Supplementary Appendix to "Detection of weak signals in high-dimensional complex-valued data"

Alexei Onatski

July 30, 2012

Abstract

This note contains an elementary proof of Corollary 1, a proof of formula (24), and a derivation of the asymptotic power of some previously proposed tests of sphericity and equality of the covariance matrix to the identity matrix.

A Elementary proof of Corollary 1.

Let $\tilde{A} = \operatorname{diag}(\tilde{a}_1, ..., \tilde{a}_p)$ and $\tilde{B} = \operatorname{diag}\left(\tilde{b}_1, ..., \tilde{b}_p\right)$ be such that $\tilde{a}_1 > ... > \tilde{a}_p > 0$ and $\tilde{b}_1 > ... > \tilde{b}_p > 0$. The right-hand side of (12)[1] is understood as

$$\lim_{\tilde{A}\to A, \tilde{B}\to B} \frac{\prod_{k=1}^{p-1} k!}{V_p\left(\tilde{A}\right) V_p\left(\tilde{B}\right)} \det_{1\le i,j\le p}\left(e^{\tilde{a}_i \tilde{b}_j}\right).$$

We will take this limit sequentially: first as $\tilde{A} \to A$ and then as $\tilde{B} \to B$. This approach is without loss of generality because the left hand side of (12) is a continuous function of $\tilde{a}_j$ and $\tilde{b}_j$ with $j = 1, ..., p$.

Let us consider a $p \times p$ matrix L with i,j-th element $L_{ij} = \sum_{t=0}^{p-1} \frac{1}{t!}\left(\tilde{a}_i \tilde{b}_j\right)^t$. Note that $e^{\tilde{a}_i \tilde{b}_j} = (1 + o(1)) L_{ij}$, where $o(1) \to 0$ when $\tilde{a}_i \to 0$. Furthermore,

$$\begin{aligned}
\det_{1\le i,j\le p}\left(e^{\tilde{a}_i \tilde{b}_j}\right) &= (1 + o(1)) \det\left([e_1, ..., e_r][u_1, ..., u_r]' + L\right) \\
&= (1 + o(1)) \det(L) \det\left([e_1, ..., e_r][u_1, ..., u_r]' L^{-1} + I_p\right) \\
&= (1 + o(1)) \det(L) \det\left(I_r + [u_1, ..., u_r]' L^{-1} [e_1, ..., e_r]\right),
\end{aligned}$$

where e_j is the j-th column of I_p and u_s' is the s-th row of $\left(e^{\tilde{a}_i \tilde{b}_j}\right)_{1\le i,j\le p}$ minus the s-th row of L. Since the i-th row of L multiplied by $L^{-1} e_j$ equals the Kronecker's

[1] Here and throughout this Supplement, numerical references are for equations in the main text.

delta δ_{ij}, we further have

$$\det_{1\leq i,j\leq p}\left(e^{\tilde{a}_i\tilde{b}_j}\right)=(1+o(1))\det(L)\det\left([w_1,...,w_r]'\,L^{-1}\,[e_1,...,e_r]\right), \tag{SA1}$$

where w_s' is the s-th row of $\left(e^{\tilde{a}_i\tilde{b}_j}\right)_{1\leq i,j\leq p}$.

Now, note that

$$L=V_{\tilde{a}}DV_{\tilde{b}}', \tag{SA2}$$

where $D=\operatorname{diag}\left(\frac{1}{0!},\frac{1}{1!},...,\frac{1}{(p-1)!}\right)$, and V_x denotes the Vandermonde matrix with $V_{x,ij}=x_i^{j-1}$ and determinant $V_p(X)$, where $X=\operatorname{diag}(x_1,...,x_p)$. As shown by Klinger (1967) (see his formula 6),

$$\left(V_x^{-1}\right)_{ij}=(-1)^{p-i}\frac{\sigma_{p-i}\left(x_{-j}\right)}{\prod_{s=1,s\neq j}^{p}\left(x_j-x_s\right)}, \tag{SA3}$$

where $\sigma_k\left(x_1,...,x_p\right)=\sum_{j_1<...<j_k}^{p}x_{j_1}...x_{j_k}$ denotes the elementary symmetric polynomial and $\sigma_k\left(x_{-j}\right)=\sigma_k\left(x_1,...,x_{j-1},x_{j+1},...,x_p\right)$.

Formula (SA3) implies that

$$\lim_{\tilde{A}\to A}\left(V_{\tilde{a}}^{-1}\right)_{ij}=0 \tag{SA4}$$

for $i\leq p-r$ and $j\leq r$. Indeed, for such i and j, $\sigma_{p-i}\left(\tilde{a}_{-j}\right)$ has degree larger than or equal to r, whereas the number of non-zero elements among $\tilde{a}_1,...,\tilde{a}_{j-1},\tilde{a}_j,...,\tilde{a}_p$ in the limit as $\tilde{A}\to A$ is $r-1$. From (SA2), (SA3), and (SA4), we get

$$\lim_{\tilde{a}\to a}\left(L^{-1}\right)_{ij}=\sum_{t=1}^{r}\frac{(p-t)!\sigma_{t-1}\left(\tilde{b}_{-i}\right)\sigma_{t-1}\left(a_{-j}\right)}{\prod_{s=1,s\neq i}^{p}\left(\tilde{b}_i-\tilde{b}_s\right)\prod_{s=1,s\neq j}^{p}\left(a_j-a_s\right)} \tag{SA5}$$

for $j\leq r$.

Using (SA1) and (SA5), we conclude that $\lim_{\tilde{A}\to A}\frac{\prod_{k=1}^{p-1}k!}{V_p\left(\tilde{A}\right)V_p\left(\tilde{B}\right)}\det_{1\leq i,j\leq p}\left(e^{\tilde{a}_i\tilde{b}_j}\right)$ equals the determinant of an $r\times r$ matrix G with

$$G_{ij}=\sum_{l=1}^{p}\sum_{t=1}^{r}\frac{e^{a_i\tilde{b}_l}(p-t)!\sigma_{t-1}\left(\tilde{b}_{-l}\right)\sigma_{t-1}\left(a_{-j}\right)}{\prod_{s=1,s\neq l}^{p}\left(\tilde{b}_l-\tilde{b}_s\right)\prod_{s=1,s\neq j}^{p}\left(a_j-a_s\right)}.$$

Using the identity $\sigma_{t-1}\left(\tilde{b}_{-l}\right)=\sigma_{t-1}\left(\tilde{b}\right)-\sigma_{t-2}\left(\tilde{b}_{-l}\right)\tilde{b}_l$ recursively, we get

$$\sigma_{t-1}\left(\tilde{b}_{-l}\right)=\sum_{u=1}^{t}(-1)^{u-1}\sigma_{t-u}\left(\tilde{b}\right)\tilde{b}_l^{u-1}.$$

