Analytic approximation to the largest eigenvalue distribution of a white Wishart matrix

J.D. Vlok and J.C. Olivier

***Abstract*—Eigenvalue distributions of Wishart matrices are given in the literature as functions or distributions defined in terms of matrix arguments requiring numerical evaluation. As a result the relationship between parameter values and statistics is not available analytically and the complexity of the numerical evaluation involved may limit the implementation, evaluation and use of eigenvalue techniques using Wishart matrices. This paper presents analytic expressions that approximate the distribution of the largest eigenvalue of white Wishart matrices and the corresponding sample covariance matrices. It is shown that the desired expression follows from an approximation to the Tracy-Widom distribution in terms of the Gamma distribution. The approximation offers largely simplified computation and provides statistics such as the mean value and region of support of the largest eigenvalue distribution. Numeric results from the literature are compared with the approximation and Monte Carlo simulation results are presented to illustrate the accuracy of the proposed analytic approximation.**

I. Introduction

The eigenvalue spectrum of noise covariance matrices plays an important role in such fields as principal component analysis (PCA) [1], singular value decomposition (SVD), multiple-input multiple-output (MIMO) communication systems [2] and signal detection [3], [4]. The behaviour of the largest eigenvalue can be used to predict the performance of MIMO systems in a fading channel and the performance of eigenvalue-based signal detection techniques. The exact distributions of individual eigenvalues can be obtained from the joint distribution, which is defined in terms of hypergeometric functions if the covariance matrix has a Wishart distribution [2], [5]. The individual distributions are then expressed in terms of Laguerre polynomials [6] which can be simplified as matrix arguments [2], [7]. These however require numerical evaluation which can be performed using extensive tables or special purpose software [8]. However, it was shown in [9] that the asymptotic distribution of the scaled largest eigenvalue of a white Wishart matrix can be described by the Tracy-Widom (TW) law [10], [11] which can be evaluated numerically [12]–[14] or approximated using a logit transform [8]. The TW distribution was also shown to be reasonably accurate for non-asymptotic cases [8], [9]. This paper presents a closed-form analytical expression to approximate the TW distribution in order to derive simple expressions for the largest eigenvalue distribution of the Wishart distributed covariance matrix and the associated sample covariance matrix, similar to an approximation given in [15]. Simple expressions describing the statistics and region of support of the largest eigenvalue distribution are also given. The rest of the paper is organised as follows. In Section II a mathematical background is given. Section III presents the approximation and Section IV the expression for the largest eigenvalue distribution. Section V provides a simulation study where numeric results from the literature are compared with Monte Carlo simulation results and finally Section VI summarises the main results and concludes the paper. The focus of this paper is on the TW law of order 1 and 2, denoted respectively by TW_1 and TW_2. TW_4 is briefly considered in the appendix.

J.D. Vlok is with Defence, Peace, Safety & Security (DPSS), Council for Scientific and Industrial Research (CSIR), Pretoria, South Africa, 0001. E-mail: jvlok@csir.co.za

J.C. Olivier is with the School of Engineering, University of Tasmania, Hobart, Australia, 7005.

II. Mathematical background

A. Noise matrix

Let $\boldsymbol{X}$ be an $M \times N$ matrix where each row of $\boldsymbol{X}$ is real and independently drawn from $\mathcal{N}_N(0, \sigma_x^2 \boldsymbol{I})$, the N-variate normal distribution with zero mean and covariance matrix $\sigma_x^2 \boldsymbol{I}$. The $N \times N$ matrix

$$\boldsymbol{Y} = \boldsymbol{X}^H \boldsymbol{X} \tag{1}$$

will then have a white Wishart distribution $\mathcal{W}_N(M, \sigma_x^2 \boldsymbol{I})$ [9], where $\boldsymbol{X}^H$ is the Hermitian transpose of $\boldsymbol{X}$. If $\boldsymbol{X}$ is complex and the complex components of each row are independently drawn from $\mathcal{N}_N(0, (\sigma_x^2/2)\boldsymbol{I})$, $\boldsymbol{Y}$ will have a complex white Wishart distribution. The largest eigenvalue λ_1 of $\boldsymbol{Y}$ in the edge scaling limit, that is when $M \to \infty$, $N \to \infty$ and $\frac{M}{N} \to \gamma \geq 1$, will obey [9]

$$\frac{(\lambda_1/\sigma_x^2) - \mu_{MN,\beta}}{\sigma_{MN,\beta}} \xrightarrow{\mathcal{D}} F_\beta \tag{2}$$

where F_β is the TW cumulative distribution function (CDF) with $\beta = 1$ if $\boldsymbol{X}$ is real and $\beta = 2$ if $\boldsymbol{X}$ is complex. The centre and scaling parameters for $\beta = 1$ are given by [9]

$$\mu_{MN,1} = \left(\sqrt{M-1} + \sqrt{N}\right)^2 \tag{3}$$

$$\sigma_{MN,1} = \sqrt{\mu_{MN,1}}\left(\frac{1}{\sqrt{M-1}} + \frac{1}{\sqrt{N}}\right)^{\frac{1}{3}} \tag{4}$$

and similarly for $\beta = 2$ are given by [9]

$$\mu_{MN,2} = \left(\sqrt{M} + \sqrt{N}\right)^2 \tag{5}$$

$$\sigma_{MN,2} = \sqrt{\mu_{MN,2}}\left(\frac{1}{\sqrt{M}} + \frac{1}{\sqrt{N}}\right)^{\frac{1}{3}}. \tag{6}$$

