SPECTRAL DISTRIBUTION OF THE EXPONENTIALLY WINDOWED SAMPLE COVARIANCE MATRIX

Ehsan Yazdian, Mohammad Hassan Bastani

Electrical Engineering Department
Sharif University of Technology
Tehran, IRAN

Saeed Gazor

Electrical and Computer Engineering
Queens University
Kingston , Ontario

ABSTRACT

In this paper, we investigate the effect of applying an exponential window on the limiting spectral distribution (l.s.d.) of the exponentially windowed sample covariance matrix (SCM) of complex array data. We use recent advances in random matrix theory which describe the distribution of eigenvalues of the doubly correlated Wishart matrices. We derive an explicit expression for the l.s.d. of the noise-only data. Simulations are performed to support our theoretical claims.

Index Terms— Array signal processing, Limiting Spectral Distribution, Random Matrix Theory

1. INTRODUCTION

The distribution of the eigenvalues of the sample covariance matrix of data has important impact on the performance of signal processing algorithms. Our knowledge about the distribution of eigenvalues and eigenvectors of complex Wishart matrices and their limiting behavior is emerging as a key tool in a number of applications such as source identification and analysis of wireless MIMO channels [1–4].

Let $\mathbf{X} = [X_1, \cdots, X_N]$ contains N independent zero mean Gaussian random M-dimensional snapshot vectors with covariance matrix of $\mathbf{A}$, i.e., $\mathcal{N}_M(0, \mathbf{A})$, where $\mathbf{A}$ is a nonnegative $M \times M$ Hermitian matrix. Some signal processing algorithms process a batch of data together using a rectangular window. However in a number of practical signal processing algorithms, the SCM is estimated by applying a window as follows

$$\mathbf{R}_N = \frac{1}{N} \sum_{i=1}^{N} w_i X_i X_i^H, \tag{1}$$

where $\{w_i \geq 0, i = 1, \cdots, N\}$ is a non-negative sequence. These weights allows to flexibly emphasize or deemphasize some of the observations. Using smaller weights for old data samples allows to improve the agility of the algorithms in many applications such as cognitive radio, where it is important to detect the activities and idle channels as fast as possible. Among all windows, the exponential window, $w_i = w_0 p^i$, is commonly used because

1. this window allows to develop and implement fast recursive algorithms in real-time applications which are considerably less expensive in terms of computational complexity (e.g. see [5, 6]), and
2. allows to forget the old data, thereby improving the tracking ability in non-stationary environments.

Most of the existing results in literature about the behavior of the eigenvalues mainly consider the rectangular window. In this case the SCM of the observations has a Wishart distribution and the joint probability density function of the eigenvalues of SCM can be expressed in terms of hyper-geometric functions [7] which is applicable in small array cases. An alternative approach is to employ the following empirical spectral distribution (e.s.d.) of an arbitrary square matrix $\mathbf{A} \in \mathbb{C}^{M\times M}$

$$F^{\mathbf{A}}(x) \triangleq \frac{1}{M} \sum_{i=1}^{M} u(x - \lambda_i), \tag{2}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M$ are eigenvalues of $\mathbf{A}$ and $u(.)$ denotes the unit step function. Note that, in this definition all eigenvalues of $\mathbf{A}$ are assumed to be real. Although this formulation is less explicit than the joint pdf of eigenvalues, it describes the behavior of the eigenvalues. For a random matrix $\mathbf{A}$, the e.s.d. $F^{\mathbf{A}}(x)$ is a random function, and in many practical cases, converges almost surely to a deterministic cumulative distribution function as the dimension of the system grows. In such cases, it referred to as the limiting spectral distribution (l.s.d.) of $\mathbf{A}$.

