbra, has g.f. which is the composition of $f(\alpha, z)$ and $f(\gamma, z)$, that is $f(\gamma \cdot (\beta \cdot \alpha), z) = f(\gamma, f(\alpha, z) - 1)$. Parenthesis can be avoided since $\gamma \cdot (\beta \cdot \alpha) \equiv (\gamma \cdot \beta) \cdot \alpha$. From (7) the moments are [10]

$$E[(\gamma \cdot \beta \cdot \alpha)^i] = \sum_{\lambda \vdash i} g_{l(\lambda)} d_\lambda a_\lambda, \tag{13}$$

with $\{g_i\}$ umbrally represented by γ . For $f(z) = f(\alpha, z)$, the coefficients of the compositional inverse of $f(z)$ are represented by the symbol $\alpha^{(-1)}$ such that $\alpha^{(-1)} \cdot \beta \cdot \alpha \equiv \alpha \cdot \beta \cdot \alpha^{(-1)} \equiv \chi$. The umbra $\alpha^{(-1)}$ is the compositional inverse of α .

Also polynomial sequences in the indeterminate x can be represented by suitable umbrae, if the complex field \mathbb{C} is replaced by $\mathbb{C}[x]$. The uncorrelation property (4) becomes $E[x^i \alpha^k \gamma^m \dots] = x^i E[\alpha^k] E[\gamma^m] \dots$ for any set of distinct umbrae in \mathcal{A} and for nonnegative integers i, k, m, \dots . In $\mathbb{C}[x][\mathcal{A}]$, an umbra is said to be a scalar umbra when its moments are elements of \mathbb{C} , while it is said to be a polynomial umbra if its moments are polynomials of $\mathbb{C}[x]$. A sequence of polynomials $q_0, q_1, \dots \in \mathbb{C}[x]$ is umbrally represented by a polynomial umbra if $q_0 = 1$ and q_i is of degree i for all nonnegative integers i . A special family of polynomials, including many classical polynomial sequences, is the Sheffer polynomial sequence. In [7], it has been proved that any Sheffer polynomial sequence is umbrally represented by the so-called Sheffer polynomial umbra $\alpha + x \cdot \beta \cdot \gamma^{(-1)}$, with γ an umbra possessing compositional inverse, that follows if $E[\gamma] \neq 0$. If α is replaced by the augmentation umbra ε then the Sheffer polynomial sequence reduces to its associated sequence $x \cdot \beta \cdot \gamma^{(-1)}$, satisfying the binomial property [22].

The polynomial ring $\mathbb{C}[x]$ can be further replaced by $\mathbb{C}[x_1, \dots, x_p]$ with x_1, \dots, x_p indeterminates [4]. As example, but also because employed in the following, let us recall that the complete homogeneous symmetric polynomials $\{h_i\}$ are umbrally represented by the umbral polynomial $x_1 \bar{u}_1 + \dots + x_p \bar{u}_p$, with $\bar{u}_1, \dots, \bar{u}_p$ uncorrelated umbrae similar to the boolean unity \bar{u} . Indeed we have

$$h_i(x_1, \dots, x_p) = \sum_{1 \leq j_1 \leq \dots \leq j_i \leq p} x_{j_1} \dots x_{j_i} = \frac{1}{i!} E[(x_1 \bar{u}_1 + \dots + x_p \bar{u}_p)^i]. \tag{14}$$

3. Symbolic representation of Wishart distributions

According to the moment method, the trace of $W(n)$ is represented by an umbral polynomial $\mu_1 + \dots + \mu_p \in \mathbb{C}[\mathcal{A}]$, where $\{\mu_1, \dots, \mu_p\}$ are umbral monomials having not necessarily disjoint supports and corresponding to the eigenvalues of $W(n)$. For the sake of brevity, we denote the summation $\mu_1 + \dots + \mu_p$ with $\text{Tr}[W(n)]$ as if it was an element of $\mathbb{C}[\mathcal{A}]$. This notation will be employed in the rest of the paper.

³ For all nonnegative integers i , cumulants $\mathfrak{C}_i(Y)$ of a r.v. Y have the following properties: (Homogeneity) $\mathfrak{C}_i(aY) = a^i \mathfrak{C}_i(Y)$ for $a \in \mathbb{C}$, (Semi-invariance) $\mathfrak{C}_1(Y + a) = a + \mathfrak{C}_1(Y)$, $\mathfrak{C}_i(Y + a) = \mathfrak{C}_i(Y)$ for $i \geq 2$, (Additivity) $\mathfrak{C}_i(Y_1 + Y_2) = \mathfrak{C}_i(Y_1) + \mathfrak{C}_i(Y_2)$, if Y_1 and Y_2 are independent r.v.'s. Formal cumulants are the coefficients of $\log[f(z)]$ with $f(z)$ a formal power series. They share the same properties of cumulants of a r.v. When numerical values are considered, we need to check that the so-called "problem of cumulants" admits solution, cf. [4] and references therein.

⁴ The Bell umbra represents the sequence of Bell numbers. Its g.f. is $f(\beta, z) = \exp[e^z - 1]$.

Theorem 1. If $\text{diag}(\theta_1, \dots, \theta_p) = \Lambda$ is the eigenvalue matrix of the covariance $\Sigma = Q\Lambda Q^T$, with Q the eigenvector matrix, $\{\bar{u}_1, \dots, \bar{u}_p\}$ uncorrelated umbrae similar to the boolean unity umbra \bar{u} and $\{\nu_1, \dots, \nu_p\}$ umbral monomials such that

$$\nu_k = \chi \cdot b_{kk} \cdot \beta \quad \text{for } k = 1, 2, \dots, p, \tag{15}$$

with $b_{11} + \dots + b_{pp} = \text{Tr}(Q^T \Omega Q)$ and Ω the non-centrality matrix $\Omega = \Sigma^{-1}M$, then

$$\text{Tr}[W(n)] \equiv \sum_{k=1}^p [-1 \cdot \beta \cdot (\nu_k \theta_k \bar{u}_k) + n \cdot (\theta_k \bar{u}_k)] \equiv -1 \cdot \beta \cdot \left(\dot{+}_{k=1}^p \nu_k \theta_k \bar{u}_k \right) + n \cdot \left(+_{k=1}^p \theta_k \bar{u}_k \right), \tag{16}$$

where the symbol $+_{k=1}^p \theta_k \bar{u}_k$ denotes $\theta_1 \bar{u}_1 + \dots + \theta_p \bar{u}_p$.

Proof. The moment generating function (m.g.f.) of $W(n)$ is [20]

$$\begin{aligned} f(\text{Tr}[W(n)], z) &= E \{ \exp(\text{Tr}[W(n)]z) \} = \frac{\exp \{ -\text{Tr}[(I_p - z \Sigma)^{-1} M z] \}}{[\det(I_p - z \Sigma)]^n} \\ &= \frac{\exp \{ -\text{Tr}[(I_p - z \Sigma)^{-1} \Sigma \Omega z] \}}{[\det(I_p - z \Sigma)]^n}. \end{aligned} \tag{17}$$

If $\theta_1, \dots, \theta_p$ are the eigenvalues of Σ , then

$$\det(I - z \Sigma)^{-1} = \prod_{i=1}^p (1 - z \theta_i)^{-1} \quad \text{and} \quad [\det(I - z \Sigma)]^{-n} = f[n \cdot (\theta_1 \bar{u}_1 + \dots + \theta_p \bar{u}_p), z]$$

from (10). From the first equivalence in (8)

$$n \cdot \left(+_{k=1}^p \theta_k \bar{u}_k \right) \equiv n \cdot (\theta_1 \bar{u}_1) + \dots + n \cdot (\theta_p \bar{u}_p)$$

as given in (16). Moreover, in (17) if $\Sigma = Q\Lambda Q^T$, then $\text{Tr}[(I_p - z \Sigma)^{-1} \Sigma \Omega z] = \text{Tr}[Q(I - z \Lambda)^{-1} Q^T Q \Lambda Q^T \Omega z] = \text{Tr}[(I - z \Lambda)^{-1} \Lambda B z]$, with $B = Q^T \Omega Q$. Therefore we have

$$\text{Tr}[(I - z \Lambda)^{-1} \Lambda B z] = \sum_{k=1}^p \frac{b_{kk} \theta_k z}{(1 - z \theta_k)}. \tag{18}$$

The i th coefficient of the formal power series in (18) is $i! (b_{11} \theta_1^i + \dots + b_{pp} \theta_p^i)$. Since $i! \theta_k^i = E[(\theta_k \bar{u}_k)^i]$ and $b_{kk} = E[(\chi \cdot b_{kk} \cdot \beta)^i]$, for all nonnegative integers i , then

$$i! \mathcal{I}_i = i! (b_{11} \theta_1^i + \dots + b_{pp} \theta_p^i) = E \left[\left(\dot{+}_{k=1}^p \nu_k \theta_k \bar{u}_k \right)^i \right], \tag{19}$$

with ν_k given in (15) and \mathcal{I}_i in (12). Note that $\mathcal{I}_i = \text{Tr}(B \Lambda^i) = \text{Tr}(\Omega \Sigma^i)$. Since [10]

$$-1 \cdot \beta \cdot \left(\dot{+}_{k=1}^p \nu_k \theta_k \bar{u}_k \right) \equiv -1 \cdot \beta \cdot (\nu_1 \theta_1 \bar{u}_1) + \dots + -1 \cdot \beta \cdot (\nu_p \theta_p \bar{u}_p),$$

both the contributions in (16) follow from (11). \square

In the following we always assume that the hypothesis of Theorem 1 holds. Moments of $\text{Tr}[W(n)]$ can be computed by binomial expansion:

$$\begin{aligned} E(\text{Tr}[W(n)]^i) &= \sum_{j=0}^i \binom{i}{j} E \left[\left\{ -1 \cdot \beta \cdot \left(\dot{+}_{k=1}^p \nu_k \theta_k \bar{u}_k \right) \right\}^j \right] E \left[\left\{ n \cdot \left(+_{k=1}^p \theta_k \bar{u}_k \right) \right\}^{i-j} \right] \\ &= i! \sum_{j=0}^i \left(\sum_{\lambda \vdash j} \frac{(-1)^{l(\lambda)}}{r_1! r_2! \dots} \prod_{i=1}^{l(\lambda)} \text{Tr}(\Omega \Sigma^i)^{r_i} \right) \left(\sum_{\lambda' \vdash i-j} \frac{n^{l(\lambda')}}{r_1! r_2! \dots} \prod_{i=1}^{l(\lambda')} \text{Tr}(\Sigma^i)^{r_i} \right). \end{aligned} \tag{20}$$

