The Eigenvectors of Single-Spiked Complex Wishart Matrices: Finite and Asymptotic Analyses

Prathapasinghe Dharmawansa, Pasan Dissanayake, and Yang Chen

Abstract—Let \( W \in \mathbb{C}^{n \times n} \) be a single-spiked Wishart matrix in the class \( W \sim \mathcal{W}_n(m, I_n + \theta vv^\dagger) \) with \( m \geq n \), where \( I_n \) is the \( n \times n \) identity matrix, \( v \in \mathbb{C}^{n \times 1} \) is an arbitrary vector with unit Euclidean norm, \( \theta \geq 0 \) is a non-random parameter, and \( (\cdot)^\dagger \) represents the conjugate-transpose operator. Let \( u_1 \) and \( u_n \) denote the eigenvectors corresponding to the smallest and the largest eigenvalues of \( W \), respectively. This paper investigates the probability density function (p.d.f.) of the random quantity \( Z_\ell^{(n)} = |v^\dagger u_\ell|^2 \) for \( \ell = 1, n \). In particular, we derive a finite dimensional closed-form p.d.f. for \( Z_\ell^{(n)} \) which is amenable to asymptotic analysis as \( m, n \) diverges with \( m - n \) fixed. It turns out that, in this asymptotic regime, the scaled random variable \( nZ_\ell^{(n)} \) converges in distribution to \( \chi_2^2/(1 + \theta) \), where \( \chi_2^2 \) denotes a chi-squared random variable with two degrees of freedom. This reveals that \( u_1 \) can be used to infer information about the spike. On the other hand, the finite dimensional p.d.f. of \( Z_\ell^{(n)} \) is expressed as a double integral in which the integrand contains a determinant of a square matrix of dimension \((n-2)\). Although a simple solution to this double integral seems intractable, for special configurations of \( n = 2, 3, \) and \( 4 \), we obtain closed-form expressions.

I. INTRODUCTION

Consider a set of \( m \geq n \) independent noisy observation vectors \( x_j \in \mathbb{C}^{n \times 1} \) modeled as

\[
x_j = \sqrt{\theta} s_j v + n_j, \quad j = 1, 2, \ldots, m,
\]

where \( \theta \geq 0 \) is a non-random parameter, \( s_j \sim \mathcal{CN}(0, 1) \) is the signal, \( v \in \mathbb{C}^{n \times 1} \) is an unknown non-random vector with \( ||v||^2 = 1 \), \( u_j \sim \mathcal{CN}(0, I_n) \) denotes the additive white Gaussian noise vector which is independent of \( s_j \), \( ||\cdot|| \) denotes the Euclidean norm, and \( I_n \) stands for the \( n \times n \) identity matrix. Given the above observations, one would seek to infer reliable information about the vector \( v \). To this end, it is customary to consider the sample covariance matrix \( S = \frac{1}{m} \sum_{j=1}^{m} x_j x_j^\dagger \) which is the simplest estimator for the population covariance matrix \( \Sigma = I_n + \theta vv^\dagger \), where \( (\cdot)^\dagger \) denotes the conjugate transpose operator. Since \( S \) admits the eigen-decomposition \( S = U \Lambda U^\dagger \), where \( U = (u_1 u_2 \ldots u_n) \in \mathbb{C}^{n \times n} \) is a unitary matrix and \( \Lambda_s = \text{diag}(\lambda_1/m, \lambda_2/m, \ldots, \lambda_n/m) \) is a diagonal matrix with \( 0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n < \infty \) denoting the ordered eigenvalues of \( nS \), the squared modulus of the projection of the true vector \( v \) (i.e., spike) onto \( u_n \) (i.e., \( |v^\dagger u_n|^2 \)) is commonly used to infer information about the latent vector \( v \) [1]–[6]. On the other hand, in the finite dimensional setting, one would expect \( u_1 \) to carry the least amount of information about \( v \) (i.e., via leaked signal energy into the noise subspace).

The utility of the general random quantity \( |v^\dagger u_\ell| \), \( \ell = 1, 2, \ldots, n \), in various application domains has been highlighted in [7]–[42]. In particular, depending on the availability of the knowledge of \( v \), these application domains can be broadly classified into two groups: applications involving detection and estimation. In this respect, to use \( |v^\dagger u_\ell| \) as a test statistic for the detection of signals with certain orientation, one requires to have the full knowledge of \( v \). For instance, in array signal processing, \( v \) is referred to as the nominal array steering vector which is known to the receiver [7]. As such, to detect a desired signal arriving from the direction of the array steering vector, based on the eigenvectors of the sample covariance matrix, the test statistic \( |v^\dagger u_{n}| \) has been used in the literature, see e.g., [7, Section 7.8.2], [8]–[11], and references therein.

Be that as it may, from the estimation perspective, the metric \( |v^\dagger u_{n}| \) has been frequently used to measure the closeness of the population eigenvector \( v \) to the sample eigenvector \( u_{n} \) (or a measure of informativeness of \( u_{n} \) with respect to the latent vector \( v \) [30]) when the unknown \( v \) is approximated by \( u_{n} \) in various studies, see e.g., [3], [12], [13], [17], [28]–[33], [35], [37]–[42], and references therein. For instance, a concrete example in this respect is the principal component analysis (PCA) in which the eigenvectors of the unknown
The population covariance matrix is approximated by the eigenvectors of the sample covariance matrix (i.e., empirical principal components). This metric has further been used in the covariance estimation based on the optimal shrinkage of the eigenvalues of the sample covariance matrix in the high dimensional setting when the unobserved population covariance matrix assumes the spiked structure [16], [20], [23], [31], [36]. In particular, in the spiked population covariance setting, optimal shrinkage functions, which depend on the former metric, corresponding to various orthogonally invariant loss functions, such as the operator loss, Frobenius loss, entropy loss, Fréchet loss, and Stein’s loss, have been derived in [19]. Recently, Yang et al. [17] have used this metric to assess the performance of doing PCA after taking various random projections (a.k.a. sketches) of the observed noisy data having spiked population covariance structure in the high dimensional setting. A more recent extension of these concepts can be found in [18]. Various versions of such combined-algorithms have been proposed in the early works of [43]–[45]. It is worth mentioning that the random projection based techniques have been studied related to certain problems arising in linear regression [46]–[48], ridge regression [47], [49], classification [50], and convex optimization [51]–[53]. Asymptotically optimal bias corrections for sample eigenvalues in certain covariance estimation problems have been evaluated in [27] using the metrics $|v^\dagger u_\ell|$, $\ell = 1, 2, \ldots, n$, and more general extensions of these results can be found in [22]. The application of $|v^\dagger u_\ell|$ in matrix denoising, again in the high dimensional setting, can be found in [23], [29], [31]–[35]. Moreover, the correlation between the sample and population eigenvectors in the high dimensional setting has been instrumental in certain radar/array signal processing applications, see e.g., [37]–[42] and references therein.

It is common to estimate the unobservable population spike $v$ (i.e., the first population PC) by its empirical counterpart $u_n$ (i.e., the leading eigenvector of the sample covariance matrix or the first sample PC). Since $v$ and $u_n$ span a plane, it is trivial that some signal energy has been leaked into the noise subspace. Therefore, the lower sample eigenvectors (i.e., $u_{n-1}, \ldots, u_1$) are informative about $v$ as well. For instance, a systematic asymptotic analysis of this phenomena based on the metric $|v^\dagger u_\ell|$, $\ell = 1, 2, \ldots, n$, has used information-theoretic and random matrix tools in [21]. To shed some light into this issue, noting that $u_\ell$, $\ell = 1, 2, \ldots, n$, forms an orthonormal basis, let us write

$$v = \sum_{\ell=1}^{n} c_\ell u_\ell$$

from which we obtain $c_\ell = u_\ell^\dagger v$. Therefore, the fraction of signal energy contained in the signal subspace (i.e., first sample PC) is $|u_n^\dagger v|^2$, whereas the noise subspace (i.e., other sample PCs) carries $1 - |u_n^\dagger v|^2$ of the total energy. Of $1 - |u_n^\dagger v|^2$, the worst eigenmode (i.e., last sample PC) carries $|u_1^\dagger v|^2$ of the total signal energy. Clearly, the distributions of the random quantities $|u_n^\dagger v|^2$ and $|u_1^\dagger v|^2$ can shed some light into how the spiked energy is statistically distributed between the signal and the noise subspaces in non-asymptotic setting. These facts further highlight the utility of the distributions of $|u_n^\dagger v|^2$ and $|u_1^\dagger v|^2$ in the presence of an unknown spike $v$.

Since the high dimensional statistical characteristics of $|v^\dagger u_\ell|^2$, for $mS \sim CN_n(m, I_n + \theta vv^\dagger)$ with $\ell = 1, n$, have been analyzed extensively in the literature, see e.g., [1]–[6] and references therein, in this study, we mainly focus on the finite dimensional distributions of these random quantities when the population spike $v$ is unknown.

Real and complex Wishart matrices play a central role in various scientific disciplines including multivariate analysis and high dimensional statistics [54]–[56], random matrix theory [57]–[63], physics [64], [65], and finance [66] with numerous engineering applications [67], particularly in signal processing and communications [68]–[71]. In this respect, of prominent interest are the eigenvalues and eigenvectors (more generally spectral projectors) of a Wishart matrix, especially with a certain nontrivial covariance structure. Among various population covariance structures, the spiked model introduced by Johnstone [76] has been widely used in the literature to analyze the effects of having a few dominant trends or correlations in the covariance matrix. Capitalizing on this model for a single dominant correlation (i.e., $\Sigma = I_n + \theta vv^\dagger$ or the rank-one perturbation of the identity), in the high dimensional setting, as $m, n \rightarrow \infty$ such that $n/m \rightarrow \gamma \in (0, 1]$, Baik et al. [77] proved that if $\theta > \sqrt{\gamma}$ (i.e., the supercritical regime) then the maximum eigenvalue of $S$ converges to a Gaussian distribution, whereas it converges to a Tracy-Widom distribution [78], [79] if $\theta < \sqrt{\gamma}$ (i.e., the subcritical regime). To be specific, in the supercritical regime, the population spike generates the top sample eigenvalue of $S$ outside the bulk, which follows the Marchenko-Pastur or “quarter-circle” law [80], supported on a single interval $[1 - (1 - \sqrt{\gamma})^2, 1 + (1 + \sqrt{\gamma})^2]$. Moreover, the top sample eigenvalue converges almost surely to a limiting position $(1 + \theta) (1 + \frac{1}{2})$ outside the upper edge of the bulk, thereby demonstrating an upward top sample eigenvalue bias. In contrast, in the subcritical regime, the top sample eigenvalue converges almost surely to the upper support of the bulk (i.e., $(1 + \sqrt{\gamma})^2$). This phenomenon is commonly known as the Baik, Ben Arous, Péché (BBP) phase transition because of their seminal contribution in [77]. Subsequently, this analysis has been extended to various other random matrix ensembles, see e.g., [1], [2], [83]–[88] and references therein. These results reveal that, in the high dimensional setting, the maximum eigenvalue of the sample covariance matrix can be used to detect a supercritical spike, whereas it cannot be used to detect a subcritical spike. However, this high dimensional behavior is not necessarily true in the finite dimensional setting.

These spikes arise in various practical settings in different scientific disciplines. For instance, they correspond to the first few dominant factors in factor models arising in financial economics [72], [73], the number of clusters in gene expression data [74], and the number of signals in detection and estimation [68], [75].

The signal processing analogy of this phenomenon is known as the “subspace swap” [16], [81], [82].
Despite its utility, the high dimensional characteristics of the eigenvectors of Wishart matrices with spiked covariance matrices have received less attention in the early literature with notable exceptions in [1], [2], [28], [89]. In particular, it was shown in [2] that, in the high dimensional setting as outlined above, $|v^\dagger u_n|^2$ converges almost surely to $\frac{1+\gamma}{2}$ when $\theta > \sqrt{\gamma}$, whereas it converges almost surely to 0 when $\theta < \sqrt{\gamma}$. This phase transition result reveals that no information about a subcritical spike can be inferred from the leading sample eigenvector. In a sharp contrast to this observation, a key recent result in [3] has established that a subcritical spike very close to the critical threshold can even cause $u_n$ to have a bias of small order towards $v$. Capitalizing on this fact, the authors in [3] have concluded that a subcritical spike can be observed using the dominant eigenvector of the sample covariance matrix. The asymptotic behavior of the extreme eigenvectors of not necessarily the Wishart type matrices has been studied on the level of the first order limit in [2], [90],[92]. A comprehensive study on the eigenvector behavior of the sample covariance matrix in the full subcritical and supercritical regimes can be found in [3]. In particular, as shown in [3], the eigenvector distribution is similar to (up to appropriate scaling) that of the bulk and edge regimes of Wigner matrices without spikes; see e.g., [93]–[97] and references therein. Moreover, in the subcritical regime, the limiting distribution of the square of the projected maximum eigenvector component (after proper scaling) is given by a chi-squared distribution, which implies the asymptotic Gaussianity of the eigenvector components [3]. Notwithstanding that the result was established for sample covariance matrix only in [3], it can well be extended to deformed Wigner random matrices without essential difference [88]. In the supercritical regime, the fluctuation of the eigenvectors was recently studied in [5], [34], [98], [99] for various other random matrix models. It is noteworthy that the aforementioned studies are restricted either to subcritical or to supercritical regime only, thereby leaving the critical regime unattended. To fill in this gap, recently, Bao and Wang [88] have focused on the critical regime of the BBP phase transition and established the distribution of the eigenvectors associated with the leading eigenvalues, however, for the unitary Gaussian ensemble with spiked external source only. In a sharp contrast, the behavior of the leading eigenvector, for $\gamma \to \infty$, has been analyzed in [101]. A different kind of asymptotic analysis based on small noise perturbation approach [102] is used in [103] to derive accurate stochastic approximations to $|v^\dagger u_n|^2$ for various complex Wishart matrices.

Whereas the above technical results assume either diverging matrix dimensions (i.e., $m, n \to \infty$) or large signal-to-noise ratio (i.e., small noise power), these quantities assume relatively small values in many practical applications. In this respect, the above mentioned results pertaining to the sample eigenvectors may provide a poor approximation to the true behavior of the population eigenvectors. For instance, as depicted in [68, Fig. 9.5], the convergence rate of the empirical average of $|v^\dagger u_n|^2$ is very slow below the phase transition. Therefore, despite the fact that, below the phase transition, $u_n$ is asymptotically orthogonal to the population spike $v$ (i.e., uninformative first order behavior of $u_n$ with respect to $v$), there exists a large range of values of $m$ and $n$ for which this orthogonality result does not hold. Against this backdrop, in the finite dimensional setting, it is plausible that $u_1$ is also informative about the spike; no matter how small the amount of information is. Having motivated with these observations, in this paper, we characterize the finite dimensional distributions of the leading (i.e., $u_n$) and least eigenvectors (i.e., $u_1$) of the matrix $mS = W \sim CV_n(m, I_n + \theta vv^\dagger)$ via the distributions of the squared modulus of the eigen-projectors $|v^\dagger u_n|^2$ and $|v^\dagger u_1|^2$. Since a closed-form analytical probability density function (p.d.f.) for the joint density of the eigenvectors of $W$ seems intractable, here we adopt a moment generating function (m.g.f.) based approach which requires the joint density of the eigenvalues and eigenvectors of $W$ instead. The resultant matrix integral over the unitary manifold is further simplified using a contour integral approach due to [104]. This key step transformed our problem into an equivalent form involving only the eigenvalues of $W$ which is amenable to further analysis.

