Dimension Boundary Between Finite and Infinite Random Matrices in Cognitive Radio Networks

Wensheng Zhang, *Member, IEEE*, Cheng-Xiang Wang, *Fellow, IEEE*, Jian Sun, *Member, IEEE*, George K. Karagiannidis, *Fellow, IEEE*, and Yang Yang, *Senior Member, IEEE*

***Abstract*— The dimension boundary between finite random matrices and infinite random matrices is originally defined in this letter. The proposed boundary provides a theoretical approach to classify random matrices based on their dimensions. Two methods are proposed to determine the dimension boundary. One is based on the eigenvalue distribution and the other is based on the eigenvalue interval. In particular, a boundary-based threshold generation scheme in cognitive radio networks is studied. The theoretical analysis and numerical results verify the proposed dimension boundary and the corresponding boundary decision methods.**

***Index Terms*— Finite random matrix theory, infinite random matrix theory, dimension boundary, cognitive radio networks.**

I. Introduction

RANDOM matrix theory (RMT) has found several applications in CRNs [1], [2] and wireless communication systems [3], [4]. According to the size of the dimension, random matrices can be generally divided into two categories: finite random matrices with finite dimension and infinite random matrices with infinite dimension. Accordingly, the finite random matrix theory (FRMT) and infinite random matrix theory (IRMT) can be applied to analyze the corresponding random matrices.

For FRMT, exact and analytical results can be used to evaluate the characteristics of finite random matrices. The distributions of eigenvalues [5], [6], standard condition number (SCN) [7], and demmel condition number (DCN) [8] can be generated with the exact and analytical formulations under the condition that the dimension of random matrix is not very large. However, for IRMT, only asymptotical and limiting results can be achieved when the dimension K and the degree of freedom (DoF) N go to infinity, i.e., $K, N \to \infty$. In the paradigm of the IRMT, lots of significant results including the Marchenko-Pastur (MP) law [9] and Tracy-Widom (TW) law [10] can be used to evaluate the theoretical eigenvalue distribution (TED). All the existing TEDs both in the FRMT and IRMT can be verified by the empirical eigenvalue distributions (EEDs). A challenging problem is how to classify a specific random matrix based on its dimension and which theory (FRMT or IRMT) is more appropriate to be explored. Therefore, the dimension boundary should be defined to determine a random matrix with the dimension K and the DoF N. When the dimension is larger than the boundary, the random matrix can be considered as infinite and the IRMT can be used. Otherwise, such a matrix is finite and the FRMT should be useable. In particular, based on the extreme eigenvalue distributions of the IRMT [1] and the SCN distributions of the FRMT [2], [7], the cooperative spectrum sensing (CSS) schemes have been studied in CRNs.

The eigenvalue distribution and eigenvalue interval, as two main eigenvalue characteristics, can be used to determine the dimension boundary. For the eigenvalue distribution, the TEDs of the FRMT and IRMT should follow the corresponding EEDs and the convergence between TED and EED can be used to define the dimension boundary. Moreover, the eigenvalue interval between the largest and smallest eigenvalues can also be utilized to calculate the dimension boundary. For the IRMT, there are no eigenvalues outside the eigenvalue interval of the random matrices with large dimensions [11] and such eigenvalue interval can be statistically determined [12]. The MP law can be used to calculate the TED and the theoretical eigenvalue interval (TEI) in the IRMT. The EED and the empirical eigenvalue interval (EEI) are determined in a numerical way.

The main contributions of this letter are summarized as follows. First, the dimension boundary between FRMT and IRMT is presented to theoretically classify random matrices into finite and infinite. To the best of our knowledge, the dimension boundary is initially defined in the RMT. The boundary provides an efficient way for CRNs to properly select the FRMT based threshold generation scheme is illustrated to generate exact theoreting thresholds, leading to superior sensing performance. Moreover, the dimension boundary can potentially point out which theory (FRMT or IRMT) is more suitable for other RMT-based sensing schemes. Second, two boundary decision methods based on the eigenvalue

Manuscript received February 25, 2017; revised April 10, 2017; accepted April 18, 2017. Date of publication April 25, 2017; date of current version August 10, 2017. This work is supported by the Natural Science Foundation of China (No. 61371110), Key R&D Program of Shandong Province (No. 2016GGX101014), EU FP7 QUICK Project (No. PIRSES-GA-2013-612652), EU H2020 RISE TESTBED Project (No. 734325), MOST 863 Hi-Tech Program (No. 2015AA01A702). The associate editor coordinating the review of this paper and approving it for publication was T. Ngatched. *(Corresponding author: Jian Sun.)*

W. Zhang and J. Sun are with the Shandong Provincial Key Laboratory of Wireless Communication Technologies, Shandong University, Shandong 250100, China (e-mail: zhangwsh@sdu.edu.cn; sunjian@sdu.edu.cn).