In order to take advantage of equivalence (16), we denote with A the matrix of umbral monomials such that

$$\text{Tr}(A) = -1 \cdot \beta \cdot \left(\dot{+}_{k=1}^p \nu_k \theta_k \bar{u}_k \right). \tag{21}$$

Remark 1 (Central Wishart Distribution). When $M = 0$ in (1), then $W(n) = \widehat{W}(n) = X_{(1)}^T X_{(1)} + \dots + X_{(n)}^T X_{(n)}$ with $\{X_{(1)}^T X_{(1)}, \dots, X_{(n)}^T X_{(n)}\}$ i.i.d. random matrices of order p . We write $\widehat{W}(1)$ to denote one of the random matrices $X_{(k)}^T X_{(k)}$. From Theorem 1, if $b_{kk} = 0$ for $k = 1, 2, \dots, p$ then $\nu_1 \equiv \dots \equiv \nu_p \equiv \varepsilon$, and

$$\text{Tr}[\widehat{W}(n)] \equiv n \cdot \left(+_{k=1}^p \theta_k \bar{u}_k \right) \equiv n \cdot (\theta_1 \bar{u}_1) + \dots + n \cdot (\theta_p \bar{u}_p). \tag{22}$$

Then equivalence (16) separates the contribution of the non-centrality matrix Ω included in A from the central Wishart distribution $W(n)$. This remark was already done in [1] without giving an explicit expression to the convolution. This because the symbol in (16) does not have a probabilistic counterpart. Indeed if it is true that $n \cdot (+_{k=1}^p \theta_k \bar{u}_k)$ corresponds to a central Wishart distribution, $\text{Tr}(A)$ in (21) is just a formal compound Poisson r.v. [9], since its parameter -1 is negative.

Proposition 2 states the connection between central Wishart distributions (22) and partition umbrae.

Proposition 2. $\text{Tr}[\widehat{W}(n)] \equiv n \cdot \beta \cdot (+_{k=1}^p \chi_{\theta_k \bar{u}_k})$.

Proof. Since $\beta \cdot \chi \equiv u$ [4], then $n \cdot (+_{k=1}^p \theta_k \bar{u}_k) \equiv n \cdot \beta \cdot \chi \cdot (+_{k=1}^p \theta_k \bar{u}_k)$. The result follows from Theorem 1 by recalling the additivity property of cumulants

$$\chi \cdot (+_{k=1}^p \theta_k \bar{u}_k) \equiv \dot{+}_{i=1}^p \chi \cdot (\theta_k \bar{u}_k) \equiv \dot{+}_{i=1}^p \chi_{\theta_k \bar{u}_k}. \quad \square$$

Theorem 3 (Sheffer Polynomial Sequence⁵). The sequence $\{E[\text{Tr}(W(n))^k]\}$ is of Sheffer polynomial type with x replaced by the nonnegative integer n .

Proof. The umbra $\alpha + x \cdot \beta \cdot \gamma^{(-1)}$ is a Sheffer umbra [7]. The result follows from Theorem 1 and Proposition 2, choosing as umbra α the umbra $-1 \cdot \beta \cdot (+_{k=1}^p \nu_k \theta_k \bar{u}_k)$ and as umbra $\gamma^{(-1)}$ the umbra $\dot{+}_{k=1}^p \chi_{\theta_k \bar{u}_k}$, given in Proposition 2. Observe that this umbra has compositional inverse since $E[\dot{+}_{k=1}^p \chi_{\theta_k \bar{u}_k}] = \text{Tr}(\Sigma) \neq 0$. \square

If $b_{ii} = 0$ for $i = 1, 2, \dots, p$, then $-1 \cdot \beta \cdot (+_{k=1}^p \nu_k \theta_k \bar{u}_k) \equiv \varepsilon$ and $\{E[\text{Tr}(W(n))^k]\}$ is a binomial sequence: $E(\text{Tr}[\widehat{W}(n+m)^i]) = \sum_{k=0}^i \binom{i}{k} E(\text{Tr}[\widehat{W}(n)]^k) E(\text{Tr}[\widehat{W}(m)]^{i-k})$.

Corollary 4 (Binomial Polynomial Sequence). The sequence $\{E(\text{Tr}[\widehat{W}(n)]^k)\}$ is the binomial sequence associated to $\{E(\text{Tr}[W(n)]^k)\}$.

Since the m.g.f. of $\text{Tr}[\widehat{W}(n)]$ is convergent in a suitable neighborhood of 0, the distribution of $\text{Tr}[\widehat{W}(n)]$ is univocally determined by its moments. Thanks to the latter equivalence in (8), with the umbra α replaced by $+_{k=1}^p \theta_k \bar{u}_k$, a further consequence of Corollary 4 is the following result.

Corollary 5. $\text{Tr}[\widehat{W}(n+m)] \stackrel{d}{=} \text{Tr}[\widehat{W}(n)] + \text{Tr}[\widehat{W}(m)]$.

Proposition 6 gives a Sheffer identity (cf. [22]) for $\text{Tr}[W(n)]$.

Proposition 6. $\text{Tr}[W(n+m, \Sigma, M)] \stackrel{d}{=} \text{Tr}[W(n, \Sigma, M)] + \text{Tr}[\widehat{W}(m)]$.

Proof. The result follows from equivalence (22) by observing that

$$\begin{aligned} \text{Tr}[W(n+m)] &\equiv -1 \cdot \beta \cdot (+_{k=1}^p \nu_k \theta_k \bar{u}_k) + (n+m) \cdot (+_{k=1}^p \theta_k \bar{u}_k) \\ &\equiv [-1 \cdot \beta \cdot (+_{k=1}^p \nu_k \theta_k \bar{u}_k) + n \cdot (+_{k=1}^p \theta_k \bar{u}_k)] + m \cdot (+_{k=1}^p \theta_k \bar{u}_k). \quad \square \end{aligned}$$

Proposition 7. If $M = M_1 + M_2$ with $M_1 = \sum_{i=1}^n \mathbf{m}_i^T \mathbf{m}_i$ and $M_2 = \sum_{i=n+1}^{n+m} \mathbf{m}_i^T \mathbf{m}_i$, then $\text{Tr}[W(n+m, \Sigma, M)] \stackrel{d}{=} \text{Tr}[W(n, \Sigma, M_1)] + \text{Tr}[W(m, \Sigma, M_2)]$.

Proof. Assume $\text{Tr}(M_1) = b_{11} + \dots + b_{pp}$ and $\text{Tr}(M_2) = \tilde{b}_{11} + \dots + \tilde{b}_{pp}$ with $\text{Tr}(M_1) + \text{Tr}(M_2) = \text{Tr}(M)$. The result follows from Theorem 1 by observing that the umbral monomials ν_k in (15) can be splitted in the summation of the umbral monomials $\tilde{\nu}_k \equiv (\chi \cdot b_{kk} \cdot \beta)$ and $\bar{\nu}_k \equiv (\chi \cdot \tilde{b}_{kk} \cdot \beta)$, since $\chi \cdot (b_{kk} + \tilde{b}_{kk}) \cdot \beta \equiv (\chi \cdot b_{kk} + \chi \cdot \tilde{b}_{kk}) \cdot \beta \equiv \chi \cdot b_{kk} \cdot \beta + \chi \cdot \tilde{b}_{kk} \cdot \beta$. This last equivalence is a consequence of (8) which holds also when n and m are replaced by some umbrae δ and γ , cf. [10]. \square

One feature of the umbral calculus is the chance to replace the nonnegative integer n in the dot-product $n \cdot \alpha$ with an umbra. Then the umbral counterpart of a randomized non-central Wishart distribution $W(N) = W_p(N, \Sigma, M)$ can be recovered from Theorem 1, with

$$W(N) = \sum_{i=1}^N (X_{(i)} - \mathbf{m})^T (X_{(i)} - \mathbf{m}), \tag{23}$$

\mathbf{m} a vector of complex numbers and $M = \mathbf{m}^T \mathbf{m}$.

⁵ In terms of g.f.'s, Sheffer polynomial sequences $\{s_i(x)\}$ are such that $\sum_{i \geq 1} s_i(x) \frac{z^i}{i!} = g(z) \exp\{xf(z)\}$ with $g(z)$ and $f(z)$ formal power series such that $g(0) = 1$ and $f(0) = 0$. For further references on Sheffer polynomial sequences see [22].