In particular, we leverage the powerful contour integral representation of unitary integrals [104] and orthogonal polynomial techniques developed in [64] to derive a closed-form expression for the p.d.f. of $|v^\dagger u_1|^2$ which is valid for arbitrary $m, n$ and $\theta$. This result further indicates that the least eigenvector contains a certain amount of information about the spike. The resultant p.d.f. expression consists of a determinant of a square matrix whose dimensions depend on the relative difference between $m$ and $n$ (i.e., $m - n$). This key feature further facilitates the asymptotic analysis of $|v^\dagger u_1|^2$ in the regime in which $m, n \to \infty$ such that $m - n$ is fixed.\footnote{This is also known as the microscopic limit in the literature of theoretical physics [65], [105]–[107].} It turns out that, in this particular regime, $|v^\dagger u_1|^2$ scales on the order $1/n$; therefore, $n|v^\dagger u_1|^2$ converges in distribution to $\frac{\chi_2^2}{2(1 + \theta)}$, with $\chi_2^2$ denoting a chi-squared random variable with two degrees of freedom. This simple stochastic convergence result reveals that the least eigenvector contains information about the spike in this particular asymptotic domain irrespective of the value of $\theta$. Moreover, our numerical experiments indicate that this stochastic convergence result compares favourably with finite values of $m$ and $n$. Apart from these outcomes, we also derive an exact finite dimensional p.d.f. expression for $|v^\dagger u_n|^2$. This key p.d.f. result is expressed in terms of a double integral in which the integrand contains a determinant of a square matrix of size $n - 2$. Although, an analytical closed-form solution to this integral seems intractable for general $n$, we have derived closed-form expressions for the special configurations of $n = 2, 3,$ and $n = 4$. Nevertheless, the double integral form of the p.d.f. can be evaluated numerically for arbitrary $m, n$, and $\theta$. To further illustrate the utility of the m.g.f. machinery developed in this manuscript beyond the
The eigenvalues of the argument matrix, \( \kappa = (k_1, \ldots, k_n) \), with \( k_i \)'s being non-negative integers, is a partition of \( k \) such that \( k_1 \geq \cdots \geq k_n \geq 0 \) and \( \sum_{i=1}^{n} k_i = k \). Moreover, the Harish-Chandra-Itzykson-Zuber integral [117], [119], [120] can be written as

\[
o \tilde{F}_0 (V, T) = \int dU \text{etr} \left( VUTU^\dagger \right) \quad (4)
\]

where \( dU \) is the invariant measure on the unitary group \( U_n \) normalized to make the total measure unity (i.e., \( \int_{U_n} dU = 1 \)). If at least one of the argument matrices assumes a rank-one structure, the hypergeometric function of two matrix arguments simplifies as shown in the following theorem which is due to [104].

**Theorem 1:** Let \( V = vv^\dagger \) with \( v \in \mathbb{C}^{n \times 1} \) and \( \|v\| = 1 \). Then we have [104]

\[
o \tilde{F}_0 (vv^\dagger, T) = \frac{(n-1)!}{2\pi i \prod_j (n - \tau_j)} \int d\omega \quad (5)
\]

where \( \tau_1, \tau_2, \ldots, \tau_n \) are the eigenvalues of the \( T \), the contour \( \mathcal{C} \) is large enough so that all \( \tau_j \)'s are in its interior, and \( i = \sqrt{-1} \).

The following statistical characterization of the eigen-decomposition of \( W \sim CW(n, \Sigma) \) is also useful in the sequel.

**Theorem 2:** Let the Hermitian positive definite matrix \( W \) assume the eigen-decomposition \( W = U\Lambda U^\dagger \), where \( \Lambda = \text{diag} (\lambda_1, \lambda_2, \ldots, \lambda_n) \) with \( 0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n < \infty \) denoting the ordered eigenvalues of \( W \) and \( U = (u_1, u_2, \ldots, u_n) \) contains the corresponding eigenvectors. Then the Jacobian of matrix transformation can be written as [117]

\[
dW = 2^n(n-1)\prod_{i<k} \Delta^2 (\lambda_i) d\Lambda dU \quad (6)
\]

**Definition 3:** For \( \rho > -1 \), the generalized Laguerre polynomial of degree \( M \), \( L_M^{(\rho)} (z) \), is given by [123]

\[
L_M^{(\rho)} (z) = \frac{(-1)^M M! \sum_{j=0}^{M} \binom{-\rho}{j} (\rho + 1)_j z^j}{M!} \quad (7)
\]

with its \( k \)th derivative satisfying

\[
\frac{d^k}{dz^k} L_M^{(\rho)} (z) = (-1)^k L_M^{(\rho+k)} (z), \quad (8)
\]

The exact algebraic definition of the zonal polynomial is tacitly avoided here, since it is not required in the subsequent analysis. More details of the zonal polynomials can be found in [117], [118].

A more generalized contour integral representation in this respect can be found in [121], [122]. However, the above form is sufficient for our work in this paper.

The uniqueness of this representation and related implications are discussed in [57].

---

**II. PRELIMINARIES**

Here we present some fundamental results pertaining to the finite dimensional representation of correlated Wishart matrices, related unitary integrals and generalized Laguerre polynomials which are instrumental in our main derivations.

**Definition 1:** Let \( X \in \mathbb{C}^{n \times m} (m \geq n) \) be distributed as \( \mathcal{CN}_{n,m} (0, \Sigma \otimes I_m) \). Then the matrix \( W = XX^\dagger \) is said to follow a complex correlated Wishart distribution \( \sim CW(n, \Sigma) \) with p.d.f. [117]

\[
f(W)dW = \frac{\det^{m-n}(W)}{\Gamma_n(m) \det^m(\Sigma)} \text{etr} \left(-\Sigma^{-1}W\right) dW \quad (3)
\]

where \( dW \) is Lebesgue measure on the space of \( n \times n \) Hermitian matrices, identifiable unambiguously with \( \mathbb{R}^{n^2} (= \mathbb{R}^n \times \mathbb{C}^{n(n-1)/2}) \), defined by taking or-on-above diagonal entries as coordinates [57], \( \Gamma_n(a) = \pi^{\frac{n(n-1)}{2}} \prod_{j=1}^{n} \Gamma(a-j+1) \)

represents the complex multivariate gamma function with \( \Gamma(\cdot) \) denoting the classical gamma function, \( \det(\cdot) \) is the determinant of a square matrix, and \( \text{etr}(\cdot) \equiv e^{\text{tr}(\cdot)} \) with \( \text{tr}(\cdot) \) denoting the trace of a square matrix.

**Definition 2:** Let \( V \in \mathbb{C}^{n \times n} \) and \( T \in \mathbb{C}^{n \times n} \) be two Hermitian non-negative definite matrices. Then the hypergeometric function of two matrix arguments is defined as [117]

\[
o \tilde{F}_0 (V, T) = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{\kappa} \frac{C_k(V)C_k(T)}{C_k(1_n)} \quad (4)
\]

where \( C_k(\cdot) \) is the complex zonal polynomial which is a symmetric, homogeneous polynomial of degree \( k \) in the
where \((a)_j = a(a+1) \ldots (a+j-1)\) with \((a)_0 = 1\) denotes the Pochhammer symbol.

It is also noteworthy that the Pochhammer symbol admits, for \(M \in \mathbb{Z}_+,\)

\[
(-M)_j = \begin{cases} \frac{(-1)^j M!}{(M-j)!} & \text{if } 0 \leq j \leq M \\ 0 & \text{if } j > M. \end{cases} \tag{9}
\]

Finally, we use the following compact notation to represent the determinants of \(N \times N\) block matrices:

\[
\det \left[ a_{i,j} \ b_{i,j-1} \ c_{i,k-3} \right]_{i=1, \ldots, N}^{j=2, \ldots, N} = \begin{vmatrix} a_1 & b_{1,1} & \cdots & b_{1,N-1} \\
 & a_2 & b_{2,1} & \cdots & b_{2,N-1} \\
 & & \ddots & \ddots & \vdots \\
 & & & a_N & b_{N,1} & \cdots & b_{N,N-1} \end{vmatrix},
\]

\[
\det \left[ a_{i,j} \ b_{i,j-1} \ c_{i,k-3} \right]_{i=1, \ldots, N}^{j=2, \ldots, N} = \begin{vmatrix} a_1 & b_{1,1} & b_{1,2} & c_{1,1} & \cdots & c_{1,N-1} \\
 & a_2 & b_{2,1} & b_{2,2} & c_{2,1} & \cdots & c_{2,N-1} \\
 & & \ddots & \ddots & \ddots & \vdots & \vdots \\
 & & & a_N & b_{N,1} & b_{N,2} & c_{N,1} & \cdots & c_{N,N-1} \end{vmatrix},
\]

and \(\det [a_{i,j}]_{i=1, \ldots, N}^{j=1, \ldots, N}\) denotes the determinant of an \(N \times N\) square matrix with its \((i,j)\)th element given by \(a_{i,j}\).

### III. Finite Dimensional Results for the Distributions of the Eigenvectors

Here we adopt an m.g.f. based approach to determine the p.d.f.s of the random variables \(Z^{(n)}_1, Z^{(n)}_2,\) and \(Z^{(n)}_3\). To this end, by definition, the m.g.f. can be written as

\[
\mathcal{M}_{Z^{(n)}_i}(s) = \mathbb{E} \left\{ e^{-s \|v^\dagger U e_i^T \|^2} \right\}, \quad \ell = 1, 2, n \tag{10}
\]

where \(\mathbb{E} \{ \cdot \} \) denotes the mathematical expectation evaluated with respect to the density of \(u_i\). To facilitate further analysis, noting that \(\|v^\dagger U e_i^T \|^2 = tr \left( v v^\dagger U e_i e_i^T U^\dagger \right)\) with \(v_i\) denoting the \(i\)th column of the \(n \times n\) identity matrix, we may rewrite the m.g.f. as

\[
\mathcal{M}_{Z^{(n)}_i}(s) = \mathbb{E} \left\{ \text{etr} \left( -sv v^\dagger U e_i e_i^T U^\dagger \right) \right\} \tag{11}
\]

where the expectation is now taken with respect to the joint distribution of \(U\) and \(\Lambda\) and \((\cdot)^T\) denotes the transpose operator. Following (6), the above expectation assumes

\[
\mathcal{M}_{Z^{(n)}_i}(s) = \frac{\pi^{n(n-1)}}{2} \int_{\mathcal{R}} \Delta^2_n(\lambda) \phi_{\lambda} \left\{ \text{etr} \left( -sv v^\dagger U e_i e_i^T U^\dagger \right) \right\} \left( \frac{1}{2\pi i} \int_{\mathcal{I}} \frac{e^{\omega} \pi^{n-1}}{c} \left( s + \omega - \beta \lambda_1 \right) \prod_{j=2}^{n} (\omega - \beta \lambda_j) \right) d\lambda \tag{12}
\]

where \(\mathcal{R} = \{ 0 < \lambda_1 < \lambda_2 < \ldots < \lambda_n < \infty \}\). Since, for a single-spiked Wishart matrix, \(\Sigma = I_n + \theta vv^\dagger\) with \(\Sigma^{-1} = I_n - \frac{\theta}{\theta + 1} vv^\dagger\), we use (3) to further simplify the above matrix integral as

\[
\mathcal{M}_{Z^{(n)}_i}(s) = K_{n,\alpha} (1 - \beta)^{n+\alpha} \times \int_{\mathcal{R}} \Delta^2_n(\lambda) \prod_{j=1}^{n} \lambda_j^\alpha e^{-\lambda_j} \Psi_\ell(\lambda, s) d\lambda \tag{13}
\]

where \(K_{n,\alpha} = 1/\prod_{j=1}^{n} (n-j)!/(n+\alpha-j)!\), \(\alpha = m - n\), and

\[
\Psi_\ell(\lambda, s) = \int_{\mathcal{U}_n} \text{etr} \left\{ vv^\dagger U (\beta \Lambda - se_i e_i^T) U^\dagger \right\} dU \tag{14}
\]

with \(\beta = \theta/(1 + \theta)\). In view of the fact that the matrix \(\beta \Lambda - se_i e_i^T\) assumes three different diagonal structures depending on the value of \(\ell\) as

\[
\beta \Lambda - se_i e_i^T
\]

\[
= \begin{cases} \text{diag} (\beta \lambda_1 - s, \beta \lambda_2 - s, \ldots, \beta \lambda_n) & \text{for } \ell = 1 \\
\text{diag} (\beta \lambda_1, \beta \lambda_2 - s, \ldots, \beta \lambda_n) & \text{for } \ell = 2 \\
\text{diag} (\beta \lambda_1, \beta \lambda_2, \ldots, \beta \lambda_n - s) & \text{for } \ell = n, 
\end{cases}
\]

we find it convenient to consider the three cases separately from this point onward.

### A. The P.D.F. of \(Z^{(n)}_1\)

For \(\ell = 1\), (14) specializes to

\[
\Psi_1(\lambda, s) = \int_{\mathcal{U}_n} \text{etr} \left\{ vv^\dagger U (\beta \Lambda - se_1 e_1^T) U^\dagger \right\} dU, \tag{15}
\]

which can be simplified using (4) and (5) to yield

\[
\Psi_1(\lambda, s) = \frac{(n-1)!}{2\pi i} \int_{\mathcal{I}} \frac{e^{\omega} \pi^{n-1}}{c} \left( s + \omega - \beta \lambda_1 \right) \prod_{j=2}^{n} (\omega - \beta \lambda_j) d\omega.
\]

This in turn enables us to express (13) as

\[
\mathcal{M}_{Z^{(n)}_1}(s) = C_{n,\alpha}^{\beta} \int_{\mathcal{R}} \Delta^2_n(\lambda) \prod_{j=1}^{n} \lambda_j^\alpha e^{-\lambda_j} \times \text{etr} \left\{ vv^\dagger U (\beta \Lambda - se_1 e_1^T) U^\dagger \right\} \left( \frac{1}{2\pi i} \int_{\mathcal{I}} \frac{e^{\omega} \pi^{n-1}}{c} \left( s + \omega - \beta \lambda_1 \right) \prod_{j=2}^{n} (\omega - \beta \lambda_j) \right) d\lambda \tag{12}
\]

where \(C_{n,\alpha}^{\beta} = (n-1)!K_{n,\alpha}(1 - \beta)^{n+\alpha}\). Now we take the inverse Laplace transform of both sides to yield

\[
f^{(a)}_{Z^{(n)}_1}(z) = C_{n,\alpha}^{\beta} \int_{\mathcal{R}} e^{\beta \lambda_1 z} \Delta^2_n(\lambda) \prod_{j=1}^{n} \lambda_j^\alpha e^{-\lambda_j} \times \text{etr} \left\{ vv^\dagger U (\beta \Lambda - se_1 e_1^T) U^\dagger \right\} \left( \frac{1}{2\pi i} \int_{\mathcal{I}} \frac{e^{(1-z)\omega} \pi^{n-1}}{c} \left( s + \omega - \beta \lambda_1 \right) \prod_{j=2}^{n} (\omega - \beta \lambda_j) \right) d\lambda \tag{16}
\]
in which the innermost contour integral can be evaluated, for \( n \geq 3 \), with the help of the Cauchy’s residue theorem to obtain

\[
f^{(n)}_{Z^{(')}(z)} = \frac{C^{(n)}_{\beta^{m-2}}} {\beta^{n-2}} \int_{\mathbb{R}} e^{(\beta z) \lambda^2} (x) \prod_{i=1}^{n} x_i^{(1-\beta z) \lambda_i} d\lambda.
\]