C.-X. Wang is with the Shandong Provincial Key Laboratory of Wireless Communication Technologies, Shandong University, Shandong, China, and also with the Institute of Sensors, Signals and Systems, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, U.K. (e-mail: cheng-xiang.wang@hw.ac.uk).

G. K. Karagiannidis is with the Electrical and Computer Engineering Department, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece (e-mail: geokarag@auth.gr).

Y. Yang is with the CAS Key Lab of Wireless Sensor Network and Communication, Shanghai Research Center for Wireless Communications, SIMIT, Shanghai 201210, China (e-mail: yang.yang@wico.sh).

Digital Object Identifier 10.1109/LCOMM.2017.2698474

distribution and eigenvalue interval are designed. The dimension step and dimension threshold are also provided as two key parameters.

II. System Model and Mathematical Preliminaries

The CSS model in CRNs is considered, i.e., the fusion center (FC) senses the target frequency bands by checking PU signal samples, which are gathered from distributed secondary users (SUs). Assuming the number of SUs is K and the number of PU signal samples per SU is N, a $K \times N$ sample matrix $\mathbf{X}$ and the covariance matrix $\mathbf{M} = \mathbf{X}\mathbf{X}^{\mathrm{H}}$ can be generated at the FC. The CSS scheme based on the extreme eigenvalues [1] or the SCN [2] can be designed. According to hypothesis test, the sensing performance can be evaluated by the probability of detection, P_d, which can be expressed as

$$P_d = \Pr\{T \geq \rho | \mathcal{H}_1\} \tag{1}$$

where T denotes the test statistic, ρ is a given threshold, $\Pr\{\cdot\}$ is the probability operation, and the condition $\mathcal{H}_1$ denotes the presence of PU signal.

The $K \times K$ covariance matrix $\mathbf{M}$ ($\mathbf{M} \sim \mathcal{W}_K(\mathbf{V}, N)$) can be regarded as a central and complex Wishart matrix with the DoF N and the matrix $\mathbf{V}$ if the columns of $\mathbf{X}$ are zero mean and independent complex Gaussian vectors. The joint probability density function (PDF) of the K ordered eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_K$ of $\mathbf{M}$ is expressed as [3], [5]

$$f_\Lambda(\lambda_1, ..., \lambda_K) = \prod_{i=1}^{K} \frac{\lambda_i^{N-K} \exp(-\lambda_i)}{(K-i)!(N-i)!} \prod_{i<j} (\lambda_j - \lambda_i)^2 \tag{2}$$

where $\Lambda = [\lambda_1 \lambda_2 \cdots \lambda_K]$ denotes the ordered eigenvalue set. For finite Wishart matrices, the exact eigenvalue distributions can be achieved based on the joint PDF of all K ordered eigenvalues. However, when the dimension of $\mathbf{M}$ becomes very large, the exact eigenvalue distributions cannot be calculated due to high computational complexity. In this case, the TED (PDF) of $\mathbf{M}$ is given by the MP law [9]

$$f_\lambda(x) = \left(1 - \frac{1}{\gamma}\right)^+ \delta(x) + \frac{\sqrt{(x-\alpha)^+(\beta-x)^+}}{2\pi\gamma x} \tag{3}$$

where $K, N \to \infty$, $\gamma = K/N$ denotes the dimension factor, $\alpha = \left(1 - \sqrt{\gamma}\right)^2$, $\beta = \left(1 + \sqrt{\gamma}\right)^2$, and $(x)^+ = \max(x, 0)$. The covariance of each entry of $\mathbf{X}$ is $1/N$ and the eigenvalue λ is located in the interval $[\alpha, \beta]$.