Proposition 8. If N is a r.v. with moment sequence umbrally represented by an umbra α , then $\{E(\text{Tr}[W(N)]^k)\}$ is umbrally represented by

$$\text{Tr}[W(\alpha)] \equiv \alpha \cdot \beta \cdot \left[-1 \cdot \beta \cdot \left(\dot{+}_{k=1}^p \nu_k \theta_k \bar{u}_k \right) + \left(\dot{+}_{i=1}^p \theta_i \bar{u}_i \right) \right].$$

Proof. The m.g.f. of a random sum $S_N = X_1 + X_2 + \dots + X_N$ is the composition of the m.g.f. $g(z)$ corresponding to the index N and of the m.g.f. $h(z)$ corresponding to $\{X_i\}$. The umbral counterpart of S_N is the composition umbra $\alpha \cdot \beta \cdot \gamma$, where α represents the sequence of moments of N and γ represents the sequence of moments of $\{X_i\}$. From (23) $\text{Tr}[W(N)] = \sum_{i=1}^N \text{Tr}[(X_{(i)} - \mathbf{m})^T (X_{(i)} - \mathbf{m})]$, so that the sequence $\{E(\text{Tr}[W(N)]^k)\}$ is represented by the composition umbra $\alpha \cdot \beta \cdot \gamma$, with α the umbral counterpart of N and γ the umbral counterpart of $\text{Tr}[(X_{(i)} - \mathbf{m})^T (X_{(i)} - \mathbf{m})]$. The result follows from (16) and Remark 1 since $\gamma \equiv \dot{+}_{i=1}^p \theta_i \bar{u}_i$. \square

3.1. Moments and cumulants of Wishart distributions

3.1.1. Central distributions

Moments of central Wishart distributions can be expressed by means of complete homogeneous symmetric polynomials $\{h_i\}$ given in (14). Indeed, from (14) and equivalence (22) we have $E(\text{Tr}[\widehat{W}(1)]^i) = i! h_i(\theta_1, \dots, \theta_p)$ and the generalization to $\text{Tr}[\widehat{W}(n)]$ follows by using the dot-product with n . From (7) we have

$$E(\text{Tr}[\widehat{W}(n)]^i) = \sum_{\lambda \vdash i} (n)_{l(\lambda)} \tilde{d}_\lambda h_\lambda(\theta_1, \dots, \theta_p) \quad \text{with } h_\lambda = h_1^{r_1} h_2^{r_2} \dots \quad \text{and } \tilde{d}_\lambda = \frac{i!}{r_1! r_2! \dots}. \quad (24)$$

Remark 2 (Cyclic Polynomials). Thanks to (24), moments of $\text{Tr}[\widehat{W}(n)]$ can be expressed in terms of cyclic polynomials

$$\mathcal{C}_i(x_1, \dots, x_i) = \sum_{\lambda \vdash i} c_\lambda x_1^{r_1} \dots x_i^{r_i} \quad \text{with } c_\lambda = \frac{i!}{1^{r_1} r_1! 2^{r_2} r_2! \dots}. \quad (25)$$

A well-known relation between cyclic polynomials and complete homogeneous polynomials is $\mathcal{C}_i(s_1, \dots, s_i) = i! h_i(x_1, \dots, x_p)$, with $\{s_j\}$ power sum symmetric polynomials in the indeterminates x_1, \dots, x_p , cf. [18]. By replacing the indeterminates $\{x_k\}$ with the eigenvalues of Σ , we have $\mathcal{C}_i[\text{Tr}(\Sigma), \dots, \text{Tr}(\Sigma^i)] = i! h_i(\theta_1, \dots, \theta_p) = E(\text{Tr}[\widehat{W}(1)]^i)$ and from (24)

$$E(\text{Tr}[\widehat{W}(n)]^i) = \sum_{\lambda \vdash i} (n)_{l(\lambda)} d_\lambda \mathcal{C}_\lambda[\text{Tr}(\Sigma), \dots, \text{Tr}(\Sigma^i)] \quad \text{with } \mathcal{C}_\lambda = \mathcal{C}_1^{r_1} \mathcal{C}_2^{r_2} \dots. \quad (26)$$

Proposition 9. If $\{\text{Tr}(\Sigma^i)\}$ is umbrally represented by σ , then $\text{Tr}[\widehat{W}(n)] \equiv n \cdot \beta \cdot (\sigma \varkappa_{\bar{u}})$.

Proof. From (13), with γ replaced by the unity umbra u , we have $E[(\beta \cdot (\alpha \varkappa_{\bar{u}}))^i] = \sum_{\lambda \vdash i} d_\lambda a_\lambda u_\lambda$ with $u_\lambda = (2-1)!^{r_2} (3-1)!^{r_3} \dots$. By observing that $c_\lambda = d_\lambda u_\lambda$ in (25), we have

$$E[(\beta \cdot (\alpha \varkappa_{\bar{u}}))^i] = \mathcal{C}_i(a_1, \dots, a_i). \quad (27)$$

Then $\text{Tr}[\widehat{W}(1)] \equiv \beta \cdot (\sigma \varkappa_{\bar{u}})$ and the result follows since $f(\text{Tr}[\widehat{W}(n)], z) = f(\text{Tr}[\widehat{W}(1)], z)^n$. \square

As a consequence, the symbolic representation of a randomized central Wishart distribution is $\alpha \cdot \beta \cdot (\sigma \varkappa_{\bar{u}})$ with moments given in (13).

Remark 3. A different expression for moments of central Wishart distributions is given in [28] where vectors and matrix exponential Bell polynomials are employed and introduced as derivatives of a function of a vector or matrix function of a vector or matrix. On the other hand there is a well-known relation between cyclic polynomials and complete Bell exponential polynomials [18], that is $\mathcal{C}_i(a_1, a_2, \dots, a_i) = Y_i[a_1, a_2 1!, a_3 2!, \dots, a_i (i-1)!]$. Due to (27), the partition umbra is the bridge between the two families of polynomials, since $E[(\beta \cdot \alpha)^i] = Y_i(a_1, \dots, a_i)$, cf. [9].

Proposition 9 is in agreement with Eq. (2.4) of [12] where permutations⁶ are employed in place of integer partitions. Indeed indexing Eq. (26) in permutations,

$$E[(n \cdot \beta \cdot (\sigma \varkappa_{\bar{u}}))^i] = \sum_{\sigma \in \mathfrak{S}_i} n^{l(\sigma)} \prod_{c \in C(\sigma)} \mathcal{C}_{l(c)}(s_1, \dots, s_{l(c)})$$

and Eq. (2.4) of [12] follows for $h_1 = \dots = h_k = I_p$ since $\mathcal{C}_i(s_1, \dots, s_i) = \sum_{\sigma \in \mathfrak{S}_i} \prod_{c \in C(\sigma)} s_{l(c)}$.

⁶ A permutation $\sigma \in \mathfrak{S}_i$ of $[i]$, with \mathfrak{S}_i the symmetric group, can be decomposed into disjoint cycles. The length of the cycle $c \in C(\sigma)$ is its cardinality, denoted by $l(c)$. The number of cycles is denoted by $|C(\sigma)|$. A permutation σ with r_1 1-cycles, r_2 2-cycles and so on is said to be of cycle class $\lambda = (1^{r_1}, 2^{r_2}, \dots) \vdash i$. The integer c_λ in (25) denotes the number of permutations $\sigma \in \mathfrak{S}_i$ of cycle class $\lambda = (1^{r_1}, 2^{r_2}, \dots) \vdash i$. When misunderstandings occur, we refer to the cycle class of a permutation σ as λ_σ . In particular we have $l(\lambda_\sigma) = |C(\sigma)|$.

As often happens with r.v.'s, the sequence of cumulants of central Wishart distributions is more manageable of its moments. In the following, $\text{Cum}_i(\text{Tr}[\widehat{W}(n)])$ denotes its i th cumulant for all positive integer i , and $\varkappa_{\text{Tr}[\widehat{W}(n)]}$ denotes the $\text{Tr}[\widehat{W}(n)]$ -cumulant umbra, that is the umbra whose moments are $\text{Cum}_i(\text{Tr}[\widehat{W}(n)])$.

Proposition 10. $\text{Cum}_i(\text{Tr}[\widehat{W}(n)]) = n(i - 1)!\text{Tr}(\Sigma^i)$.

Proof. From Proposition 9, $\text{Cum}_i(\text{Tr}[\widehat{W}(n)]) = E[(\chi \cdot n \cdot \beta \cdot (\sigma \varkappa_{\widehat{u}}))^i]$. The result follows from (13), with γ replaced by the umbra $\chi \cdot n$. Indeed $E[(\chi \cdot n)^{l(\lambda)}] \neq 0$ only with the partition λ of i having length 1. In this case, $E[(\chi \cdot n)^{l(\lambda)}] = n$ and only the i th moment $E[(\sigma \varkappa_{\widehat{u}})^i] = (i - 1)!\text{Tr}(\Sigma^i)$ gives contribution in (13). \square

Moments from cumulants can be recovered by using the partition umbra since if $\varkappa \equiv \chi \cdot \alpha$ then $\alpha \equiv \beta \cdot \varkappa$, cf. [4].

Corollary 11. $\text{Tr}[\widehat{W}(n)] \equiv \beta \cdot \varkappa_{\text{Tr}[\widehat{W}(n)]} \equiv n \cdot \beta \cdot \varkappa_{\text{Tr}[\widehat{W}(1)]}$.

3.2. Non-central distributions

From the additivity property of cumulants and Theorem 1, cumulants of non-central Wishart distributions are such that $\text{Cum}_i(\text{Tr}[W(n)]) = \text{Cum}_i(\text{Tr}[\widehat{W}(n)]) + \text{Cum}_i(\text{Tr}(A))$. This is the key to prove the following theorem.

Theorem 12. $\text{Cum}_i(\text{Tr}[W(n)]) = n(i - 1)!\text{Tr}(\Sigma^i) - i!\delta_i$, with $\{\delta_i\}$ given in (19).

Proof. From Theorem 1, cumulants of $\text{Tr}[W(n)]$ are umbrally represented by

$$\chi \cdot [-1 \cdot \beta \cdot (\nu_1 \theta_1 \bar{u}_1 + \dots + \nu_p \theta_p \bar{u}_p) + n \cdot (\theta_1 \bar{u}_1 + \dots + \theta_p \bar{u}_p)].$$

From the additivity property of cumulants this umbra is similar to

$$\chi \cdot [-1 \cdot \beta \cdot (\nu_1 \theta_1 \bar{u}_1 + \dots + \nu_p \theta_p \bar{u}_p)] + \chi \cdot n \cdot (\theta_1 \bar{u}_1 + \dots + \theta_p \bar{u}_p).$$

Since $\chi \cdot n \cdot (\theta_1 \bar{u}_1 + \dots + \theta_p \bar{u}_p) \equiv \chi \cdot n \cdot \beta \cdot (\sigma \varkappa_{\widehat{u}})$, its moments are given in Proposition 10. Moreover $E\{(\chi \cdot [-1 \cdot \beta \cdot (\nu_1 \theta_1 \bar{u}_1 + \dots + \nu_p \theta_p \bar{u}_p)])^i\} = \sum_{\lambda \vdash i} E[(-\chi)^{l(\lambda)}] d_\lambda \lambda! \delta_\lambda = -i! \delta_i$ with $\lambda! = \lambda_1! \lambda_2! \dots$ since $E[(-\chi)^{l(\lambda)}] \neq 0$ only for $l(\lambda) = 1$. \square

Corollary 13. $\text{Cum}_i(\text{Tr}[W(n)]) = (i - 1)! \sum_{j=1}^p [n - i b_{jj}] \theta_j^i$.