Now let us rewrite the above \( n \)-fold integral, keeping the integration with respect to \( \lambda_1 \) last, as

\[
f^{(n)}_{Z^{(')}(z)} = \frac{C^{(n)}_{\beta^{m-2}}} {\beta^{n-2}} \int_{0}^{\infty} \lambda_1^\alpha x^{(1-\beta z) \lambda_1} \Phi(\lambda_1, z) d\lambda_1
\]

where

\[
\Phi(\lambda_1, z) = \int_{S} \sum_{k=2}^{n} \frac{e^{(1-z) \lambda_k}} {\prod_{i=1}^{k} (\lambda_k - \lambda_i)} \Delta^2_{n-1}(\lambda)
\]

\[
\times \prod_{j=2}^{n} \lambda_j^\beta x^{-(1-\beta z) \lambda_j} d\lambda_j
\]

in which we have employed the decomposition \( \Delta^2_{n}(\lambda) = \prod_{j=2}^{n} (\lambda_j - \lambda_1)^2 \Delta^2_{n-1}(\lambda) \) and \( S = \{\lambda_1 < \lambda_2 < \ldots < \lambda_n < \infty\} \). For convenience, we relabel the variables as \( \lambda_1 = x \) and \( x_j-1 = \lambda_j, j = 2, 3, \ldots, n \), to arrive at

\[
f^{(n)}_{Z^{(')}(z)} = \frac{C^{(n)}_{\beta^{m-2}}} {\beta^{n-2}} \int_{0}^{\infty} x^\alpha e^{(1-\beta z) x} \Phi(x, z) dx
\]

where

\[
\Phi(x, z) = \int_{S_x} \sum_{k=2}^{n-1} \frac{e^{(1-z) \beta x_k}} {\prod_{i=1}^{k} (x_k - x_i)} \Delta^2_{n-1}(x)
\]

\[
\times \prod_{j=1}^{n-1} x_j^\alpha e^{-(1-z) x} (x_j - x)^2 dx_j
\]

with \( S_x = \{x < x_1 < x_2 < \ldots < x_{n-1} < \infty\} \). Since the integrand in the above \( (n-1) \)-fold integral is symmetric in \( x_1, x_2, \ldots, x_{n-1} \), we may remove the ordered region of integration to obtain

\[
\Phi(x, z) = \frac{1} {(n-1)!} \int_{(x, \infty)^{n-1}} \sum_{k=1}^{n-1} \frac{e^{(1-z) \beta x_k}} {\prod_{i=1}^{k} (x_k - x_i)} \Delta^2_{n-1}(x)
\]

\[
\times \prod_{j=1}^{n-1} x_j^\alpha e^{-(1-z) x} (x_j - x)^2 dx_j
\]

Consequently, we can observe that the each term in the above summation evaluates to the same amount. Therefore, capitalizing on that observation, we may simplify the above \( (n-1) \)-fold integral to yield

\[
\Phi(x, z) = \frac{1} {(n-2)!} \int_{(x, \infty)^{n-1}} \frac{e^{(1-z) \beta x}} {\prod_{i=2}^{n-1} (x_i - x_{i-1})} \Delta^2_{n-1}(x)
\]

\[
\times \prod_{j=1}^{n-1} x_j^\alpha e^{-(1-z) x} (x_j - x)^2 dx_j.
\]

To facilitate further analysis, we introduce the variable transformations \( y_j = x_j - x, j = 1, 2, \ldots, n-1 \), to the above integral to arrive at

\[
\Phi(x, z) = \frac{e^{-n(1-\beta z)(-\beta z-\beta-1)x}} {(n-2)!} \int_{(0, \infty)^{n-1}} \frac{e^{(1-z) \beta y_j}} {\prod_{i=2}^{n-1} (y_i - y_{i-1})} \Delta^2_{n-1}(y)
\]

\[
\times \prod_{j=1}^{n-1} (y_j + x)^\alpha y_j^\alpha e^{-y_j} dy_j,
\]

from which we obtain by keeping the integration with respect to \( y_1 \) last

\[
\Phi(x, z) = \frac{e^{-n(1-\beta z)(-\beta z-\beta-1)x}} {(n-2)!} \int_{0}^{\infty} e^{-(1-\beta z) y_1^2 (y_1 + x)^\alpha} \times \mathcal{Q}_{n-2}(y_1, x) dy_1
\]

where

\[
\mathcal{Q}_{n-2}(y_1, x)
\]

\[
= \int_{(0, \infty)^{n-2}} \Delta^2_{n-2}(y) \prod_{i=2}^{n-1} (y_i - y_j)^\alpha y_j^\alpha e^{-y_j} dy_j
\]

To further simplify \( \mathcal{Q}_{n-2}(y_1, x) \), noting the decomposition \( \Delta^2_{n-1}(y) = \prod_{j=2}^{n-1} (y_j - y_1)^2 \Delta^2_{n-2}(y) \), we may relabel the variables as \( z_{j-1} = y_j, j = 2, 3, \ldots, n-1 \), with some algebraic manipulation to arrive at

\[
\mathcal{Q}_{n-2}(y_1, x)
\]

\[
= \int_{(0, \infty)^{n-2}} \Delta^2_{n-2}(z) \prod_{j=1}^{n-2} (y_1 - z_j)(z_j + x)^\alpha z_j^\alpha e^{-z_j} dz_j.
\]

The above multiple integral can be solved following the orthogonal polynomial approach due to Mehta [64], as shown in (89) and (95) of Appendix A, to yield

\[
\mathcal{Q}_{n-2}(y_1, x)
\]

\[
= (-1)^n \tilde{K}_{n-2, \alpha}
\]

\[
\times \det \begin{bmatrix} L_{n+i-3}(y_1) L_{n+i-j-1}^{(j)}(\cdot) \end{bmatrix}_{i=1, \ldots, \alpha+1, j=2, \ldots, \alpha+1}
\]

where

\[
\tilde{K}_{n-2, \alpha} = \prod_{j=1}^{\alpha+1} (n+j-3)! \prod_{j=0}^{n-3} (j+1)! (j+2)! / \prod_{j=0}^{\alpha-1} j!.
\]
It is noteworthy that, when \( \alpha = 0 \), we interpret the above determinant as \( L_{n-2}^{(2)}(y_1) \), since the \( (\alpha + 1) \times (\alpha + 1) \) matrix degenerates to a scalar. Moreover, here an empty product is interpreted as unity. Now we substitute \( Q_{n-2}(y_1, x) \) in (21) into (19) with some algebraic manipulation to obtain
\[
\Phi(x, z) = \frac{(-1)^nK_{n-2, z, \beta-1}}{(n-2)!} e^{-(n+\beta z-\beta-1)x} \times \int_0^\infty e^{-(1-(1-z)\beta)ny_1^2} \times \det \begin{bmatrix}
L_{n+i-3}^{(2)}(y_1) & L_{n+1-i-j-1}^{(j)}(-x) \\
1 & 1
\end{bmatrix}_{i=1,\ldots,\alpha+1} \times 1_{j=2,\ldots,\alpha+1} dy_1.
\]
(22)

Since only the first column of the determinant depends on \( y_1 \), we can conveniently rewrite the above integral as (23), shown at the bottom of the page. Now we use [124, Eq. 7.414.8] to evaluate the lower integral and substitute back the resultant \( \Phi(x, z) \) into (17) along with some algebraic manipulations to arrive at
\[
\begin{align*}
\zeta_i(z, \beta) &= \frac{(n+i-1)! (-\beta)^{i-1} (1-z)^{n+i-3}}{(n+i-3)! [1-\beta (1-z)]^{n+i+1}}.
\end{align*}
\]
(25)

To facilitate further analysis, noting that the columns denoted by \( j = 2, 3, \ldots, \alpha + 1 \), depend only on \( x \), we use (7) to rewrite the Laguerre polynomials in the determinant with some algebraic manipulation as in (26), shown at the bottom of the page. Further manipulation in this form is highly undesirable due to the dependence of the summation upper limits on \( i \) and \( j \). To circumvent this difficulty, in view of making the summation upper limits depend only on \( j \), we use the decomposition
\[
\begin{align*}
&\det \begin{bmatrix}
\zeta_i(z, \beta) & L_{n+i-j-1}^{(j)}(-x) \\
1 & 1
\end{bmatrix}_{i=1,\ldots,\alpha+1} \\
&\quad = \frac{(n+\alpha-1)! (n+\alpha)!}{(n-1)!} \times \sum_{k_2=0}^{n+\alpha-2} \sum_{k_3=0}^{n+\alpha-3} \cdots \sum_{k_{\alpha+1}=0}^{n+\alpha-1} \frac{(n+\alpha-j)!}{(j+k_j)!} \prod_{j=2}^{\alpha+1} k_j \times \det \begin{bmatrix}
\zeta_i(z, \beta) & 1 \\
(n+i-1)! & \Gamma(n+i-j-k_j) \\
1 & 1
\end{bmatrix}_{i=1,\ldots,\alpha+1} \\
&\quad = \frac{(-n-\alpha+j)_k}{(n-i-j+k_j)_k} \frac{(n-i+j+1)_k}{(n-\alpha+j)_k} \\
&\quad = \frac{(-n-\alpha+j)_k}{(n-i-j+k_j)_k} \frac{(n-i+j+1)_k}{(n-\alpha+j)_k}.
\end{align*}
\]
(26)

Finally, we use (28) and (25) in (24) and perform the integration with respect to \( x \) followed by shifting the indices from \( i, j \) to \( i-1, j-1 \) with some algebraic manipulation to obtain p.d.f. of \( Z_1^{(n)} \), for \( n \geq 3 \), as shown in Theorem 3.

\[
\begin{align*}
\Phi(x, z) &= \frac{(-1)^nK_{n-2, z, \beta-1}}{(n-2)!} e^{-(n+\beta z-\beta-1)x} \times \int_0^\infty e^{-(1-(1-z)\beta)ny_1^2} \times \det \begin{bmatrix}
L_{n+i-3}^{(2)}(y_1) & L_{n+1-i-j-1}^{(j)}(-x) \\
1 & 1
\end{bmatrix}_{i=1,\ldots,\alpha+1} \times 1_{j=2,\ldots,\alpha+1} dy_1.
\end{align*}
\]
(23)
For $n = 2$ case, in view of the Cauchy's residue theorem, (16) specializes to
\[
\int_{0<\lambda_1<\lambda_2<\infty} e^{\beta_2 x_2} \Delta_2^2(\lambda)e^{(1-z)\beta_2 x_2} \prod_{j=1}^{2} \lambda_j^e - \lambda_i d\lambda_1 d\lambda_2,
\]
from which we obtain after interchanging the order of integration followed by some algebraic manipulation
\[
f_{Z_1}^{(\alpha)}(z) = C_{2,\alpha}^1 \int_0^{\infty} \lambda_1^e - (1 - \beta_2 x_2) \lambda_1 d\lambda_2
\]
\[
= \sum \frac{(k + 2)!}{(\alpha - k)! \lambda_1^{k+1}} \left[ 1 - 1 \right]^{k+3} (1 - \beta(1 - z))^{-1},
\]
which upon substituting into (29) followed by integration with respect to $\lambda_2$ with some algebraic manipulation gives the p.d.f. of $Z_1^{(\alpha)}$, for $n = 2$, as shown in Theorem 3.

**Theorem 3:** Let $\mathbf{W} \sim \mathcal{C} \mathcal{W}_n (\mathbf{m}, \mathbf{I}_n + \theta \mathbf{v} \mathbf{v}^t)$ with $||\mathbf{v}|| = 1$ and $\theta \geq 0$. Let $v_i$ be the eigenvector corresponding to the smallest eigenvalue of $\mathbf{W}$. Then the p.d.f. of $Z_1^{(\alpha)} = |v^t |^2 \in (0, 1)$ is given by
\[
f_{Z_1}^{(\alpha)}(z) = (1 - \beta)^{n+\alpha} (1 - z)^{n-2} 1 - 1 + \lambda_1^{(\alpha, \beta)} (z), n \geq 3
\]
\[
f_{Z_1}^{(\alpha)}(z) = \frac{(1 - \beta)^{2+\alpha}}{(\alpha + 1)!} \frac{(1 - z)^{n-2} 1 - 1 + \lambda_1^{(\alpha, \beta)} (z)}{(1 - \beta(1 - z))^{k+3}}
\]
where $\lambda_1^{(\alpha, \beta)} (z)$ is given by (32), shown at the bottom of the page, with $\alpha_i, j = 1$ and $\beta = \theta / (\theta + 1)$.

The above p.d.f. expression depends on $\theta$ via $\beta$ as we clearly see. This observation further reveals that, in the finite dimensional setting, the least eigenvector also contains a certain amount of information about the spike $v$. As we shall see below in Corollary 2, the least eigenvector retains this ability even in a particular asymptotic domain as well.

It is noteworthy that, since the number of nested summations in Theorem 3 depends only on $\alpha$, it provides an efficient way of evaluating the p.d.f. of $Z_1^{(\alpha)}$, particularly for small values of $\alpha$. As such, for some small values of $\alpha$, (30) admits the following simple forms.

**Corollary 1:** The exact p.d.f.s of $Z_1^{(\alpha)}$ corresponding to $\alpha = 0$ and $\alpha = 1$, for $n \geq 3$, are given respectively by
\[
f_{Z_1}^{(0)}(z) = \frac{n(n - 1)(1 - \beta)^n(1 - z)^n-2}{(n - \beta)(1 - (1 - z))^{n+1}}
\]
\[
f_{Z_1}^{(1)}(z) = \frac{n(n^2 - 1)(1 - \beta)^n(1 - z)^n-2}{2(n - \beta^2)(1 - (1 - z))^{n+1}}
\]
\[
+ \frac{1}{1 - \beta(1 - z)} \frac{2F_1 \left( -n + 1, 2; 3; -\frac{1}{n - \beta} \right)}
\]
where $2F_1 (a, b; c; z)$ is the Gauss hypergeometric function [125].

**Remark 1:** The p.d.f. corresponding to $\theta = 0$ (i.e., $\beta = 0$) can be obtained from (30) after some tedious algebraic manipulations. Alternatively, for $\beta = 0$, noting that $\zeta_i (z, 0) = 0, i = 2, 3, \ldots, \alpha + 1$, (24) can be simplified to yield
\[
f_{Z_1}^{(\alpha)}(z) = \frac{(n - 1)!}{(n + \alpha - 1)!} (1 - z)^{n-2}
\]
\[
\times \int_0^\infty e^{-nx} x^\alpha \det \left[ L_{n+i+j-1}(x) \right]_{i,j=2,\ldots,\alpha+1} dx
\]
from which we obtain after relabeling the indices $k = i-1$ and $\ell = j-1$.

\[
f_{Z_1}^{(\alpha)}(z) = \frac{(n - 1)!}{(n + \alpha - 1)!} (1 - z)^{n-2}
\]
\[
\times \int_0^\infty e^{-nx} x^\alpha \det \left[ L_{n+k-1}(x) \right]_{k,\ell=1,\ldots,\alpha} dx.
\]
Since the p.d.f. of $\lambda_1$, for $\beta = 0$, assumes the form [126, Eq. 3.8]
\[
f_{\lambda_1}(x) = \frac{n!}{(n + \alpha - 1)!} e^{-nx} x^\alpha
\]
\[
\times \det \left[ L_{n+k-1}(x) \right]_{k,\ell=1,\ldots,\alpha},
\]
where $\lambda_1^{(\alpha, \beta)} (z)$ is given by (32), shown at the bottom of the page, with $\alpha_i, j = 1$ and $\beta = \theta / (\theta + 1)$.

The above p.d.f. expression depends on $\theta$ via $\beta$ as we clearly see. This observation further reveals that, in the finite dimensional setting, the least eigenvector also contains a certain amount of information about the spike $v$. As we shall see below in Corollary 2, the least eigenvector retains this ability even in a particular asymptotic domain as well.
we easily obtain the desired answer $f^{(\alpha)}_{Z_1^{(n)}}(z) = (n - 1)(1 - z)^{n-2}$ as expected.\(^9\)

Figure 1 compares the analytical p.d.f. result for $Z_1^{(n)}$ with simulated data. Analytical curves are generated based on Theorem 3. As can be seen from the figure, our analytical results match with the simulated data, thus validating our theorem. Quantile-Quantile (Q-Q) plots have been provided in Fig. 2 to further compare the analytical and simulated data. Moreover, the effect of $\theta$ on the p.d.f. of $Z_1^{(n)}$ is depicted in Fig. 3, while the corresponding Q-Q plots are given in Fig. 4.

To further illustrate the utility of Theorem 3, we now focus on the asymptotic behavior of scaled $Z_1^{(n)}$ as $m$ and $n$ grow large, but their difference does not. The following corollary establishes the stochastic convergence result in this respect.

**Corollary 2:** As $m, n \to \infty$ such that $m - n$ is fixed (i.e., $\alpha$ is fixed), the scaled random variable $nZ_1^{(n)} = n|v^\dagger u_1|^2$ converges in distribution to $\frac{\chi_2^2}{2(1 + \theta)}$, where $\chi_2^2$ is a chi-squared random variable with two degrees of freedom (i.e., sum of squares of two independent standard normal random variables).