Therefore, we can write

$$
\begin{aligned}
G_{ij} &= \sum_{l=1}^{p}\sum_{t=1}^{r}\sum_{u=1}^{t}\frac{e^{a_i\tilde{b}_l}(p-t)!(-1)^{u-1}\sigma_{t-u}\left(\tilde{b}\right)\tilde{b}_l^{u-1}\sigma_{t-1}(a_{-j})}{\prod_{s=1,s\neq l}^{p}\left(\tilde{b}_l-\tilde{b}_s\right)\prod_{s=1,s\neq j}^{p}(a_j-a_s)} \\
&= \sum_{t=1}^{r}\sum_{u=1}^{t}\frac{1}{2\pi\mathrm{i}}\oint_{\mathcal{K}}\frac{(-1)^{u-1}e^{a_iz}(p-t)!\sigma_{t-u}\left(\tilde{b}\right)z^{u-1}\sigma_{t-1}(a_{-j})}{\prod_{s=1}^{p}\left(z-\tilde{b}_s\right)\prod_{s=1,s\neq j}^{p}(a_j-a_s)}dz \\
&= \sum_{u=1}^{r}\frac{(-1)^{u-1}}{2\pi\mathrm{i}}\oint_{\mathcal{K}}\frac{e^{a_iz}z^{u-1}dz}{\prod_{s=1}^{p}\left(z-\tilde{b}_s\right)}\sum_{t=u}^{r}\frac{(p-t)!\sigma_{t-u}\left(\tilde{b}\right)\sigma_{t-1}(a_{-j})}{\prod_{s=1,s\neq j}^{p}(a_j-a_s)}
\end{aligned}
$$

where $\mathcal{K}$ is a contour in the complex plane that encircles counter-clockwise $\tilde{b}_1,...,\tilde{b}_p$.

Assuming that $\mathcal{K}$ is chosen so that $\tilde{b}_1,...,\tilde{b}_p$ and $b_1,...,b_p$ remain inside $\mathcal{K}$ as $\tilde{B}\to B$, we get

$$\lim_{\tilde{B}\to B}G = HK$$

where H and K are $r\times r$ matrices with

$$
\begin{aligned}
H_{ij} &= \frac{1}{2\pi\mathrm{i}}\oint_{\mathcal{K}}\frac{e^{a_iz}z^{j-1}dz}{\prod_{s=1}^{p}(z-b_s)} \text{ and} \\
K_{ij} &= \sum_{t=i}^{r}\frac{(-1)^{i-1}(p-t)!\sigma_{t-i}(b)\sigma_{t-1}(a_{-j})}{\prod_{s=1,s\neq j}^{p}(a_j-a_s)}.
\end{aligned}
$$

Let M be an $r\times r$ matrix with i,j-th element

$$M_{ij} = \frac{(-1)^{i-1}(p-i)!\sigma_{i-1}(a_{-j})}{\prod_{s=1,s\neq j}^{p}(a_j-a_s)}.$$

Note that the rows of matrix K are obtained from the corresponding rows of matrix M by adding linear combinations of other rows of M. Therefore,

$$
\begin{aligned}
\det K &= \det M = \frac{(-1)^{r(r-1)/2}\prod_{t=1}^{r}(p-t)!}{\prod_{t=1}^{r}\prod_{s=1,s\neq t}^{p}(a_t-a_s)}\det_{1\le i,j\le r}(\sigma_{i-1}(a_{-j})) \\
&= \frac{(-1)^{r(r-1)/2}\prod_{t=1}^{r}(p-t)!}{\prod_{t=1}^{r}a_t^{p-r}(-1)^{r(r-1)/2}(V_r(A))^2}\det_{1\le i,j\le r}(\sigma_{i-1}(a_{-j}))
\end{aligned}
$$

As shown on pp.41-42 of Macdonald (1995),

$$\det_{1\le i,j\le r}(\sigma_{i-1}(a_{-j})) = (-1)^{r(r-1)/2}V_r(\mathcal{A}),$$

where $V_r(\mathcal{A})$ is the Vandermonde determinant associated with $a_1,...,a_r$. Therefore,

$$\det K = \frac{(-1)^{r(r-1)/2}\prod_{t=1}^{r}(p-t)!}{\prod_{t=1}^{r}a_t^{p-r}V_r(\mathcal{A})},$$

and finally,

$$\lim_{\tilde{A}\to A,\tilde{B}\to B}\frac{\prod_{k=1}^{p-1}k!}{V_p\left(\tilde{A}\right)V_p\left(\tilde{B}\right)}\det_{1\leq i,j\leq p}\left(e^{\tilde{a}_i\tilde{b}_j}\right)=\det H\det K$$

$$=\frac{(-1)^{r(r-1)/2}\prod_{t=1}^{r}(p-t)!}{V_r(\mathcal{A})\prod_{t=1}^{r}a_t^{p-r}}\det_{1\leq i,j\leq r}\left(\frac{1}{2\pi\mathrm{i}}\oint_{\mathcal{K}}\frac{e^{a_iz}z^{j-1}dz}{\prod_{s=1}^{p}(z-b_s)}\right).\square$$

B Proof of formula (24)

We start as in the proof of (23), which is given in the main text. Lemma 5 in Onatski, Moreira and Hallin (2012) (OMH in what follows) implies that

$$\oint_{\mathcal{K}_i}e^{-nf_i(z)}g(z)\,\mathrm{d}z=e^{-nf_{i0}}\left[\frac{g(z_{i0})\,\pi^{1/2}}{f_{i2}^{1/2}n^{1/2}}+\frac{O_p(1)}{h_in^{3/2}}\right],\tag{SA6}$$

where $g(z)=\exp\left\{-\frac{1}{2}\Delta_p(z)\right\}$ and $O_p(1)$ is uniform in $h_i\in\left(0,\bar{h}\right]$. We would like to extend (SA6) to the case where $g(z)$ is replaced by a function of several variables. Such an extension allows us to derive an asymptotic expression for the repeated contour integral in (24).