Example 1. Set $n = 3$. If

$$\Sigma = \begin{pmatrix} 0.025 & 0.0075 & 0.00175 \\ 0.0075 & 0.0070 & 0.00135 \\ 0.00175 & 0.00135 & 0.00043 \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} 0.0001 & 0.0210 & 0.3000 \\ 0.0400 & 0.0005 & 0.0200 \\ 0.0010 & 0.0100 & 0.0004 \end{pmatrix}$$

then $\Omega = \begin{pmatrix} -1.8721 & 0.1670 & 17.3410 \\ 14.7071 & -11.3087 & -5.8077 \\ -36.2288 & 58.0801 & -51.4103 \end{pmatrix}$ and $B = \begin{pmatrix} -58.8960 & 59.9910 & 19.4203 \\ -19.2629 & -7.2024 & -5.9850 \\ -10.0248 & -2.1620 & 1.5074 \end{pmatrix}$ and from Corollary 13 we have $\text{Cum}_1(\text{Tr}[W(3)]) = 0.0963$, $\text{Cum}_2(\text{Tr}[W(3)]) = 0.000327$, $\text{Cum}_3(\text{Tr}[W(3)]) = -0.00003089$ and so on.

Complete Bell exponential polynomials give moments in terms of cumulants [4]: if $\{c_i\}$ is the sequence of formal cumulants of $\{a_n\}$, then $a_n = Y_n(c_1, \dots, c_n)$. The connection with cyclic polynomials is a consequence of Remark 3.

Corollary 14. If $\tilde{c}_i = n(i - 1)!\text{Tr}(\Sigma^i)$ and $\bar{c}_i = -i! \delta_i$, then

$$E(\text{Tr}[W(n)]^i) = Y_i(\tilde{c}_1 + \bar{c}_1, \dots, \tilde{c}_i + \bar{c}_i) = \mathcal{C}_i \left(\tilde{c}_1 + \bar{c}_1, \dots, \frac{\tilde{c}_i + \bar{c}_i}{(i - 1)!} \right).$$

A completely different expression of $E(\text{Tr}[W(n)]^i)$ is given in [15], where weighted generating functions of special graphs are employed.

In [6], a different family of cumulants has been introduced, representing $\{\text{Cum}_i(\text{Tr}[W(n)])\}$ normalized to the dimension p of $W(n)$. Indeed the definition given in [6] leads to the introduction of umbral polynomials $\{c_{1,\mu}, \dots, c_{p,\mu}\}$ with $\mu = (\mu_1, \dots, \mu_p)$ and $\text{Tr}[W(n)] \equiv \mu_1 + \dots + \mu_p$, such that

$$\text{Tr}[W(n)] \equiv p \cdot \beta \cdot \text{Tr}[\mathcal{C}(n)] \tag{28}$$

where $\text{Tr}[\mathcal{C}(n)] = c_{1,\mu} + \dots + c_{p,\mu}$. Note that the order p plays a fundamental role in (28), differently from Theorem 12 where instead the integer n gives the main contribution. The following theorem highlights the connection between the two families of cumulants.

Theorem 15. $\varkappa_{\text{Tr}[W(n)]} \equiv \chi \cdot p \cdot \beta \cdot \text{Tr}[\mathcal{C}(n)]$ and $\text{Tr}[\mathcal{C}(n)] \equiv \chi \cdot \frac{1}{p} \cdot \beta \cdot \varkappa_{\text{Tr}[W(n)]}$.

Proof. The first equivalence follows from (28) by taking the dot-product with χ of both sides. The latter follows from the first by taking the dot-product of both sides with $\chi \cdot \frac{1}{p}$. \square

4. Joint moments and joint cumulants of Wishart distributions

In order to deal with joint moments (2), we need to recall the multivariate version of the symbolic method. As before, we just introduce the tools necessary to work with the object of this paper. The reader interested in a detailed exposition of the topic is referred to [5].

4.1. Multivariate symbolic method

Let $\{v_1, \dots, v_m\}$ be a set of umbral monomials with supports not necessarily disjoint. A sequence $\{g_i\}_{i \in \mathbb{N}_0^m} \in \mathbb{C}$, with $g_i = g_{i_1, i_2, \dots, i_m}$ and $g_0 = 1$, is represented by the m -tuple $\mathbf{v} = (v_1, \dots, v_m)$ if

$$E[\mathbf{v}^{\mathbf{i}}] = g_i, \tag{29}$$

for each multi-index $\mathbf{i} \in \mathbb{N}_0^m$. The elements $\{g_i\}_{i \in \mathbb{N}_0^m}$ in (29) are called *multivariate moments* of \mathbf{v} . Paralleling (5), the g.f. of \mathbf{v} is the exponential formal power series

$$e^{v_1 z_1 + \dots + v_m z_m} = u + \sum_{k \geq 1} \sum_{|\mathbf{i}|=k} \mathbf{v}^{\mathbf{i}} \frac{\mathbf{z}^{\mathbf{i}}}{\mathbf{i}!} \in \mathbb{C}[\mathcal{A}][[z_1, \dots, z_m]] \tag{30}$$

with $\mathbf{z} = (z_1, \dots, z_m)$, $|\mathbf{i}| = i_1 + \dots + i_m$ and $\mathbf{i}! = i_1! \dots i_m!$. If the sequence $\{g_i\}$ is umbrally represented by \mathbf{v} and has (exponential) g.f.

$$f(\mathbf{z}) = 1 + \sum_{k \geq 1} \sum_{|\mathbf{i}|=k} g_i \frac{\mathbf{z}^{\mathbf{i}}}{\mathbf{i}!} \tag{31}$$

then from (30) $E[e^{v_1 z_1 + \dots + v_m z_m}] = f(\mathbf{z})$. Taking into account (29), the g.f. in (31) is denoted by $f(\mathbf{v}, \mathbf{z})$. Two umbral vectors \mathbf{v}_1 and \mathbf{v}_2 are said to be *similar*, in symbols $\mathbf{v}_1 \equiv \mathbf{v}_2$, iff $f(\mathbf{v}_1, \mathbf{z}) = f(\mathbf{v}_2, \mathbf{z})$, that is $E[\mathbf{v}_1^{\mathbf{i}}] = E[\mathbf{v}_2^{\mathbf{i}}]$ for all $\mathbf{i} \in \mathbb{N}_0^m$. They are said to be *uncorrelated* iff $E[\mathbf{v}_1^{\mathbf{i}} \mathbf{v}_2^{\mathbf{j}}] = E[\mathbf{v}_1^{\mathbf{i}}] E[\mathbf{v}_2^{\mathbf{j}}]$ for all $\mathbf{i}, \mathbf{j} \in \mathbb{N}_0^m$.

An analogous of (7) holds for the multivariate case, provided to replace integer partitions with multi-index partitions⁷ [5]. The dot-product $n \cdot \mathbf{v}$ of a nonnegative integer n and a m -tuple \mathbf{v} denotes the summation $\mathbf{v}' + \mathbf{v}'' + \dots + \mathbf{v}'''$ with $\{\mathbf{v}', \mathbf{v}'', \dots, \mathbf{v}'''\}$ a set of n uncorrelated and similar m -tuples. For $\mathbf{i} \in \mathbb{N}_0^m$, its multivariate moments are

$$E[(n \cdot \mathbf{v})^{\mathbf{i}}] = \sum_{\lambda \vdash \mathbf{i}} d_\lambda (n)_{l(\lambda)} g_\lambda, \quad \text{with } d_\lambda = \frac{\mathbf{i}!}{m(\lambda)! \lambda!}, \tag{32}$$

the sum runs over all partitions $\lambda = (\lambda_1^{r_1}, \lambda_2^{r_2}, \dots)$ of the multi-index \mathbf{i} , $g_\lambda = g_{\lambda_1^{r_1} \lambda_2^{r_2} \dots}$ and $g_{\lambda_i} = E[\mathbf{v}^{\lambda_i}]$. If we replace the integer n in (32) with the dot-product $\alpha \cdot \beta$, we get the auxiliary umbra $\alpha \cdot \beta \cdot \mathbf{v}$ whose moments are

$$E[(\alpha \cdot \beta \cdot \mathbf{v})^{\mathbf{i}}] = \sum_{\lambda \vdash \mathbf{i}} d_\lambda a_{l(\lambda)} g_\lambda, \tag{33}$$

with $\{a_i\}$ in $a_{l(\lambda)}$ umbrally represented by α . In particular the g.f. of the auxiliary umbra $\alpha \cdot \beta \cdot \mathbf{v}$ turns to be $f(\alpha \cdot \beta \cdot \mathbf{v}, \mathbf{z}) = f[\alpha, f(\mathbf{v}, \mathbf{z}) - 1]$, that is the composition of the univariate g.f. $f(\alpha, z)$ and the multivariate g.f. $f(\mathbf{v}, \mathbf{z})$. More details on the composition of multivariate formal power series are given in [5]. For what we need in the following, we just recall that $f(-1 \cdot \beta \cdot \mathbf{v}, \mathbf{z}) = \exp\{-[f(\mathbf{v}, \mathbf{z}) - 1]\}$ and $f(n \cdot \beta \cdot \mathbf{v}, \mathbf{z}) = \exp\{n[f(\mathbf{v}, \mathbf{z}) - 1]\}$. If in (33) the umbra α is replaced by the umbra $\chi \cdot \chi$, then the \mathbf{v} -cumulant umbra $\chi \cdot \chi \cdot \beta \cdot \mathbf{v} \equiv \chi \cdot \mathbf{v}$ is recovered, whose moments are the multivariate cumulants of the m -tuple \mathbf{v} .