**Proof:** See Appendix B.

**Remark 2:** It is noteworthy that, since the chi-squared random variable $V = \frac{\chi_2^2}{2(1 + \theta)}$ has a continuous c.d.f., we have

$$\lim_{n \to \infty} \Pr \left\{ nZ_1^{(n)} \leq z \right\} = \Pr \{ V \leq z \} = 1 - \exp \left( -(1 + \theta)z \right)$$

in which the convergence is uniform in $z$ [127, Lemma 2.11].

This interesting result reveals that, in this particular asymptotic regime (i.e., $m, n \to \infty$ such that $m - n$ is fixed), irrespective of the value of $\theta$ (i.e., whether $\theta$ is either below or above the phase transition threshold $1$), the eigenvector corresponding to the least eigenvalue is informative with respect to the latent spike $v$. In contrast, as we are well

---

\(^9\)It is well known that, for $\theta = 0$, $U$ is Haar distributed (i.e., uniformly distributed over the unitary manifold). Consequently, some algebraic manipulations will establish the result $f^{(\alpha)}_{Z_1^{(n)}}(z) = (n - 1)(1 - z)^{n-2}$. It is also noteworthy that, since any permutation matrix is orthogonal, the same p.d.f. result holds for a general class of random variables of the form $Z_\ell^{(n)} = |v^\dagger u_\ell|^2$, $\ell = 1, 2, \ldots, n$.\]
Corollary 2, the largest eigenvalue have been proposed in [128]–[130] informative tests based on distribution [1], [12], [77]. To circumvent this ambiguity, the maximum eigenvalue (after proper centering and scaling) stochastically converges to a Tracy-Widom distribution [1], [12], [77]. To circumvent this ambiguity, informative tests based on all the eigenvalues instead of the largest eigenvalue have been proposed in [128]–[130] to detect subcritical spikes. Although not directly related to Corollary 2, \( \chi^2 \) type stochastic convergence results for the spectral projectors of various matrix ensembles have been established in [3], [5], [34], [98], [99].

The advantage of the asymptotic formula presented in Corollary 2 is that it provides a simple expression which compares favorably with finite \( n \) values. To further highlight this salient point, in Fig. 5, we compare the analytical asymptotic c.d.f. of \( V \) with simulated data points corresponding to \( \alpha = 2 \) with \( n = 15, 25 \) and 30 for various values of \( \theta \). The close agreement is clearly apparent from the figure.

Let us now focus on the leading eigenvector (i.e., \( \ell = n \)).

B. The P.D.F. of \( Z_n^{(n)} \)

For \( \ell = n \), (14) specializes to

\[
\Psi_n(\lambda, s) = \int_{U_n} \text{etr} \left\{ vv^\dagger U \left( \beta \Lambda - s e_n e_n^T \right) U^\dagger \right\} dU, \tag{36}
\]

which can be simplified using (4) and (5) to yield

\[
\Psi_n(\lambda, s) = \frac{(n-1)!}{2\pi i} \int_{C} \frac{e^{s \omega}}{(s - \beta \lambda_n) \prod_{j=1}^{n-1} (\omega - \beta \lambda_j)} d\omega.
\]

This in turn enables us to express (13) as

\[
\mathcal{M}_{Z_n^{(n)}}(s) = C_n^{\beta} \int_{R} \Delta_n^2(\lambda) \prod_{j=1}^{n} \lambda_j^\alpha e^{-\lambda_j} \times \left\{ \frac{1}{2\pi i} \int_{C} e^{s \omega} \prod_{j=1}^{n-1} (\omega - \beta \lambda_j) \right\} d\omega \right\} d\Lambda.
\]

Now we take the inverse Laplace transform of both sides to yield

\[
f_{Z_n^{(n)}}(z) = C_n^{\beta} \int_{R} e^{s \lambda_n z} \Delta_n^2(\lambda) \prod_{j=1}^{n} \lambda_j^\alpha e^{-\lambda_j} \times \left\{ \frac{1}{2\pi i} \int_{C} e^{s \omega} \prod_{j=1}^{n-1} (\omega - \beta \lambda_j) \right\} d\omega \right\} d\Lambda
\]

in which the innermost contour integral can be evaluated with the help of the Cauchy’s residue theorem to obtain

\[
f_{Z_n^{(n)}}(z) = C_n^{\beta} \int_{R} e^{s \lambda_n z} \Delta_n^2(\lambda) \prod_{j=1}^{n} \lambda_j^\alpha e^{-\lambda_j} \times \sum_{k=1}^{n-1} \frac{e^{(1-z)\beta \lambda_k}}{\prod_{i=1}^{n-1} (\lambda_k - \lambda_i)} d\Lambda.
\]

Now let us rewrite the above \( n \)-fold integral, keeping the integration with respect to \( \lambda_n \) last, as

\[
f_{Z_n^{(n)}}(z) = C_n^{\beta} \int_{R} e^{s \lambda_n z} \Delta_n^2(\lambda) \prod_{j=1}^{n} \lambda_j^\alpha e^{-\lambda_j} \times \sum_{k=1}^{n-1} \frac{e^{(1-z)\beta \lambda_k}}{\prod_{i=1}^{n-1} (\lambda_k - \lambda_i)} d\Lambda.
\]

where

\[
\Omega(\lambda_n, z) = \int_{I} \sum_{k=1}^{n-1} \frac{e^{(1-z)\beta \lambda_k \Delta_n^2(\lambda)}}{\prod_{i=1}^{n-1} (\lambda_k - \lambda_i)^2 \lambda_j^\alpha e^{-\lambda_j} (\lambda_n - \lambda_j)^2 d\lambda_j
\]

in which we have made use of the decomposition \( \Delta_n^2(\lambda) = \prod_{j=1}^{n-1} (\lambda_n - \lambda_j)^2 \Delta_{n-1}^2(\lambda) \) and \( I = \{ 0 < \lambda_1 < \lambda_2 < \ldots < \lambda_n \} \). For convenience, we relabel the variables as \( \lambda_n = x \) and \( x_j = \lambda_j, j = 1, 2, \ldots, n-1 \), to arrive at

\[
f_{Z_n^{(n)}}(z) = C_n^{\beta} \int_{R} \int_{I} x^\alpha e^{-(1-z)x} \Omega(x, z) d\lambda_n d\Lambda.
\]
Fig. 4. Quantile-Quantile plots of simulated data of $Z_{1}^{(n)}$ drawn from $f_{x_{1}^{(n)}}(z)$ for different values of $\theta$ with $\alpha = 2$ and $n = 3$.

where

$$\Omega(x, z) = \int_{x_{1}}^{x_{n-1}} \frac{e^{(1-z)\beta x_{k}}}{\prod_{i=1}^{n-1} (x_{k} - x_{i})} \Delta_{n-1}^{2}(x)$$

Fig. 5. Comparison of simulated data points and the asymptotic analytical c.d.f. $F_{V}(v)$ for different values of $n, \theta$ with $\alpha = 2$.

$$\times \prod_{j=1}^{n-1} x_{j}^{\alpha} e^{-x_{j}} (x_{j} - x)^{2} \, dx_{j}$$

with $I_{x} = \{0 < x_{1} < x_{2} < \ldots < x_{n-1} < x\}$. Since the integrand in the above $(n-1)$-fold integral is symmetric in $x_{1}, x_{2}, \ldots, x_{n-1}$, we may remove the ordered region of integration to obtain

$$\Omega(x, z) = \frac{1}{(n-1)!} \int_{(0,x)^{n-1}} \frac{e^{(1-z)\beta x_{k}}}{\prod_{i=1}^{n-1} (x_{k} - x_{i})} \Delta_{n-1}^{2}(x)$$

$$\times \prod_{j=1}^{n-1} x_{j}^{\alpha} e^{-x_{j}} (x_{j} - x)^{2} \, dx_{j}.$$

Consequently, we can observe that each term in the above summation evaluates to the same amount. Therefore, capitalizing on that observation, we may simplify the above $(n-1)$-fold integral to yield

$$\Omega(x, z) = \frac{1}{(n-2)!} \int_{(0,x)^{n-1}} \frac{e^{(1-z)\beta x_{1}}}{\prod_{i=2}^{n-1} (x_{1} - x_{i})} \Delta_{n-1}^{2}(x)$$

$$\times \prod_{j=1}^{n-1} x_{j}^{\alpha} e^{-x_{j}} (x_{j} - x)^{2} \, dx_{j}.$$

Now it is convenient to introduce the variable transformations $x_{t_{j}} = x_{j}, j = 1, 2, \ldots, n-1$, to the above integral to arrive at

$$\Omega(x, z) = \frac{x^{(n-1)(n+\alpha)+1}}{(n-2)!} \int_{(0,1)^{n-1}} \frac{e^{(1-z)\beta t_{1}}}{\prod_{i=2}^{n-1} (t_{1} - t_{i})} \Delta_{n-1}^{2}(t)$$

$$\times \prod_{j=1}^{n-1} t_{j}^{\alpha} (1-t_{j})^{2} e^{-t_{j}} \, dt_{j}.$$
with respect to \( t \) last giving

\[
\Omega(x, z) = \frac{x^{(n-1)(n+\alpha+1)}}{(n-2)!} \int_0^1 e^{-(1-(1-z))x t} (t-1)^2 t^\alpha \times \mathcal{P}_{n-2}(t, x) dt
\]

(39)

where

\[
\mathcal{P}_{n-2}(t, x)
= \int_{(0, 1)^{n-2}} \Delta_{n-2}^2(y) \prod_{j=1}^{n-2} y_j^\alpha (1 - y_j)^2 (t - y_j) e^{-xy_i} dy_j
\]

(40)

and we have used the decomposition \( \Delta_{n-2}^2(t) = \prod_{j=1}^{n-2} (t - y_j)^2 \Delta_{n-2}^2(y) \). The \((n-2)\)-fold integral in (40) can be evaluated with the help of [131, Corollary 1] and [125, Eq. 6.5.1] to yield the Hankel determinant

\[
\mathcal{P}_{n-2}(t, x)
= (n-2)! \det \left[ tA_{i,j}^{(\alpha)}(x) - A_{i,j}^{(\alpha+1)}(x) \right]_{i,j=1,2,\ldots,(n-2)}
\]

(41)

where

\[
A_{i,j}^{(\alpha)}(x) = B(3, i + j + \alpha - 1) \\
\times F_1(\alpha + i + j - 1; \alpha + i + j + 2; -x)
\]

(42)

\( \mathcal{B}(p, q) = (p-1)!(q-1)!/(p+q-1)! \), and \( F_1(a; c; z) \) is the confluent hypergeometric function of the first kind. Finally, we use (42) and (41) in (39) and substitute the resultant integral expression into (38) with some algebraic manipulation to arrive at the exact p.d.f. of \( Z_n^{(n)} \) as given in the following theorem.

\textbf{Theorem 4:} Let \( \mathbf{W} \sim \text{CVM}_n(m, \mathbf{I}_n + \mathbf{vv}^\top) \) with \( ||v|| = 1 \) and \( \theta > 0 \). Let \( u_n \) be the eigenvector corresponding to the largest eigenvalue of \( \mathbf{W} \). Then the p.d.f. of \( Z_n^{(n)} = ||v||^2 u_n^2 \) in (0, 1) is given by

\[
f^{(\alpha)}_{Z_n^{(n)}}(z) = \frac{c_{\alpha}^{(\alpha, \beta)}}{\beta^{n-2}} \int_0^\infty x^{n^2 + n\alpha - n + 1} e^{-(1-\beta)z} x f_n(x, z) dx
\]

(43)

\textbf{Corollary 3:} The exact p.d.f. of \( Z_n^{(n)} \) in (0, 1) corresponding to \( n = 2, 3 \), and \( n = 4 \) are given, respectively, by

\[
f^{(\alpha)}_{Z_2^{(2)}}(z) = \frac{2(2\alpha + 3)!}{(\alpha + 1)!(\alpha + 3)!} \left[ \frac{1 - \beta(1 - z)}{2 - \beta} \right] \]

\[
f^{(\alpha)}_{Z_3^{(3)}}(z) = \frac{4(3\alpha + 7)!}{(\alpha + 2)!(\alpha + 4)!} \left[ \frac{1 - \beta(1 - z)}{2 - \beta} \right]^3
\]

\[
f^{(\alpha)}_{Z_4^{(4)}}(z) = \frac{2\alpha!(\alpha + 1)!}{(\alpha + 2)!} \left[ \frac{1 - \beta(1 - z)}{2 - \beta} \right]^4
\]

(45)

(46)

(47)

where \( G_{\alpha, \beta}^{(\alpha, \beta)}(b, c, z) \) is given by (48), shown at the bottom of the page,

\[
F_2^{(\alpha, \beta)}(a, b, c, z) = \frac{(3\alpha + 2 - N)!}{(3 - \beta)^{3\alpha + 3 - N} F_2^{(\alpha, \beta)}(3\alpha + 3 - N, b, c, 1 - \beta/(1 - \theta))}
\]

(48)

\[
G_{\alpha, \beta}^{(\alpha, \beta)}(b, c, z) = B(3, \alpha + b - 3) B(3, \alpha + c - 3) \left[ \frac{(\alpha + N)!}{(3\alpha + 12 - N)!} \frac{1 - \beta(1 - z)}{3 - \beta} \right]^{\alpha + N}
\]

times

\[
\left[ \frac{(3\alpha + 12 - N)!}{(3 - \beta)^{3\alpha + 3 + N} F_2^{(\alpha, \beta)}(3\alpha + 3 - N, b, c, 1 - \beta/(1 - \theta))} \right]
\]

\[
- \sum_{k=0}^{\alpha + N} \frac{(3\alpha + 12 + k - N)!}{k!(4 - \beta)^{3\alpha + 3 + N - k} (1 - \beta(1 - z))^{k} F_2^{(\alpha, \beta)}(3\alpha + 3 - N + k, b, c, 1 - \beta/(1 - \theta))}
\]

(48)

This function assumes a finite series expansion involving exponentials and powers of \( x \), since \( \alpha, i, \) and \( j \) take non-negative integer values in (42).
To demonstrate this, following Theorem 4, let us rewrite the determinant of square matrix, whose size depends on $n$ of \([125, \text{Eq. 2.1.4.23}] \) and \([125, \text{Eq. 5.8.1.2}] \). However, those expressions, for the parameters of our interest here, they can be expressed as finite sums of rational functions of $z$. For each fixed $t, x$, $e^{(1-\theta)tx}$ is convex in $z$ and $\rho(x, t) \geq 0$, we conclude that $f_{Z_n^{(z)}}(z)$ is convex in $z$. This observation does not hold for $n \geq 3$, since the determinant term in the integrand changes its sign depending on the values of $x$ and $t$. However, keeping in mind that $f_{Z_n^{(z)}}(z) \leq C n \sqrt{\beta \pi} t^{n^2/2}$ where $J_n(x, z)$

\[
J_n(x, z) = \int_0^1 e^{-(1-(1-z)\beta)xt} t^{n^2/2} dt. \]

we may follow similar arguments to obtain the convex function $h_n(z)$ such that $f_{Z_n^{(z)}}(z) \leq h_n(z), n = 2, 3, \ldots, (49)$ for which the equality holds for $n = 2$.

**Remark 3:** It is noteworthy that although the hypergeometric functions $2F_1(\cdot)$ and $F_2(\cdot)$ assume infinite series expansions, for the parameters of our interest here, they can be expressed as finite sums of rational functions of $z$ with the help of \([125, \text{Eq. 2.1.4.23}] \) and \([125, \text{Eq. 5.8.1.2}] \). However, those formulas are not presented here to keep the representations clear and concise.

A careful inspection of (43) and (44) reveals that the determinant of square matrix, whose size depends on $n$, prevents us from applying the same asymptotic framework that we have developed to characterize the asymptotic behavior of $Z_1^{(n)}$ (i.e., as $m, n \to \infty$ such that $m - n$ is fixed). This stems from the fact that, as we are well aware of, in this asymptotic regime, $Z_n^{(n)}$ undergoes a phase transition at $\theta = 1$ (i.e., above it is $O(1)$, whereas below it is $o(1)$). To circumvent this difficulty, it is natural to consider other asymptotic regimes\(^{11}\) in view of obtaining relatively simple expressions for the p.d.f. of $Z_n^{(n)}$. In this respect, various stochastic convergence results for $Z_n^{(n)}$ have been established in \([1]–[6], [27], [36], [39], [58], [99].\) Nevertheless, a finite dimensional analysis has not been available in the literature to date.

