A WHITENESS TEST BASED ON THE SPECTRAL MEASURE OF LARGE NON-HERMITIAN RANDOM MATRICES

A. Bose[*]

Indian Statistical Institute, Kolkata

W. Hachem[†]

CNRS / LIGM (UMR 8049), Université Paris-Est Marne-la-Vallée, ESIEE, ENPC

ABSTRACT

In the context of multivariate time series, a whiteness test against an MA(1) correlation model is proposed. This test is built on the eigenvalue distribution (spectral measure) of the non-Hermitian one-lag sample autocovariance matrix, instead of its singular value distribution. The large dimensional limit spectral measure of this matrix is derived. To obtain this result, a control over the smallest singular value of a related random matrix is provided. Numerical simulations show the excellent performance of this test.

***Index Terms*—** Antenna array processing, Large non-Hermitian matrix theory, Limit spectral distribution, Smallest singular value, Whiteness test in multivariate time series.

1. INTRODUCTION

In the fields of wireless communications, Radar, Sonar, wideband antenna array processing, or diagnotic checking, testing whether the observed multidimensional signal is spatially and temporally a white noise is often a challenging problem. Given a positive integer N, consider the $\mathbb{C}^N$–valued time series $(y_k)_{k\in\mathbb{Z}}$ given as

$$y_k = B_0 x_k + B_1 x_{k-1},$$

where B_0 and B_1 are $\mathbb{C}^{N\times N}$ deterministic parameter matrices, and where the $\mathbb{C}^N$–valued random process $(x_k)_{k\in\mathbb{Z}}$ is such that, by writing $x_k = [x_{k,0}, \ldots, x_{k,N-1}]^\mathsf{T}$, the random variables $\{x_{k,\ell}\}_{k=-\infty,\ell=0}^{+\infty,N-1}$ are centered, independent and identically distributed. Assuming that the observer has access to a finite sequence $(y_0, \ldots, y_{n-1})$ of n samples, the problem that we tackle in this paper is to test the null (white noise) hypothesis

$$\mathbf{H0}: \quad B_0 = I_N, B_1 = 0,$$

against the alternative

$$\mathbf{H1}: \quad B_0 = I_N, B_1 \neq 0 \text{ and unknown.}$$

[*]Thanks to the Indo-French Center for Applicable Mathematics (IFCAM), project High Dimensional Random Matrix Models with Applications.

[†]Thanks to the IFCAM and to the French ANR, grant HIDITSA (ANR-17-CE40-0003).

The classical algorithms for performing whiteness tests, such as the Box-Pierce or the Ljung-Box tests in the univariate case $N = 1$, or their multivariate counterparts such as the Li-McLeod test, are built on the sample autocovariance matrices $\widehat{\Gamma}_\ell = n^{-1}\sum_{k=\ell}^{n-1} y_k y_{k-\ell}^*$, for the lags $\ell = 0, 1$. Most of these tests are well-suited to the case where the signal dimension N is much smaller than the window length n. Accordingly, their performance is usually studied in the asymptotic regime where N is fixed while $n \to \infty$.

In many modern signal processing applications such as the massive MIMO communications, or the large antenna array processing for radioastonomy, it happens that the signal dimension is large and comparable to the window length. In such situations, it appears more convenient to consider the asymptotic regime where $N, n \to \infty$ in such a way that N/n converges to a constant $\gamma > 0$, and to design tests that are adapted to this asymptotic regime by appealing to the Large Random Matrix Theory (LRMT). This route was taken in, *e.g.*, [1, 2]. We consider herein the same asymptotic regime that we shall denote as "$n \to \infty$".

Obvioulsy, the matrix $\widehat{\Gamma}_1$ is non-Hermitian. As is well known, most of the LRMT results pertaining to the spectral behavior of large random matrices concern Hermitian or symmetric matrices. So far, LRMT results related with our problem are known for certain symmetrized versions of these matrices, such as $\widehat{\Gamma}_1\widehat{\Gamma}_1^*$. The tests that are proposed in [1, 2] and in the other papers in the same strain, rely on the spectral behavior of such Hermitian matrices.