4.2. Joint moments and necklaces

The aim of this paragraph is to find a symbolic representation of joint moment (2) as the \mathbf{i} -th coefficient of the g.f. $E(\exp\{\text{Tr}[W(n)(H_1 z_1 + \dots + H_m z_m)]\})$, with $\mathbf{i} = (i_1, \dots, i_m)$. It is well-known that if Z is a $p \times p$ Hermitian parameter matrix, then

$$E(\exp\{\text{Tr}[W(n) Z]\}) = \frac{\exp\{-\text{Tr}[(I_p - \Sigma Z)^{-1} \Omega \Sigma Z]\}}{[\det(I_p - \Sigma Z)]^n} \tag{34}$$

⁷ A partition of a multi-index \mathbf{i} , in symbols $\lambda \vdash \mathbf{i}$, is a matrix $\lambda = (\lambda_{ij})$ of nonnegative integers and with no zero columns in lexicographic order such that $\lambda_{r1} + \lambda_{r2} + \dots + \lambda_{rk} = i_r$ for $r = 1, 2, \dots, n$. The number of columns of λ is called the length of λ and denoted by $l(\lambda)$. As for integer partitions, the notation $\lambda = (\lambda_1^{r_1}, \lambda_2^{r_2}, \dots)$ means that in the matrix λ there are r_1 columns equal to λ_1 , r_2 columns equal to λ_2 and so on, with $\lambda_1 < \lambda_2 < \dots$. We call r_i multiplicity of λ_i and set $m(\lambda) = (r_1, r_2, \dots)$.

with Ω the non-centrality matrix [15]. Within formal power series, $f(W(n)H, Z) = f(W(n), HZ)$ with $H \in \mathbb{C}^{p \times p}$. Then choose H and Z such that $\text{Tr}[W(n)HZ] = \text{Tr}[W(n)H_1]z_1 + \dots + \text{Tr}[W(n)H_m]z_m$. From (34)

$$E[\exp(\text{Tr}[W(n)HZ])] = \frac{\exp\{-\text{Tr}[(I_p - \Sigma HZ)^{-1}\Omega\Sigma HZ]\}}{[\det(I_p - \Sigma HZ)]^n}. \tag{35}$$

Eq. (35) is the starting point to prove Theorem 16, where the symbolic representation of a non-central Wishart distribution is generalized to the multivariate case. The form of this symbolic representation is very similar to the univariate case, provided scalar umbrae are replaced by suitable m -tuples of umbral monomials.

In order to take advantage of the cyclic property of traces, we resort the notion of necklace. A m -ary necklace is an equivalence class of m -ary strings equivalent under rotation (the cyclic group) [2]. Let the m -ary strings be elements of the set $\{1, \dots, m\}^{|\mathbf{i}|}$, then having length $|\mathbf{i}|$, with $\mathbf{i} = (i_1, \dots, i_m)$. Denote the set of all necklaces of length $|\mathbf{i}|$ over the m -ary alphabet of indexes $\{1, \dots, m\}$ with $\mathbf{N}_m(|\mathbf{i}|)$. It is natural to choose as representative α of a necklace its lexicographically smallest string. Then we denote the corresponding necklace with $[\alpha]$.

Definition 1. In the set $\{1, \dots, m\}^{|\mathbf{i}|}$, a necklace is said to be of kind \mathbf{i} , if the elements in its strings are chosen in the multiset⁸ $\{1^{i_1}, 2^{i_2}, \dots, m^{i_m}\}$.

To the best of our knowledge, this definition is given here for the first time. We denote the set of representatives of necklaces of kind \mathbf{i} with $\mathbf{N}_m[\mathbf{i}]$.

Example 2. If $m = 3$ and $|\mathbf{i}| = 3$, then $\mathbf{N}_3(3) = \{111, 222, 333, 112, 122, 113, 133, 223, 233, 123, 132\}$. The string 123 is the representative of the necklace $[123] = \{123, 312, 231\}$ and the string 112 is the representative of the necklace $[112] = \{112, 121, 211\}$. Moreover $\mathbf{N}_3[(3, 0, 0)] = \{111\}$ and $\mathbf{N}_3[(1, 2, 0)] = \{122\}$. We have $|\mathbf{N}_3[\mathbf{i}]| = 1$, for all \mathbf{i} with $|\mathbf{i}| = 3$, except for $\mathbf{i} = (1, 1, 1)$ as $\mathbf{N}_3[(1, 1, 1)] = \{123, 132\}$.

In most cases, a necklace has full size, which is equal to the length of a string. But, there are periodic necklaces of period $d < m$, a divisor of m . In the example, 111 is a necklace of period 1. The aperiodic necklaces are also known as *Lyndon words*. Denote the set of Lyndon words of length $|\mathbf{i}|$ over the m -ary alphabet of indexes $\{1, \dots, m\}$ with $\mathbf{M}_m(|\mathbf{i}|)$. As before, denote the set of aperiodic necklaces of kind \mathbf{i} with $\mathbf{M}_m[\mathbf{i}]$. The closed form formula for joint moments (2) is given in the following theorem.

Theorem 16. For $\mathbf{i} = (i_1, \dots, i_m) \in \mathbb{N}_0^m$

$$E\{\text{Tr}[W(n)H_1]^{i_1} \dots \text{Tr}[W(n)H_m]^{i_m}\} = E[(-1 \cdot \beta \cdot \tilde{\eta} + n \cdot \beta \cdot \tilde{\rho})^{\mathbf{i}}], \tag{36}$$

with

(i) $\tilde{\rho} = (\bar{u}_1 \rho_1, \dots, \bar{u}_m \rho_m)$ and $\rho = (\rho_1, \dots, \rho_m)$ such that

$$E[\rho^{\mathbf{i}}] = \sum_{\alpha \in \mathbf{M}_m[\mathbf{i}]} \text{Tr} \left[\prod_{k \in \alpha} (\Sigma H_k) \right] + \sum_{\alpha \in \mathbf{N}_m[\mathbf{i}] - \mathbf{M}_m[\mathbf{i}]} \frac{1}{d} \text{Tr} \left[\prod_{k \in \alpha} (\Sigma H_k) \right]; \tag{37}$$

(ii) $\tilde{\eta} = (\bar{u}_1 \eta_1, \dots, \bar{u}_m \eta_m)$ and $\eta = (\eta_1, \dots, \eta_m)$ such that

$$E[\eta^{\mathbf{i}}] = \sum_{\alpha \in \mathbf{N}_m[\mathbf{i}]} \left(\sum_{b \in [\alpha]} \text{Tr} \left[\Omega \prod_{k \in b} (\Sigma H_k) \right] \right). \tag{38}$$

If $|\mathbf{i}|$ is prime or $\mathbf{N}_m[\mathbf{i}] = \mathbf{M}_m[\mathbf{i}]$, then (37) reduces to the first summation.

Proof. Within formal power series, for a matrix A the following identity holds $[\det(I - A)]^{-1} = \exp(\sum_{j \geq 1} \text{Tr}(A^j)/j)$, cf. [24]. Then taking into account (31) and (35), with HZ replaced by $H_1z_1 + \dots + H_mz_m$, we have $(\det[I_p - \Sigma(H_1z_1 + \dots + H_mz_m)])^{-n} = \exp\{n[f(\tilde{\rho}, \mathbf{z}) - 1]\}$ with

$$f(\tilde{\rho}, \mathbf{z}) - 1 = \sum_{j \geq 1} \frac{1}{j} \sum_{|\mathbf{i}|=j} \left[\sum_{\alpha \in \{1, \dots, m\}^{|\mathbf{i}|}} \text{Tr} \left(\prod_{k \in \alpha} \Sigma H_k \right) \right] z_1^{i_1} \dots z_m^{i_m}. \tag{39}$$

Eq. (37) follows by indexing the inner summation in (39) with the necklaces in $\mathbf{N}_m[\mathbf{i}]$ for each \mathbf{i} and by multiplying the corresponding traces with the cardinality of necklaces, due to their cyclic property. In (35) since $(I - A)^{-1} = I + \sum_{j \geq 1} A^j$, then

⁸ Informally, a multiset is a generalization of the notion of set where elements are allowed to appear more than once. The notation $\{1^{i_1}, 2^{i_2}, \dots, m^{i_m}\}$ denotes that 1 appears i_1 times, 2 appears i_2 times and so on. For a formal definition, cf. [4].

$\text{Tr}[(I_p - \Sigma H Z)^{-1} \Omega \Sigma H Z] = \text{Tr}[\Sigma H Z (I_p - \Sigma H Z)^{-1} \Omega] = \sum_{j \geq 1} \text{Tr}[\Omega (\Sigma H Z)^j]$ and for HZ replaced by $H_1 z_1 + \dots + H_m z_m$ a characterization of $f(\tilde{\eta}, \mathbf{z})$ similar to (39) follows

$$f(\tilde{\eta}, \mathbf{z}) - 1 = \sum_{j \geq 1} \sum_{|\mathbf{i}|=j} \left[\sum_{\mathbf{a} \in \{1, \dots, m\}^{|\mathbf{i}|}} \text{Tr} \left(\Omega \prod_{k \in \mathbf{a}} \Sigma H_k \right) \right] z_1^{i_1} \dots z_m^{i_m}.$$

Differently from (39), we could not group the inner summation with respect to the elements of necklaces, since we could not use the cyclic property of the trace, due to the position of the non-centrality matrix Ω . \square

A multinomial expansion is applied to the right-hand-side of (36) in order to get

$$E \left\{ \prod_{j=1}^m \text{Tr} [W(n)H_j]^{i_j} \right\} = \sum_{\substack{\mathbf{t}_1, \mathbf{t}_2 \in \mathbb{N}_0^m \\ \mathbf{t}_1 + \mathbf{t}_2 = \mathbf{i}}} \frac{\mathbf{i}!}{\mathbf{t}_1! \mathbf{t}_2!} E[(-1 \cdot \beta \cdot \tilde{\eta})^{\mathbf{t}_1}] E[(n \cdot \beta \cdot \tilde{\rho})^{\mathbf{t}_2}] \tag{40}$$

$$= \mathbf{i}! \sum_{\substack{\mathbf{t}_1, \mathbf{t}_2 \in \mathbb{N}_0^m \\ \mathbf{t}_1 + \mathbf{t}_2 = \mathbf{i}}} \left(\sum_{\lambda \vdash \mathbf{t}_1} \frac{(-1)^{l(\lambda)}}{m(\lambda)} \prod_{\lambda_i} E[\eta^{\lambda_i}]^{r_i} \right) \left(\sum_{\lambda \vdash \mathbf{t}_2} \frac{(n)^{l(\lambda)}}{m(\lambda)} \prod_{\lambda_i} E[\rho^{\lambda_i}]^{r_i} \right) \tag{41}$$

where $E[\eta^{\lambda_i}]$ and $E[\rho^{\lambda_i}]$ are given in (37) and (38) respectively. The function `MakeTab` in [5] has been employed for computing all the multi-indexes $\mathbf{t}_1, \mathbf{t}_2$ in (41). A separate `Maple` procedure [3] has been set up in order to expand (41), see the following example.