Figures 6, 7, 8, and 9 verify the accuracy of our formulation. In particular, Fig. 6 shows the effect of $n$ on the p.d.f. of $Z_n^{(n)}$ for $\alpha = 2$ and $\theta = 3$, whereas Fig. 8 demonstrates the effect of $\theta$ for $n = 3$ and $\alpha = 2$. The corresponding Q-Q plots are shown in 7 and 9, respectively. The theoretical curves corresponding to $n = 5, 6$ and $n = 7$ in Fig. 6 have been generated by numerically evaluating the integrals in (43) and (44).

As can be seen from the figures, our theoretical formulations are corroborated by the simulation results.

### C. The P.D.F. of $Z_2^{(n)}$

The m.g.f. based machinery, which we have developed, can be readily extended to determine the p.d.f. of $Z_2^{(n)}$, for $\ell \neq 1, n$, however, at the expense of increased algebraic complexity. To further strengthen this claim, in what follows, we present a summary of technical arguments which lead to the exact p.d.f. of $Z_2^{(n)}$.

For $\ell = 2$ and $n \geq 3$, (14) specializes to

\[
\Psi_2(\lambda, s) = \frac{(n - 1)!}{2\pi i} \int \frac{e^{\omega}}{c (\omega - \beta \lambda_1) (s + \omega - \beta \lambda_2) \prod_{j=3}^{n} (\omega - \beta \lambda_j)} d\omega, \]

which in turn enables us to rewrite (13) as

\[
\mathcal{M}_{Z_2^{(n)}}(s) = C_{n, \alpha} \int_{R} \Delta_n^2(\lambda) \prod_{j=1}^{n} \lambda_j^{\alpha} e^{-\lambda_j}\]

\(^{11}\)The most prevalent asymptotic domain assumes $m, n \to \infty$ such that $n/m \to \gamma \in (0, 1)$.\)
Consequently, the inverse Laplace transform yields
\[
f_{Z_2^{(2)}}(z) = \frac{C_{\beta, n-2}^{\alpha}}{\beta n-2} \int \mathcal{R} e^{\beta \lambda z} \Delta_n^{(2)}(\lambda) \prod_{j=1}^{n} \lambda_j^\alpha e^{-\lambda_j} \times \prod_{k=3}^{n} \frac{e^{(1-z)\beta \lambda_k}}{(\lambda_k - \lambda_1)} d\lambda.
\]
where we have also used the Cauchy’s residue theorem. To facilitate further analysis, we split the integral into two parts as
\[
f_{Z_2^{(2)}}(z) = \frac{C_{\beta, n-2}^{\alpha}}{\beta n-2} \left( \mathcal{A}_n^{(2)}(z) + \mathcal{B}_n^{(2)}(z) \right)
\]
where
\[
\mathcal{A}_n^{(2)}(z) = \int \mathcal{R} e^{\beta \lambda z} \Delta_n^{(2)}(\lambda) \prod_{j=1}^{n} \lambda_j^\alpha e^{-\lambda_j} \frac{e^{(1-z)\beta \lambda_1}}{\prod_{i=3}^{n} (\lambda_i - \lambda_1)} d\lambda
\]
and
\[
\mathcal{B}_n^{(2)}(z) = \int \mathcal{R} e^{\beta \lambda z} \Delta_n^{(2)}(\lambda) \prod_{j=1}^{n} \lambda_j^\alpha e^{-\lambda_j} \times \prod_{k=3}^{n} \frac{e^{(1-z)\beta \lambda_k}}{(\lambda_k - \lambda_1)} \prod_{i=3}^{n} (\lambda_k - \lambda_i) d\lambda.
\]

Let us first consider \(\mathcal{B}_n^{(2)}(z)\). Employing the decomposition \(\Delta_n^{(2)}(\lambda) = (\lambda_2 - \lambda_1)^2 \prod_{i=3}^{n} (\lambda_i - \lambda_1)^2 (\lambda_j - \lambda_1)^2 \Delta_{n-2}^{(2)}(\lambda)\) along with relabelling the variables as \(t = \lambda_1, x = \lambda_2\) and \(t_{j-2} = \lambda_j, j = 3, \ldots, n\), gives us
\[
\mathcal{B}_n^{(2)}(z) = \int_{0}^{\infty} x^{\alpha} e^{-x(1-\beta z)} \times \left\{ \int_{0}^{x} (x-t)^2 t^{\alpha} e^{-t} \Upsilon(t, x, z) dt \right\} dx
\]
where
\[
\Upsilon(t, x, z) = \int_{D} \sum_{k=1}^{n-2} \frac{e^{(1-z)\beta t_k}}{(t_k - t) \prod_{i=1}^{n-2} (t_k - t_i)} \times \prod_{j=1}^{n-2} (t_j - t)^2 (t_j - x)^2 t_j^{\alpha} e^{-t_j} \Delta_{n-2}^{2}(t) dt_j
\]
and \(D = \{x < t_1 < \cdots < t_{n-2} < \infty\}\). An important observation that facilitates further simplification of \(\Upsilon(t, x, z)\)

(52)
can be made at this point. Since the integrand is symmetric with respect to \(t_j\)’s, the ordered region of integration can be removed. Consequently, all the terms in the summation will
Now we introduce the variable transformations

\[ Q_{\tilde{t}, x, z} = \frac{w}{t} \]

and

\[ L_{n+i-4}^{(2)}(\vartheta(z)u) = (3)_{n+i-4} \sum_{\ell=0}^{n+i-4} \frac{(1 - \vartheta(z))^\ell u^\ell}{(n + i - 4 - \ell)!\ell!3^{\ell}} \]

Following [123] and [125, Eq. 6.15.1.1], we obtain

\[ L_{n+i-4}^{(2)}(\vartheta(z)u) = (3)_{n+i-4} \sum_{\ell=0}^{n+i-4} \frac{(1 - \vartheta(z))^\ell u^\ell}{(n + i - 4 - \ell)!\ell!3^{\ell}} \]

Substituting the above result back into (57) gives us

\[ \varphi_i^{(b)}(x, t, z) = \frac{1}{(n - 3)!} \int_0^\infty e^{-w(1 - \beta[1 - z])} \left( 1 - \vartheta(z) \right) \frac{du}{u + (x - t)/\vartheta(z)} \]

The integral in the above expressions can be evaluated with the help of [125, Eq. 6.15.2.16] to get

\[ \int_0^\infty \frac{u^{2+\ell}e^{-u}}{u + (x - t)/\vartheta(z)} \]

which along with some algebraic manipulation yields

\[ \varphi_i^{(b)}(x, t, z) = \frac{1}{(n - 3)!} \int_0^\infty e^{-w(1 - \beta[1 - z])} \left( 1 - \vartheta(z) \right) \frac{du}{u + (x - t)/\vartheta(z)} \]

Following similar steps as in Appendix A, \( \tilde{Q}_{n-3}(w, t, x) \) can be evaluated to obtain (54), shown at the bottom of the page.

Now, by observing that only the first column of the above determinant depends on \( w \), we can re-write (53) with the help of (54) to get (55), shown at the bottom of the page, with

\[ \varphi_i^{(b)}(x, t, z) = \int_0^\infty \frac{e^{-w(1 - \beta[1 - z])} w^2}{w + x - t} L_{n+i-4}^{(2)}(w) \]

With the intention of simplifying \( \varphi_i^{(b)}(x, t, z) \) further, we employ the variable transformation \( u = w/\vartheta(z) \), where \( \vartheta(z) = 1/(1 - \beta + \beta z) \). Consequently, \( \varphi_i^{(b)}(x, t, z) \) becomes

\[ \varphi_i^{(b)}(x, t, z) = \vartheta(z) \int_0^\infty \frac{u^2 e^{-u} L_{n+i-4}^{(2)}(\vartheta(z)u)}{u + (x - t)/\vartheta(z)} \]

Following similar steps as in Appendix A, \( \tilde{Q}_{n-3}(w, t, x) \) can be evaluated to obtain (54), shown at the bottom of the page.

Now, by observing that only the first column of the above determinant depends on \( w \), we can re-write (53) with the help of (54) to get (55), shown at the bottom of the page, with

\[ \varphi_i^{(b)}(x, t, z) = \int_0^\infty \frac{e^{-w(1 - \beta[1 - z])} w^2}{w + x - t} L_{n+i-4}^{(2)}(w) \]

With the intention of simplifying \( \varphi_i^{(b)}(x, t, z) \) further, we employ the variable transformation \( u = w/\vartheta(z) \), where \( \vartheta(z) = 1/(1 - \beta + \beta z) \). Consequently, \( \varphi_i^{(b)}(x, t, z) \) becomes

\[ \varphi_i^{(b)}(x, t, z) = \vartheta(z) \int_0^\infty \frac{u^2 e^{-u} L_{n+i-4}^{(2)}(\vartheta(z)u)}{u + (x - t)/\vartheta(z)} \]

Following similar steps as in Appendix A, \( \tilde{Q}_{n-3}(w, t, x) \) can be evaluated to obtain (54), shown at the bottom of the page.

Now, by observing that only the first column of the above determinant depends on \( w \), we can re-write (53) with the help of (54) to get (55), shown at the bottom of the page, with

\[ \varphi_i^{(b)}(x, t, z) = \int_0^\infty \frac{e^{-w(1 - \beta[1 - z])} w^2}{w + x - t} L_{n+i-4}^{(2)}(w) \]

With the intention of simplifying \( \varphi_i^{(b)}(x, t, z) \) further, we employ the variable transformation \( u = w/\vartheta(z) \), where \( \vartheta(z) = 1/(1 - \beta + \beta z) \). Consequently, \( \varphi_i^{(b)}(x, t, z) \) becomes

\[ \varphi_i^{(b)}(x, t, z) = \vartheta(z) \int_0^\infty \frac{u^2 e^{-u} L_{n+i-4}^{(2)}(\vartheta(z)u)}{u + (x - t)/\vartheta(z)} \]

Following similar steps as in Appendix A, \( \tilde{Q}_{n-3}(w, t, x) \) can be evaluated to obtain (54), shown at the bottom of the page.

Now, by observing that only the first column of the above determinant depends on \( w \), we can re-write (53) with the help of (54) to get (55), shown at the bottom of the page, with

\[ \varphi_i^{(b)}(x, t, z) = \int_0^\infty \frac{e^{-w(1 - \beta[1 - z])} w^2}{w + x - t} L_{n+i-4}^{(2)}(w) \]

With the intention of simplifying \( \varphi_i^{(b)}(x, t, z) \) further, we employ the variable transformation \( u = w/\vartheta(z) \), where \( \vartheta(z) = 1/(1 - \beta + \beta z) \). Consequently, \( \varphi_i^{(b)}(x, t, z) \) becomes

\[ \varphi_i^{(b)}(x, t, z) = \vartheta(z) \int_0^\infty \frac{u^2 e^{-u} L_{n+i-4}^{(2)}(\vartheta(z)u)}{u + (x - t)/\vartheta(z)} \]

Following similar steps as in Appendix A, \( \tilde{Q}_{n-3}(w, t, x) \) can be evaluated to obtain (54), shown at the bottom of the page.

Now, by observing that only the first column of the above determinant depends on \( w \), we can re-write (53) with the help of (54) to get (55), shown at the bottom of the page, with

\[ \varphi_i^{(b)}(x, t, z) = \int_0^\infty \frac{e^{-w(1 - \beta[1 - z])} w^2}{w + x - t} L_{n+i-4}^{(2)}(w) \]

With the intention of simplifying \( \varphi_i^{(b)}(x, t, z) \) further, we employ the variable transformation \( u = w/\vartheta(z) \), where \( \vartheta(z) = 1/(1 - \beta + \beta z) \). Consequently, \( \varphi_i^{(b)}(x, t, z) \) becomes

\[ \varphi_i^{(b)}(x, t, z) = \vartheta(z) \int_0^\infty \frac{u^2 e^{-u} L_{n+i-4}^{(2)}(\vartheta(z)u)}{u + (x - t)/\vartheta(z)} \]

Following similar steps as in Appendix A, \( \tilde{Q}_{n-3}(w, t, x) \) can be evaluated to obtain (54), shown at the bottom of the page.

Now, by observing that only the first column of the above determinant depends on \( w \), we can re-write (53) with the help of (54) to get (55), shown at the bottom of the page, with

\[ \varphi_i^{(b)}(x, t, z) = \int_0^\infty \frac{e^{-w(1 - \beta[1 - z])} w^2}{w + x - t} L_{n+i-4}^{(2)}(w) \]

With the intention of simplifying \( \varphi_i^{(b)}(x, t, z) \) further, we employ the variable transformation \( u = w/\vartheta(z) \), where \( \vartheta(z) = 1/(1 - \beta + \beta z) \). Consequently, \( \varphi_i^{(b)}(x, t, z) \) becomes

\[ \varphi_i^{(b)}(x, t, z) = \vartheta(z) \int_0^\infty \frac{u^2 e^{-u} L_{n+i-4}^{(2)}(\vartheta(z)u)}{u + (x - t)/\vartheta(z)} \]
where \( U(a; c; x) \) being the confluent hypergeometric function of the second kind [125]. Having simplified \( \varphi_1(x, t, z) \), we can now substitute (55) into (51) to obtain (58), shown at the bottom of the next page.

The evaluation of \( \hat{A}_n^{(\alpha)}(z) \) follows along similar lines of arguments as before and hence is omitted. As such, we obtain

\[
\hat{A}_n^{(\alpha)}(z) = \frac{(-1)^n K_{n-2, \alpha}}{(n-2)!} \times \int_0^\infty x^n e^{-x(n-1-\beta z)} \left\{ \int_0^x (x-t)^{2(1-\beta \alpha)} dt \right\} dx.
\]

(59)

By substituting (58) and (59) into (50) along with the variable transformation \( y = 1 - \frac{x}{\alpha} \), we arrive at the final result which is stated in Theorem 5.

**Theorem 5:** Let \( \mathbf{W} \sim \mathcal{C}W_n (m, \mathbf{I}_n + \theta \mathbf{v} \mathbf{v}^H) \) with \( \| \mathbf{v} \| = 1 \) and \( \theta > 0 \). Let \( \mathbf{u}_2 \) be the eigenvector corresponding to the second smallest eigenvalue of \( \mathbf{W} \). Then the p.d.f. of \( Z_2^{(n)} = |\mathbf{v}^H \mathbf{u}_2|^2 \in (0, 1) \), for \( n \geq 3 \), is given by

\[
f_{Z_2^{(n)}}(z) = (-1)^n \beta^{\alpha+2} \left( \frac{1}{\beta} - 1 \right)^{n+\alpha} \times \int_0^\infty x^n e^{-x(n-\beta)} \int_0^1 y^2 e^{xy} Z_n^{(\alpha, \beta)}(x, y, z) dy dx.
\]

(60)

where

\[
Z_n^{(\alpha, \beta)}(x, y, z) = \frac{(n-1)!}{(n+\alpha-1)!} e^{-\beta xy(1-z)} x^{\alpha-2} \mathcal{V}_n^{(\alpha)}(x, y)
\]

\[
\mathcal{V}_n^{(\alpha)}(x, y) = \det \left[ L_{n+i-j}^{(\alpha)}(-xy) \right]_{i=1}^{n+1} \quad j=1, \ldots, n+1,
\]

\[
\varphi_1^{(\alpha)}(x, y, z) = (n+i-2)! \times \sum_{\ell=0}^{n+i-4} \frac{(-\beta xy(1-z))^{\ell}}{\ell!} U(n+i-1; 3+\ell; xy(1-\beta + \beta z)),
\]

and \( U^{(\alpha, \beta)}(x, y, z) \) is given by (61), shown at the bottom of the next page.

One of the most important features in the above formula is that the dimensions of the square matrices, whose determinants are in the integrand, depend only on the relative difference \( m - n \). Consequently, it is plausible that this further facilities an asymptotic analysis akin to what we have demonstrated in Corollary 2; nevertheless, we do not pursue it here.