III. Dimension Boundary and Its Application to Threshold Generation

For a central and complex Wishart matrix $\mathbf{M}$ with dimension K and DoF N ($N \geq K$), let K_b denote the dimension boundary. When K is larger than K_b, the random matrix $\mathbf{M}$ can be considered as infinite and the IRMT is therefore applicable. Otherwise, $\mathbf{M}$ is finite and the FRMT can be used. If the dimension of the Wishart matrix is infinite, i.e., $K, N \to \infty$ and $K/N = \gamma$, the TED converges to the corresponding EED and all the eigenvalues are located in the TEI with high probability [11], [12]. In this paper letter, we use the convergence between TED (TEI) and EED (EEI) to determine the dimension boundary.

A. Eigenvalue Distribution Based Dimension Boundary

For a $K \times K$ Wishart matrix $\mathbf{M}$ with the DoF N, $N \geq K$, the dimension boundary K_b is defined as

$$K_b = \min\left\{\underset{K}{\operatorname{argmin}}\left[\sum_{\lambda} \left(|D_t(K, N; \lambda) - D_e(K, N; \lambda)|\right) \leq \Phi_d\right]\right\} \tag{4}$$

$$\text{s.t.} \quad 0 < K \leq N < \infty \tag{5}$$

$$K/N = \gamma \tag{6}$$

$$\lambda \in [a, b] \tag{7}$$

where $D_t(K, N; \lambda)$ is the TED, which can be implemented by the MP law in (3) and $D_e(K, N; \lambda)$ is the EED, which can be evaluated numerically. The given factor Φ_d is used to evaluate the convergence between the TED and EED. The lower and upper bounds of λ are denoted by a and b, respectively. The eigenvalue interval is $[a, b]$, which is usually larger than the TEI determined by the MP law $[\alpha, \beta]$, i.e., $a \leq \alpha$ and $b \geq \beta$.

The sum of the distances between two distributions for J ($J \leq K$) eigenvalues can be calculated by

$$\Omega(K, N; \Lambda_J) = \sum_{j=1}^{J} \left|D_t(K, N; \lambda_j) - D_e(K, N; \lambda_j)\right| \tag{8}$$

where Λ_J denotes a J-length eigenvalue set taken from the EED. For a fixed J, $\Omega(K, N; \Lambda_J)$ decreases with the increase of K and N. The similarity between the TED and EED is calibrated as

$$\Gamma(K, \gamma; \Theta_d; \Lambda_J) = \frac{\Omega(K + \Theta_d, \lfloor (K + \Theta_d)/\gamma \rfloor; \Lambda_J)}{\Omega(K, K/\gamma; \Lambda_J)} \tag{9}$$

where Θ_d is an integer denoting the given dimension step, and $\lfloor \cdot \rfloor$ denotes the floor function. The dimension boundary K_b can be determined when the following condition is satisfied

$$\Gamma(K_b, \gamma; \Theta_d; \Lambda_J) \geq \phi_d \tag{10}$$

where ϕ_d denotes a given dimension threshold. The dimension step Θ_d can be determined by the convergence rate of the EED and the dimension threshold $\phi_d \in (0, 1)$ is given in an empirical way. Compared with the given factor Φ_d in (4), the dimension threshold ϕ_d is used to numerically evaluate the similarity between the EED and TED.

The boundary decision method based on the eigenvalue distribution is summarized as follows:

Algorithm 1 : Generating the Dimension Boundary K_b

1: Given Θ_d and γ, initialize $K = 2$, $N = \lfloor K/\gamma \rfloor$;
2: Generate $\mathbf{M}_i \sim \mathcal{W}_K(\mathbf{V}, N)$, $i = 1, \cdots, I$;
3: Calculate $D_e(K, N; \lambda)$, $D_t(K, N; \lambda)$, and $\Omega(K, N; \Lambda_J)$;
4: Let $K = K + \Theta_d$, $N = \lfloor K/\gamma \rfloor$, generate $\mathbf{M}_i^{'} \sim \mathcal{W}_K(\mathbf{V}, N)$, $i = 1, \cdots, I$;
5: Calculate $D_e^{'}(K, N; \lambda)$, $D_t^{'}(K, N; \lambda)$, and $\Omega^{'}(K, N; \Lambda_J)$;
6: **If** $\Gamma(K - \Theta_d, \gamma; \Theta_d; \Lambda_J) < \phi_d$, $\Omega = \Omega^{'}$, **goto** 4;
7: **Else** $K_b = K - \Theta_d$, **return** K_b.