Example 3. For $m = 2$ and $\mathbf{i} = (1, 2)$, an explicit expression of (41) is given in the following:

$$\begin{aligned} E \{ \text{Tr} [W(n)H_1] \text{Tr} [W(n)H_2]^2 \} &= n \text{Tr} (H_2) \text{Tr} (\Omega H_1 H_2) - n \text{Tr} (H_2) \text{Tr} (\Omega H_2 H_1) + n \text{Tr} (H_2) \text{Tr} (\Omega H_1) \text{Tr} (\Omega H_2) \\ &\quad - n \text{Tr} (\Omega H_2) \text{Tr} (H_1 H_2) - n^2 \text{Tr} (\Omega H_2) \text{Tr} (H_1) \text{Tr} (H_2) + 2 \text{Tr} (\Omega H_2) \text{Tr} (\Omega H_1 H_2) \\ &\quad + 2 \text{Tr} (\Omega H_2) \text{Tr} (\Omega H_2 H_1) - \text{Tr} (\Omega H_1) (\text{Tr} (\Omega H_2))^2 - \text{Tr} (\Omega H_1 H_2^2) \\ &\quad - \text{Tr} (\Omega H_2 H_1 H_2) + \text{Tr} (\Omega H_1) \text{Tr} (\Omega H_2^2) + 2 n^2 \text{Tr} (H_1 H_2) \text{Tr} (H_2) \\ &\quad + n^3 \text{Tr} (H_1) (\text{Tr} (H_2))^2 + n \text{Tr} (H_1 H_2^2) + (n^2/2) \text{Tr} (H_1) \text{Tr} (H_2^2) \\ &\quad + n^2 \text{Tr} (\Omega H_1) (\text{Tr} (H_2))^2 - (n/2) \text{Tr} (\Omega H_1) \text{Tr} (H_2^2) \\ &\quad + n \text{Tr} (H_1) (\text{Tr} (\Omega H_2))^2 - n \text{Tr} (H_1) \text{Tr} (\Omega H_2^2). \end{aligned}$$

Remark 4. If we replace the nonnegative integer n with a scalar umbra α in (36) then

$$E \{ \text{Tr} [W(\alpha)H_1]^{i_1} \dots \text{Tr} [W(\alpha)H_m]^{i_m} \} = E[(-1 \cdot \beta \cdot \tilde{\eta} + \alpha \cdot \beta \cdot \tilde{\rho})^{\mathbf{i}}] \tag{42}$$

joint moments of randomized non-central Wishart distributions. Although not explicitly mentioned, this replacement could be performed for all the results given in the following.

Joint moments of $\tilde{\rho}$ correspond to joint moments $E\{\text{Tr}[\widehat{W}(n)H_1]^{i_1} \dots \text{Tr}[\widehat{W}(n)H_m]^{i_m}\}$ of central Wishart matrices and are a generalization of formulae given in [12]. Joint moments of $\tilde{\eta}$ correspond to joint moments of $E\{\text{Tr}[AH_1]^{i_1} \dots \text{Tr}[AH_m]^{i_m}\}$ with A given in (21). Differently from [12,16], where the representation theory of symmetric group is resorted in order to compute the moments of non-central Wishart distributions, Theorem 16 involves integer partitions. Proposition 17 shows the connection between the two methods.

Proposition 17. For $\mathbf{i} = (1, \dots, 1)$

$$E[(n \cdot \beta \cdot \tilde{\rho})^{\mathbf{i}}] = \sum_{\sigma \in \mathfrak{S}_m} n^{|\mathcal{C}(\sigma)|} \prod_{c \in \mathcal{C}(\sigma)} \text{Tr} \left(\prod_{j \in c} \Sigma H_j \right), \tag{43}$$

$$E[(-1 \cdot \beta \cdot \tilde{\eta})^{\mathbf{i}}] = \sum_{\sigma \in \mathfrak{S}_m} (-1)^{|\mathcal{C}(\sigma)|} \prod_{c \in \mathcal{C}(\sigma)} l(c) \text{Tr} \left(\Omega \prod_{j \in c} \Sigma H_j \right). \tag{44}$$

Proof. Partitions of the multi-index $(1, \dots, 1)$ are matrices with entries equal to 0 and 1. When \mathbf{i} has entries equal to 0 or 1, the multivariate moment of ρ involves cyclic permutations of $(\Sigma H_1)^{i_1} \dots (\Sigma H_k)^{i_k}$ with some powers equal to 0 and the remaining equal to 1. The elements in the strings \mathbf{a} in (37) and (38) are all different and chosen in $\{1, \dots, m\}$, so that

$M_m[\mathbf{i}] = N_m[\mathbf{i}]$. Therefore for these multi-indexes, Eq. (37) reduces to $E[\rho^{\mathbf{i}}] = \text{Tr}[(\Sigma H_1)^{i_1} \cdots (\Sigma H_k)^{i_k}]$. Eq. (43) follows by observing that for $\mathbf{i} = (1, \dots, 1)$ and from (33)

$$E[(n \cdot \beta \cdot \tilde{\rho})^{\mathbf{i}}] = \sum_{\pi \in \Pi_{|\mathbf{i}|}} p_{\pi} n^{|\pi|} \prod_{B \in \pi} \text{Tr} \left[\left(\prod_{j \in B} \Sigma H_j \right) \right],$$

with $\Pi_{|\mathbf{i}|}$ the set of all partitions⁹ of $\{1, 2, \dots, |\mathbf{i}|\}$ and $p_{\pi} = \prod_{B \in \pi} (|B| - 1)!$ the number of permutations of $\{1, 2, \dots, |\mathbf{i}|\}$ corresponding to the partition π . The result follows by indexing the summation with permutations instead of partitions. Eq. (44) follows by similar arguments. \square

For $\mathbf{i} = (1, \dots, 1)$ and by using the permutation identity $e \in \mathfrak{S}_m$, we have

$$E[(n \cdot \beta \cdot \tilde{\rho})^{\mathbf{i}}] = E \left\{ \prod_{c \in C(e)} \text{Tr} [\widehat{W}(n)H_c] \right\} \quad \text{and} \quad E[(-1 \cdot \beta \cdot \tilde{\eta})^{\mathbf{i}}] = E \left\{ \prod_{c \in C(e)} \text{Tr} [AH_c] \right\}.$$

By the group action of \mathfrak{S}_m on \mathbb{C} , from (43) and (44) we have

$$E \left\{ \prod_{c \in C(\sigma)} \text{Tr} \left[\prod_{j \in c} \widehat{W}(n)H_j \right] \right\} = \sum_{\tau \in \mathfrak{S}_m} n^{l(\sigma \tau^{-1})} \prod_{c \in C(\tau)} \text{Tr} \left(\prod_{j \in c} \Sigma H_j \right), \tag{45}$$

$$E \left\{ \prod_{c \in C(\sigma)} \text{Tr} \left[\prod_{j \in c} AH_j \right] \right\} = \sum_{\tau \in \mathfrak{S}_m} (-1)^{l(\sigma \tau^{-1})} \prod_{c \in C(\tau)} l(c) \text{Tr} \left(\Omega \prod_{j \in c} \Sigma H_j \right). \tag{46}$$

4.3. Joint cumulants

Joint cumulants of non-central Wishart distributions can be recovered from Theorem 16, by using the additivity property given in Proposition 4.1 of [5]. Denote the m -tuple $(\text{Tr}[W(n)H_1], \dots, \text{Tr}[W(n)H_m])$ with μ . Its \mathbf{i} -th joint cumulant $\text{Cum}_{\mathbf{i}}(\text{Tr}[W(n)H_1], \dots, \text{Tr}[W(n)H_m])$ is the multivariate moment of $(\chi \cdot \mu)^{\mathbf{i}}$. From (36) we have

$$E[(\chi \cdot \mu)^{\mathbf{i}}] = E[(\chi \cdot (-1) \cdot \beta \cdot \tilde{\eta})^{\mathbf{i}}] + E[(\chi \cdot n \cdot \beta \cdot \tilde{\rho})^{\mathbf{i}}]. \tag{47}$$

Eq. (33) allows us to compute $E[(\chi \cdot (-1) \cdot \beta \cdot \tilde{\eta})^{\mathbf{i}}]$ and $E[(\chi \cdot n \cdot \beta \cdot \tilde{\rho})^{\mathbf{i}}]$, with α replaced by $\chi \cdot (-1)$ and $\chi \cdot n$ respectively, and ν replaced by $\tilde{\eta}$ and $\tilde{\rho}$ too. Since the moments of $\chi \cdot (-1)$ and $\chi \cdot n$ are all zero except the first, the only contribution in (33) is for $\lambda = \mathbf{i}$ for which $d_{\lambda} = 1$. Then the following theorem is proved, which generalizes Theorem 12 and Proposition 10.

Theorem 18. $\text{Cum}_{\mathbf{i}}(\text{Tr}[W(n)H_1], \dots, \text{Tr}[W(n)H_m]) = \mathbf{i}!(nE[\rho^{\mathbf{i}}] - E[\eta^{\mathbf{i}}])$, with $E[\rho^{\mathbf{i}}]$ and $E[\eta^{\mathbf{i}}]$ given in (37) and (38) respectively.