According to the limit $\frac{M}{N} \to \gamma \geq 1$, (2) to (6) hold only for $M \geq N$. It is however stated in [9] that (2) applies equally well if $M < N$ when $M \to \infty$, $N \to \infty$, and the roles of M and N are reversed in (3) and (4). Following the same argument, (5) and (6) can also be used for $M < N$ since reversing the roles of M and N has no effect in this case. Although (2) is true in the limit, [9] showed that it can provide a satisfactory approximation for matrix dimensions M and N as small as 10.

Note that (2) is usually stated for the unit variance case $\sigma_x^2 = 1$ (as in [1], [9]). The normalisation of λ_1 is required to develop expressions for the largest eigenvalue distribution and the associated statistics for the general case of σ_x^2. To explain the normalisation and show how a given eigenvalue λ of $\boldsymbol{Y}$ scale in comparison with the unit variance case, suppose $\boldsymbol{X}_{(u)}$ represents $\boldsymbol{X}$ when $\sigma_x^2 = 1$ and $\boldsymbol{Y}_{(u)} = \boldsymbol{X}_{(u)}^H \boldsymbol{X}_{(u)}$ from (1). The corresponding eigenvalue of $\boldsymbol{Y}_{(u)}$ is $\lambda_{(u)}$. By substituting $\boldsymbol{X} = \sigma_x \boldsymbol{X}_{(u)}$ for the general case into (1), it follows that $\boldsymbol{Y} = \sigma_x^2 \boldsymbol{Y}_{(u)}$. From the definition of eigenvalues and eigenvectors ($\boldsymbol{Yv} = \lambda \boldsymbol{v}$ with $\boldsymbol{v}$ an eigenvector of $\boldsymbol{Y}$) it can then be shown that

$$\lambda = \sigma_x^2 \lambda_{(u)}. \tag{7}$$

The eigenvalues of $\boldsymbol{Y}$ therefore scale with σ_x^2 compared with the unit variance case $\boldsymbol{Y}_{(u)}$. The eigenvalue λ_1 can therefore be normalised by dividing it with σ_x^2 as is done in (2).

B. Sample covariance matrix and relation to noise matrix

The sample covariance matrix of $\boldsymbol{X}$ is given as [1]

$$\boldsymbol{R} = \frac{1}{M} \sum_{m=1}^{M} \boldsymbol{x}_m^H \boldsymbol{x}_m \tag{8}$$

with $\boldsymbol{x}_m$ the m^{th} $1 \times N$ row of $\boldsymbol{X}$. The ij^{th} element of the $N \times N$ matrix formed by $\boldsymbol{x}_m^H \boldsymbol{x}_m$ in (8) can by definition be expressed in terms of the elements of $\boldsymbol{X}$ as

$$\left[\boldsymbol{x}_m^H \boldsymbol{x}_m\right]_{ij} = \boldsymbol{X}_{mi}^* \boldsymbol{X}_{mj}. \tag{9}$$

From (8) and (9), each element of $\boldsymbol{R}$ can be expressed as

$$\boldsymbol{R}_{ij} = \frac{1}{M} \sum_{m=1}^{M} \boldsymbol{X}_{mi}^* \boldsymbol{X}_{mj}. \tag{10}$$

Likewise, each element of $\boldsymbol{Y}$ can be expressed from (1) as

$$\boldsymbol{Y}_{ij} = \sum_{m=1}^{M} \boldsymbol{X}_{mi}^* \boldsymbol{X}_{mj} \tag{11}$$

which is the scalar product of the i^{th} row of $\boldsymbol{X}^H$ and the j^{th} column of $\boldsymbol{X}$. From (10) and (11) it is clear that

$$\boldsymbol{R} = \frac{1}{M} \boldsymbol{Y}. \tag{12}$$

The largest eigenvalue of $\boldsymbol{R}$ denoted by l_1 is therefore related to λ_1 as

$$l_1 = \frac{1}{M} \lambda_1. \tag{13}$$

Note that both λ_1 and l_1 are always real and nonnegative since $\boldsymbol{Y}$ and $\boldsymbol{R}$ are always Hermitian (or symmetric if $\beta = 1$) and positive semidefinite.