In this paper, we get rid of the symmetrization and rather develop a test that is based on the eigenvalue distribution of the one-lag sample autocovariance matrix $\widehat{R}_1 = n^{-1}\sum_{k=0}^{n-1} y_k y_{k-1}^*$, where the sum is taken this time modulo-n (this choice will be justified below). Denote as

$$\mu_n = \frac{1}{N}\sum_{i=0}^{N-1} \delta_{\lambda_i(\widehat{R}_1)}$$

the spectral measure of the matrix $\widehat{R}_1$, where $\{\lambda_i(M)\}_{i=0}^{N-1}$ denote the eigenvalues of a matrix $M \in \mathbb{C}^{N\times N}$. Note that this random probability measure is supported by $\mathbb{C}$, due to the non-Hermitian nature of $\widehat{R}_1$. One of the main results of this paper, Theorem 1 below, shows that under **H0**, the measure μ_n converges weakly in probability as $n \to \infty$ towards a deterministic probability measure $\boldsymbol{\mu}$ that can be determined (the weak convergence in probability amounts to the convergence $\int \varphi d\mu_n \to_{n\to\infty} \int \varphi d\boldsymbol{\mu}$ in probability for each continuous and bounded function $\varphi : \mathbb{C} \to \mathbb{R}$; we use the notation $\mu_n \Rightarrow \boldsymbol{\mu}$ in probability to refer to this fact).

Based on this result, we propose a whiteness test that is based on the 2-Wasserstein distance between the spectral measure of $\widehat{R}_1$ and $\boldsymbol{\mu}$. By simulation, this test reveals a much better performance than more classical tests which are based on Hermitian matrices built from the observations. A non-rigorous justification of this performance improvement is that when performing an eigenvalue-based test, we take advantage of the higher sensitivity of the eigenvalues of a matrix with respect to perturbations as compared to its singular values. This intuition needs to be corroborated rigorously by first establishing the consistency of our test in the regime $n \to \infty$, and then, by attempting to evaluate its power. This is left for a future work.

Theorem 1 is stated and discussed in Section 2. Simulations are provided in Section 3. The main steps of the proof of Theorem 1 are provided in Section 4. Applying a now well-known methodology for studying the spectral behavior of large non Hermitian matrices (see [3]), one crucial step in the proof consists in controlling the smallest singular value of the matrix $\widehat{R}_1 - zI_N$ under **H0** for almost all $z \in \mathbb{C}$. Inspired by the technique developed in [4], we provide a more general result that amounts to controlling the smallest singular value of the matrix $n^{-1}XAX^* - zI_N$, where $X = \begin{bmatrix} x_0 & \cdots & x_{n-1} \end{bmatrix} \in \mathbb{C}^{N\times n}$, the matrix A is an arbitrary deterministic matrix such that $\|A\|$ and $\|A^{-1}\|$ are bounded as $n \to \infty$, with $\|\cdot\|$ being the spectral norm, and $z \in \mathbb{C} \setminus \{0\}$ (Theorem 2). Indeed, it is easy to see that $\widehat{R}_1 = n^{-1}XJX^*$ under **H0**, where J is the circulant matrix

$$J = \begin{bmatrix} 0 & & & 1 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \in \mathbb{R}^{n\times n},$$

that obviously satisfies the assumptions that we put on A. The control of the smallest singular value of $n^{-1}XAX^* - zI_N$ has an interest of its own, and can be used among other things to study the spectral measures of non Hermitian matrices of the type $n^{-1}XAX^*$. The details of the proof are provided in [5].

2. LIMIT SPECTRAL MEASURE OF $\widehat{R}_1$ UNDER H0

Our standing assumptions on the elements of the matrix $X = \begin{bmatrix} x_{ij} \end{bmatrix}_{i,j=0}^{N-1,n-1}$, which can depend on n in general, are the following:

Assumption 1. The random variables $\{x_{ij}\}_{i,j=0}^{N-1,n-1}$ are complex, independent and identically distributed with $\mathbb{E}x_{00} = 0$, $\mathbb{E}|x_{00}|^2 = 1$, $\mathbb{E}|x_{00}|^4 \le \boldsymbol{m}_4$ where $\boldsymbol{m}_4$ is independent of n, and $\sup_n |\mathbb{E}x_{00}^2| < 1$.

Note that the last statement excludes the random variables x_{ij} from being real. The generalization of the results of Theorem 1 to the real case is under study.

To state our result regarding the limit spectral measure of $n^{-1}XJX^*$, we need the following function. Recall that $N/n \to \gamma > 0$. Define

$$g(y) = \frac{y}{y+1}(1-\gamma+2y)^2, \ (0 \vee (\gamma - 1)) \le y \le \gamma.$$

Then g^{-1} exists on the interval $[0 \vee ((\gamma-1)^3/\gamma), \gamma(\gamma+1)]$ and maps it to $[0 \vee (\gamma - \gamma^{-1}), \gamma]$. It is an analytic increasing function on the interior of the interval.