Example 4. If $m = 2$ and $\mathbf{i} = (1, 2)$, then

$$\begin{aligned} \text{Cum}_{(1,2)}(\text{Tr}[W(n)H_1], \text{Tr}[W(n)H_2]) &= 2! \{ n \text{Tr} [(\Sigma H_1)(\Sigma H_2)^2] - \text{Tr} [\Omega(\Sigma H_1)(\Sigma H_2)^2] \\ &\quad - \text{Tr} [\Omega(\Sigma H_2)^2(\Sigma H_1)] - \text{Tr} [\Omega(\Sigma H_1)(\Sigma H_2)(\Sigma H_1)] \}. \end{aligned}$$

Depending on the symbolic representation of non-central Wishart distributions (42), cumulants of the randomized version are obtained from (47) by replacing $\chi \cdot n$ with $\chi \cdot \alpha$, which is the α -cumulant umbra. Then Theorem 18 needs to be updated as follows.

Theorem 19. If $\{c_i\}$ is the sequence of cumulants of α , then

$$\text{Cum}_{\mathbf{i}}(\text{Tr}[W(\alpha)H_1], \dots, \text{Tr}[W(\alpha)H_m]) = \mathbf{i}! \left(\sum_{\lambda \vdash \mathbf{i}} \frac{c_{l(\lambda)}}{m(\lambda)} \prod_{\lambda_i} E[\rho^{\lambda_i}]^{r_i} - E[\eta^{\mathbf{i}}] \right). \tag{48}$$

⁹ A partition π of $\{v_1, \dots, v_i\}$ is a collection $\pi = \{B_1, \dots, B_k\}$ of $k \leq i$ disjoint non-empty subsets of $\{v_1, \dots, v_i\}$ whose union is $\{v_1, \dots, v_i\}$. We denote the set of all partitions of the set $\{v_1, \dots, v_i\}$ with Π_i .

Example 5. If $m = 2$ and $\mathbf{i} = (1, 2)$, in order to compute $\text{Cum}_{(1,2)}(\text{Tr}[W(\alpha)H_1], \text{Tr}[W(\alpha)H_2])$ and to better understand the power of the symbolic method, we could employ the results of **Example 3** for the part not involving the non-centrality matrix Ω . Indeed, the summation in (48) corresponds to the latter in (41) with $n^{l(\lambda)}$ replaced by $c_{l(\lambda)}$. Then in **Example 3**, we can select the terms not involving Ω and replace the occurrences of n^k with c_k . The result is the following:

$$\begin{aligned} \text{Cum}_{(1,2)}(\text{Tr}[W(\alpha)H_1], \text{Tr}[W(\alpha)H_2]) &= 4c_2 \text{Tr}(H_1H_2) \text{Tr}(H_2) + 2c_3 \text{Tr}(H_1) (\text{Tr}(H_2))^2 \\ &\quad + 2c_1 \text{Tr}(H_1H_2^2) + c_2 \text{Tr}(H_1) \text{Tr}(H_2^2) - 2\text{Tr}[\Omega (\Sigma H_1) (\Sigma H_2)^2] \\ &\quad - 2\text{Tr}[\Omega (\Sigma H_2)^2 (\Sigma H_1)] - 2\text{Tr}[\Omega (\Sigma H_1) (\Sigma H_2) (\Sigma H_1)]. \end{aligned}$$

5. Applications

5.1. Permanents

Let $Y = (y_{ij}) \in \mathbb{C}^{p \times p}$. For $d \in \mathbb{C}$, the d -permanent $\text{per}_d(Y)$ of the matrix Y is the function

$$\text{per}_d(Y) := \sum_{\sigma \in \mathfrak{S}_p} d^{|\mathcal{C}(\sigma)|} \prod_{j=1}^p y_{j, \sigma(j)}. \tag{49}$$

The determinant function corresponds to $d = -1$ while $d = 1$ gives the classical permanent function. The d -extension of the master theorem [17] states that if $Z = \text{diag}(z_1, \dots, z_m)$ and $T = (t_{i,j})$ is a $m \times m$ matrix such that $\|ZT\| < 1$ then $\text{per}_d[T(\mathbf{i})]$ is the \mathbf{i} -th coefficient of $\det(I - ZT)^{-d}$,

$$\text{per}_d[T(\mathbf{i})] = \frac{\partial^{|\mathbf{i}|}}{\partial z_1^{i_1} \dots \partial z_m^{i_m}} [\det(I - ZT)^{-d}]_{z=0} \quad \text{with } T(\mathbf{i}) = \begin{pmatrix} t_{i_1, i_1} & \dots & t_{i_1, i_m} \\ \vdots & \vdots & \vdots \\ t_{i_m, i_1} & \dots & t_{i_m, i_m} \end{pmatrix}.$$

From (34), $\det(I - ZT)^{-d}$ is the g.f. of a class of central Wishart distributions. Then the following result holds when ZT can be decomposed as $\Sigma (H_1 z_1 + \dots + H_m z_m)$ for suitable matrices Σ and H_1, \dots, H_m .

Proposition 20. *If there exists a symmetric matrix $\Sigma \in \mathbb{C}^{p \times p}$ and $H_1, \dots, H_m \in \mathbb{C}^{p \times p}$ such that $ZT = \Sigma(H_1 z_1 + \dots + H_m z_m)$, then $\text{per}_d[T(\mathbf{i})] = E[(d \cdot \beta \cdot \tilde{\rho})^{\mathbf{i}}]$ with $\tilde{\rho}$ given in (37).*

A generalization of **Proposition 20** consists in replacing the complex d with an umbra α umbrally representing $\{a_i\}$, that is

$$\text{per}_\alpha[T(\mathbf{i})] := \sum_{\sigma \in \mathfrak{S}_p} a_{|\mathcal{C}(\sigma)|} \prod_{j=1}^p [T(\mathbf{i})]_{j, \sigma(j)} = E[(\alpha \cdot \beta \cdot \tilde{\rho})^{\mathbf{i}}].$$

5.2. Spectral polykays

Cumulants of Wishart distributions have been employed in approximating the density of sample correlation matrix [13] or within wireless communications [27]. Moreover, due to **Theorem 12** estimations of cumulants can add more information on the structure of the covariance matrix and of the non-centrality matrix. A classical way to estimate cumulants and their products is by using k -statistics and polykays [4]. They rely on power sum symmetric polynomials involving i.i.d. r.v.'s of random samples. A different way has been recently proposed in [6] by introducing the notion of spectral sample. Consider the matrix $H_{[m \times p]}$ obtained by deleting the last $p - m$ rows of a random unitary matrix H uniformly distributed with respect to Haar measure [21]. The set of eigenvalues $\mathbf{y} = (y_1, \dots, y_m) \in \mathbb{R}^m$ of the $m \times m$ Hermitian random matrix $Y = H_{[m \times p]} X H_{[p \times m]}^\dagger$ is called a *spectral sample* of size m from $(x_1, \dots, x_p) \in \mathbb{R}^p$ with $X = \text{diag}(x_1, \dots, x_p)$. If $\lambda = (1^{r_1}, 2^{r_2}, \dots)$ is a partition of some nonnegative integers,¹¹ spectral polykays $\mathfrak{R}_\lambda(\mathbf{y})$ are symmetric polynomials expressed in terms of $\text{Tr}(Y^i)$, such that

$$E[\mathfrak{R}_\lambda(\mathbf{y})] = \prod_{i=1}^{l(\lambda)} E[\text{Tr}[\mathcal{C}(m)]^{\lambda_i}] \tag{50}$$

¹⁰ If A is a $m \times m$ matrix, then $\|A\| := \max\{|\theta_1|, \dots, |\theta_m|\}$, where $\theta_1, \dots, \theta_m$ are the eigenvalues of A .

¹¹ Here λ runs through all partitions of length less than or equal to m [18].

with $\text{Tr}[\mathfrak{C}(m)]$ given in (28). For non-central Wishart distributions and from Theorem 15, spectral polykays (50) are unbiased estimators of normalized cumulants

$$\text{Cum}_\lambda(\text{Tr}[W(m)]) = \prod_{i=1}^{l(\lambda)} (\text{Cum}_i(\text{Tr}[W(m)]))^{r_i} = p^{l(\lambda)} E[\mathfrak{R}_\lambda(W(m))]. \tag{51}$$

For central Wishart distributions, which are unitarily invariant¹² the spectral sampling is a way to extract information from the set of its eigenvalues when $X = \Lambda$. In this case, the matrix Y is a principal submatrix [11] of $\widehat{W}(n)$, and $\mathfrak{R}_\lambda(\mathbf{y})$ are unbiased estimators of their normalized cumulants. These estimators do not depend on the size of the submatrix, since $E[\mathfrak{R}_\lambda(W(n)_{mm})|W(n)] = \mathfrak{R}_\lambda(W(n))$, for $m \leq n$, where $W(n)_{mm}$ denotes the principal submatrix involving the first m rows of $W(n)$, cf. [6]. The spectral estimators of mean, variance, asymmetry and kurtosis are given in the following:

$$\begin{aligned} \kappa_{(1)} &= \frac{S_1}{m}, & \kappa_{(2)} &= \frac{mS_2 - S_1^2}{m(m^2 - 1)}, & \kappa_{(3)} &= 2 \frac{2S_1^3 - 3mS_1S_2 + m^2S_3}{m(m^2 - 1)(m^2 - 4)} \\ \kappa_{(4)} &= 6 \frac{-5S_1^4 + 10mS_1^2S_2 + (3 - 2m^2)S_2^2 - (4 + 4m^2)S_1S_3 + (m + m^3)S_4}{m^2(m^2 - 1)(m^2 - 4)(m^2 - 9)}. \end{aligned}$$

5.3. Generalized moments and open problems

Similarly to the product (45) involving central Wishart matrices, a generalization of joint moments (2) is given by

$$E \left\{ \prod_{c \in \mathfrak{C}(\sigma)} \text{Tr} \left[\prod_{j \in c} W(n)H_j \right] \right\}, \quad \sigma \in \mathfrak{S}_k. \tag{52}$$

Differently from (45), this computation could not be performed by using the group action of \mathfrak{S}_k on \mathfrak{C} . A different strategy consists in using the convolution given in Theorem 1. The first step is to transform (52) in a summation of joint moments involving $\widehat{W}(n)$ and A . This is the result of the following proposition.