In this paper, we study the effects of exponentially windowing on the distribution of the eigenvalues of the SCM. In this case, the SCM has a doubly correlated Wishart distribution [8–10]. We must note that, there are numerous research results for the case of Wishart matrices, however, the spectral properties in the doubly correlated case has not been sufficiently studied. In recent years, some results have been

This work is supported in part by Iran Telecommunication Research Center (ITRC).

obtained on the limiting behavior of the e.s.d. of correlated Wishart matrices. For the white noise case, we study the behavior of eigenvalues of the exponentially windowed SCM. The demonstrated results provides a key step toward characterization of the distribution of eigenvalues in the general Covariance matrix of windowed data. The results of this work are useful in the design and implementation of robust algorithms using windowed snapshots.

The remainder of this paper is organized as follows: Section II introduces the system model and some important mathematical tools. Asymptotic spectrum of the eigenvalues in noise-only data case is analyzed in section III. Section IV provides simulation results. Finally, we conclude this work and suggest future works in Section V.

2. SYSTEM MODEL

Using the same assumptions as in eq. (1), windowed SCM can be rewritten as

$$\mathbf{R}_N = \frac{1}{N}\sum_{i=1}^{N} w_i \mathbf{A}^{\frac{1}{2}} U_i U_i^H \mathbf{A}^{\frac{1}{2}} = \frac{1}{N}\mathbf{A}^{\frac{1}{2}}\mathbf{U}\mathbf{W}_N\mathbf{U}^H\mathbf{A}^{\frac{1}{2}}, \quad (3)$$

where $\mathbf{U} = [U_1, \cdots, U_N]$ is an $M \times N$ matrix contains i.i.d. zero-mean unit-variance complex Gaussian entries, $\mathbf{A}^{\frac{1}{2}}$ is a nonnegative definite square root of the covariance matrix $\mathbf{A}$ and $\mathbf{W}_N \triangleq \mathrm{diag}(w_1, \cdots, w_N)$. The matrix $\mathbf{R}_N$ has a doubly correlated Wishart distribution. In practice, it is very complex to directly characterize the e.s.d. of $\mathbf{R}_N$, thus, we use the Stieltjes transform of this distribution and indirectly characterize the behavior of the eigenvalues. Then, in the asymptotic regime as $M, N \to \infty$ given $\frac{M}{N} \to c > 0$, the inverse transform of the limit gives the l.s.d. of SCM.

Definition 1 *Stieltjes transform, $m(z)$, $z \in \mathbb{C}^+ \equiv \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ of a distribution function $F^{\mathbf{R}}(x)$ is defined as*

$$m(z) = \int \frac{1}{\lambda - z} dF^{\mathbf{R}}(\lambda). \quad (4)$$

The inverse Stieltjes transform formula is as follows:

$$F^{\mathbf{R}}(x) = \frac{1}{\pi}\lim_{y\to 0^+}\int_{-\infty}^{x} \mathrm{Im}\{m(t+iy)\}dt, \quad \forall x \in \mathbb{R}. \quad (5)$$

We use the following theorem which gives the Stieltjes transform of the correlated Wishart matrix [10] and is the basis for derivations in this paper.

Theorem 1 *For a window with length of N, consider the matrix defined by $\mathbf{R}_N = \frac{1}{N}\mathbf{A}_N{}^{\frac{1}{2}}\mathbf{U}\mathbf{W}_N\mathbf{U}^H\mathbf{A}_N{}^{\frac{1}{2}}$. Assume that all elements of $\mathbf{U} \in \mathbb{C}^{M\times M}$ are i.i.d. random variables with zero-mean, unit variance and $E\{|U_{ij}|^4\} < \infty$. In addition, suppose that $\mathbf{A}_N \in \mathbb{C}^{M\times M}$ is a Hermitian nonnegative definite matrix, $\mathbf{W}_N = \mathrm{diag}(w_1, \cdots, w_N)$, $F^{\mathbf{A}_N} \xrightarrow{D} F^{\mathbf{A}}$, $F^{\mathbf{W}_N} \xrightarrow{D} F^{\mathbf{W}}$ when $M, N \to \infty$ with $\frac{M}{N} \to c > 0$. In this case, the empirical distribution $F^{\mathbf{R}_N}$, with probability 1, converges weakly to a probability distribution function $F^{\mathbf{R}}$ whose Stieltjes transform $m(z)$, for $z \in \mathbb{C}^+$, is given by*