It is convenient to rewrite the repeated integral on the left hand side of (24) in the form

$$\oint_{\mathcal{K}_{\rho(1)}}e^{-nf_{\rho(1)}(z_1)}\psi_{\rho1}(z_1)\,\gamma_{\rho1}(z_1)\ldots\oint_{\mathcal{K}_{\rho(r)}}e^{-nf_{\rho(r)}(z_r)}\psi_{\rho r}(z_r)\,\gamma_{\rho r}(z_1,...,z_r)\,\mathrm{d}z_r\ldots\mathrm{d}z_1,\tag{SA7}$$

where, denoting $\frac{h_i}{1+h_i}$ as χ_i, we have

$$\psi_{\rho j}(z)=z^{j-1}\left(1-\chi_{\rho(j)}\frac{z}{S}\right)^{-p(n-r)-r(r+1)/2}\exp\left\{-\Delta_p(z)-n\chi_{\rho(j)}z\right\},$$

and

$$\gamma_{\rho j}(z_1,...,z_j)=\left(1-\frac{\chi_{\rho(j)}z_j}{S-\chi_{\rho(j)}z_j}\frac{\sum_{i=1}^{j-1}\chi_{\rho(i)}z_i}{S-\sum_{i=1}^{j-1}\chi_{\rho(i)}z_i}\right)^{-p(n-r)-r(r+1)/2}.$$

Note that the innermost integral in (SA7) has form $\oint_{\mathcal{K}_i}e^{-nf_i(z_r)}g(z_r)\,\mathrm{d}z_r$ with $i=\rho(r)$ and $g(z_r)=\psi_{\rho r}(z_r)\,\gamma_{\rho r}(z_1,...,z_r)$ so that $g(z_r)$ now depends on the "additional variables" $z_1,...,z_{r-1}$.

A careful reading of OMH's proof of their Lemma 5 reveals that a version of (SA6) remains valid for general functions $g(z)$ that are analytic in the open ball $B(z_{i0},r_i)$ with center at z_{i0} and radius $r_i=\min\left\{z_{i0}-\max\left\{\bar{b}_p,\lambda_1\right\},\frac{1+h_i}{h_i}S-z_{i0}\right\}$ with probability approaching 1 as $n,p\to\infty$. Precisely, for such general $g(z)$ we have

$$\oint_{\mathcal{K}_i}e^{-nf_i(z)}g(z)\,\mathrm{d}z=e^{-nf_{i0}}\frac{g(z_{i0})\,\pi^{1/2}}{f_{i2}^{1/2}n^{1/2}}+\Psi_1+\Psi_2+\Psi_3\tag{SA8}$$

with

$$
\begin{aligned}
|\Psi_1| &< C_1 e^{-nf_{i0}} h_i^{-1} n^{-3/2} \sup_{z\in\bar{B}} |g(z)|, && \text{(SA9)}\\
|\Psi_2| &< C_1 e^{-nf_{i0}} e^{-nC_2} h_i^{-1} \sup_{z\in\mathcal{K}_{i1}\cup\bar{\mathcal{K}}_{i1}} |g(z)|, \text{ and} && \text{(SA10)}\\
|\Psi_3| &< C_1 \left| \oint_{\mathcal{K}_{i2}\cup\bar{\mathcal{K}}_{i2}} e^{-nf_i(z)} g(z)\,\mathrm{d}z \right|, && \text{(SA11)}
\end{aligned}
$$

where $\bar{B}$ is a closed ball with center at z_{i0} and radius $\frac{1}{2}r_i$, and C_1 and C_2 are positive constants. Moreover, for $g(z_r) = \psi_{\rho r}(z_r)\gamma_{\rho r}(z_1,...,z_r)$, Ψ_1, Ψ_2 and Ψ_3 are functions of the "additional variables" $z_1,...,z_{r-1}$ that are analytic in the region represented by a direct product of $r-1$ open balls $B\left(z_{\rho(i)0}, r_{\rho(i)}\right)$ with $i=1,...,r-1$ with probability approaching 1 as $n,p\to\infty$.

Now, Lemma A2 in OMH implies that

$$
\sup_{z_r\in\bar{B}\cup\mathcal{K}_{i1}\cup\bar{\mathcal{K}}_{i1}} \left|\psi_{\rho r}(z_r)\right| = h_{\rho(r)}^{1-r} O_p(1) \qquad \text{(SA12)}
$$

uniformly in $h_{\rho(r)} \in \left(0,\bar{h}\right]$. Further, an elementary analysis shows that

$$
\sup_{z_r\in\bar{B}\cup\mathcal{K}_{i1}\cup\bar{\mathcal{K}}_{i1}} \left|\gamma_{\rho r}(z_1,...,z_r)\right| = O_p(1) \qquad \text{(SA13)}
$$

uniformly in $h\in\left(0,\bar{h}\right]^r$ and in $(z_1,...,z_{r-1})\in\Omega_{r-1}$, where

$$
\Omega_{r-1} = \left\{z_1,...,z_{r-1} : \operatorname{Re} z_j < z_{\rho(j)0} + r_{\rho(j)} \text{ and } |\operatorname{Im} z_j| \le 3z_{\rho(j)0} \text{ for all } j\le r-1\right\}.
$$

Indeed, for $z_r\in\bar{B}\cup\mathcal{K}_{i1}\cup\bar{\mathcal{K}}_{i1}$, $h\in\left(0,\bar{h}\right]^r$, and for sufficiently large n and p, we have: $\operatorname{Re}\chi_{\rho(r)}z_r > 0$, $\operatorname{Re}\chi_{\rho(r)}z_r < 2\chi_{\rho(r)}z_{\rho(r)0} = 2\left(h_{\rho(r)}+c_p\right) < 2\left(1+\sqrt{c}\right)^2$, and $\left|\operatorname{Im}\chi_{\rho(r)}z_r\right| < 3\chi_{\rho(r)}z_{\rho(r)0} < 3\left(1+\sqrt{c}\right)^2$. Therefore, and since under the null hypothesis S is asymptotically equivalent to p, $\operatorname{Re}\frac{\chi_{\rho(r)}z_r}{S-\chi_{\rho(r)}z_r}$ is positive, and is of order $O_p\left(p^{-1}\right)$, whereas $\operatorname{Im}\frac{\chi_{\rho(r)}z_r}{S-\chi_{\rho(r)}z_r}$ is of order $O_p\left(p^{-1}\right)$ by absolute value. Similarly, for $(z_1,...,z_{r-1})\in\Omega_{r-1}$, $h\in\left(0,\bar{h}\right]^r$, and for sufficiently large n and p, $\operatorname{Re}\frac{\sum_{i=1}^{j-1}\chi_{\rho(i)}z_i}{S-\sum_{i=1}^{j-1}\chi_{\rho(i)}z_i}$ is either negative, or positive, and then, is of order $O_p\left(p^{-1}\right)$, whereas $\operatorname{Im}\frac{\sum_{i=1}^{j-1}\chi_{\rho(i)}z_i}{S-\sum_{i=1}^{j-1}\chi_{\rho(i)}z_i}$ is of order $O_p\left(p^{-1}\right)$ by absolute value. These estimates imply that, $\operatorname{Re}\left(\frac{\chi_{\rho(j)}z_j}{S-\chi_{\rho(j)}z_j}\frac{\sum_{i=1}^{j-1}\chi_{\rho(i)}z_i}{S-\sum_{i=1}^{j-1}\chi_{\rho(i)}z_i}\right)$ is smaller than a positive quantity of order $O_p\left(p^{-2}\right)$ (although, when negative, it may be large by absolute value), which implies (SA13).