B. Eigenvalue Interval Based Dimension Boundary

For the Wishart matrix $\mathbf{M}$ with sufficiently large dimension K and DoF N, there are no eigenvalues outside the TEI [11]. Moreover, the probability that all eigenvalues of $\mathbf{M}$ are located in the EEI is relatively high [12]. Therefore, the EEI and TEI tend to be consistent under the condition that $K, N \to \infty$ and $K/N = \gamma$. Based on the consistence between the TEI and EEI, the dimension boundary is defined as

$$K_b = \min\left\{\underset{K}{\operatorname{argmin}} \ [|I_t(K, N; \Lambda) - I_e(K, N; \Lambda)| \le \Phi_i]\right\} \tag{11}$$

$$\text{s.t. } 0 < K \le N < \infty \tag{12}$$

$$K/N = \gamma \tag{13}$$

where $I_t(K, N; \Lambda)$ and $I_e(K, N; \Lambda)$ denote the TEI and EEI, respectively. The given factor Φ_i denotes the convergence between two eigenvalue intervals. The TEI can be taken from the MP law $[\alpha, \beta]$, i.e., $I_t(K, N; \Lambda) = \beta - \alpha$. The EEI can be calculated with the average distance between the largest eigenvalue and smallest eigenvalue, i.e.,

$$I_e(K, N; \Lambda) = \bar{\lambda}_{max} - \bar{\lambda}_{min} \tag{14}$$

where $\bar{\lambda}_{max}$ and $\bar{\lambda}_{min}$ denote the average largest eigenvalue and average smallest eigenvalue, respectively. For a specific Wishart matrix, the distance between two eigenvalue intervals is given by

$$\Psi(K, N; \Lambda) = |I_t(K, N; \Lambda) - I_e(K, N; \Lambda)| \tag{15}$$

where Ψ decreases with the increase of K and N due to the convergence between the TEI and EEI. Similar to the eigenvalue distribution based scheme, for the given dimension step Θ_i, the difference between two interval distances is calculated by

$$\Delta(K, \gamma; \Theta_i; \Lambda) = \Psi(K, N; \Lambda) - \Psi(K + \Theta_i, \lfloor (K + \Theta_i)/\gamma \rfloor; \Lambda). \tag{16}$$

The dimension boundary K_b can be determined when the following condition is satisfied

$$\Delta(K_b, \gamma; \Theta_i; \Lambda) \le \phi_i \tag{17}$$

where ϕ_i is the dimension threshold for the eigenvalue interval based scheme. The dimension boundary decision method based on the eigenvalue interval can also be designed similar to the eigenvalue distribution based scheme.

C. Boundary-Based Threshold Generation

The sensing performance of RMT-based CSS system is evaluated by P_d shown in (1), in which the threshold ρ should be generated in advance. For a given probability of false alarm, P_f, the threshold can be calculated by

$$\rho = \mathrm{F}_T^{-1}(1 - P_f) \tag{18}$$

where $\mathrm{F}_T^{-1}(\cdot)$ denotes the inverse CDF of the test statistic T. The accurate threshold is essentially required for the CSS system to achieve high sensing performance. It should be noted that the CDF of the proposed test statistic T can be achieved in an empirical way if there are no the corresponding theoretical expressions. For the RMT-based CSS system in CRNs, the test statistic T is formulated with the eigenvalue characteristics of the covariance matrix $\mathbf{M}$, e.g., largest eigenvalue [1] and SCN [2]. Therefore, the proper distributions (especially CDFs) of the eigenvalue characteristics should be selected to calculate the threshold based on the proposed dimension boundary. If the dimension of $\mathbf{M}$ is larger than K_b, the asymptotic distributions in the IRMT, e.g., the MP-law and the TW-law, can be selected. Otherwise, the exact distributions in the FRMT should be selected.

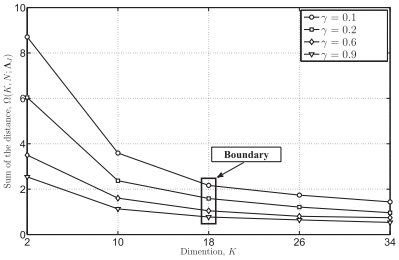


Fig. 1. Eigenvalue distribution based dimension boundary.

