A Simple Formula for the Moments of
Unitarily Invariant Matrix Distributions

Stephen Howard\textsuperscript{1} and Ali Pezeshki\textsuperscript{2}

\textsuperscript{1}Defence Science & Technology Group, Australia
\textsuperscript{2}Colorado State University, USA

CoE Meeting
June 2022
Complex Unitarily Invariant Distributions

- **Complex multivariate Gaussian**: Modeling baseband signal measurements.

- Complex Wishart and complex inverse Wishart: Sample covariances and inverse sample covariances

\[ R = XX^H \text{ and } R^{-1} = (XX^H)^{-1} \]

- Complex beta: Likelihood ratios for detection

\[ y = \mu H x + J z + n \sim \begin{cases} \mathcal{CN}[J z, \sigma^2 I], & \mathcal{H}_0 (\mu = 0) \\ \mathcal{CN}[\mu H x + Jz, \sigma^2 I], & \mathcal{H}_1 (\mu > 0) \end{cases} \]

\[ \ell = \frac{y^H P_J^\perp E_J H P_J^\perp y}{y^H P_J^\perp (I - E_J H) P_J^\perp y} \]
Complex Unitarily Invariant Distributions

- **Complex multivariate Gaussian:** Modeling baseband signal measurements.

- **Complex Wishart and complex inverse Wishart:** Sample covariances and inverse sample covariances

\[ R = XX^H \text{ and } R^{-1} = (XX^H)^{-1} \]

- **Complex beta:** Likelihood ratios for detection

\[ y = \mu H x + J z + n \sim \begin{cases} 
\mathcal{CN}[J z, \sigma^2 I], & \mathcal{H}_0 (\mu = 0) \\
\mathcal{CN}[\mu H x + J z, \sigma^2 I], & \mathcal{H}_1 (\mu > 0) 
\end{cases} \]

\[ \ell = \frac{y^H P_J^\perp E_J H P_J^\perp y}{y^H P_J^\perp (I - E_J H) P_J^\perp y} \]
Complex Unitarily Invariant Distributions

- **Complex multivariate Gaussian**: Modeling baseband signal measurements.

- **Complex Wishart and complex inverse Wishart**: Sample covariances and inverse sample covariances

  \( R = XX^H \) and \( R^{-1} = (XX^H)^{-1} \)

- **Complex beta**: Likelihood ratios for detection

  \[ y = \mu H x + J z + n \sim \begin{cases} 
  \mathcal{CN}[J z, \sigma^2 I], & H_0 (\mu = 0) \\
  \mathcal{CN}[\mu H x + J z, \sigma^2 I], & H_1 (\mu > 0) 
\end{cases} \]

  \[ \ell = \frac{y^H P^\perp J E J^H P^\perp_J y}{y^H P^\perp_J (I - E J^H) P^\perp_J y} \]
Complex matrix beta:

Fisher Information in compressed sensing (Pakrooh et al. 2015)

\[ y = x(\theta) + n \quad \text{vs} \quad z = \Phi y \]

\[ G = \begin{bmatrix} \frac{\partial}{\partial \theta_1} x(\theta) & \cdots & \frac{\partial}{\partial \theta_p} x(\theta) \end{bmatrix} \]

\[ F = \frac{1}{\sigma^2} G^H P_{\Phi H} G \quad \text{and} \quad \text{CRLB} = F^{-1}. \]

Detection statistics in passive radar (Howard et al. 2015)

\[ \gamma = \frac{|G|}{\prod_{i=1}^{M} |G_{ii}|}; \quad G = X^H X \]
Complex matrix beta:

- Fisher Information in compressed sensing (Pakrooh et al. 2015)

\[ y = x(\theta) + n \quad \text{vs} \quad z = \Phi y \]

\[ G = \left[ \frac{\partial}{\partial \theta_1} x(\theta) \mid \cdots \mid \frac{\partial}{\partial \theta_p} x(\theta) \right] \]

\[ F = \frac{1}{\sigma^2} G^H P_{\Phi H} G \quad \text{and} \quad \text{CRLB} = F^{-1}. \]

- Detection statistics in passive radar (Howard et al. 2015)

\[ \gamma = \frac{|G|}{\prod_{i=1}^{M} |G_{ii}|}; \quad G = X^H X \]
Select prior work:

- Multivariate normal: Isserlis (1918)
- Complex multivariate normal: James (1953); Muirhead (book)
- Real Wishart and matrix beta: Muirhead (book)
- Complex Wishart: Maiwald & Kraus (2000); Graczyk et al. (2003); Nagar & Gupta (2011).
- Complex matrix beta: Gupta et al. (2009), Pakrooh et al. (2015)

Goal: Derive a formula for moments of all complex matrix distributions that can be transformed to a unitarily invariant distribution via conjugation by a fixed matrix.
Methodology: Exploiting the relationship between the moments of unitarily invariant distributions and the joint action of the unitary and symmetric groups on $(\mathbb{C}^N)^{\otimes d}$.