The definition $g(z_r) = \psi_{\rho r}(z_r)\gamma_{\rho r}(z_1,...,z_r)$, along with (SA12) and (SA13) imply that

$$
\sup_{z_r\in\bar{B}\cup\mathcal{K}_{i1}\cup\bar{\mathcal{K}}_{i1}} |g(z_r)| = h_{\rho(r)}^{1-r} O_p(1), \qquad \text{(SA14)}
$$

where $O_p(1)$ is uniform in $h \in \left(0, \bar{h}\right]^r$ and in $(z_1, ..., z_{r-1}) \in \Omega_{r-1}$.

Next, by definition of $f_i(\cdot), \psi_{\rho r}(\cdot)$, and $\gamma_{\rho r}(\cdot)$,

$$e^{-nf_{\rho(r)}(z_r)} g(z_r) = \left(1 - \frac{\chi_{\rho(r)} z_r}{S - \sum_{i=1}^{r-1} \chi_{\rho(i)} z_i}\right)^{-p(n-r)-r(r+1)/2} z_r^{r-1} \prod_{j=1}^{p} (z_r - \lambda_j)^{-1}.$$

Since S is asymptotically equivalent to p, the base of the power representing the first term on the right hand side of the above equality is larger than or equal to 1 by absolute value for all z_r such that $\operatorname{Re} z_r < -3z_{\rho(j)0}$ and $|\operatorname{Im} z_r| \leq 3z_{\rho(j)0}$, uniformly in $h \in \left(0, \bar{h}\right]^r$ and in $(z_1, ..., z_{r-1}) \in \Omega_{r-1}$, for sufficiently large n and p with probability approaching 1. Since the exponent of the power is negative, the first term itself is no larger than 1 by absolute value for all z_r such that $\operatorname{Re} z_r < -3z_{\rho(j)0}$ and $|\operatorname{Im} z_r| \leq 3z_{\rho(j)0}$, uniformly in $h \in \left(0, \bar{h}\right]^r$ and in $(z_1, ..., z_{r-1}) \in \Omega_{r-1}$.

Let us split the contour $\mathcal{K}_{\rho(r)2}$ into parts $\mathcal{K}_{\rho(r)2-}$ and $\mathcal{K}_{\rho(r)2+}$, where $\mathcal{K}_{\rho(r)2-}$ includes all points of $\mathcal{K}_{\rho(r)2}$ with real part smaller than $-3z_{\rho(j)0}$. For sufficiently large n and p, we have with probability approaching 1,

$$\begin{aligned}
&\left| \oint_{\mathcal{K}_{\rho(r)2-} \cup \bar{\mathcal{K}}_{\rho(r)2-}} e^{-nf_{\rho(r)}(z_r)} g(z_r) \,\mathrm{d}z_r \right| \leq \\
&\oint_{\mathcal{K}_{\rho(r)2-} \cup \bar{\mathcal{K}}_{\rho(r)2-}} \left| z_r^{r-1} \prod_{j=1}^{p} (z_r - \lambda_j)^{-1} \,\mathrm{d}z_r \right| \leq 2^r \int_{-\infty}^{-3z_{\rho(j)0}} |x|^{r-1-p} \,\mathrm{d}x = \\
&\frac{2^r}{p-r} e^{(r-p)\ln\left(3z_{\rho(j)0}\right)} = e^{-nf_{\rho(j)0}} \frac{\left(6z_{\rho(j)0}\right)^r}{p-r} e^{-n\left[c_p \ln\left(3z_{\rho(j)0}\right) - f_{\rho(j)0}\right]} \leq \\
&e^{-nf_{\rho(j)0}} \frac{\left(6z_{\rho(j)0}\right)^r}{p-r} e^{-n\left[c_p \ln\left(z_{\rho(j)0}\right) - f_{\rho(j)0}\right]}.
\end{aligned}$$

On the other hand, as shown in the main text,

$$c_p \ln\left(z_{\rho(j)0}\right) - f_{\rho(j)0} > h_{\rho(j)} + c_p \tag{SA15}$$

for $h_{\rho(r)} \in \left(0, \bar{h}\right]$. Therefore,

$$\left| \oint_{\mathcal{K}_{\rho(r)2-} \cup \bar{\mathcal{K}}_{\rho(r)2-}} e^{-nf_{\rho(r)}(z_r)} g(z_r) \,\mathrm{d}z_r \right| = e^{-nf_{\rho(j)0}} h_{\rho(j)}^r O_p\left(e^{-\frac{n}{2}c}\right). \tag{SA16}$$

where $O_p\left(e^{-\frac{n}{2}c}\right)$ is uniform in $h \in \left(0, \bar{h}\right]^r$ and in $(z_1, ..., z_{r-1}) \in \Omega_{r-1}$.

For the integral over $\mathcal{K}_{\rho(r)2+} \cup \bar{\mathcal{K}}_{\rho(r)2+}$, we have

$$\begin{aligned}
&\left| \oint_{\mathcal{K}_{\rho(r)2+} \cup \bar{\mathcal{K}}_{\rho(r)2+}} e^{-nf_{\rho(r)}(z_r)} g(z_r) \,\mathrm{d}z_r \right| \leq 8z_{\rho(j)0} \sup_{z_r \in \mathcal{K}_{\rho(r)2+} \cup \bar{\mathcal{K}}_{\rho(r)2+}} \left| e^{-nf_{\rho(r)}(z_r)} g(z_r) \right| \leq \\
&8z_{\rho(j)0} \left(6z_{\rho(j)0}\right)^{r-1} e^{-p\ln\left(3z_{\rho(j)0}\right)} \sup_{z_r \in \mathcal{K}_{\rho(r)2+} \cup \bar{\mathcal{K}}_{\rho(r)2+}} \left| 1 - \frac{\chi_{\rho(r)} z_r}{S - \sum_{i=1}^{r-1} \chi_{\rho(i)} z_i} \right|^{-p(n-r)-r(r+1)/2}.
\end{aligned} \tag{SA17}$$

But for sufficiently large n and p, for $(z_1, ..., z_{r-1}) \in \Omega_{r-1}$ and an arbitrary small positive ε,

$$\inf_{z_r \in \mathcal{K}_{\rho(r)2+} \cup \bar{\mathcal{K}}_{\rho(r)2+}} \left| 1 - \frac{\chi_{\rho(r)} z_r}{S - \sum_{i=1}^{r-1} \chi_{\rho(i)} z_i} \right| = \min_{z_r = z_{\rho(r)0} \pm \mathrm{i} 3 z_{\rho(r)0}} \left| 1 - \frac{\chi_{\rho(r)} z_r}{S - \sum_{i=1}^{r-1} \chi_{\rho(i)} z_i} \right| > 1 - \frac{\chi_{\rho(r)} z_{\rho(j)0}}{(1-\varepsilon) S}.$$