C. Tracy-Widom law

The Tracy-Widom law [11] or distribution TW_β refers to a family of CDFs F_β and related probability density functions (PDFs) f_β describing the limiting distributions of the largest eigenvalues of symmetric ($\beta = 1$), Hermitian ($\beta = 2$) or self-dual ($\beta = 4$) random matrices in the Gaussian ensembles[1]. The three TW CDFs are defined as [11], [12]

$$F_1(x) = \exp\left(-\frac{1}{2}\int_x^\infty q(w)\, dw\right)\sqrt{F_2(x)} \tag{14}$$

$$F_2(x) = \exp\left(-\int_x^\infty (w - x)\, q^2(w)\, dw\right) \tag{15}$$

$$F_4\left(\frac{x}{\sqrt{2}}\right) = \cosh\left(-\frac{1}{2}\int_x^\infty q(w)\, dw\right)\sqrt{F_2(x)} \tag{16}$$

with $q(w)$ the solution to the Painlevé II differential equation

$$q''(w) = wq(w) + 2q^3(w) \tag{17}$$

with the boundary condition $q(w) \sim \text{Ai}(w)$ as $w \to \infty$ where $\text{Ai}(w)$ is the Airy function. Calculation of F_β therefore requires evaluation of the Painlevé II differential equation which can be performed numerically and tabulated (see [14] for a review on the numerical evaluation of distributions defined in terms of Painlevé transcendents). A number of authors [12]–[14] developed and made available software modules to calculate double precision solutions of TW_β. Tables containing solutions of Painlevé II, TW_1 and TW_2 over $x \in [-40, 200]$ with step size $\Delta x = 0.0625$ as described in [13] are available at [16]. The numeric solutions of f_β obtained from [16] for $\beta = 1$ and 2 are shown in Figure 1 and are used in this paper to develop the approximation.

III. Tracy-Widom approximation

In this section an approximation to TW_β using the Gamma distribution is proposed and the goodness-of-fit of the approximation is evaluated against the double precision numeric values of [16], which are exact to sixteen significant decimal digits. The numeric values of the PDF and CDF of TW_β obtained from [16] are denoted respectively by f_β and F_β. Likewise, the PDF and CDF of the Gamma approximation are denoted by g_β and G_β. Whereas only $\beta = 1$ and 2 are considered in this section, $\beta = 4$ is considered in the appendix.

A. Proposed Gamma approximation

By observing the numeric solutions of f_β in Figure 1, the functions appear to resemble slightly asymmetric Gaussian density functions shifted on the *x*-axis. To incorporate the asymmetry, f_β could therefore be approximated using the Gamma PDF given by

$$g_\beta(x) = \frac{(x - x_0)^{k-1}}{\theta^k \Gamma(k)} \exp\left[\frac{-(x - x_0)}{\theta}\right] \tag{18}$$

with x_0 the location or shift parameter, k the shape, θ the scale and $\Gamma(k)$ the Gamma function. Values for these parameters

[1] The Gaussian ensembles include the orthogonal (GOE), unitary (GUE) and symplectic (GSE) ensembles corresponding respectively to real ($\beta = 1$), complex ($\beta = 2$) and quaternion ($\beta = 4$) random matrices [5].

(which are given in Table I) were obtained by fitting g_β to the numeric values of f_β and minimising the sum of squared differences (SSD)

$$\epsilon_\beta^2 = \sum_{i=1}^{L} [f_\beta(x_i) - g_\beta(x_i)]^2 \tag{19}$$

over the full range of x in [16] such that $x_1 = -40$ and $x_L = 200$ with sample step size $\Delta x = 0.0625$. The statistics of the resultant Gamma approximation are given in Table I, which resemble the TW statistics given in Table 1 of [11], [14]. In addition to the numeric solutions, Figure 1 also shows the Gamma approximations using (18) with the parameter values from Table I. The SSD values obtained using (19) are also given in Table I.

TABLE I
PARAMETER VALUES AND RELATED RESULTS FOR THE GAMMA APPROXIMATION TO TW$_\beta$.

Parameter	Symbol	$\beta = 1$	$\beta = 2$
Shape	k	46.5651	79.3694
Scale	θ	0.1850	0.1010
Location	x_0	-9.8209	-9.7874
Mean	$k\theta + x_0$	-1.2064	-1.7711
Variance	$k\theta^2$	1.5937	0.8096
Skewness	$2/\sqrt{k}$	0.2931	0.2245
SSD	ϵ_β^2	2.8270×10^{-5}	9.3883×10^{-6}
SCvM statistic	W_β^2	1.0547×10^{-7}	4.7651×10^{-8}
Kolmogorov statistic	K_β	8.0577×10^{-4}	4.0428×10^{-4}

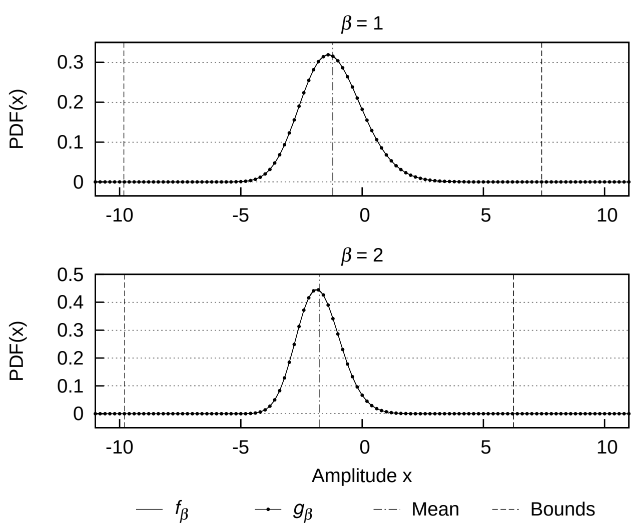

Fig. 1. Numeric and approximated PDFs for $\beta = 1$ and $\beta = 2$. f_β is the numeric solution of the TW PDF obtained from [16] and g_β is the Gamma PDF given in (18).