To further strengthen our claim in the above theorem, we compare the simulated data points and the analytical p.d.f. \( f_{Z_2^{(n)}}(z) \) in Fig. 10 for different values of \( n \) with \( \theta = 3 \) and \( \alpha = 1 \). Here the analytical p.d.f. is obtained by numerically evaluating the double integral in (60). Q-Q plots in this respect are depicted in Fig. 11. The agreement between the analytical and simulation results is clearly evident from the figures; moreover, this verifies the accuracy of Theorem 5.
IV. REAL AND SINGULAR WISHART EXTENSIONS

Here we provide a detailed analysis of how to extend the results derived in the former sections to real and singular (i.e., \( m < n \)) Wishart scenarios. As shown in the sequel, the main technical challenge in this respect is to evaluate the corresponding contour integrals in closed-form which are amenable to further analysis. Although, in general, no closed-form solutions are possible, we demonstrate a few important special cases for which closed-form solutions lead to exact p.d.f.s.

A. Real Wishart Case

To shed some light on this scenario, following [117], [104], [132], [128], let us rewrite the real analogy of Theorem 1 as

\[
\mathbb{F}_0(vv^T, T) = \int_{\mathcal{O}_n} \operatorname{etr} \left( vv^T \mathbf{O} \mathbf{T} \mathbf{O}^T \right) \, d\mathbf{O}
\]

\[
= \frac{\Gamma(n/2)}{2\pi i} \int_{\mathcal{K}} \prod_{j=1}^n (\omega - \tau_j)^\frac{n}{2} e^{\omega} \, d\omega
\]  

(62)

where \( d\mathbf{O} \) is the invariant measure on the orthogonal group \( \mathcal{O}_n \) normalized to make the total measure unity, \( \mathbb{F}_0(\cdot, \cdot) \) is the real hypergeometric function of two matrix arguments [55], [117], the contour \( \mathcal{K} \) starts from \(-\infty\) and encircles the eigenvalues of \( T \in \mathbb{R}^{n \times n} \) given by \( \tau_1, \tau_2, \ldots, \tau_n \) in the positive direction (i.e., counter-clockwise) and goes back to \(-\infty\), \( v \in \mathbb{R}^n \) with \( ||v|| = 1 \), and \((\cdot)^T\) represents the transpose operator. Now, for \( W \sim W_n \left( m, I_n + \theta vv^T \right) \), we have from [117]

\[
f(W) dW = K_{m,n}^\theta \prod_{k=1}^n \lambda_k^\frac{(\alpha-1)}{2} e^{-\frac{\lambda_k}{2}} \Delta_n(\lambda) \]

\[
\times \operatorname{etr} \left( \frac{\beta}{2} vv^T O \Lambda O^T \right) \, d\Lambda d\mathbf{O}
\]  

(63)

\[
\mathbb{F}_n^{(\alpha)}(z) = (-1)^{n-1} K_{n-3, \alpha} \frac{(n + \alpha - 1)! (n + \alpha - 2)!}{(n - 3)!} \int_0^\infty x^{\alpha} e^{-x(n-1-\beta)}
\]

\[
\times \left\{ \int_0^x (x-t)^2 t^{-\alpha} e^{-t} \det \left[ \varphi^{(j)}_i(x, t, z) L_{n+i-2-j}^{(j)}(t-x) L_{n+i-k}^{(k-2)}(-x) \right]_{i=1, \ldots, \alpha+3}^{j=2,3} \right\} dx.
\]  

(58)

\[
\mathbb{U}_n^{(\alpha, \beta)}(x, y, z) = \det \left[ \varphi^{(j)}_i(x, y, z) L_{n+i-2-j}^{(j)}(-xy) L_{n+i-k}^{(k-2)}(-x) \right]_{i=1, \ldots, \alpha+3}^{j=2,3} \]

(61)
where
\[ K_{m,n}^\theta = \frac{\pi^{\frac{m}{2}} 2^{-\frac{m-2n}{2}} (1 + \theta)^{-\frac{m}{2}}}{\prod_{k=1}^n \Gamma \left( \frac{m-k+1}{2} \right) \Gamma \left( \frac{n-k+1}{2} \right)}. \]
Let us denote the \( \ell \)-th column of \( \mathbf{O} \) as \( \mathbf{o}_\ell \). Consequently, the m.g.f. of \( W_n^{(n)} \) can be written as
\[
\mathcal{M}_{W_n^{(n)}}(s) = K_{m,n}^\theta \int_{\mathbb{R}^n} \prod_{k=1}^n \lambda_k^{\frac{m}{2}-(1+\theta)} e^{-\frac{\lambda_k}{2} \Delta_n(\lambda)} \times \Psi(\lambda, s) \, d\Lambda,
\]
where
\[
\Psi(\lambda, s) = \frac{1}{2\pi i} \int_{\mathbb{C}} \frac{e^{\omega}}{(s + \omega - \frac{\theta}{2} \lambda_1)} \prod_{k=2}^n (\omega - \frac{\beta}{2} \lambda_k) \, d\omega.
\]
To facilitate further analysis, following [133, Eq. 5.3.21], let us take the inverse Laplace transform of the above m.g.f. to yield
\[
f_{W_n^{(n)}}(z) = \frac{K_{m,n}^\theta \Gamma \left( \frac{n}{2} \right)}{\sqrt{\pi}} z^{-\frac{1}{2} (1 - z)} \int_{\mathbb{R}^n} \prod_{k=1}^n \lambda_k^{\frac{m}{2}-(1+\theta)} e^{-\frac{\lambda_k}{2} \Delta_n(\lambda)} e^{\frac{\theta}{2} \lambda_1 z} \times \frac{1}{2\pi i} \int_{\mathbb{C}} \frac{e^{\omega(1-z)}}{\prod_{k=2}^n (\omega - \frac{\beta}{2} \lambda_k)} \, d\omega \, d\Lambda,
\]
from which we obtain, in view of (62), the m.g.f. of \( W_1^{(n)} \) as
\[
\mathcal{M}_{W_1^{(n)}}(s) = K_{m,n}^\theta \Gamma \left( \frac{n}{2} \right) \int_{\mathbb{R}^n} \prod_{k=1}^n \lambda_k^{\frac{m}{2}-(1+\theta)} e^{-\frac{\lambda_k}{2} \Delta_n(\lambda)} \times \Psi(\lambda, s) \, d\Lambda,
\]
where
\[
\Psi(\lambda, s) = \frac{1}{2\pi i} \int_{\mathbb{C}} \frac{e^{\omega}}{(s + \omega - \frac{\theta}{2} \lambda_1)} \prod_{k=2}^n (\omega - \frac{\beta}{2} \lambda_k) \, d\omega.
\]
Similarly, we can write the p.d.f. of \( W_n^{(n)} \) as (65), shown at the bottom of the page.

Further manipulation of the above multiple integrals is highly challenging due to the presence of \( \Phi_2^{(n-1)} \) in each of the integrands. Therefore, unlike in the complex case, a closed-form p.d.f. for \( W_n^{(n)} \) seems intractable for general values of \( m \) and \( n \). Nevertheless, it is noteworthy that in the important case of \( n = 2 \), \( \Phi_2^{(1)} \) degenerates to an exponential function. Capitalizing on this observation, we may write the p.d.f. of \( W_1^{(2)} \) as
\[
f_{W_1^{(2)}}(z) = \frac{K_{m,n}^\theta \pi}{z^{-\frac{1}{2} (1 - z)} \int_{0<\lambda_1<\lambda_2<\infty} \prod_{k=1}^2 \lambda_k^{\frac{m}{2}-(1+\theta)} e^{-\frac{\lambda_k}{2} \Delta_2(\lambda)} \times e^{\frac{\theta}{2} (\lambda_1 z + \lambda_2 (1-z))} \, d\lambda_1 \, d\lambda_2,
\]
from which we obtain, after some algebraic manipulation,
\[
f_{W_1^{(2)}}(z) = \frac{K_{m,n}^\theta \pi}{z^{-\frac{1}{2} (1 - z)} \int_{0<\lambda_1<\lambda_2<\infty} \prod_{k=1}^2 \lambda_k^{\frac{m}{2}-(1+\theta)} e^{-\frac{\lambda_k}{2} \Delta_2(\lambda)} \times e^{-\frac{\theta}{2} (1-z)} e^{-\frac{\theta}{2} (1-\beta (1-z))} \, d\lambda_1 \, d\lambda_2.
\]
Now we may evaluate the above double integral with the help of [124, Eq. 3.194.1] to yield
\[
f_{W_1^{(2)}}(z) = \frac{2^{m-1} (m - 1)}{\pi (1 + \theta)^{\frac{m-1}{2}}} z^{-\frac{1}{2} (1 - z)} \int_{0<\lambda_1<\lambda_2<\infty} \prod_{k=1}^2 \lambda_k^{\frac{m}{2}-(1+\theta)} e^{-\frac{\lambda_k}{2} \Delta_2(\lambda)} \times \Phi_2^{(1)}(b_1, b_2, \ldots, b_n; \beta_1, \beta_2, \ldots, \beta_n) \, d\lambda_1 \, d\lambda_2.
\]
\[13\]The p.d.f. of \( W_n^{(n)} \), for \( \theta = 0 \), can be obtained by setting \( \beta = 0 \) in (64) as \( f_{W_n^{(n)}}(z) = \frac{1}{\sqrt{\pi (2n-1)}} z^{-\frac{1}{2} (1 - z)} \Phi_2^{(n-1)} \). Moreover, due to the orthogonality of permutation matrices, the same p.d.f. result remains valid for the p.d.f. of \( W_{\ell}^{(n)} \) for \( \ell = 2, 3, \ldots, n \).
of our result for different values of $\theta$ and the accompanying Q-Q plot in Fig. 13 verify the accuracy of our result for different values of $\theta$ with $n = 2$ and $\alpha = 3$.

### B. Singular Complex Wishart Case

Another scenario of technical interest is when the number of observations $m$ is less than the dimension $n$ of the random vectors (i.e., $m < n$). In this situation, the matrix $W$ degenerates to the so-called singular Wishart matrix whose rank is $m$ almost surely. As such, the density of $W$ is defined on the space of $m \times m$ Hermitian positive semi-definite matrices of rank $m$ [110]. Consequently, $W$ assumes the eigen-decomposition $W = U_1 \Lambda U_1^\dagger$ with $U_1^\dagger U_1 = I_m$ and $\Lambda = \text{diag} (\lambda_1, \lambda_2, \ldots, \lambda_m)$ are the non-zero eigenvalues of $W$ ordered such that $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_m < \infty$. It is noteworthy that the set of all $n \times m$ complex matrices $U_1$ such that $U_1^\dagger U_1 = I_m$ (i.e., with orthonormal columns), denoted by $\mathcal{V}_{m,n}$, is known as the complex Stiefel manifold.

Following [110, Eqs. 2, 22], the joint density of the eigenvalues and eigenvectors of $W \sim \mathcal{CW}_n (m, I_n + \theta vv^\dagger)$, for $m < n$, can be written as

$$f(\Lambda, U_1) = \frac{\pi^{m(m-n-1)/2-m}}{\Gamma_m(m)(1+\theta)^m} \prod_{k=1}^{m} \lambda_k^{n-m} e^{-\lambda_k} \Delta^2_m (\Lambda)$$

$$\times \text{etr} \left( \beta vv^\dagger U_1 \Lambda U_1^\dagger \right).$$

Therefore, the m.g.f. of $Y^{(n)}_\ell = \|v^\dagger u_\ell\|$, $\ell = 1, 2, \ldots, m$, can be written as

$$M_{Y^{(n)}_\ell} (s) = \frac{\pi^{m(m-n-1)/2-m}}{\Gamma_m(m)(1+\theta)^m} \int \prod_{k=1}^{m} \lambda_k^{n-m} e^{-\lambda_k} \Delta^2_m (\Lambda)$$

$$\times \Psi_\ell (\Lambda, s) d\Lambda \quad (66)$$

where

$$\Psi_\ell (\Lambda, s) = \int_{\mathcal{V}_{m,n}} \text{etr} \left\{ vv^\dagger U_1 (\beta \Lambda - s g_\ell g_\ell^T) U_1^\dagger \right\} (U_1^\dagger dU_1) \quad (67)$$

with $(U_1^\dagger dU_1)$ denoting the exterior differential form representing the uniform measure on the complex Stiefel manifold [108], [110] and $g_\ell$ denotes the $\ell$th column of the $m \times m$ identity matrix. The next technical challenge is to evaluate the matrix integral in (67). To this end, for convenience, let us consider an equivalent matrix integral given by

$$J(T) = \int_{\mathcal{V}_{m,n}} \text{etr} \left\{ vv^\dagger U_1 T U_1^\dagger \right\} (U_1^\dagger dU_1) \quad (68)$$

where $T = \text{diag} (\tau_1, \tau_2, \ldots, \tau_m) \in \mathbb{R}^{m \times m}$ is a diagonal matrix. Since further manipulation of this matrix integral in the current form is an arduous task, following [108, Appendix A], we may rewrite it as a matrix integral over the unitary manifold to yield

$$J(T) = \frac{\pi^m}{\Gamma_m(n)} \int_{U_n} \text{etr} \left\{ vv^\dagger U T U^\dagger \right\} dU \quad (69)$$

where we have the partitioned matrix $U = (U_1 \ U_2) \in U_n$ and $T = \text{diag} (\tau_1, \tau_2, \ldots, \tau_m, 0, 0, \ldots, 0) \in \mathbb{R}^{n \times n}$. Now, the
above matrix integral yields [117]

\[ J(T) = \frac{2^m \pi^{mn}}{\Gamma_m(n)} F_0 \left( \nu \nu^T, T \right). \]  

(70)

Keeping in mind that \( \tilde{T} \) is rank deficient, we apply Theorem 1 to arrive at

\[ \int_{\nu_{m,n}} \text{etr} \left\{ \nu \nu^T U_1 U_1^T \right\} \left( U_1^T dU_1 \right) = \frac{K_{m,n}}{2\pi i} \oint_c \frac{\omega^{m-n} e^{\omega}}{(s + \omega - \beta \lambda_1) \prod_{j=2}^N (\omega - \beta \lambda_j)} \, d\omega \]  

(71)

where \( K_{m,n} = \frac{2^m \pi^{mn} \Gamma_m(n)}{\Gamma_m(n)} \). Therefore, in view of (71), (67) specializes to

\[ \Psi_1(\lambda, s) = \frac{K_{m,n}}{2\pi i} \oint_c \frac{\omega^{m-n} e^{\omega}}{(s + \omega - \beta \lambda_1) \prod_{j=2}^N (\omega - \beta \lambda_j)} \, d\omega, \]

\[ \Psi_m(\lambda, s) = \frac{K_{m,n}}{2\pi i} \oint_c \frac{\omega^{m-n} e^{\omega}}{(s + \omega - \beta \lambda_m) \prod_{j=1}^{m-1} (\omega - \beta \lambda_j)} \, d\omega, \]

from which we obtain after taking the inverse Laplace transform

\[ \mathcal{L}^{-1} \{ \Psi_1(\lambda, s) \} = e^{\beta \lambda_1 z} \frac{K_{m,n}}{2\pi i} \oint_c \frac{\omega^{m-n} e^{\beta \lambda_1 z + (1-z)\omega}}{\prod_{j=2}^N (\omega - \beta \lambda_j)} \, d\omega, \]

\[ \mathcal{L}^{-1} \{ \Psi_m(\lambda, s) \} = e^{\beta \lambda_m z} \frac{K_{m,n}}{2\pi i} \oint_c \frac{\omega^{m-n} e^{\beta \lambda_m z + (1-z)\omega}}{\prod_{j=1}^{m-1} (\omega - \beta \lambda_j)} \, d\omega, \]

where \( \mathcal{L}^{-1} \{ \cdot \} \) denotes the inverse Laplace operator. Now we evaluate the contour integrals, for \( m \geq 2 \), to yield

\[ e^{-\beta \lambda_1 z} \mathcal{L}^{-1} \{ \Psi_1(\lambda, s) \} = K_{m,n} \left( \frac{1}{\beta^{m-n-2}} \sum_{k=2}^{m} \frac{\lambda_k^{m-n} e^{(1-z)\beta \lambda_k}}{\prod_{i=2}^{m} (\lambda_k - \lambda_i)} + J_1^{(n-m)}(z, \lambda) \right), \]