Therefore,

$$\begin{aligned} &\sup_{z_r \in \mathcal{K}_{\rho(r)2+} \cup \bar{\mathcal{K}}_{\rho(r)2+}} \left| 1 - \frac{\chi_{\rho(r)} z_r}{S - \sum_{i=1}^{r-1} \chi_{\rho(i)} z_i} \right|^{-p(n-r)-r(r+1)/2} \\ \leq\ & e^{(-p(n-r)-r(r+1)/2) \ln\left(1 - \frac{\chi_{\rho(r)} z_{\rho(j)0}}{(1-\varepsilon)S}\right)} = e^{n \frac{\chi_{\rho(r)} z_{\rho(j)0}}{(1-\varepsilon)}} O_p(1), \end{aligned} \tag{SA18}$$

where $O_p(1)$ is uniform in $h \in \left(0, \bar{h}\right]^r$ and in $(z_1, ..., z_{r-1}) \in \Omega_{r-1}$.

Combining (SA17) and (SA18), we get

$$\begin{aligned} \left| \oint_{\mathcal{K}_{\rho(r)2+} \cup \bar{\mathcal{K}}_{\rho(r)2+}} e^{-n f_{\rho(r)}(z_r)} g(z_r) \, \mathrm{d} z_r \right| &\leq 2 \left(6 z_{\rho(j)0}\right)^r e^{-p \ln\left(3 z_{\rho(j)0}\right) + n \frac{\chi_{\rho(r)} z_{\rho(j)0}}{(1-\varepsilon)}} \\ &= 2 \left(6 z_{\rho(j)0}\right)^r e^{-n\left(c_p \ln\left(3 z_{\rho(j)0}\right) - \frac{h_{\rho(j)} + c_p}{(1-\varepsilon)}\right)} O_p(1). \end{aligned}$$

Since ε can be chosen arbitrarily small, (SA15) implies then that

$$\left| \oint_{\mathcal{K}_{\rho(r)2+} \cup \bar{\mathcal{K}}_{\rho(r)2+}} e^{-n f_{\rho(r)}(z_r)} g(z_r) \, \mathrm{d} z_r \right| = h_{\rho(j)}^{-r} e^{-n f_{\rho(j)0}} O_p\left(e^{-\frac{nc}{2}}\right), \tag{SA19}$$

where $O_p\left(e^{-\frac{nc}{2}}\right)$ is uniform in $h \in \left(0, \bar{h}\right]^r$ and in $(z_1, ..., z_{r-1}) \in \Omega_{r-1}$.

Finally, (SA8-SA11), (SA14), (SA16), and (SA19) imply that

$$\begin{aligned} &\oint_{\mathcal{K}_{\rho(r)}} e^{-n f_{\rho(r)}(z_r)} \psi_{\rho r}(z_r) \gamma_{\rho r}(z_1, ..., z_r) \, \mathrm{d} z_r = \\ &e^{-n f_{\rho(j)0}} \left(\frac{\psi_{\rho r}\left(z_{\rho(r)0}\right) \gamma_{\rho r}\left(z_1, ..., z_{r-1}, z_{\rho(r)0}\right) \pi^{1/2}}{f_{\rho(r)2}^{1/2} n^{1/2}} + \frac{O_p(1)}{h_{\rho(r)}^r n^{3/2}} \right), \end{aligned}$$

where $\gamma_{\rho r}\left(z_1, ..., z_{r-1}, z_{\rho(r)0}\right)$ and $O_p(1)$ are functions of $z_1, ..., z_{r-1}$ that are analytic in $(z_1, ..., z_{r-1}) \in \otimes_{i=1}^{r-1} B\left(z_{\rho(i)0}, r_{\rho(i)}\right)$ and bounded in probability uniformly in $h \in \left(0, \bar{h}\right]^r$ and in $(z_1, ..., z_{r-1}) \in \Omega_{r-1}$.

We now can use the uniform boundedness of $\gamma_{\rho r}\left(z_1, ..., z_{r-1}, z_{\rho(r)0}\right)$ and of $O_p(1)$ to repeat the above analysis for the second, third, etc inner integrals in (SA7). In the end, we obtain an asymptotic representation of (SA7) with the highest order term

$$\prod_{j=1}^{r} \left\{ e^{-n f_{\rho(j)0}} \frac{\psi_{\rho j}\left(z_{\rho(r)0}\right) \gamma_{\rho j}\left(z_{\rho(r)0}, ..., z_{\rho(r)0}\right) \pi^{1/2}}{f_{\rho(j)2}^{1/2} n^{1/2}} \right\},$$

and the second highest order term

$$\frac{1}{n}\prod_{j=1}^{r}\left\{e^{-nf_{\rho(j)0}}\frac{O_{pj}(1)}{h_{\rho(j)}^{j}n^{1/2}}\right\},$$

where $O_{pj}(1)$ are random variables that are bounded in probability uniformly in $h\in\left(0,\bar{h}\right]^{r}$.□

C Asymptotic power of some previously proposed tests

In this section, we consider examples of some of the tests of sphericity and of the equality of the covariance matrix to the identity matrix that have been proposed previously in the literature, and, in Proposition SA1, derive their asymptotic power functions.

Example 1 (John's (1971) test of sphericity) *John (1971) proposes testing the sphericity hypothesis $\theta=0$ against general alternatives using the test statistic*

$$U=\frac{1}{p}\operatorname{tr}\left[\left(\frac{\hat{\Sigma}}{(1/p)\operatorname{tr}\left(\hat{\Sigma}\right)}-I_p\right)^2\right], \tag{SA20}$$

where $\hat{\Sigma}$ is the sample covariance matrix of the data. He shows that, when $n>p$, such a test is locally most powerful invariant. Ledoit and Wolf (2002) study John's test in the case of real-valued data when $p/n\to c\in(0,\infty)$. They prove that, under the null, $nU-p\xrightarrow{d}N(1,4)$. In Proposition SA1, we show that, in the case of complex-valued data, under the null, $nU-p\xrightarrow{d}N(0,2)$. Hence the test with asymptotic size α rejects the null of sphericity whenever $\frac{1}{\sqrt{2}}(nU-p)>\Phi^{-1}(1-\alpha)$.