B. Support region

Although the support of TW$_\beta$ is not bounded, both left and right tails of f_β exhibit exponential decay [9]. It is therefore possible to truncate the support region to certain bounds $[b_-, b_+]$ without losing much probability mass. This section proposes a truncated support region for TW$_\beta$ based on the Gamma approximation presented in Section III-A. The probability mass lost when using the truncated support region is also considered.

The Gamma PDF given in (18) has support $[x_0, \infty)$ and the location parameter x_0 can therefore be used as the lower bound b_-. The upper bound b_+ is chosen such that the mean value of the Gamma distribution is also the mean of the lower and upper bounds[2]. The support of the truncated Gamma approximation is then

$$[b_-, b_+] = [x_0, 2k\theta + x_0] \tag{20}$$

which is also displayed in Figure 1.

To illustrate the effect of the truncation, the PDFs and CDFs of TW$_\beta$ and the associated Gamma approximations are shown with logarithmic ordinate axes in Figures 2 and 3. To quantify the loss in probability mass due to the truncation, values from Figure 3 for the mass of each tail in terms of the cumulative distribution outside the bounded region of (20) are given in Table II. Interpolated values of [16] are used as reference solutions for F_β. The total probability mass lost in the truncation is the mass outside the support region.

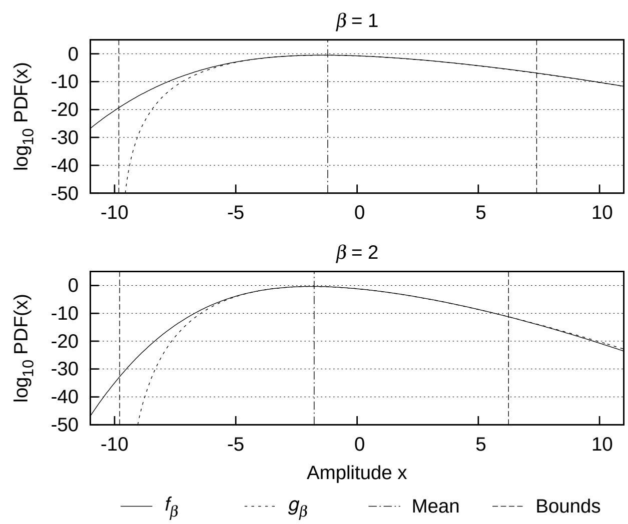

Fig. 2. Numeric and approximated PDFs for $\beta = 1$ and $\beta = 2$. f_β is the numeric solution of the TW PDF obtained from [16] and g_β is the Gamma PDF given in (18).

C. Goodness-of-fit

The approximation accuracy can also be measured using goodness-of-fit tests which indicate how close an empirical CDF is to a theoretical CDF. The difference or distance between the two CDFs is given for the purpose of this study as

$$D_\beta(x) = F_\beta(x) - G_\beta(x). \tag{21}$$

[2]For the purpose of choosing bounds the TW PDFs can be assumed to be approximately symmetric - which can be seen clearly in Figure 1.

TABLE II
PARAMETER VALUES RELATED TO TRUNCATED SUPPORT AND LOSS IN PROBABILITY MASS.

Parameter	**Expression**	$\beta = 1$	$\beta = 2$
Lower bound	b_-	-9.8209	-9.7874
Upper bound	b_+	7.4082	6.2452
Left tail mass (reference)	$F_\beta(b_-)$	3.4799×10^{-21}	7.6093×10^{-35}
Right tail mass (reference)	$1 - F_\beta(b_+)$	4.3875×10^{-8}	1.0734×10^{-12}
Total mass lost (reference)	$F_\beta(b_-) + 1 - F_\beta(b_+)$	4.3875×10^{-8}	1.0734×10^{-12}
Left tail mass (approximation)	$G_\beta(b_-)$	0	0
Right tail mass (approximation)	$1 - G_\beta(b_+)$	3.4942×10^{-8}	1.1563×10^{-12}
Total mass lost (approximation)	$G_\beta(b_-) + 1 - G_\beta(b_+)$	3.4942×10^{-8}	1.1563×10^{-12}