Key: Schur-Weyl duality, which relates the irreducible representations of unitary and symmetric groups.

Advantage: A moment formula for all complex multivariate unitarily invariant distribution that separates the combinatorial aspect of moment computation from the calculation of a small number of specific distribution dependent moments.

Closest prior work: Graczyk et al. (2003); Exploits the relationship between moments of the Wishart distribution and the symmetric group.
The Moment Problem

- $Q$ is a complex $N \times N$ matrix-valued random variable with probability distribution $F(Q)$.

- The distribution is unitarily invariant, if for all $V \in U(N)$, the group of $N \times N$ unitary matrices, $VQV^\dagger$ has the same distribution as $Q$.

- The notation $Q \otimes^d$ denotes the tensor power $Q \otimes Q \otimes \cdots \otimes Q$, with $d$ factors.

- View $Q$ as a random linear operator on $\mathbb{C}^N$ and $Q \otimes^d$ as a random linear operator on $(\mathbb{C}^N) \otimes^d$.

- $E(Q \otimes^d)$ is a fixed linear operator on $(\mathbb{C}^N) \otimes^d$. Its matrix elements include the moments corresponding to all products of matrix elements of $Q$ with $d$ factors.

- **Problem**: Compute $E(Q \otimes^d)$. 
Two important group actions can be defined on $(\mathbb{C}^N)^\otimes d$.

The first is the action of the symmetric group $S_d$ (group of permutations on $d$ letters) on $(\mathbb{C}^N)^\otimes d$ defined by

$$L_\sigma(v_1 \otimes \cdots \otimes v_d) = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(d)}.$$ 

The other is the action of the unitary group $U(N)$ on $(\mathbb{C}^N)^\otimes d$ defined by

$$U^{\otimes d}(v_1 \otimes \cdots \otimes v_d) = Uv_1 \otimes \cdots \otimes Uv_d.$$ 

Under these two group actions $(\mathbb{C}^N)^\otimes d$ becomes a unitary representation of the product group $U(N) \times S_d$. 
Decomposition into Irreducible Representations

- This representation decomposes into irreducible representations \( \{ W_\lambda \} \):

\[
(\mathbb{C}^N)^d = \bigoplus_{\lambda} W_\lambda
\]

where the sum is over all partitions \( \lambda \) of \( d \) with no more than \( N \) parts. That is, \( \lambda = (\lambda_1, \ldots, \lambda_\ell), \ \ell \leq N \), with the \( \lambda_i \) integers satisfying \( \lambda_i \geq \lambda_{i+1} \) and \( \lambda_1 + \cdots + \lambda_\ell = d \).

- Each \( \lambda \) can be associated with a **Young diagram**:

\[
\lambda = (3, 2^2, 1) \quad \lambda' = (4, 3, 1)
\]
The decomposition \((\mathbb{C}^N)^\otimes d = \bigoplus_{\lambda} W_\lambda\) means that linear operators in \(U(N) \times S_d\) can be simultaneously block diagonalized with the blocks labeled by \(\lambda\).

Schur-Weyl duality states that

\[
W_\lambda = V_\lambda \otimes S_\lambda
\]

where \(V_\lambda\) are non-equivalent unitary irreducible representations of \(U(N)\) and \(S_\lambda\) are non-equivalent unitary irreducible representations of \(S_d\).
Associated with each $W_\lambda$ there is a Young orthogonal projection operator $Y_\lambda$ onto that subspace. These satisfy

$$\sum_\lambda Y_\lambda = I, \quad Y_{\lambda_1} Y_{\lambda_2} = \delta_{\lambda_1,\lambda_2} Y_{\lambda_1}$$

and take the form

$$Y_\lambda = \frac{\chi_\lambda(e)}{d!} \sum_{\sigma \in S_d} \chi_\lambda(\sigma)L_\sigma.$$

The $\chi_\lambda(\sigma)$ are real integer coefficients which comprise the irreducible characters of $S_d$. These can be constructed by standard methods (Hamermesh (1962)).
A Key Observation

- The linear operator \( E(Q \otimes d) \) commutes with the action of \( \mathcal{U}(N) \times S_d \) on \((\mathbb{C}^N)^{\otimes d}\).

- So if \( E(Q \otimes d) \) is restricted to one of the irreducible representations \( W_\lambda \), then Schur’s Lemma implies that \( E(Q \otimes d) \) acts as a multiple, \( m_\lambda(Q) \), of the identity.

- Then we simply need to calculate this \( m_\lambda(Q) \), which is a coefficient moment of \( Q \).