Example 2 (The Ledoit-Wolf (2002) test of $\Sigma=I$.) *Ledoit and Wolf (2002) propose to use*

$$W=\frac{1}{p}\operatorname{tr}\left[\left(\hat{\Sigma}-I\right)^2\right]-\frac{p}{n}\left[\frac{1}{p}tr\hat{\Sigma}\right]^2+\frac{p}{n} \tag{SA21}$$

as a test statistic for testing the hypothesis that the population covariance matrix is unity. They show that, in the case of real-valued data, under the null, $nW-p\xrightarrow{d}N(1,4)$. In Proposition SA1, we show that, in the case of complex-valued data, under the null, $nW-p\xrightarrow{d}N(0,2)$. As with the previous example, the null is rejected at asymptotic size α whenever $\frac{1}{\sqrt{2}}(nW-p)>\Phi^{-1}(1-\alpha)$.

Example 3 (The "corrected" LRT of Bai et al. (2009).) *When $n > p$, Bai et al. (2009) propose to use a corrected version $CLR = \operatorname{tr}\hat{\Sigma} - \ln\det\hat{\Sigma} - p - p\left(1 - \left(1 - \frac{n}{p}\right)\ln\left(1 - \frac{p}{n}\right)\right)$ of the likelihood ratio statistic based on the entire data, as opposed to λ or μ only, to test the equality of the population covariance matrix to the identity matrix against general alternatives. Under the null, for the case of real-valued data, $CLR \xrightarrow{d} N\left(-\frac{1}{2}\ln(1-c), -2\ln(1-c) - 2c\right)$, and for the case of complex-valued data, $CLR \xrightarrow{d} N(0, -\ln(1-c) - c)$ (still, as both n and p go to infinity, with p/n converging to c). The null hypothesis is rejected at asymptotic level α whenever CLR is larger than $(-\ln(1-c) - c)^{1/2}\,\Phi^{-1}(1-\alpha)$.*

Consider the tests described in Examples 1, 2 and 3, and denote by $\beta_J(h)$, $\beta_{LW}(h)$ and $\beta_{CLR}(h)$ their respective asymptotic powers at asymptotic level α.

Proposition SA1. *Let U and W be the test statistics defined in (SA20) and (SA21). Under the null,*

$$nU - p \xrightarrow{d} N(0,2) \text{ and } nW - p \xrightarrow{d} N(0,2). \tag{SA22}$$

Further, the asymptotic power functions of the tests described in Examples 1-3 satisfy

$$\beta_J(h) = \beta_{LW}(h) = 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \frac{1}{\sqrt{2}}\sum_{i=1}^{r}\frac{h_i^2}{c}\right), \text{ and} \tag{SA23}$$

$$\beta_{CLR}(h) = 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \sum_{j=1}^{r}\frac{h_j - \ln(1+h_j)}{\sqrt{-\ln(1-c) - c}}\right), \tag{SA24}$$

for any $h = (h_1, ..., h_r)$ such that $h_j < \sqrt{c}$ for $j = 1, ..., r$.

The asymptotic power functions of the tests from Examples 1, 2, and 3 are non-trivial. Figure 1 compares these power functions to the corresponding power envelopes for $r = 2$. Note that $\beta_{CLR}(h)$ depends on c. As c converges to one, $\beta_{CLR}(h)$ converges to α, which corresponds to the case of trivial power. As c converges to zero, $\beta_{CLR}(h)$ converges to $\beta_{LW}(h)$. In Figure 1, we provide a plot of $\beta_{CLR}(h)$ that correspond to $c = 0.5$.

Since John's test is invariant with respect to rotations and scalings of real-valued data, $\beta_J(h)$ is compared to the power envelope $\beta_\mu(h)$. The asymptotic power functions $\beta_{LW}(h)$ and $\beta_{CLR}(h)$ are compared to the power envelope $\beta_\lambda(h)$ because the Ledoit-Wolf test of $\Sigma = I$ and the "corrected" likelihood ratio test for real-valued data are invariant only with respect to the unitary transformations of the data. We see that the power of the tests in examples 1-3 is increasing very slowly and is very far below the corresponding power envelope. As discussed in the main test, such a comparison is somewhat unfair to the tests from examples 1-3 because these tests are designed to test the null hypothesis against general alternatives, as opposed to the spiked covariance alternative.