$$m(z) = \int \frac{1}{a\left(\int \frac{w}{1+cwe(z)}dF^{\mathbf{W}}(w)\right) - z} dF^{\mathbf{A}}(a), \quad (6)$$

where $e(z)$ is the unique solution of the following equation in $\mathbb{C}^+$

$$e(z) = \int \frac{a}{a\left(\int \frac{w}{1+cwe(z)}dF^{\mathbf{W}}(w)\right) - z} dF^{\mathbf{A}}(a). \quad (7)$$

Proof 1 *See [10] for proof. Similar results are also demonstrated in [8, 9] with some differences in the assumptions on correlation matrices.*

Since in practice, the array dimension and/or sample size are usually finite numbers, this method gives a deterministic approximation for the actual sample eigenvalue distribution. For an arbitrary window and white noise we have $\mathbf{A} = \sigma^2\mathbf{I}_{M\times M}$ and thus $dF^{\mathbf{A}}(x) = \delta(x-\sigma^2)dx$. In this case from (6), (7), we obtain $m(z) = \frac{1}{\sigma^2}e(z)$ and

$$m(z) = \frac{1}{\sigma^2\int \frac{w}{1+cw\sigma^2 m}dF^{\mathbf{W}}(w) - z}. \quad (8)$$

Let S_F denotes the support of the distribution function $F^{\mathbf{R}}(x)$ and S_F^c shows its complement. From (5), we see that S_F consists of points on the real axis where the imaginary part of $m(z)$ is positive. In [11], it is shown that for one sided correlated Wishart matrix, $\lim_{y\to 0^+} m_F(x+iy)$ exists for all $x \neq 0$, and therefore we can define

$$m_F(x) = \lim_{y\to 0^+} m_F(x+iy), x \in \mathbb{R}\setminus\{0\}. \quad (9)$$

The following lemma allows to determine these points on real axis.

Lemma 1 ([11, Lemma 6.1.]) *For any c.d.f. F, let S_F denote its support and S_F^c be the complement of S_F. For $x \in S_F^c$, $m = m_F(x)$ is the only real solution of $x = z(m)$ which satisfies $\frac{dz(m)}{dm} > 0$, where $z(m)$ is the inverse function of $m(z)$. Also conversely, for any real m in the domain of $z(m)$ if $\frac{dz(m)}{dm} > 0$ then $x = z(m)$ is outside the support of F.*

This means that S_F^c, is the intervals on the vertical axis where $z(m)$ is increasing for real values of m. According to (8), for noise only data $z(m)$ can be written as follows

$$z(m) = \int \frac{\sigma^2 w}{1+cw\sigma^2 m}dF^{\mathbf{W}}(w) - \frac{1}{m}. \quad (10)$$

3. SPECTRAL ANALYSIS OF NOISE-ONLY DATA

First we should determine $F^{\mathbf{W}}(w)$ as a continuous function for the wights of window in order to derive some explicit closed form expressions for the Stieltjes transform. We note that the exponential window is inherently an infinite length window, however, in Theorem 1 the length of the window and array dimension jointly tend to infinity where $\lim_{M,N\to\infty} \frac{M}{N} = c > 0$. This approximation of a real exponential window, is only accurate if N is large enough such that the omitted coefficients are negligible. For the truncated exponential window ($w_i = w_0 p^i = w_0 \gamma^{\frac{i}{N}},\ i = 1, \cdots, N$), where γ is the ratio of smallest to largest weights of the window, we define $F^{\mathbf{W}_N}(w) = \sum_{i=1}^{N} \frac{1}{N} u(w - w_i)$. From $i = \lfloor \frac{N}{\ln\gamma} \ln(\frac{w_i}{w_0}) \rfloor$, it is easy to show that