(72)

\[ e^{-\beta \lambda_m z} \mathcal{L}^{-1} \{ \Psi_m(\lambda, s) \} = K_{m,n} \left( \frac{1}{\beta^{m-n-2}} \sum_{k=1}^{m-1} \frac{\lambda_k^{m-n} e^{(1-z)\beta \lambda_k}}{\prod_{i=1}^{m-1} (\lambda_k - \lambda_i)} + J_m^{(n-m)}(z, \lambda) \right), \]

(73)

where

\[ J_\ell^{(n-m)}(z, \lambda) = \begin{cases} \frac{1}{2\pi i} \oint_0^{2\pi} \frac{\omega^{m-n} e^{(1-z)\omega}}{\prod_{j=2}^N (\omega - \beta \lambda_j)} \, d\omega & \text{for } \ell = 1, \\ \frac{1}{2\pi i} \oint_0^{2\pi} \frac{\omega^{m-n} e^{(1-z)\omega}}{\prod_{j=1}^{m-1} (\omega - \beta \lambda_j)} \, d\omega & \text{for } \ell = m, \end{cases} \]

(74)

in which the contour 0 is taken to be a small circle around the origin such that all \( \lambda_j, j = 1, 2, \ldots, m \), are in the exterior of the contour. A careful inspection of (74) reveals that the pole of order \( n - m \) at the origin does not facilitate the direct use of Cauchy’s integral formula to evaluate these contour integrals\(^{14}\) for general values of \( m \) and \( n \) such that \( n > m \). To circumvent this difficulty, noting the power series expansion [104, Eq. 247]

\[ \frac{1}{N} \prod_{j=1}^{N} (\omega - \beta \lambda_j) = \sum_{k=0}^{\infty} \frac{\omega^k}{(-\beta)^k} C_k \left( \text{diag} \left( \lambda_1^{-1}, \ldots, \lambda_N^{-1} \right) \right), \]

where \( k \equiv (k, 0, 0, \ldots, 0) \) denotes the length \( N \) partition of \( k \) into not more than one part and \( C_k(\cdot) \) is the complex zonal polynomial, we may rewrite (74) as

\[ J_1^{(n-m)}(z, \lambda) = \frac{1}{(-\beta)^{m-1}} \prod_{j=2}^{m} \frac{C_k \left( \Lambda_1^{-1} \right)}{\omega^{n-m-k}} \omega^{(1-z)} \]

\[ J_m^{(n-m)}(z, \lambda) = \frac{1}{(-\beta)^{m-1}} \prod_{j=1}^{m-1} \frac{C_k \left( \Lambda_m^{-1} \right)}{\omega^{n-m-k}} \omega^{(1-z)} \]

where the diagonal matrices \( \Lambda_1 \) and \( \Lambda_2 \) are given, respectively, by \( \Lambda_1 = \text{diag} \left( \lambda_2, \ldots, \lambda_m \right) \) and \( \Lambda_m = \text{diag} \left( \lambda_1, \ldots, \lambda_{m-1} \right) \). Now, keeping in mind that \( \int_0^{2\pi} \frac{\omega^{(1-z)\omega}}{\prod_{j=2}^N (\omega - \beta \lambda_j)} \, d\omega = 0 \) for \( k \geq n - m, \)

\(^{14}\)To be precise, for \( \ell = 1 \), in light of Cauchy integral theorem, one can obtain the relation \( J_1^{(n-m)}(z, \lambda) = g(0, z, \lambda) \), where \( g(\omega, z, \lambda) = \frac{1}{2\pi i} \oint_0^{2\pi} \frac{\omega^{m-n} e^{(1-z)\omega}}{\prod_{j=2}^N (\omega - \beta \lambda_j)} \, d\omega \). Here it is worth noting that obtaining a general expression for the 4th derivative of \( e^{(1-z)\omega}/\prod_{j=2}^N (\omega - \beta \lambda_j) \) is intractable. Similar arguments apply to the case corresponding to \( \ell = m \) as well.
we may evaluate the above contour integrals to yield
\[
J_1^{(n-m)}(z, \lambda) = \frac{1}{(-\beta)^{m-1}} \prod_{j=2}^{m} \lambda_j \sum_{k=0}^{n-m-1} C_k \left(\Delta^{-1}_m\right) (1-z)^{n-m-1-k} (-\beta)^{k(n-m-1-k)!},
\]
\[
J_m^{(n-m)}(z, \lambda) = \frac{1}{(-\beta)^{m-1}} \prod_{j=1}^{m-1} \lambda_j \sum_{k=0}^{n-m-1} C_k \left(\Delta^{-1}_m\right) (1-z)^{n-m-1-k} (-\beta)^{k(n-m-1-k)!}.
\]
(75) (76)

The above general forms are not amenable to further analysis due to the availability of the zonal polynomials. However, for \(n-m=1,2\), the above formulas simplify to
\[
J_1^{(1)}(z, \lambda) = \begin{cases} \lambda_1 \prod_{j=1}^{m} A_j & \text{for } \ell = 1 \\ (-\beta)^{m-1} \prod_{j=1}^{m} A_j & \text{for } \ell = m \end{cases}
\]
(77)
and
\[
J_1^{(2)}(z, \lambda) = \begin{cases} \frac{1}{(-\beta)^{m-1}} \prod_{j=2}^{m} \lambda_j - \frac{(-\beta)^{-m}}{\prod_{j=2}^{m} \lambda_j} \sum_{k=2}^{m} \frac{1}{\lambda_k} & \text{for } \ell = 1 \\ \frac{1}{(-\beta)^{m-1}} \prod_{j=1}^{m-1} \lambda_j - \frac{(-\beta)^{-m}}{\prod_{j=1}^{m-1} \lambda_j} \sum_{k=1}^{m-1} \frac{1}{\lambda_k} & \text{for } \ell = m. \end{cases}
\]
(78)

Since we are primarily interested in the p.d.f. of \(Y_{\ell}^{(1)}\), for \(\ell = 1, m\), we use (72), (73) and (74) to take the inverse Laplace transform of (66) to obtain
\[
f_{Y_{\ell}^{(1)}}(z) = \frac{\pi^{m(m-1)-1}}{\Gamma_m(m)(1+\theta)^m} \int_\mathbb{R} \prod_{k=1}^{m} \lambda_k^{-m-\lambda_k} \Delta^2_m(\lambda) \times L^{-1} \{\Psi_m(\lambda, s)\} \, d\lambda.
\]
(79)

Consequently, the p.d.f.s corresponding to \(n > m \geq 2\) can be determined, in principle, by capitalizing on the framework developed in the previous sections along with (72), (73), (74), (75), and (76). However, the ensuing algebraic complexity prevents us from obtaining general closed-form solutions.

Nevertheless, for the trivial case of \(m = 1\), it can be shown that
\[
L^{-1} \{\Psi_1(\lambda, s)\} = \frac{K_{1,n}}{\Gamma(n-1)(1+\theta)} e^{\beta\lambda_1 z} (1-z)^{n-2},
\]
which upon substituting into (79) with some algebraic manipulation gives
\[
f_{Y_{1}^{(1)}}(z) = \frac{(1-z)^{n-2}}{(1+\theta)(1-\beta z)^n} \lambda_1^{-n} e^{-\lambda_1 (1-\beta z)} d\lambda_1 = \frac{(n-1)(1-z)^{n-2}}{(1+\theta)(1-\beta z)^n}, \ z \in (0,1).
\]

Moreover, for \(n-m=1\), the p.d.f. of \(Y_{1}^{(n)}\) in (79) can be simplified in view of (72), (74) and (77) as
\[
f_{Y_{1}^{(n)}}(z) = \frac{\pi^{m(m-1)}}{\Gamma_m(m)(m+1)(1+\theta)^m \beta^{m-1}} \int_\mathbb{R} \prod_{j=1}^{m} \lambda_j e^{-\lambda_j} \sum_{k=1}^{m} e^{(1-z)\beta \lambda_k} \, d\Lambda
\]
(82)
where
\[
A_m(z) = \int_\mathbb{R} e^{\lambda_1 z} \Delta^2_m(\lambda) \prod_{j=1}^{m} \lambda_j e^{-\lambda_j} \sum_{k=1}^{m} e^{(1-z)\beta \lambda_k} \, d\Lambda.
\]
(83)

and
\[
B_m(z) = (-1)^{m-1} \int_\mathbb{R} \lambda_1 e^{\lambda_1 z} \Delta^2_m(\lambda) \prod_{j=1}^{m} e^{-\lambda_j} \, d\Lambda.
\]
(84)

Let us first focus on \(A_m(z)\). To this end, following the similar arguments as in Subsection II. A, we may decompose, after some algebraic manipulation, the multiple integral in (83) as
\[
A_m(z) = \frac{1}{(m-2)!} \int_0^\infty x e^{-(m-\beta)x} \times \left\{ \int_0^\infty y^2 e^{-[1-(1-z)\beta]y} Q_{m-2}(y, x) \, dy \right\} \, dx
\]
(85)
Fig. 15. Quantile-Quantile plots of simulated data of $Y_{1}^{(n)}(z)$ for different values of $n$ with $n - m = 1$ and $\theta = 0.3$.

where

$$Q_{m-2}(y, x) = (-1)^m m \hat{K}_{m-2,1} (x + y)^{-1} \times \det \begin{bmatrix} L_{m+i-3}^{(2)}(y) & L_{m+i-j-1}^{(j)}(-x) \\ L_{m+i-2}^{(2)} & -1 \end{bmatrix}_{i=1,2},$$

To facilitate further analysis, noting the term $(x + y)$ in the denominator of $Q_{m-2}(y, x)$, we may use the Christoffel-Darboux formula \[135, Eq. 22.12.1\] to obtain

$$Q_{m-2}(y, x) = (-1)^m m \hat{K}_{m-2,1} \times \sum_{\ell=0}^{m-2} \frac{(m - 2 - \ell)!}{(m - \ell)!} f_{m-2-\ell}^{(2)}(-x) f_{m-2-\ell}^{(2)}(y).$$

Now we substitute $Q_{m-2}(y, x)$ into (85) and solve the resultant double integral with the help of \[124, Eq. 7.414.8\] and the Definition 3 with some algebraic manipulation to arrive at

$$A_m(z) = m \hat{K}_{m-2,1} \sum_{\ell=0}^{m-2} \sum_{k=0}^{m-2-\ell} a_{\ell,k}(m, \beta) \frac{(1 - z)^{m-2-\ell}}{[1 - (1 - z)\beta]^{m-\ell+1}}$$

where

$$a_{\ell,k}(m, \beta) = \frac{(-1)^{\ell}(m - \ell)!\beta^{m-2-\ell}}{(k + 2)![(m - 2 - \ell - k)!(m - \beta)k + 2]]}.$$ 

Having evaluated $A_m(z)$, we now focus $B_m(z)$. As such, we may decompose the corresponding multiple integral, after some algebraic manipulation, to yield

$$B_m(z) = \frac{(-1)^{m-1}}{(m - 1)!} \int_0^{\infty} x e^{(m-\beta)z} dx \times \int_{(0,\infty)^{m-1}} \Delta^2_{m-1}(\lambda) \prod_{k=2}^{m} \lambda_k e^{-\lambda_k} d\lambda_k,$$

from which we obtain

$$B_m(z) = \frac{(-1)^{m-1} \prod_{k=0}^{m-2} (k + 1)! (k + 2)!}{(m - 1)! (m - \beta)^2}. \ (87)$$

Finally, we may substitute (86) and (87) into (82) with some algebraic manipulation to arrive at the p.d.f. of $Y_{1}^{(n)}$.
corresponding to $n - m = 1$ as

$$f_{Y_{m}^{(n)}}(z) = \frac{m}{\beta^{m-1}}(1 - \beta)^m$$

Similarly, following the algebraic machinery shown in Subsection II. B, we can obtain the p.d.f. of $Y_m^{(n)}$, for $n - m = 1$, as

$$f_{Y_m^{(n)}}(z) = \frac{(1 - \beta)^{m-1-m}}{\prod_{j=1}^{m}(m-j)!(m-j)!} 	imes \left( \int_0^\infty x^{m-2}(1-\beta)x^{m-2} J_m^s(x, z)dx \right) + (-1)^{m-1} \int_0^\infty x^{m-2} e^{-(1-\beta)x} \times \det \left[ A_{i,j}^{(0)}(x) \right]_{i,j=1,2,...,(m-1)} dx$$

where

$$J_m^s(x, z) = \int_0^1 (1-t)^2e^{-[1-(1-z)\beta]xt} \times \det \left[ tA_{i,j}^{(1)}(x) - A_{i,j}^{(2)}(x) \right]_{i,j=1,2,...,(m-2)} dt.$$  

Here, an empty determinant (i.e., when $m = 2$) is interpreted as unity. It is noteworthy that, although further simplification of the above integrals seems an arduous task, they can be evaluated numerically for not so large values of $m$.

Figures 14 and 15 compare the theoretical and simulated data of $Y_{1}^{(n)}$ with respect to the p.d.f. and Q-Q plots for different values of $n$ with $n - m = 1$ and $\theta = 0.3$. A close match between the results verifies the accuracy of our formulation. Moreover, Figs. 16 and 17 depict the respective plots corresponding to $Y_m^{(n)}$. Again, a close match between the analytical and simulation results validates the accuracy of our expressions.

V. CONCLUSION

This paper investigates the finite dimensional distributions of the eigenvectors corresponding to the extreme eigenvalues (i.e., the minimum and the maximum) of $W \sim \mathcal{CW}_n [m, I_n + \theta v v^H]$. In this respect, the exact p.d.f.s of $|v^H u_1|^2$ and $|v^H u_n|^2$ have been derived. In particular, the p.d.f. of $|v^H u_1|^2$ assumes a closed-form involving the determinant of a square matrix of dimension $m - n$, whereas the p.d.f. of $|v^H u_n|^2$ is expressed as a double integral with its integrand containing the determinant of a square matrix of dimension $n - 2$. Our analytical p.d.f.s reveal that, in the finite dimensional setting, the least eigenvector of $W$ contains information about the dominant eigenvector (i.e., the spike $v$) of the population covariance matrix. In this respect, a somewhat interesting stochastic convergence result derived here shows that, as $m, n \to \infty$ such that $m - n$ is fixed, $n|v^H u_1|^2$ converges in distribution to $\chi_n^2/2(1+\theta)$. This further highlights the discrimination power of the least eigenvector of $W$ in certain asymptotic domains. Moreover, our numerical
results show that, although derived in the asymptotic setting, this analytical relationship remains valid for small values of \(m\) and \(n\) as well. On the other hand, the double integral form of p.d.f. derived for \(|\mathbf{v}^\top \mathbf{u}_n|^2\) does not seem to admit a simple form except in the case of \(n = 2, 3,\) and \(4.\) Nevertheless, this double integral can be evaluated numerically as shown in our numerical results.

The m.g.f. framework developed here can be easily adapted to derive the p.d.f.s pertaining to the other eigenmodes, however with an extra algebraic complexity, as shown in Theorem 5. Moreover, the same framework has been extended to derive the corresponding p.d.f.s for real and singular Wishart scenarios; however, with closed-form solutions limited to a few special configurations of \(m\) and \(n.\) This stems from the fact that the analogous contour integrals, in general, do not admit tractable forms, which are amenable to further analysis, except in those special cases. Be that as it may, one of the other major technical challenges is to extend the current finite dimensional analysis to rank-\(r\) (\(\leq n\)) spiked Wishart matrices (i.e., \(\Sigma = \mathbf{I}_n + \sum_{k=1}^r \theta_k \mathbf{v}_k \mathbf{v}_k^\top\)) with \((\mathbf{v}_1 \mathbf{v}_2 \ldots \mathbf{v}_r) \in \mathcal{V}_{r,n},\) which remains as an open problem.