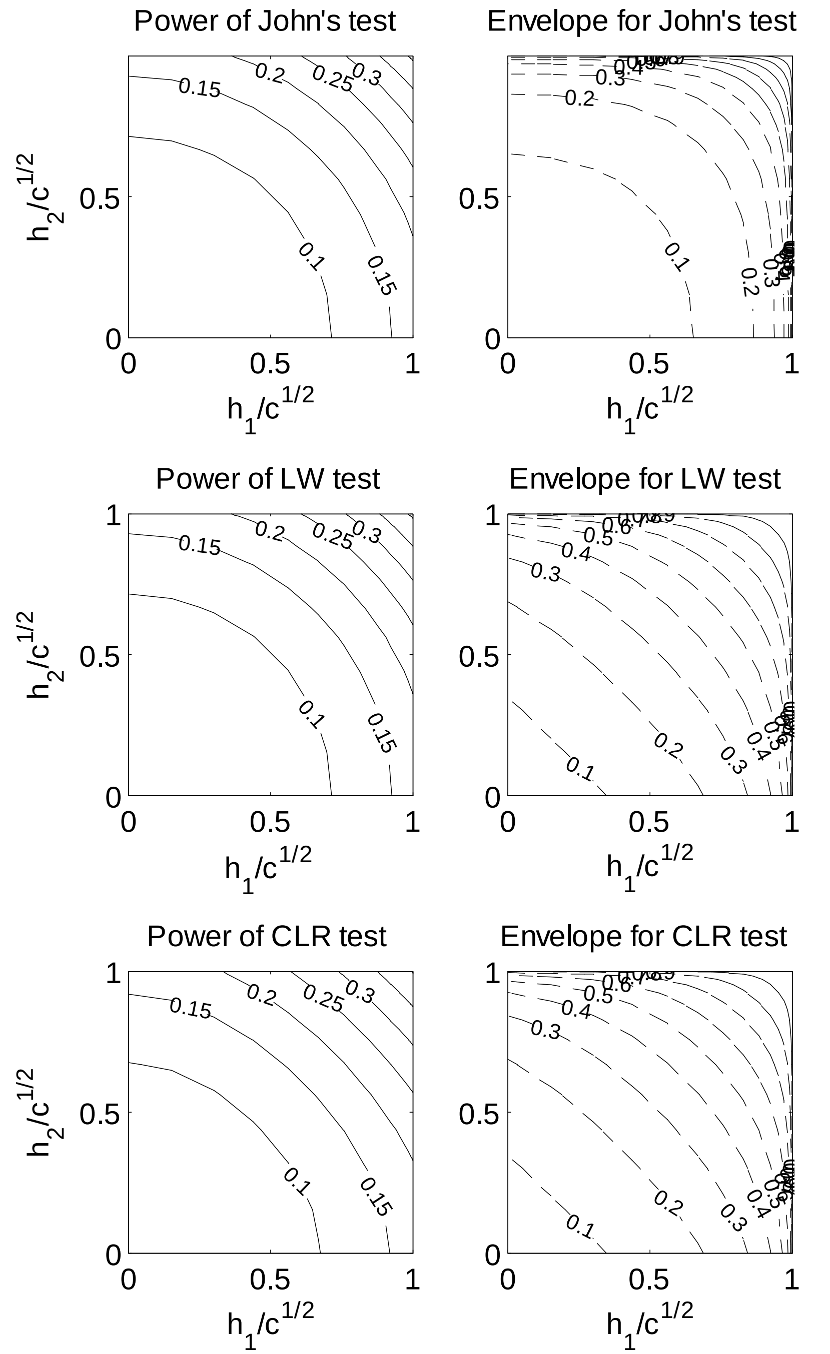

Figure 1: Asymptotic power of John's, Ledoit-Wolf and CLR tests (left panel) and the corresponding power envelopes (right panel); $r = 2, \alpha = 0.05$.

Proof of proposition SA1

Note that

$$nU - p = \frac{1}{c_p (S/p)^2} \left(T - (1 + c_p)\, p - (1 + c_p) \left(\frac{S}{p} + 1 \right) (S - p) \right) \tag{SA25}$$

and

$$nW - p = \frac{1}{c_p} \left(T - (1 + c_p)\, p - \left(2 + c_p \left(\frac{S}{p} + 1 \right) \right) (S - p) \right). \tag{SA26}$$

Using these expressions and Lemma A3, we get, using notations of Lemma A3,

$$nU - p \xrightarrow{d} N \left(0, \frac{1}{c^2} \left[\mathrm{Var}\zeta + 4(1+c)^2 \mathrm{Var}\eta - 4\,(1+c)\, \mathrm{Cov}\,(\zeta, \eta) \right] \right) \text{ and}$$
$$nW - p \xrightarrow{d} N \left(0, \frac{1}{c^2} \left[\mathrm{Var}\zeta + 4(1+c)^2 \mathrm{Var}\eta - 4\,(1+c)\, \mathrm{Cov}\,(\zeta, \eta) \right] \right).$$

Hence, $nU - p \xrightarrow{d} N(0, 2)$ and $nW - p \xrightarrow{d} N(0, 2)$.

Let us now derive (SA23). Theorem 1, Lemma A3 and (SA25) imply that the joint asymptotic distribution of $\frac{1}{\sqrt{2}}(nU - p)$ and $\mathcal{L}_\mu(h)$ under the null is Gaussian and, using notations of Lemma A3,

$$\begin{aligned} \mathrm{Cov}\left(\frac{nU - p}{\sqrt{2}}, \mathcal{L}_\mu(h) \right) &= -\frac{1}{c\sqrt{2}} \sum_{i=1}^{r} \mathrm{Cov}\,(\zeta, \xi_i) + \frac{2(1+c)}{c\sqrt{2}} \sum_{i=1}^{r} \mathrm{Cov}\,(\eta, \xi_i) \\ &\quad -\frac{1}{c^2\sqrt{2}} \mathrm{Cov}\,(\zeta, \eta) \sum_{j=1}^{r} h_j + \frac{2(1+c)}{c^2\sqrt{2}} \mathrm{Var}\,(\eta) \sum_{j=1}^{r} h_j \\ &= \frac{1}{c\sqrt{2}} \sum_{i=1}^{r} h_i^2 \end{aligned}$$

As we show above, under the null, $\frac{nU-p}{\sqrt{2}} \xrightarrow{d} N(0, 1)$. By Le Cam's third lemma, under the alternative $h = (h_1, ..., h_r)$, $\frac{nU-p}{\sqrt{2}}$ converges to a Gaussian random variable with the same variance but with mean equal to $\mathrm{Cov}\left(\frac{nU-p}{\sqrt{2}}, \mathcal{L}_\mu(h) \right)$. Hence, under the alternative, $\frac{nU-p}{\sqrt{2}} \xrightarrow{d} N\left(\frac{1}{c\sqrt{2}} \sum_{i=1}^{r} h_i^2, 1 \right)$, and the asymptotic power of John's test equals $1 - \Phi\left(\Phi^{-1}(1-\alpha) - \frac{1}{\sqrt{2}} \sum_{i=1}^{r} \frac{h_i^2}{c} \right)$.

Note that John's test is invariant with respect to both unitary transformations and arbitrary scalings of the data. Therefore, above we consider the joint distribution of $\frac{nU-p}{\sqrt{2}}$ and $\mathcal{L}_\mu(h)$, rather than that of $\frac{nU-p}{\sqrt{2}}$ and $\mathcal{L}_\lambda(h)$. In fact, as is easy to check, the asymptotic covariance between $\frac{nU-p}{\sqrt{2}}$ and $\mathcal{L}_\lambda(h)$ is the same as that between $\frac{nU-p}{\sqrt{2}}$ and $\mathcal{L}_\mu(h)$. So the asymptotic power of John's test does not depend on whether we specify σ^2 or not, as should be the case.

In contrast to John's test, Ledoit and Wolf's test of $\Sigma = I$ is not invariant with respect to scalings of the data. Therefore, we will first find the joint asymptotic distribution of $\frac{nW-p}{\sqrt{2}}$ and $\mathcal{L}_\lambda(h)$ under the null. Theorem 1, Lemma A3 and (SA26) imply that this distribution is Gaussian and

$$\begin{aligned} \operatorname{Cov}\left(\frac{nW-p}{\sqrt{2}}, \mathcal{L}_\lambda(h)\right) &= -\frac{1}{c\sqrt{2}} \sum_{i=1}^{r} \operatorname{Cov}(\zeta, \xi_i) + \frac{2(1+c)}{c\sqrt{2}} \sum_{i=1}^{r} \operatorname{Cov}(\eta, \xi_i) \\ &= \frac{1}{c\sqrt{2}} \sum_{i=1}^{r} h_i^2. \end{aligned}$$

Hence, similarly to the case of $\frac{nU-p}{\sqrt{2}}$, under the alternative, $\frac{nW-p}{\sqrt{2}} \xrightarrow{d} N\left(\frac{1}{c\sqrt{2}} \sum_{i=1}^{r} h_i^2, 1\right)$, and the asymptotic power of the Ledoit-Wold test equals $1-\Phi\left(\Phi^{-1}(1-\alpha) - \frac{1}{\sqrt{2}} \sum_{i=1}^{r} \frac{h_i^2}{c}\right)$.