$$F^{\mathbf{W}_N}(w) = \begin{cases} 0 & w < \gamma w_0, \\ 1 - \frac{1}{N}\left\lfloor \frac{N}{\ln\gamma} \ln(\frac{w}{w_0}) \right\rfloor & \gamma w_0 \le w \le w_0 \gamma^{\frac{1}{N}}, \\ 1 & w_0 \gamma^{\frac{1}{N}} \le w, \end{cases} \quad (11)$$

where $\lfloor . \rfloor$ is the floor function. This increasing staircase function takes values on $\{0, \frac{1}{N}, \frac{2}{N}, \cdots, 1\}$. To satisfy the constraints of Theorem 1 for the exponential window, we assume that $\gamma = p^N > 0$, is an arbitrary small real constant. In other words, the forgetting factor of the window $p = \gamma^{\frac{1}{N}} \in (0, 1)$ approaches to 1, as $M, N \to \infty$. The smaller γ, the better this truncated exponential model fits the exponential window with the forgetting factor p. From, $\lim_{N\to\infty} w_0 = \frac{\ln\gamma}{\gamma - 1}$, we conclude that $\lim_{N\to\infty} F^{\mathbf{W}_N} = F^{\mathbf{W}}(w)$ where

$$F^{\mathbf{W}}(w) = \begin{cases} 0 & w < \frac{\gamma\ln\gamma}{\gamma-1}, \\ 1 - \frac{1}{\ln\gamma} \ln(\frac{w(\gamma-1)}{\ln\gamma}) & \frac{\gamma\ln\gamma}{\gamma-1} < w < \frac{\ln\gamma}{\gamma-1}, \\ 1 & w > \frac{\ln\gamma}{\gamma-1}. \end{cases} \quad (12)$$

is a continuous function, independent of window size N and satisfies the assumptions of Theorem 1. Thus, Theorem 1 is applicable to the exponential window truncated at some large integer N.

Substituting (12) in (10) as the arbitrary constant $\gamma \to 0$, in the asymptotic regime of Theorem 1, as $M, N \to \infty$ such that $\frac{M}{n_0} \to c_0$, $z(m)$ satisfies

$$z(m) = \frac{1}{c_0 m} \ln(1 + c_0 \sigma^2 m) - \frac{1}{m}, \quad (13)$$

for $m \in (-\frac{1}{c_0\sigma^2}, \infty) - \{0\}$, where $n_0 = -\frac{1}{\ln(p)}$.

Theorem 2 gives the explicit distribution.

Theorem 2 *For the exponentially weighted window, the limiting spectral distribution of SCM is given by*

$$f^{\mathbf{R}}(x) = \frac{e^{c_0 - \frac{x}{\sigma^2}}}{\pi c_0 \sigma^2} \mathrm{Im}\left(e^{-\omega_{-1}(-\frac{x}{\sigma^2}\exp\{c_0 - \frac{x}{\sigma^2}\})}\right) \Pi_{x_-,x_+}(x), \quad (14)$$

where $\omega_{-1}(x)$ is the branch of Lambert W function[1] *[12] with $k = -1$ and*

$$x_- = \sigma^2 \frac{\omega_0(-e^{-c_0-1}) + 1}{\exp\{\omega_0(-\exp(-c_0 - 1)) + c_0 + 1\} - 1}, \quad (15)$$

$$x_+ = \sigma^2 \frac{\omega_{-1}(-e^{-c_0-1}) + 1}{\exp\{\omega_{-1}(-\exp(-c_0 - 1)) + c_0 + 1\} - 1}. \quad (16)$$

are upper and lower boundaries of the support, respectively.