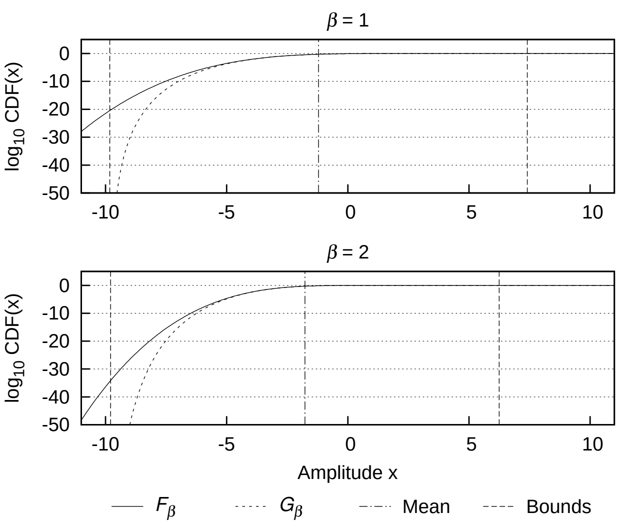

Fig. 3. Numeric and approximated CDFs for $\beta = 1$ and $\beta = 2$. F_β is the numeric solution of the TW CDF obtained from [16] and G_β is the Gamma CDF derived from (18).

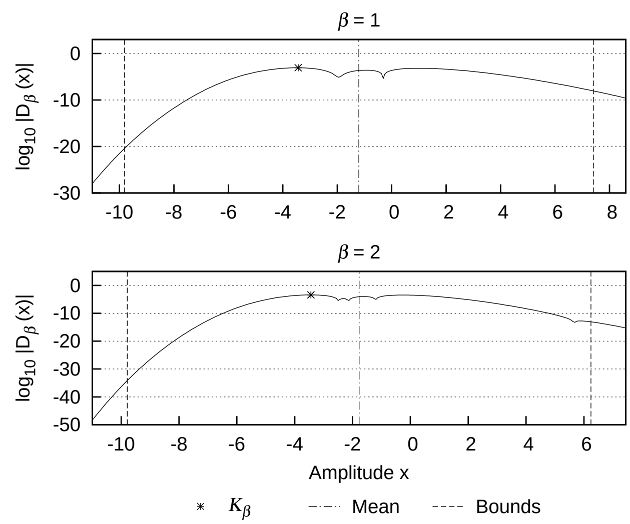

Fig. 4. Absolute difference between the CDFs F_β and G_β as defined in (21) for $\beta = 1$ and $\beta = 2$. The Kolmogorov statistic defined in (23) is also shown on each graph.

Two tests from [17] based on (21) are used in this paper to evaluate the approximation accuracy. The first test is the Smirnov-Cramér-Von-Mises (SCvM) test with test statistic

$$W_\beta^2 = \int_{b_-}^{b_+} D_\beta^2(x) g_\beta(x) \; dx. \tag{22}$$

The second test is the Kolmogorov test with test statistic

$$K_\beta = \max |D_\beta(x)| \, ; \quad x \in [b_-, b_+]. \tag{23}$$

Both the SCvM and Kolmogorov test statistics are indications of how well the numeric values from [16] fit the analytic expression in (18). These statistics will approach zero as the goodness-of-fit improves. The values of (22) obtained through numerical integration with step size $\Delta x = 0.0625$ and (23) are given in Table I. The values of the test statistics remain unchanged whether they are evaluated over $[-40, 200]$ or $[b_-, b_+]$ given in (20), confirming that the truncation has a negligible effect on the accuracy of the approximation. Graphs depicting the absolute value of (21) over x and the associated Kolmogorov statistics are shown in Figure 4.

IV. AN EXPRESSION FOR THE LARGEST EIGENVALUE DISTRIBUTION

This section provides expressions for the largest eigenvalue distributions of the noise matrix ***Y*** and the sample covariance matrix ***R*** based on the TW approximation presented in Section III. Other approximation methods are also briefly considered.

A. *Noise matrix* **Y**

Using (2) and (18) and linear random variable transformations [18] the PDF of λ_1 can be expressed as

$$p_{\lambda_1}(x) = \frac{1}{\sigma_x^2 \; \sigma_{MN,\beta}} g_\beta \left\{ \frac{(x/\sigma_x^2) - \mu_{MN,\beta}}{\sigma_{MN,\beta}} \right\} \tag{24}$$

which can be written in the form of (18) as

$$p_{\lambda_1}(x) = \frac{(x - x_0')^{k-1}}{\theta'^k \Gamma(k)} \exp\left[\frac{-(x - x_0')}{\theta'} \right] \tag{25}$$

with updated parameters

$$\theta' = \sigma_x^2 \; \sigma_{MN,\beta} \; \theta \tag{26}$$
$$x_0' = \sigma_x^2 \left(\mu_{MN,\beta} + x_0 \; \sigma_{MN,\beta} \right). \tag{27}$$