Theorem 1. Suppose Assumption 1 holds true. Then, there exists a deterministic probability measure $\boldsymbol{\mu}$ such that the spectral measure μ_n of $n^{-1}XJX^*$ satisfies $\mu_n \Rightarrow \boldsymbol{\mu}$ in probability as $n \to \infty$. The limit measure $\boldsymbol{\mu}$ is rotationally invariant on $\mathbb{C}$. Let $F(r) = \boldsymbol{\mu}(\{z \in \mathbb{C} : |z| \le r\}), \ 0 \le r < \infty$ be the distribution function of the radial component.
If $\gamma \le 1$, then

$$F(r) = \begin{cases} \gamma^{-1}g^{-1}(r^2) & \text{if } 0 \le r \le \sqrt{\gamma(\gamma+1)}, \\ 1 & \text{if } r > \sqrt{\gamma(\gamma+1)}. \end{cases}$$

If $\gamma > 1$, then

$$F(r) = \begin{cases} 1-\gamma^{-1} & \text{if } 0 \le r \le (\gamma-1)^{3/2}/\sqrt{\gamma}, \\ \gamma^{-1}g^{-1}(r^2) & \text{if } \frac{(\gamma-1)^{3/2}}{\sqrt{\gamma}} < r \le \sqrt{\gamma(\gamma+1)}, \\ 1 & \text{if } r > \sqrt{\gamma(\gamma+1)}. \end{cases}$$

The theorem implies that the support of $\boldsymbol{\mu}$ is the disc $\{z : |z| \le \sqrt{\gamma(\gamma+1)}\}$ when $\gamma \le 1$, and when $\gamma > 1$, it is the ring $\{z : (\gamma-1)^{3/2}/\sqrt{\gamma} \le |z| \le \sqrt{\gamma(\gamma+1)}\}$ together with the point $\{0\}$ where there is a mass $1-\gamma^{-1}$.

Moreover, $F(r)$ has a positive and analytical density on the open interval $(0 \vee \text{sign}(\gamma-1)|\gamma-1|^{3/2}/\sqrt{\gamma}, \sqrt{\gamma(\gamma+1)})$. A closer inspection of g shows that this density is bounded if $\gamma \ne 1$. If $\gamma = 1$, then the density is bounded everywhere except when $r \downarrow 0$. A cumbersome closed form expression for g^{-1} (and hence for $F(\cdot)$) can be obtained by calculating the root of a third degree polynomial.

3. THE WHITENESS TEST

By Theorem 1, the spectral measure of $\widehat{R}_1$ converges weakly in probability under **H0** to the probability measure $\boldsymbol{\mu}$. This fact can be used to design a whiteness test that is based on a distance between the spectral measure of $\widehat{R}_1$ and $\boldsymbol{\mu}$. We chose here the 2-Wasserstein distance between these two distributions. For the sake of comparison with more classical tests, we also considered the test which consists in comparing $N^{-1} \operatorname{tr} \widehat{R}_1 \widehat{R}_1^*$ to a threshold. We denote these two tests as T1 and T2 respectively. Note that T2 is a singular value based test. To get a more complete picture of the problem, we also considered a third test which is based on the eigenvalue distribution of the Hermitian sample covariance matrix

$$\widehat{R}_{0,1} = \frac{1}{n} \sum_{k=0}^{n-1} \begin{bmatrix} y_k \\ y_{k-1} \end{bmatrix} \begin{bmatrix} y_k^* & y_{k-1}^* \end{bmatrix} .$$

It is known that under **H0**, the spectral distribution of this matrix, whose support lies in $[0, \infty)$, converges weakly in the almost sure sense to the well-known Marchenko-Pastur distribution $\mathrm{MP}_{2\gamma}$ with parameter 2γ. This suggests the use of the 2-Wasserstein distance between the spectral measure of $\widehat{R}_{0,1}$ and $\mathrm{MP}_{2\gamma}$. We denote the resulting test as T3.

The ROC curves for these three tests are plotted on Figures 1 and 2. The tests T1 and T3 were implemented by sampling $\boldsymbol{\mu}$ and $\mathrm{MP}_{2\gamma}$ from the spectra of two large random matrices and by using the `transport` library of the `R` software. For Figure 1, we chose $B_1^{(n)} = \alpha I_N$, while for Figure 2, the elements b_{ij} of $B_1^{(n)}$ are chosen as $b_{ij} = \alpha' \exp(-8|i-j|/N)$, where α and α' are non-zero real numbers.