Proof 2 *Lemma 1 states that boundaries of the support of eigenvalues are the real solutions of $z'(m) = 0$. Denoting $y = \ln(1 + c_0\sigma^2 m) - c_0 - 1$, we obtain*

$$y e^y = -e^{-c_0-1} \in [-e^{-1}, 0), \quad \forall c_0 > 0. \quad (17)$$

This equation has two real solutions m_- and m_+ expressed using Lambert W function as

$$m_- = \frac{1}{c_0\sigma^2}(\exp\{\omega_0(-e^{-c_0-1}) + c_0 + 1\} - 1), \quad (18)$$

$$m_+ = \frac{1}{c_0\sigma^2}(\exp\{\omega_{-1}(-e^{-c_0-1}) + c_0 + 1\} - 1). \quad (19)$$

Using (13)*, the boundaries $z(m_-)$ and $z(m_+)$ are obtained as in* (15) *and* (16) *which determine the support of eigenvalues as the interval $[z(m_-), z(m_+)] \subset \mathbb{R}$.*

To obtain the l.s.d. of SCM, we should find $m(z)$ with positive imaginary part for all $z \in [z(m_-), z(m_+)]$. Denoting $v = -\ln(1 + c_0\sigma^2 m) - \frac{z}{\sigma^2} + c_0$, we obtain

$$v e^v = (-\frac{z}{\sigma^2} e^{c_0 - \frac{z}{\sigma^2}}). \quad (20)$$

Therefore, the solutions are

$$v_k = \omega_k(-\frac{z}{\sigma^2} e^{c_0 - \frac{z}{\sigma^2}}), \quad \forall k \in \mathbb{Z}. \quad (21)$$

According to (9) *and given the obtained support, only the branches with $k = 0$ and $k = -1$ are acceptable solutions. For other values of k the real part of $m(z)$ is not continuous at the boundaries. It is easy to see that for $z \in [z(m_-), z(m_+)]$, the expression on the right-hand side of* (20) *belongs to $[-e^{c_0-1}, -e^{-1}]$. From the propoerties of Lambert W function, we also deduce that $\mathrm{Im}\{m\}$ and $\sin(-\mathrm{Im}\{v\})$ have the same signs, and for $x \in [-e^{c_0-1}, -e^{-1}]$ the function $\sin(-\mathrm{Im}\{\omega_k(x)\})$ is positive for $k = -1$ and is negative for $k = 0$. Therefore, the Stieltjes transform of the l.s.d. of SCM is obtained from* (21) *as*

$$m = \frac{1}{c_0\sigma^2}(e^{c_0 - \frac{z}{\sigma^2} - \omega_{-1}(-\frac{z}{\sigma^2} e^{c_0 - \frac{z}{\sigma^2}})} - 1). \quad (22)$$

[1] The Lambert W function [12], $\omega(x)$ is also called the Omega function and is the solution of $\omega e^{\omega} = z$ for any complex number z. This equation is not injective, thus the function $\omega(z)$ is multivalued and has a set of different branches named $\omega_k(z)$ for any integer k. For real values of z, there exist two real valued branches of Lambert W function $\omega_0(z)$ and $\omega_{-1}(z)$ which take on real values for $z \in [-\frac{1}{e}, \infty) \cup [-\frac{1}{e}, 0)$ and complex values, otherwise. The function $\omega_0(z)$ is referred to as the principal branch of the Lambert W function and shown by $\omega(z)$ for simplicity.

Using the inverse formula in (10), the limiting SD of SCM is

$$f^{\mathbf{R}}(x) = \frac{1}{c_0\sigma^2\pi}\mathrm{Im}\Big(e^{c_0 - \frac{x}{\sigma^2} - \omega_{-1}(-\frac{x}{\sigma^2}\exp\{c_0 - \frac{x}{\sigma^2}\})} - 1\Big). \quad (23)$$

Dropping the real terms inside the brackets and some simplifications, we obtain (14).

4. SIMULATION RESULTS

In Figure 1, we plot the density functions and a histogram to show the accuracy of the derived l.s.d. in this paper for an array with a finite dimension $M = 20$ and an exponential window with $p = 0.97$ and $p = 0.995$. The histogram of the eigenvalues is generated by 2000 samples of SCMs computed from 2000 independent data sets. It can be observed that the histogram of the eigenvalues accurately fits the derived l.s.d. of the exponentially weighted windowed data in (14). In some array signal processing applications, the effective length of the exponential window has been considered to be $N_e = \frac{1}{1-p}$ [5, 6]. To evaluate the accuracy of this approximation, the M-P density function using this effective length is also plotted which shows a large deviation from the simulated data.