Now, let us turn to the proof of (SA24). Note that $CLR = \sum_{j=1}^{p} q(\lambda_j) - p\int q(x)\mathrm{d}\mathcal{F}_p(x)$, where $q(x) = x - \ln x - 1$. Therefore, using the arguments of the proof of Lemma A3, we find that CLR and $-\Delta_p(z_{j0})$ jointly converge in distribution to a Gaussian vector with covariance

$$R_j = \frac{1}{4\pi^2} \oint \oint \frac{\ln(\bar{z}_{j0} - z_1)\, q(z_2)}{(\underline{m}(z_1) - \underline{m}(z_2))^2} \frac{\mathrm{d}\underline{m}(z_1)}{\mathrm{d}z_1} \frac{\mathrm{d}\underline{m}(z_2)}{\mathrm{d}z_2} \mathrm{d}z_1 \mathrm{d}z_2. \tag{SA27}$$

Here $\underline{m}(z)$ is as defined in (62), and the contours of integration are closed, oriented counterclockwise, enclose the support of the Marchenko-Pastur distribution with parameter $c < 1$, and do not enclose $\bar{z}_{j0}$. Further, we will choose such contours so that the z_1-contour encloses 0, but the z_2-contour does not.

From our Theorem 1, the asymptotic covariance between CLR and $\mathcal{L}_\lambda(h)$ under the null equals $\sum_{j=1}^{r} R_j$. Let us find the value of R_j as a function of h_j. Using Formula 1.16 of Bai and Silverstein (2004) we can simplify (SA27) to get

$$R = \frac{1}{4\pi^2} \oint \oint \frac{\ln(\bar{z}_{j0} - z(m_1))\,(z(m_2) - \ln z(m_2) - 1)}{(m_1 - m_2)^2} \mathrm{d}m_1 \mathrm{d}m_2,$$

where $z(m) = -\frac{1}{m} + \frac{c}{1+m}$ and the contours of integration over m_1 and over m_2 are obtained from the contours of integration over z_1 and z_2 in (SA27) by the transformation $\underline{m}(z)$. In particular, m_1-contour is oriented clockwise and encloses $-\frac{h_j}{h_j+c}$ and 0 but not -1 and $-\frac{1}{1+h_j}$, whereas m_2-contour is oriented counterclockwise and encloses $\frac{1}{c-1}$ and -1 but not $-\frac{h_j}{h_j+c}$ and 0.

Using (67), we can write: $R_j = R_{1j} + R_{2j} + R_{3j}$, where

$$\begin{aligned}
R_{1j} &= \frac{2\pi \mathrm{i}}{4\pi^2} \oint \left(-\frac{1}{m_2} + \frac{1}{m_2 + h_j (h_j + c)^{-1}} \right) z(m_2)\, \mathrm{d}m_2, \\
R_{2j} &= \frac{-2\pi \mathrm{i}}{4\pi^2} \oint \left(-\frac{1}{m_2} + \frac{1}{m_2 + h_j (h_j + c)^{-1}} \right) \ln z(m_2)\, \mathrm{d}m_2, \text{ and} \\
R_{3j} &= \frac{-2\pi \mathrm{i}}{4\pi^2} \oint \left(-\frac{1}{m_2} + \frac{1}{m_2 + h_j (h_j + c)^{-1}} \right) \mathrm{d}m_2.
\end{aligned}$$

Since $-\frac{1}{m_2} + \frac{1}{m_2 + h_j(h_j+c)^{-1}}$ is analytic in the area enclosed by the m_2-contour, $R_{3j} = 0$. Further, using Cauchy's theorem and the fact that $z(m_2) = -\frac{1}{m_2} + \frac{c}{1+m_2}$, we get: $R_{1j} = h_j$. Finally, applying integration by parts formula to R_{2j}, and using the fact that $\ln z(m_2)$ is a single-valued function on the m_2-contour, we get

$$R_{2j} = \frac{2\pi \mathrm{i}}{4\pi^2} \oint \frac{\frac{1}{m_2^2} - \frac{c}{(1+m_2)^2}}{-\frac{1}{m_2} + \frac{c}{m_2+1}} \left(-\ln m_2 + \ln \left(m_2 + h_j (h_j + c)^{-1} \right) \right) \mathrm{d}m_2.$$

The integrand in the above integral has only two singularities in the area enclosed by the m_2-contour: a pole at $\frac{1}{c-1}$ and a pole at -1. Therefore, by Cauchy's residue theorem, we get $R_{2j} = -\ln(1 + h_j)$. To summarize, $R_j = R_{1j} + R_{2j} + R_{3j} = h_j - \ln(1 + h_j)$.

Now, from Bai et al. (2009), we know that under the null,

$$CLR \xrightarrow{d} N(0, -\ln(1-c) - c).$$

By Le Cam's third lemma, under the alternative $h = (h_1, ..., h_r)$, CLR converges to a Gaussian random variable with the same variance but with mean equal to $\sum_{j=1}^{r} R_j$. Hence, under the alternative,

$$CLR \xrightarrow{d} N \left(\sum_{j=1}^{r} [h_j - \ln(1 + h_j)], -\ln(1-c) - c \right).$$

Therefore, the power of the "corrected" likelihood ratio test of asymptotic size α equals $1 - \Phi\left(\Phi^{-1}(1-\alpha) - \sum_{j=1}^{r} \frac{h_j - \ln(1+h_j)}{\sqrt{-\ln(1-c) - c}} \right)$.$\square$

References

[1] Bai, Z.D., Jiang, D., Yao, J.F., and S. Zheng (2009) "Corrections to LRT on Large-dimensional Covariance Matrix by RMT", Annals of Statistics 37, 3822-3840.

[2] Bai, Z.D. and J.W. Silverstein (2004) "CLT for Linear Spectral Statistics of Large-Dimensional Sample Covariance Matrices", Annals of Probability 32, 553-605.

[3] John, S. (1971). "Some optimal multivariate tests", Biometrika 58, 123–127.

[4] Klinger, A. (1967) "The Vandermonde Matrix", *The American Mathematical Monthly* 74, 571-574.

[5] Ledoit, O. and M. Wolf (2002) "Some Hypothesis Tests for the Covariance Matrix when the Dimension is Large Compared to the Sample Size", Annals of Statistics 30, 1081-1102.

[6] Macdonald, I. G. (1995) Symmetric functions and Hall polynomials. Second edition. Oxford Mathematical Monographs. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York.

[7] Onatski, A., Moreira, M. J., and Hallin, M. (2012) "Asymptotic Power of Sphericity Tests for High-dimensional Data", manuscript, Economics Faculty, University of Cambridge.