\section{Appendix A}

\textbf{The Evaluation of} \(Q_{n-2}(y_1, x)\) \textbf{in (20)}

For clarity, let us consider the following integral

\[
T_n(y, x) = \int_{(0, \infty)^n} \Delta_2^n(z) \prod_{j=1}^{n} (y - z_j)(x - z_j)^\alpha z_j^2 e^{-z_j} \, dz_j
\]

which in turn gives

\[
Q_{n-2}(y_1, x) = (-1)^{\alpha} T_n(y, x).
\]

Therefore, in the sequel, we evaluate \(T_n(y, x).\) To this end, following [64, Eqs. 22.4.2, 22.4.11], we focus on the related integral

\[
\int_{(0, \infty)^n} \prod_{j=1}^{n} e^{-z_j} \prod_{i=1}^{\alpha+1} (r_i - z_j) \Delta_2^n(z) \, dz_1 \cdots dz_n
\]

\[
= \prod_{i=0}^{\alpha-1} (i + 1)! (i + 2)! \frac{\det [\mathcal{P}_{n-i}(r_j)]_{i,j=1,\ldots,\alpha+1}}{\Delta_{\alpha+1}(r)},
\]

where \(\mathcal{P}_k(x)\)’s are monic polynomials orthogonal with respect to \(x^2 e^{-x},\) over \(0 \leq x < \infty.\) Consequently, we choose \(\mathcal{P}_k(x) = (-1)^k k! L_k^{(2)}(x),\) which upon substituting into (90) gives

\[
\int_{(0, \infty)^n} \prod_{j=1}^{n} z_j^2 e^{-z_j} \prod_{i=1}^{\alpha+1} (r_i - z_j) \Delta_2^n(z) \, dz_1 \cdots dz_n
\]

\[
= \tilde{K}_{n,\alpha} \frac{\det [L_{n-i-1}(r_j)]_{i,j=1,\ldots,\alpha+1}}{\Delta_{\alpha+1}(r)},
\]

where

\[
\tilde{K}_{n,\alpha} = (-1)^{(n-1)(\alpha+1)} \prod_{i=0}^{n-1} (i + 1)! (i + 2)! \prod_{i=1}^{\alpha+1} (-1)^i (n + i - 1)!
\]

Now we need to choose the \(r_i\)’s in the above relation such that the multiple integral on the left coincides with (88). In this respect, we select \(r_i\)’s as follows:

\[
r_i = \begin{cases} y & \text{if } i = 1 \\ x & \text{if } i = 2, \ldots, \alpha + 1. \end{cases}
\]

However, since the \(r_i\)’s in (91) are distinct in general, the above choice of parameters drives the right side to \(0/0\) indeterminate form. To alleviate this technical difficulty, capitalizing on an approach given in [136], instead of direct substitution, we use the limiting argument

\[
T_n(y, x) = \tilde{K}_{n,\alpha} \frac{\det [L_{n-i-1}(y) \, L^{(2)}_{n-i-1}(r_j)]_{i=1,\ldots,\alpha+1}}{\det[y^{i-1} L^{(2)}_{n-i}(y)]_{i=1,\ldots,\alpha+1}}
\]

to yield

\[
T_n(y, x) = \tilde{K}_{n,\alpha} \frac{\det [L_{n+i-1}(y) \, d^{j-2} L^{(2)}_{n+i-1}(x)]_{j=1,\ldots,\alpha+1}}{\det[y^{i-1} d^{j-2} x^{j-1}]_{j=1,\ldots,\alpha+1}}.
\]

Now the denominator of (92) gives

\[
\det[y^{i-1} d^{j-2} x^{j-1}]_{i=1,\ldots,\alpha+1} = \prod_{i=1}^{\alpha-1} (i + 1)! (i + 2)! (x - y)^\alpha.
\]

The numerator can be simplified using (8) to yield

\[
\det[L^{(2)}_{n+i-1}(y) \, d^{j-2} L^{(2)}_{n+i-1}(x)]_{j=1,\ldots,\alpha+1} = (-1)^{\frac{\alpha}{2}} \det [L^{(2)}_{n+i-1}(y) L^{(j)}_{n+i-1}(x)]_{j=1,\ldots,\alpha+1}.
\]

Therefore, we substitute (93) and (94) into (92) with some algebraic manipulation to arrive at

\[
T_n(y, x) = (-1)^{n+\alpha(n+\alpha)} \tilde{K}_{n,\alpha} \frac{\det [L^{(2)}_{n+i-1}(y) L^{(j)}_{n+i-1}(x)]_{i=1,\ldots,\alpha+1}}{\det y^{i-1} d^{j-2} x^{j-1}}
\]
where
\[
\hat{K}_{n,\alpha} = \prod_{j=1}^{n} (n + j - 1) \prod_{j=0}^{n-1} (j + 1)!(j + 2)! \prod_{j=0}^{\alpha-1} j!.
\]
Finally, we use (95) in (89) with some simple algebraic manipulation to obtain (21).

**APPENDIX B**

**Proof of Corollary 2**

Our strategy is to first show that the p.d.f. of \( nZ_1^{(n)} \) converges almost everywhere to the p.d.f. of a certain scaled chi-squared random variable. Subsequent application of the Scheffé’s theorem [127, Corollary 2.3] then establishes the convergence in distribution result for \( nZ_1^{(n)} \).

Let us evaluate the limiting p.d.f. of \( nZ_1^{(n)} \). To this end, we write
\[
f_{nZ_1^{(n)}} (z) = \frac{1}{n} f_{Z_1^{(n)}} \left( \frac{z}{n} \right),
\]
from which we obtain, in view of (30) and by employing elementary limiting arguments,
\[
\lim_{n \to \infty} f_{nZ_1^{(n)}} (z) = (1 - \beta)^{-1} e^{-\beta z/n} \left( \prod_{k=0}^{\alpha} \prod_{j=0}^{n} (j + k + 1)k! \right) \lim_{n \to \infty} \left\{ \zeta^{(\alpha)}_n (z, k_j, \beta) \right\}
\]
where \( \zeta^{(\alpha)}_n (z, k_j, \beta) \) given by (98), shown at the bottom of the page. Now multiplying and dividing \( a_{i,j}(\ell) \) by \((n + \alpha - j - \ell - 1)!\) along with some algebraic manipulation gives us (99), shown at the bottom of the page, where \( \tilde{c}_j = n + \alpha - j - k_j - 1 \) and an empty product is interpreted as 1. To further simplify the determinant, we apply the following row operations
\[
(i+1)^{th} \text{ row} \leftarrow (i+1)^{th} \text{ row} + \sum_{k=0}^{\alpha-1} S^{(k)}_{\alpha-i} \times (\alpha-k+1)^{th} \text{ row},
\]
for \( i = 0, 1, \ldots, \alpha - 1 \), where \( S^{(m)}_n \) is the Stirling number of the second kind [125], to yield
\[
\xi^{(\alpha)}_n (z, k_j, \beta) = \prod_{j=1}^{\alpha} \frac{(n + \alpha - j - 1)!}{(n - \beta)^{k_j}(n + \alpha - j - k_j - 1)!} \times \det \left[ \frac{(n + \alpha)!(-\beta)^i(1 - \tilde{z}/n)^i}{n(n+i-2)!(1 - \beta(1 - \tilde{z}/n)^i)^i} \right]_{j=0,\ldots,\alpha}.
\]

In which \( \mu_n (z, i, \beta) \) given by (101), shown at the bottom of the page. To facilitate further analysis, we expand the determinant in (100) with its first column to obtain
\[
\det \left[ \mu_n (z, i, \beta) \right]_{j=0,\ldots,\alpha} = \sum_{i=0}^{\alpha} (-1)^{i} \mu_n (z, i, \beta) M_{i}(n)
\]
where \( M_i(n) \) is the minor corresponding to the \((i+1)^{th}\) element of the first column. Since it is clear that \( \mu_n (z, i, \beta) = 1/n^i + o(1/n^i) \) as \( n \) grows large, we need to determine the highest power of \( n \) in \( M_{i}(n) \) and corresponding coefficient to determine the limit of (102) for large \( n \). To this end, following the definition of Schur-polynomials \( s_{\nu_i} \) [118], for \( i > 0 \), we obtain
\[
\xi^{(\alpha)}_n (z, k_j, \beta) = \frac{\Delta_{\alpha} (\tilde{c})}{\Delta_{\alpha} (\tilde{c})} \delta_{i=0} + s_{\nu_i} (\tilde{c}_1, \ldots, \tilde{c}_\alpha) \Delta_{\alpha} (\tilde{c})
\]
where \( \nu_i = \{1, 1, \ldots, 1, 0, \ldots, 0\} \) and \( \Delta_{\alpha} (\tilde{c}) \) is the Vandermonde determinant in terms of \( \tilde{c} = \{\tilde{c}_1, \ldots, \tilde{c}_\alpha\} \). Moreover, for \( i = 0 \), it is easy to show that \( M_{0}(n) = \Delta_{\alpha} (\tilde{c}) \). Consequently, noting the fact that \( \Delta_{\alpha} (\tilde{c}) = \Delta_{\alpha} (c) \) for \( \tilde{c} = \{c_1(k_1), \ldots, c_\alpha(k_\alpha)\} \) with \( c_{\ell}(j) = j + \ell \), we rewrite \( M_i(n) \) as
\[
M_{i}(n) = \begin{cases} 
\Delta_{\alpha} (c) & i = 0 \\
& \\
\sum_{\nu_i} (c_1, \ldots, c_\alpha) \Delta_{\alpha} (c) & i > 0 
\end{cases}
\]
Therefore, the problem boils down to determining the highest power of \( n \) in the expansion of \( s_{\nu_i} (c_1, \ldots, c_\alpha) \) which is a
symmetric homogeneous polynomial of degree $i$. To this end, following [118, Eq. 5.3.5], we may write
\begin{equation}
\xi_n^{(z)}(z, k, \beta) = \prod_{j=1}^{\alpha} \frac{(n - \alpha - j)!}{(n - \beta)^k(j)(n + \alpha - j)} - 1 + o(1)
\end{equation}
where the sum includes $\binom{n}{i}$ different terms of degree $i$, where $\binom{n}{i} = a^i / i!(\alpha - i)!$ denotes the binomial coefficient. Since we are interested in the highest power of $n$, we may write $s_{\nu_i}(\tilde{c}_1, \ldots, \tilde{c}_\alpha)$ as a polynomial of $n$ as
\begin{equation}
h_i(n) = \left( \frac{\alpha}{i} \right) n^i + \text{other lower order terms.}
\end{equation}
As such, (100) assumes
\begin{equation}
\xi_n^{(z)}(z, k, \beta) = \prod_{j=1}^{\alpha} \frac{(n - \alpha - j)!}{(n - \beta)^k j} (n + \alpha - j - 1)! 
\end{equation}
\begin{equation}
\times \sum_{i=0}^{\alpha} (-1)^i \mu_n(z, i, \beta) h_i(n) \Delta_\alpha(e).
\end{equation}
Noting that
\begin{equation}
\frac{(n + \alpha - j)!}{(n - \beta)^k j (n + \alpha - j - 1)!} = 1 + o(1)
\end{equation}
and $\mu_n(z, i, \beta) = \frac{1}{n^i} + o(\frac{1}{n^i})$, we take the limit as $n \to \infty$ with the help of (106) to yield
\begin{equation}
\lim_{n \to \infty} \xi_n^{(z)}(z, k, \beta) = \sum_{i=0}^{\alpha} (-1)^i \left( \frac{-1}{1 - \beta} \right)^i \left( \frac{\alpha}{i} \right) \Delta_\alpha(e)
\end{equation}
\begin{equation}
= \frac{\Delta_\alpha(e)}{(1 - \beta)^\alpha}.
\end{equation}
Therefore, noting that $\beta = \theta / (1 + \theta)$, the limiting p.d.f. (97) specializes to
\begin{equation}
\lim_{n \to \infty} f_n^{(z)}(z) = \mathcal{K}_\alpha(1 + \theta) e^{-(1 + \theta)z}
\end{equation}
where
\begin{equation}
\mathcal{K}_\alpha = \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \cdots \sum_{k_\alpha=0}^{\infty} \frac{(\alpha + \sum_{j=1}^{\alpha} k_j)!}{\prod_{j=1}^{\alpha} (j + k_j + 1)!} \Delta_\alpha(e)
\end{equation}
is a constant. The direct evaluation of $\mathcal{K}_\alpha$ seems to be an arduous task due to the presence of the term $(\alpha + \sum_{j=1}^{\alpha} k_j)!$. To circumvent this difficulty, we replace it with an equivalent integral to yield
\begin{equation}
\mathcal{K}_\alpha = \int_0^{\infty} \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \cdots \sum_{k_\alpha=0}^{\infty} \frac{x^{\alpha + \sum_{j=1}^{\alpha} k_j}}{\prod_{j=1}^{\alpha} (j + k_j + 1)!} \Delta_\alpha(e) dx
\end{equation}
from which we obtain
\begin{equation}
\mathcal{K}_\alpha = \int_0^{\infty} \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \cdots \sum_{k_\alpha=0}^{\infty} \frac{x^{\alpha + \sum_{j=1}^{\alpha} k_j}}{\prod_{j=1}^{\alpha} (j + k_j + 1)!} \Delta_\alpha(e) dx
\end{equation}
Now we adopt the similar procedure as in [137, Eqs. 21-B.23] to obtain (112), shown at the top of the page, where $\rho > 0$ is an arbitrary number. Consequently, we choose $\rho = 2$ to further simplify the above integral as
\begin{equation}
\mathcal{K}_\alpha = \int_0^{\infty} x^\alpha e^{-x} \times \det \frac{1}{(1 + 2 - i)} \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \cdots \sum_{k_\alpha=0}^{\infty} \frac{x^{\alpha + \sum_{j=1}^{\alpha} k_j}}{\prod_{j=1}^{\alpha} (j + k_j + 1)!} \Delta_\alpha(e) dx
\end{equation}
where $I_p(z) = \sum_{k=0}^{\infty} \frac{(z/2)^{2k+p}}{(k+p)!}$ denotes the modified Bessel function of the first kind and order $p$. A careful inspection of the above integral and [138, Corollary 4] reveals that the function $e^{-x} \times \det \{I_{j+i+2} (2\sqrt{x})\}_{i,j=1,...,\alpha}$ is another p.d.f, thereby $\mathcal{K}_\alpha = 1$. Therefore, (108) simplifies to
\begin{equation}
\lim_{n \to \infty} f_n^{(z)}(z) = (1 + \theta) e^{-(1 + \theta)z}
\end{equation}
which is the p.d.f. of a $\frac{\lambda^2}{2(1 + \theta)}$ random variable. Finally, we invoke the Scheffé’s theorem [127, Corollary 2.3] to conclude the proof.

**Acknowledgment**

The authors would like to thank the anonymous reviewers and the editor for their valuable comments.

**References**

[1] D. Paul, “Asymptotics of sample eigenstructure for a large dimensional spiked covariance model,” *Statist. Sinica*, vol. 17, no. 4, pp. 1617–1642, 2007.

[2] F. Benaych-Georges and R. R. Nadakuditi, “The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices,” *Adv. Math.*, vol. 227, no. 1, pp. 494–521, May 2011.

[3] A. Bloemendal, A. Knowles, H.-T. Yau, and J. Yin, “On the principal components of spiked covariance matrices,” *Probab. Theory Rel. Fields*, vol. 164, nos. 1–2, pp. 459–552, Feb. 2016.

[4] H. Xi, F. Yang, and J. Yin, “Convergence of eigenvector empirical spectral distribution of sample covariance matrices,” *Ann. Statist.*, vol. 48, no. 2, pp. 953–982, Apr. 2020.

[5] Z. Bao, X. Ding, J. Wang, and K. Wang, “Statistical inference for principal components of spiked covariance matrices,” *Ann. Statist.*, vol. 50, no. 2, pp. 1144–1169, Apr. 2022.

[6] W. Wang and J. Fan, “Asymptotics of empirical eigenstructure for high dimensional spiked covariance,” *Ann. Statist.*, vol. 45, no. 3, pp. 1342–1374, Jun. 2017.

[7] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. New York, NY, USA: Wiley, 2004.