Proposition 21.

$$E \left[\prod_{c \in \mathfrak{C}(\sigma)} \text{Tr} \left(\prod_{j \in c} W(n)H_j \right) \right] = \sum_{(B_1, \dots, B_k) \in \{\widehat{W}(n), A\}^k} \prod_{c \in \mathfrak{C}(\sigma)} E \left[\text{Tr} \left(\prod_{j \in c} B_j H_j \right) \right]. \tag{53}$$

Proof. In (52), replace $W(n)$ with $\widehat{W}(n) + A$ and observe that

$$\prod_{c \in \mathfrak{C}(\sigma)} \text{Tr} \left(\prod_{j \in c} (\widehat{W}(n) + A)H_j \right) = \prod_{c \in \mathfrak{C}(\sigma)} \sum_{(B_{j_1}, \dots, B_{j_{l(c)}}) \in \{\widehat{W}(n), A\}^k} \text{Tr}(B_{j_1}H_{j_1} \cdots B_{j_{l(c)}}H_{j_{l(c)}}).$$

The result follows by multiplying the summations and by applying the operator E . \square

The second step is to compute the right-hand-side of (53) by separating the contribution of $\widehat{W}(n)$ and A , in order to employ (45) and (46). A way to perform such computation is to resort free probability [21], since $\widehat{W}(n)$ and A are elements of a noncommutative probability space¹³ $(M_n(\mathbb{C}[\mathcal{A}]), \text{Tr})$, with $M_n(\mathbb{C}[\mathcal{A}])$ the algebra of umbral matrices with the usual matrix multiplication. In the following, a way involving free cumulants is suggested to perform such computation: it would be convenient to characterize some closed form formula involving the functional Tr without the employment of free cumulants. This is still an open problem.

Let \mathcal{NC} be the lattice of all noncrossing partitions¹⁴ of $[i]$. For any noncrossing partition π , set

$$\text{Tr}_\pi(B_1, \dots, B_i) = \prod_{D \in \pi} \text{Tr}(B_{j_1} \cdots B_{j_s})$$

for $D = (j_1 < \dots < j_s)$. Free cumulants are defined as multilinear functionals such that

$$c_\pi(B_1, \dots, B_i) = \prod_{D \in \pi} c_{|D|}(B_{j_1} \cdots B_{j_s}) \quad \text{with} \quad c_i(B_1, \dots, B_i) = \sum_{\pi \in \mathcal{NC}} m(\pi, 1_i) \text{Tr}_\pi(B_1, \dots, B_i), \tag{54}$$

¹² A unitarily invariant central Wishart matrix is such that $\widehat{W}(n) = U\Lambda U^\dagger$, with U a Haar matrix independent of the diagonal matrix Λ of the eigenvalues of $\widehat{W}(n)$, and the superscript \dagger denoting the conjugate transpose. The columns of U are the eigenvectors of $\widehat{W}(n)$.

¹³ A noncommutative probability space is a pair (\mathbb{A}, Φ) , where \mathbb{A} is a unital noncommutative algebra and $\Phi : \mathbb{A} \rightarrow \mathbb{C}$ is a unital linear functional.

¹⁴ A noncrossing partition $\pi = \{B_1, B_2, \dots, B_k\}$ of the set $[i]$ is a partition such that if $1 \leq h < l < s < k \leq i$, with $h, s \in B_n$ and $l, k \in B_{n'}$, then $n = n'$.

where $m(\pi, 1_j)$ is the Moebius function on the lattice of noncrossing partitions [21]. For two sets $\{S_1, \dots, S_k\}$ and $\{T_1, \dots, T_k\}$ of random matrices satisfying the hypothesis of Theorem 14.4 in [21], we have

$$\text{Tr}(S_1 T_1 \cdots S_k T_k) = \sum_{\pi_S \in \mathcal{NC}(1, 3, \dots, 2n-1)} c_{\pi_S}[S_1, \dots, S_k] \left(\sum_{\substack{\pi_T \in \mathcal{NC}(2, 4, \dots, 2n) \\ \pi_S \cup \pi_T \in \mathcal{NC}(n)}} c_{\pi_T}[T_1, \dots, T_k] \right). \quad (55)$$

The products on the right-hand-side of (53) could be computed by using (55) with a suitable choice of the matrices S_i and T_i . For example, in order to compute $\text{Tr}(A H_1 A H_2 \widehat{W}(n) H_3 \widehat{W}(n) H_4)$ choose $S_1 = A H_1, T_1 = I, S_2 = A H_2, T_2 = \widehat{W}(n) H_3, S_3 = I, T_3 = \widehat{W}(n) H_4$. An expression of (53) in terms of (45) and (46) follows by using (54) and by characterizing the permutations σ corresponding to the non-crossing partitions π involved in (55). The hypothesis of Theorem 14.4 in [21] rely on the notion of freely independence of S_i and T_i . This theorem could be still employed if S_i and T_i are asymptotical free independent, that is if A and $\widehat{W}(n)$ are asymptotical free independent. This is true if the fluctuation of A is not too large, behaving like a constant matrix. Of course, this fluctuation depends on the non-centrality matrix Ω . So, if A admits an eigenvalue distribution and all normalized traces of A would have $1/p^2$ estimates [21], then A and $\widehat{W}(n)$ are asymptotical free independent.

References

- [1] T.W. Anderson, The non-central Wishart distribution and certain problems for multivariate statistics, *Ann. Math. Stat.* 17 (4) (1946) 409–431.
- [2] K. Cattell, F. Ruskey, J. Sawada, M. Serra, Fast algorithms to generate necklaces, unlabeled necklaces, and irreducible polynomials over $\text{GF}(2)$, *J. Algorithms* 37 (2000) 267–282.
- [3] E. Di Nardo, G. Guarino, A new algorithm for computing moments of complex non-central Wishart distributions, *Maple Worksheet*, 2013, <http://www.maplesoft.com/applications/view.aspx?SID=143890>.
- [4] E. Di Nardo, G. Guarino, D. Senato, A unifying framework for k -statistics, polykays and their multivariate generalizations, *Bernoulli* 14 (2) (2008) 440–468.
- [5] E. Di Nardo, G. Guarino, D. Senato, A new algorithm for computing the multivariate Faà di Bruno's formula, *Appl. Math. Comput.* 217 (2011) 6286–6295.
- [6] E. Di Nardo, P. McCullagh, D. Senato, Natural statistics for spectral samples, *Ann. Statist.* 41 (2) (2013) 982–1004.
- [7] E. Di Nardo, H. Niederhausen, D. Senato, A symbolic handling of Sheffer sequences, *Ann. Mat. Pura Appl.* 190 (2011) 489–506.
- [8] E. Di Nardo, P. Petruccio, D. Senato, Cumulants and convolutions via Abel polynomials, *European J. Combin.* 31 (2010) 1792–1804.
- [9] E. Di Nardo, D. Senato, Umbral nature of the Poisson random variables, in: *Algebraic Combinatorics and Computer Science* (Crapo, Senato Eds.) A Tribute to Gian-Carlo Rota, Springer-Verlag, 2001, pp. 245–266.
- [10] E. Di Nardo, D. Senato, An umbral setting for cumulants and factorial moments, *European J. Combin.* 27 (2006) 394–413.
- [11] M. Drton, H. Massam, I. Olkin, Moments of minors of Wishart matrices, *Ann. Statist.* 36 (5) (2008) 2261–2283.
- [12] P. Graczyk, G. Letac, H. Massam, The complex Wishart distribution and the symmetric group, *Ann. Statist.* 31 (1) (2003) 287–309.
- [13] T. Kollo, K. Ruul, Approximations to the distribution of the sample correlation matrix, *J. Multivariate Anal.* 85 (2) (2003) 318–334.
- [14] T. Kollo, D. Von Rosen, Approximating by the Wishart distribution, *Ann. Inst. Statist. Math.* 47 (4) (1995) 767–783.
- [15] A. Kuriki, Y. Numata, Graph presentations for moments of noncentral Wishart distributions and their applications, *Ann. Inst. Statist. Math.* 62 (4) (2010) 645–672.
- [16] G. Letac, H. Massam, The noncentral Wishart as an exponential family and its moments, *J. Multivariate Anal.* 99 (1) (2008) 1393–1417.
- [17] I. Li-Lu, D. Richards, MacMahon's master theorem, representation theory and moments of Wishart distributions, *Adv. Appl. Math.* 27 (2001) 531–547.
- [18] I.G. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford University Press, 1998.
- [19] D. Maiwald, D. Kraus, Calculation of moments of complex Wishart and complex inverse Wishart distributed matrices, *IEE Proc., Radar Sonar Navig.* 147 (4) (2000) 162–168.
- [20] R.J. Muirhead, *Aspects of multivariate statistical theory*, Wiley Series in Probability and Statistics (1982) 673.
- [21] A. Nica, R. Speicher, Lectures on the combinatorics of free probability, London mathematical society, in: *Lecture Note Series*, vol. 335, Cambridge University Press, 2006, p. 417.
- [22] H. Niederhausen, *Finite Operator Calculus With Applications to Linear Recursions*, 2010, <http://math.fau.edu/niederhausen/HTML/Research/UmbralCalculus/bookS2010.pdf>.
- [23] G.-C. Rota, B.D. Taylor, The classical umbral calculus, *SIAM J. Math. Anal.* 25 (2) (1994) 694–711.
- [24] R.P. Stanley, *Enumerative combinatorics*, Vol. I, Wadsworth & Brooks/Cole Advanced Books & Software, 1986.
- [25] J.A. Tague, C.I. Caldwell, Expectations of useful complex Wishart forms, *Multidimens. Syst. Signal Process.* 5 (3) (1994) 263–279.
- [26] T. Tao, Topics in random matrix theory, *Graduate Studies in Mathematics*. 132 (2012) 282.
- [27] A.M. Tulino, S. Verdú, *Random Matrix Theory and Wireless Communications*, Now Publishers Inc., 2004, p. 192.
- [28] C.S. Withers, S. Nadarajah, Moments and cumulants for the complex Wishart, *J. Multivariate Anal.* 112 (2012) 242–247.