IV. Numerical Results and Discussions

In this section, the two proposed boundary decision schemes are illustrated with numerical results. The dimension boundary including the dimension step and dimension threshold for the Wishart matrix is determined. The $K \times K$ central and complex Wishart matrix $\mathbf{M}$ with the DoF N and covariance $1/N$ is used in simulations. The number of trials I is 10000 and the number of sample eigenvalues J is 1000. Note that the J sample eigenvalues are averagely taken from the TEI. The TED is calculated with the MP-law in (3) and the TEI is determined by $[\alpha, \beta]$. The EED is generated in an empirical way and the EEI is determined by the averaged distance between the empirical largest and smallest eigenvalues. According to the convergence rate of random matrices, we find that both dimension steps Θ_d and Θ_i can be set to 8. For a larger dimension step, the EED (EEI) will converge to the corresponding TED (TEI) with a higher convergence rate. However, the fine dimension boundary cannot be determined. There is a tradeoff between the convergence rate and the accuracy for different dimension steps. Based on the convergence between the theoretical and empirical eigenvalue distributions and intervals, the dimension thresholds ϕ_d and ϕ_i are set to 0.7 and 0.1, respectively. Note that there is a tradeoff between the boundary step and boundary threshold. A larger boundary can be generated if a tighter threshold is set.

The dimension boundary of the Wishart matrix based on the eigenvalue distribution is shown in Fig. 1, in which the sum of the distribution distance $\Omega(K, N; \Lambda_J)$ for varying dimensions is indicated. The dimension step Θ_d and the dimension threshold ϕ_d are set to 8 and 0.7, respectively. The dimension boundary can be determined as $K_b = 18$. The corresponding distribution similarity $\Gamma(K, \gamma; \Theta_d; \Lambda_J)$ is indicated in Fig. 2, in which we can observe that $\Gamma(10, \gamma; 8; \Lambda_J)$ for varying γ is less than the given dimension threshold

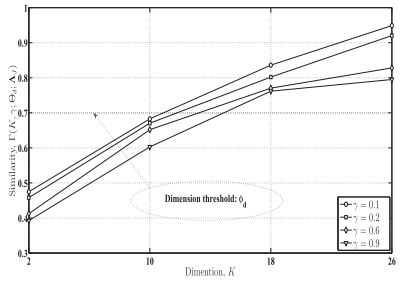


Fig. 2. Distribution similarity vs. dimension K with different dimension factor γ.

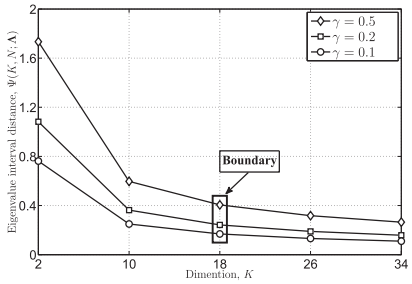
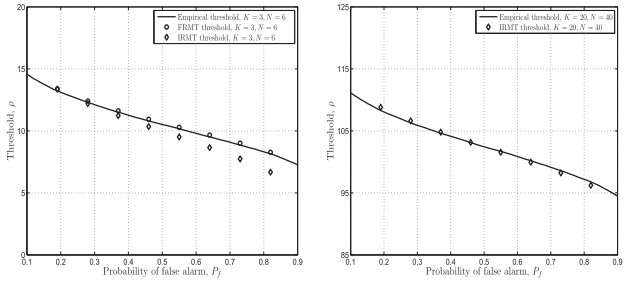


Fig. 3. Eigenvalue interval based dimension boundary.



(a) For finite matrices. (b) For infinite matrices.

Fig. 4. Boundary-based threshold generation with FRMT and IRMT. T is constructed with the largest eigenvalue.

$\phi_d = 0.7$ and therefore $K = 10$ cannot be the dimension boundary.