- Choose a convenient unit vector \( \mathbf{w}_\lambda \in W_\lambda \) in each irreducible representation and compute

\[
m_\lambda = \langle \mathbf{w}_\lambda, E(Q \otimes d) \mathbf{w}_\lambda \rangle.
\]
The Main Result

The $d^{\text{th}}$ order moments of $Q$ are given by

$$E (Q^{\otimes d}) = \sum_{\lambda} m_{\lambda} Y_{\lambda}$$

where the coefficient moments $m_{\lambda}(Q)$ are given by

$$m_{\lambda} = E \left( M_{\chi_1}^{\alpha_1'} M_{\chi_2}^{\alpha_2'} \ldots M_{\chi_k}^{\alpha_k'} \right)$$

where $M_{\ell}$ denotes the leading principal minor of order $\ell$ for $Q$, i.e., the determinant of its top left $\ell \times \ell$ block:

$$\det[\langle e_i, Q e_j \rangle]_{i,j=1,\ldots,\ell},$$

and $\lambda' = (\lambda_1'^{\alpha_1'}, \lambda_2'^{\alpha_2'}, \ldots, \lambda_k'^{\alpha_k'})$ is the transpose of $\lambda$. 

Howard & Pezeshki
Moments of Unitarily Invariant Matrix Distributions
If $X$ is a complex matrix-valued random variable with the property that $Q = G^{-1}X(G^{-1})^\dagger$ has a unitarily invariant probability distribution for some fixed non-singular matrix $G$, then

$$E(X \otimes_d) = \sum_\lambda m_\lambda(Q) \ R_\lambda$$

where $R_\lambda = Y_\lambda R^{\otimes d} Y_\lambda$, with $R = GG^\dagger$. 

Howard & Pezeshki
Moments of Unitarily Invariant Matrix Distributions
Computing Moments

- Calculate a \( d^{th} \) order moment of elements of \( X = GQG^\dagger \).

- For \( R = GG^\dagger \), matrix elements of \( R_\lambda \) are

\[
(R_\lambda)_{IJ} = \frac{\chi_\lambda(e)}{d!} \sum_{\sigma \in S_d} \chi_\lambda(\sigma)(R^{\otimes d})_{I\sigma(J)}
\]

where \( I = (i_1, \ldots, i_d) \) and \( J = (j_1, \ldots, j_d) \).

- We have

\[
E(x_{i_1j_1} \cdots x_{i_dj_d}) = E(X^{\otimes d})_{IJ} = \sum_{\lambda} m_\lambda(Q) \ (R_\lambda)_{IJ}
\]

- For \( R = I \),

\[
E(q_{i_1j_1} \cdots q_{i_dj_d}) = E(Q^{\otimes d})_{IJ} = \sum_{\lambda} m_\lambda(Q) \frac{\chi_\lambda(e)}{d!} \sum_{\sigma: I=\sigma(J)} \chi_\lambda(\sigma)
\]
For some complex matrix distributions, the coefficient moments $m_\lambda(Q)$ can be computed from just a knowledge of the normalization factors of the distributions.

Two important examples:

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$m_\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta</td>
<td>$\prod_{\ell=1}^{k} \frac{B_{k_\ell}'(a+\sum_{j=1}^{\ell} \alpha_j', b)}{B_{k_\ell}'(a+\sum_{j=1}^{\ell-1} \alpha_j', b)}$</td>
</tr>
<tr>
<td>Wishart/Gamma</td>
<td>$\prod_{\ell=1}^{k} \frac{\Gamma_{k_\ell}'(a+\sum_{j=1}^{\ell} \alpha_j')}{\Gamma_{k_\ell}'(a+\sum_{j=1}^{\ell-1} \alpha_j')}$</td>
</tr>
</tbody>
</table>

Also applies to the type II complex beta distribution, the inverse complex matrix beta distribution, and the inverse complex gamma/Wishart distributions.
Suppose $z$ is a unitarily invariant vector-valued random variable.

If $y = R^{1/2}z$, with $R$ a positive-definite hermitian matrix, and $Q = zz^\dagger$, then there is only a single term $\lambda = (d)$ in the expansion with $m_d(Q) = E(|z_1|^{2d})$:

$$E(y_{i_1} \bar{y}_{j_1} \cdots y_{i_d} \bar{y}_{j_d}) = R^{\otimes d} E(Q^{\otimes d})$$

$$= \frac{1}{d!} E(|z_1|^{2d}) \sum_{\sigma \in S_d} (R^{\otimes d})_{I\sigma(J)}.$$

This is a generalization of Isserlis’ Theorem (1918) to unitarily invariant vector distributions.
The formula for the moments of complex multivariate unitarily invariant distributions derived here provides a powerful tool for

- dealing with the combinatorial aspect of moment computation
- and separating them from calculations of a small number of specific distribution dependent moments.

Our approach could be generalized to compute moments of other classes of distributions.