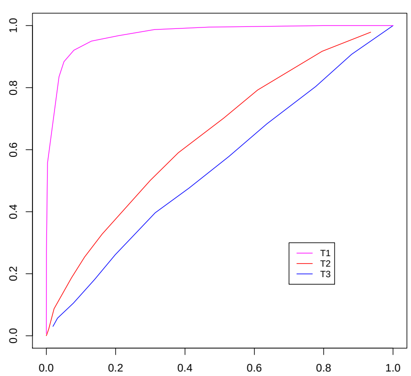

Fig. 1. ROC curves. Setting: $B_1^{(n)} = \alpha I_N$ with $\alpha^2 = 10^{-2.5}$, $(N, n) = (50, 100)$.

These figures show that T1 clearly outperforms T2 and T3. This tends to corroborate the intuition that the eigenvalue sensitivity alluded to earlier, can be beneficial when it comes to designing white noise tests.

To better understand the behavior of the eigenvalue-based tests, the next step would be to study the spectral distribution of $\widehat{R}_1$ under **H1**. This quite non-trivial task is left for future research.

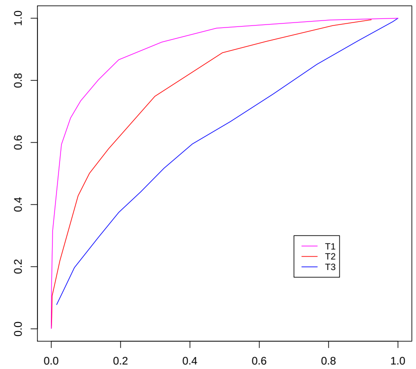

Fig. 2. ROC curves. Setting: $B_1^{(n)}$ is a Toeplitz matrix with $\operatorname{tr} B_1^{(n)} (B_1^{(n)})^* / N = 10^{-2}$, $(N, n) = (50, 100)$.

4. MAIN STEPS OF THE PROOF OF THEOREM 1

4.1. General approach

The now well-established general approach for characterizing the asymptotic behavior of the spectral measure of a random non-Hermitian matrix can be summarized as follows. The reader is referred to, *e.g.*, the tutorial [3] for a detailed exposition of this technique.

A probability measure π on $\mathbb{C}$ that integrates $\log |\cdot|$ near infinity can be identified through its logarithmic potential, which is the function

$$U_\pi : \mathbb{C} \to (-\infty, \infty], \quad z \mapsto U_\pi(z) = -\int \log |\lambda - z| \, \pi(d\lambda)$$

Denote as $s_0(M) \geq \cdots \geq s_{n-1}(M)$ the singular values of a matrix $M \in \mathbb{C}^{n \times n}$. The logarithmic potential U_{μ_n} of the spectral measure μ_n is given as

$$\begin{aligned} & U_{\mu_n}(z) \\ & = -\frac{1}{N} \sum \log |\lambda_i(\widehat{R}_1) - zI| = -\frac{\log |\det(\widehat{R}_1 - zI)|}{N} \\ & = -\frac{\log \det(\widehat{R}_1 - zI)(\widehat{R}_1 - zI)^*}{2N} = -\int \log \lambda \, \nu_{n,z}(d\lambda), \end{aligned}$$

where the random probability measure

$$\nu_{n,z} = \frac{1}{N} \sum_{i=0}^{N-1} \delta_{s_i(\widehat{R}_1 - zI)}$$

is the singular value distribution of $\widehat{R}_1 - zI$. This passage from the eigenvalues of $\widehat{R}_1$ to the singular values of the matrices $\widehat{R}_1 - zI$, which can be studied more easily, is known as the hermitization technique.

To evaluate the asymptotic behavior of μ_n under **H0**, we need to study the asymptotic behavior of $U_{\mu_n}(z)$ for almost every $z \in \mathbb{C}$. To that end, the three following steps are required:

1. For almost every $z \in \mathbb{C}$, there exists a deterministic probability measure $\boldsymbol{\nu}_z$ such that $\nu_{n,z} \Rightarrow \boldsymbol{\nu}_z$ in probability.

2. $\log$ is uniformly integrable in probability with respect to the sequence $\nu_{n,z}$ (see [3] for a definition). Since $\log$ is unbounded near zero, this step is the most challenging one.

Once these two steps are performed, it is known that there exists a probability measure $\boldsymbol{\mu}$ with the logarithmic potential $U_{\boldsymbol{\mu}}(z) = \int \log \lambda \, \boldsymbol{\nu}_z(d\lambda)$, and such that $\mu_n \Rightarrow \boldsymbol{\mu}$ in probability.