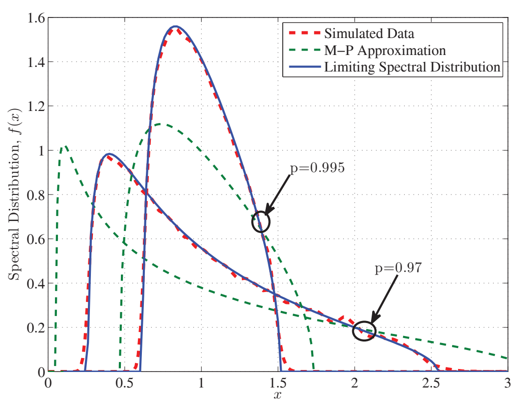

Fig. 1. Distribution of eigenvalues using the exponential window for $M = 20$ and $p \in \{0.97, 0.995\}$.

Also in Figure 1 we observe that for larger values of p (closer to one), the eigenvalues become more concentrated around their true values. This is because the effective length of the window increases as p approaches 1 and this is analogous with the results for rectangular window.

5. CONCLUSION

In this paper the limiting spectral distribution of sample covariance matrix in the case of exponentially weighted windowed data has been studied. We have derived an exact expression for the l.s.d. of the SCM in terms of Lambert W function, which has excellent agreement with the simulation results. The results of this work could be used in design and improvement of detectors and estimators using an exponentially weighted window.

6. REFERENCES

[1] J.W. Silverstein and A.M. Tulino, "Theory of large dimensional random matrices for engineers," in *Spread Spectrum Techniques and Applications, 2006 IEEE Ninth Int. Symp. on*. IEEE, 2006, pp. 458–464.

[2] E. Yazdian, S. Gazor, and H. Bastani, "Source enumeration in large arrays using moments of eigenvalues and relatively few samples," *Submitted to IET Sig. Pro.*

[3] X. Mestre, "On the asymptotic behavior of the sample estimates of eigenvalues and eigenvectors of covariance matrices," *Signal Processing, IEEE Transactions on*, vol. 56, no. 11, pp. 5353–5368, 2008.

[4] R. Couillet, J.W. Silverstein, Z. Bai, and M. Debbah, "Eigen-inference for energy estimation of multiple sources," *Information Theory, IEEE Transactions on*, vol. 57, no. 4, pp. 2420–2439, 2011.

[5] S. Valaee and P. Kabal, "An information theoretic approach to source enumeration in array signal processing," *Signal Processing, IEEE Transactions on*, vol. 52, no. 5, pp. 1171–1178, 2004.

[6] B. Champagne, "Adaptive eigendecomposition of data covariance matrices based on first-order perturbations," *Signal Processing, IEEE Transactions on*, vol. 42, no. 10, pp. 2758–2770, 1994.

[7] A.T. James, "Distributions of matrix variates and latent roots derived from normal samples," *The Annals of Mathematical Statistics*, vol. 35, pp. 475–501, 1964.

[8] Z. Burda, J. Jurkiewicz, and B. Wacław, "Spectral moments of correlated wishart matrices.," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 71, no. 2 Pt 2, pp. 026111, 2005.

[9] L. Zhang, "Spectral analysis of large dimensional random matrices," *Ph.D. Thesis*, 2006.

[10] D. Paul and J.W. Silverstein, "No eigenvalues outside the support of the limiting empirical spectral distribution of a separable covariance matrix," *Journal of Multivariate Analysis*, vol. 100, no. 1, pp. 37–57, 2009.

[11] Z. Bai and J.W. Silverstein, *Spectral analysis of large dimensional random matrices*, Springer Verlag, 2010.

[12] R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey, and D.E. Knuth, "On the lambertw function," *Advances in Comp. mathematics*, vol. 5, pp. 329–359, 1996.