The support of $p_{\lambda_1}(x)$ can then be written from (20) with the updated parameters given in (26) and (27) as

$$[b_-, b_+] = [x_0', 2k\theta' + x_0']. \tag{28}$$

B. Sample covariance matrix $\mathbf{R}$

Using (13) the PDF of l_1 can be written from (24) as [18]

$$p_{l_1}(x) = Mp_{\lambda_1}(Mx) \tag{29}$$

which can also be written in the form of (18) or (25) as

$$p_{l_1}(x) = \frac{(x - x_0'')^{k-1}}{\theta''^k \Gamma(k)} \exp\left[\frac{-(x - x_0'')}{\theta''}\right] \tag{30}$$

with parameters updated again, giving

$$\theta'' = \frac{\theta'}{M} \tag{31}$$

$$x_0'' = \frac{x_0'}{M}. \tag{32}$$

The support of $p_{l_1}(x)$ can then be written as (28) by replacing the updated parameters with the twice-updated parameters given in (31) and (32).

C. Other approximations

Other related approximations include the logit transform approximation to the TW law presented in [8] and a Gamma approximation describing the largest eigenvalue distribution in [15]. The logit transform approach considers only $\beta = 1$ and is computationally more complex than the approximation proposed in this paper. The approximation of [15] calculates the shape k and scale θ of the Gamma distribution by matching the first two moments of the largest eigenvalue and Gamma distributions using an equivalent of (2) and the TW law. The TW distribution is however not approximated directly and the shift parameter x_0 is not used. The focus of [15] is on spectrum sensing applicable to cognitive radio and only $\beta = 2$ is considered for matrix $\boldsymbol{Y}$. The approximation of [15] is however evaluated in the simulation study in Section V against the approximation presented in this paper for both $\beta = 1$ and 2 using the same scaling parameters given in Section II-A and the values of the first two TW moments given in [14].

V. Simulation study and results

A Monte Carlo computer simulation study was conducted with the aim of evaluating how accurate the proposed Gamma approximations can predict actual largest eigenvalue distributions. Empirical distributions of the largest eigenvalues of matrices $\boldsymbol{Y}$ and $\boldsymbol{R}$ for both $\beta = 1$ and 2 were obtained through simulation using 10^6 replications of these matrices for a given set of matrix dimensions (M, N) with $\sigma_x^2 = 1$. Every simulation set was started using identical random seed values. The empirical PDF for a given set was obtained from the simulated data by calculating the histogram over the support region given in (20) with the number of bins fixed to 100. To measure the approximation accuracy, the SCvM criterion given by (21) and (22) was used with $F_\beta(x)$ corresponding to the empirical CDF obtained through the Monte Carlo simulations. Likewise, $G_\beta(x)$ and $g_\beta(x)$ correspond to the Gamma approximations with densities defined by (25) or (30) depending on whether λ_1 or l_1 is concerned. Subsequently the SCvM results are presented. Section V-A considers an example set $(M, N) = (20, 40)$ and Section V-B a range of matrix dimensions. For the purpose of comparison, the SCvM statistics calculated for the approximation method given in [15] (see Section IV-C) are also given in Section V-B. The results are discussed in Section V-C.

A. Example set

Figures 5 and 6 show the predicted and simulated distributions of the largest eigenvalues for $\beta = 1$ and $(M, N) = (20, 40)$. The predicted curves correspond to the Gamma approximations based on (18) and the simulated curves to the empirical data. Figure 5 shows the results for λ_1 using (25) as prediction and Figure 6 shows the results for l_1 using (30) as prediction. Table III shows parameter values for $\beta = 1$ (corresponding to Figures 5 and 6) and $\beta = 2$. The goodness-of-fit statistics are identical for λ_1 and l_1 for each case of β since the random seed values used are identical and the number of histogram bins used in determining the empirical CDF is constant.

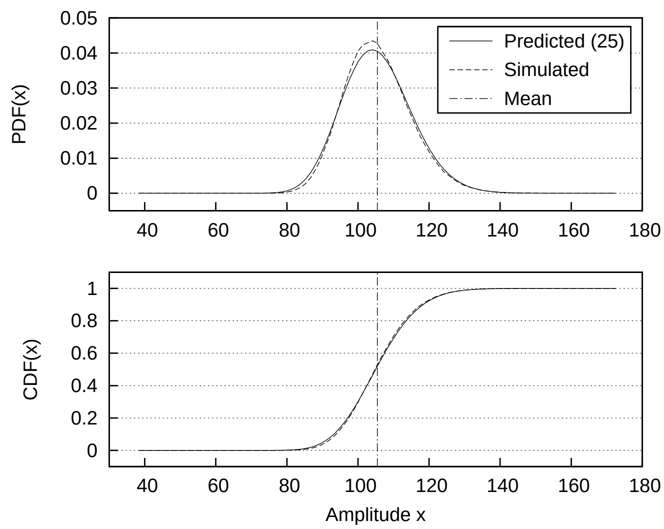

Fig. 5. Predicted and simulated PDFs and CDFs of λ_1 for $\beta = 1, (M, N) = (20, 40)$ and $\sigma_x^2 = 1$.