The distribution similarity improves with the increase of the dimension K. Then, for $K = 18$, $\Gamma(18, \gamma; 8; \Lambda_J)$ for all γ is larger than the given dimension threshold ϕ_d, therefore $K_b = 18$ is the required dimension boundary. For example, $\Gamma(10, 0.1; 8; \Lambda_J)$ is about 0.68 and $\Gamma(18, 0.1; 8; \Lambda_J)$ is 0.83, which is larger than ϕ_d. From the distribution based scheme shown in Fig. 1 and Fig. 2, we can also observe that for the Wishart matrix with large dimension, the similarity between the TED and EED increases. Moreover, the accuracy of the distribution based scheme can be indicated with Fig. 2.

The dimension boundary method based on the eigenvalue interval is indicated in Fig. 3. Here, the dimension step Θ_i is 8 and the dimension threshold ϕ_i is 0.1. The dimension factor γ is set to 0.1, 0.2, and 0.5. Based on K and γ, the corresponding DoF N and the TEI $[\alpha, \beta]$ can be generated. The eigenvalue interval distance $\Psi(K, N; \Lambda)$ for varying dimension K with different dimension factor γ is shown in this figure. We can observe that with the larger dimension, the distance decreases. It means that the TEI and EEI converge with the increase of the matrix dimension. Moreover, when the DoF N increases, i.e., the smaller γ, the distance decreases. When the difference $\Delta(K, \gamma; \Theta_i; \Lambda)$ in (16) is less than the given dimension threshold ϕ_i, the dimension boundary is determined. In this case, the dimension boundary K_b of the Wishart matrix is still 18. From the numerical results shown in Figs. 1 – 3, we can find that the dimension boundary can be determined with the eigenvalue distribution and eigenvalue interval. For both boundary decision methods, there is a tradeoff between the dimension step and boundary threshold. We use two methods shown in (10) and (17) to generate the corresponding boundaries. It should be noted that both methods are inherently based on the convergence between the theoretical eigenvalues and empirical eigenvalues. Boundary-based threshold generation schemes of CRNs are shown in Fig. 4, in which the impact of the correct selection of dimension boundary is evaluated. If the IRMT or FRMT is properly selected, the consistency between theoretical thresholds and empirical ones can be achieved. As shown in Fig. 4(a), the thresholds calculated by the FRMT can properly fit the empirical thresholds. However, the thresholds generated by the IRMT cannot fit the empirical results especially for high P_f. For the infinite matrices shown in Fig. 4(b), only the IRMT can generate the exact thresholds and the FRMT cannot formulate the exact distribution due to high computational complexity. Based on the boundary, the exact theoretical thresholds can be generated with the IRMT or FRMT for CRNs, leading to superior sensing performance.

REFERENCES

[1] Y. Zeng and Y.-C. Liang, "Eigenvalue-based spectrum sensing algorithms for cognitive radio," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1784–1793, Jun. 2009.

[2] W. Zhang, G. Abreu, M. Inamori, and Y. Sanada, "Spectrum sensing algorithms via finite random matrices," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 164–175, Jan. 2012.

[3] P. A. Dighe, R. K. Mallik, and S. S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading," *IEEE Trans. Commun.*, vol. 51, no. 4, pp. 694–703, Apr. 2003.

[4] C.-X. Wang, S. Wu, L. Bai, X. You, J. Wang, and C.-L. I, "Recent advances and future challenges for massive MIMO channel measurements and models," *Sci China Inf. Sci*, vol. 59, no. 2, pp. 1–16, Feb. 2016.

[5] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, no. 2, pp. 475–501, 1964.

[6] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.

[7] M. Matthaiou, M. R. McKay, P. J. Smith, and J. A. Nossek, "On the condition number distribution of complex wishart matrices," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1705–1717, Jun. 2010.

[8] C. Zhong, M. R. McKay, T. Ratnarajah, and K.-K. Wong, "Distribution of the demmel condition number of Wishart matrices," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 1309–1320, May 2011.

[9] V. A. Marčenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Math. USSR Sbornik*, vol. 1, no. 4, pp. 457–483, 1967.

[10] C. A. Tracy and H. Widom, "On orthogonal and symplectic matrix ensembles," *Commun. Math. Phys.*, vol. 177, no. 3, pp. 727–754, Apr. 1996.

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

[12] M. Chiani, "On the probability that all eigenvalues of Gaussian, Wishart, and double Wishart random matrices lie within an interval," *IEEE Trans. Inf. Theory*, to be published.