3. Identify $\boldsymbol{\mu}$ from $U_{\boldsymbol{\mu}}$.

Following a standard procedure in the field of LRMT, Step 1 mainly consists in studying the resolvent $Q_z(\eta) = (\Sigma_z - \eta I_{2N})^{-1}$, where Σ_z is the Hermitian matrix

$$\Sigma_z = \begin{bmatrix} & \widehat{R}_1 - zI_N \\ \widehat{R}_1^* - \bar{z}I_N & \end{bmatrix},$$

and η is a complex number such that $\Im\eta > 0$.

Due to the unboundedness of the $\log$ near zero, to perform Step 2, one essentially needs to show that for each $z \in \mathbb{C} \setminus \{0\}$, the smallest singular value $s_{N-1}(\widehat{R}_1 - zI_N)$ cannot not converge to zero too quickly. As said in the introduction, we control the behavior of this smallest singular value for a more general matrix model than $\widehat{R}_1 - zI_N$:

Theorem 2. Let Assumption 1 hold true. Let $A \in \mathbb{C}^{n\times n}$ be a deterministic matrix such that

$$\sup_n \|A\| < \infty \text{ and } \sup_n \|A^{-1}\| < \infty.$$

Then, there exist $\alpha, \beta > 0$ such that for each $C > 0$, $t > 0$, and $z \in \mathbb{C} \setminus \{0\}$,

$$\mathbb{P}\left[s_{N-1}(n^{-1}XAX^* - zI_N) \le t,\ \|X\| \le C\right] \le c\left(n^{\alpha}t^{1/2} + n^{-\beta}\right),$$

where the constant $c > 0$ depends on C, z, and $\boldsymbol{m}_4$ only.

The principle of the proof of this theorem is sketched below.

Step 3 is based on the fact that a probability measure π on $\mathbb{C}$ can be recovered from its logarithmic potential by the relation $\pi = -(2\pi)^{-1}\Delta U_\pi$, where Δ is the distributional Laplace operator.

4.2. Proof principle of Theorem 2

The general idea of the proof is the following. Defining the matrix

$$H = \begin{bmatrix} A^{-1} & n^{-1/2}X^* \\ n^{-1/2}X & zI_N \end{bmatrix} \in \mathbb{C}^{(N+n)\times(N+n)},$$

and using the inversion formula for partitioned matrices, one can check that $\|(n^{-1}XAX^* - zI_N)^{-1}\| \le \|H^{-1}\|$. Thus, the problem can be reduced to controlling the smallest singular value $s_{N+n-1}(H)$ of H. As is well known,

$$s_{N+n-1}(H) = \min_{u\in\mathbb{S}^{N+n-1}} \|Hu\|,$$

where $\mathbb{S}^{N+n-1}$ is the unit-sphere of $\mathbb{C}^{N+n}$. Invoking an idea that has been frequently used in the non-Hermitian LRMT literature, we partition $\mathbb{S}^{N+n-1}$ into two sets of *compressible*, *i.e.*, close to sparse, and *incompressible* vectors.

The infimum of $\|Hu\|$ over the set of compressible vectors is the easier one to handle. Given a fixed vector $u \in \mathbb{S}^{N+n-1}$, we first show that $\mathbb{P}[\|Hu\| \le t]$ is exponentially small in n. Then, since the compressible vectors are close to being sparse, the set of compressible vectors has an ε-net with a controlled cardinality, and we are done with a union bound argument.

The set of incompressible vectors is much harder to manage, since the ε-net fails for this set. The idea here is that when u is incompressible, Hu is close to a sum of $\mathcal{O}(n)$ columns of H with comparable weights. The problem is thus reduced to controlling the distance between an arbitrary column of H and the subspace generated by the other columns. The methodology developed in [4] can be used here with a substantial adaptation to our model, which contains more structure than the one studied in this reference.

5. REFERENCES

[1] M. Bhattacharjee and A. Bose, *Large Covariance and Autocovariance Matrices.*, Chapman & Hall/CRC, Boca Raton, London, New York, 2018.

[2] Z. Li, C. Lam, J. Yao, and Q. Yao, "On testing for high-dimensional white noise," *arXiv preprint arXiv:1808.03545*, 2018.

[3] Ch. Bordenave and D. Chafaï, "Around the circular law," *Probab. Surv.*, vol. 9, pp. 1–89, 2012.

[4] R. Vershynin, "Invertibility of symmetric random matrices," *Random Structures & Algorithms*, vol. 44, no. 2, pp. 135–182, 2014.

[5] A. Bose and W. Hachem, "Smallest singular value and limit eigenvalue distribution of a class of non-Hermitian random matrices with statistical application," *arXiv preprint arXiv:1812.07237*, 2018.