B. Range of matrix dimensions

The range of matrix dimensions from $(M, N) = (20, 20)$ to $(200, 200)$ for both cases of $M \geq N$ and $M < N$ is considered. Results for the square matrix case $M = N$ are given in Table IV and plotted in Figure 7. As in Table III, the SCvM statistics for λ_1 and l_1 are identical in Table IV. Figure 7 shows that as the matrix dimensions increase, the SCvM statistics decrease indicating an improvement in the approximation accuracy. For $\beta = 1$, the approximation given in (25) outperforms [15] up to a maximum SCvM difference of 9.2809×10^{-5} at $M = 200$. For $\beta = 2$ the two approximation methods show similar accuracies though for smaller values of

TABLE III
PARAMETER VALUES FOR THE LARGEST EIGENVALUE DISTRIBUTIONS FOR $(M, N) = (20, 40)$.

	$\beta = 1$		$\beta = 2$	
Parameter	λ_1	l_1	λ_1	l_1
Mean (Theory) $k\theta + x_0$	105.4619	5.2731	102.6974	5.1349
Mean (Measured)	105.4344	5.2717	103.1117	5.1556
Lower bound b_-	38.3724	1.9186	39.9136	1.9957
Upper bound b_+	172.5515	8.6276	165.4811	8.2741
Bin size Δx	1.3418	6.7090×10^{-2}	1.2557	6.2784×10^{-2}
SCvM statistic W_β^2	9.9414×10^{-5}	9.9414×10^{-5}	4.2237×10^{-4}	4.2237×10^{-4}
Kolmogorov statistic K_β	1.3901×10^{-2}	1.3901×10^{-2}	3.2812×10^{-2}	3.2812×10^{-2}

TABLE IV
SCVM STATISTICS FOR THE LARGEST EIGENVALUES WHEN $M = N$.

	W_1^2 $(\beta = 1)$		W_2^2 $(\beta = 2)$	
M	λ_1 (25) and l_1 (30)	λ_1 [15]	λ_1 (25) and l_1 (30)	λ_1 [15]
20	2.8308×10^{-4}	2.8493×10^{-4}	7.4245×10^{-4}	6.9550×10^{-4}
40	9.5246×10^{-5}	1.0662×10^{-4}	2.7204×10^{-4}	2.0513×10^{-4}
60	5.2480×10^{-5}	8.7108×10^{-5}	1.6838×10^{-4}	1.1199×10^{-4}
80	3.1733×10^{-5}	7.8397×10^{-5}	1.0092×10^{-4}	6.7867×10^{-5}
100	2.5417×10^{-5}	8.2993×10^{-5}	7.9621×10^{-5}	5.4370×10^{-5}
120	1.8896×10^{-5}	7.9911×10^{-5}	5.9821×10^{-5}	4.4188×10^{-5}
140	1.1512×10^{-5}	9.0566×10^{-5}	3.7689×10^{-5}	4.2799×10^{-5}
160	1.2007×10^{-5}	9.3116×10^{-5}	3.6145×10^{-5}	4.3566×10^{-5}
180	8.3928×10^{-6}	9.2033×10^{-5}	2.7315×10^{-5}	4.1183×10^{-5}
200	6.8972×10^{-6}	9.9706×10^{-5}	2.7319×10^{-5}	4.0576×10^{-5}

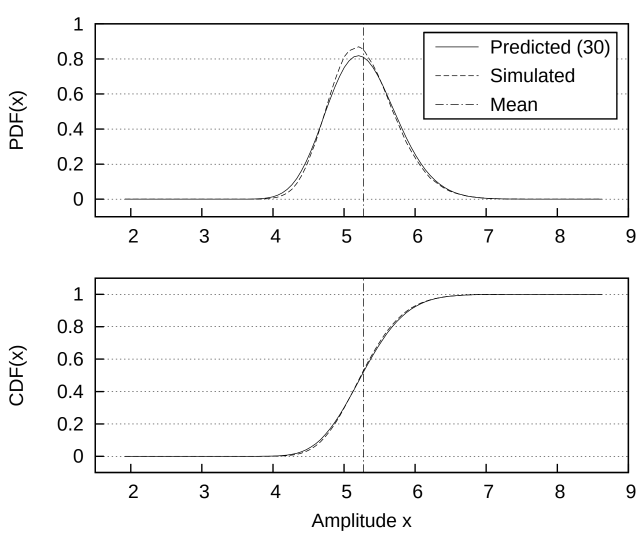

Fig. 6. Predicted and simulated PDFs and CDFs of l_1 for $\beta = 1, (M, N) = (20, 40)$ and $\sigma_x^2 = 1$.

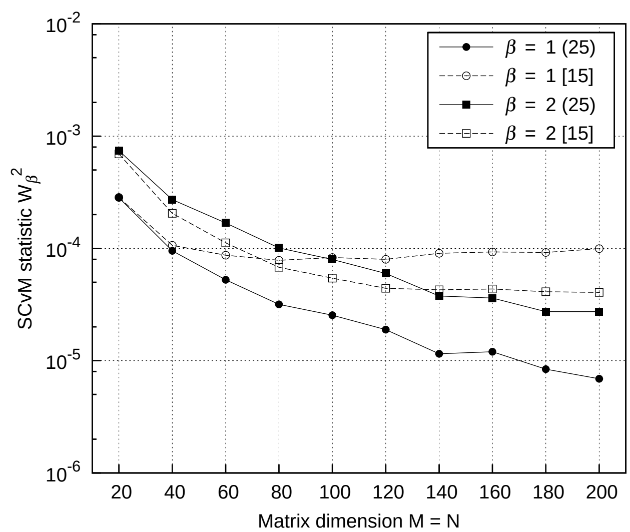

Fig. 7. SCvM statistics for λ_1 as given in Table IV. The curves labelled (25) correspond to the Gamma approximation given in (25) and the curves labelled [15] correspond to the approximation method given in [15].

M, [15] performs slightly better and for larger values of M, (25) performs slightly better.

SCvM results for fixed values of $M = 20$ and 200 (the extreme cases) over the range of $N \in [20, 200]$ when $\beta = 2$ are shown in Figure 8. Again it is evident that larger matrix dimensions result in improved approximation accuracy. Figure 8 also shows that the two methods (25) and [15] exhibit similar approximation accuracies, though [15] is slightly better for $M = 20$ and (25) is slightly better for $M = 200$.

C. Discussion of results

The presented results indicate that the Gamma approximation can provide an accurate prediction of the empiric distribution of the largest eigenvalue. It was also shown that the approximation accuracy improves as the matrix dimensions

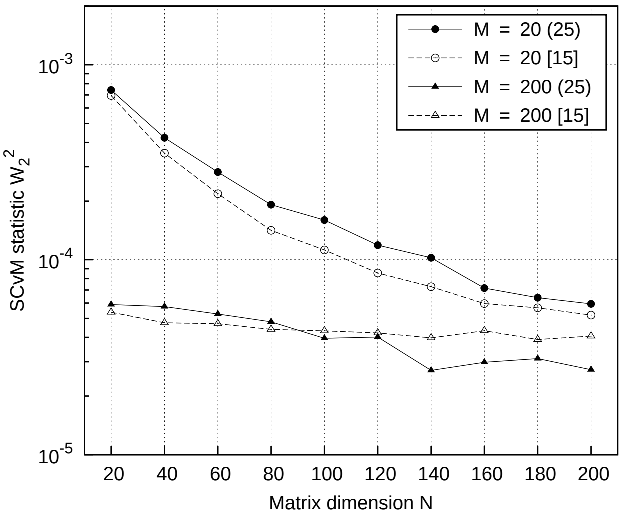

Fig. 8. SCvM statistics for λ_1 ($\beta = 2$) with M fixed over the range of N. The curves labelled (25) correspond to the Gamma approximation given in (25) and the curves labelled [15] correspond to the approximation method given in [15].

increase. This can be explained from (2) which is stated in terms of the edge scaling limits of the matrix dimensions. As the matrix dimensions increase, the TW law will provide a better prediction of the largest eigenvalue distribution. The approximation to the TW law will therefore also provide a more accurate prediction for larger matrix dimensions. Lastly, the approximation given by (25) is generally more accurate than [15] (especially for larger matrix dimensions). This can be ascribed to the different approximation methods. The method of [15] relies on matching moments (see Section IV-C) to find k and θ of the Gamma distribution. The method presented in this paper fits the Gamma distribution (k, θ and x_0) to the TW law directly and then uses (2) to derive the largest eigenvalue distribution. Using the shift parameter in the approximation provides a method to more accurately describe the TW law in terms of the Gamma distribution, which results in improved approximation accuracies.

VI. Conclusion

This paper presented an approximation to the Tracy-Widom law based on the Gamma distribution which was shown, through Monte Carlo computer simulation and an analysis of the distributions, to be able to accurately predict the largest eigenvalue distribution of white Wishart matrices and their corresponding sample covariance matrices. The approximation provides a tractable and closed-form solution and does not require numerical evaluation. Furthermore, simple equations were derived to accurately predict the statistics and support region of the principal component of a noise matrix directly from the matrix dimensions. The results of this paper can be used to develop analytic expressions where the Tracy-Widom law forms part of the argument. Such expressions will be useful in the analysis and application of detection receivers where decision thresholds in noisy environments are concerned, e.g. in MIMO, cognitive radio and signal detection systems.

VII. Acknowledgements

This work was supported by the Armaments Corporation of South Africa (Armscor) under contract KT521896. The authors would like to thank the anonymous reviewers for their valuable comments and suggestions.

VIII. Appendix

This appendix considers two approximations to TW_4. The first approximation (referred to as the indirect Gamma approximation) is based on the Gamma approximations to TW_1 and TW_2 developed in this paper. The second approximation (direct Gamma approximation) is obtained using the method presented in Section III-A.

A. Indirect Gamma approximation

The CDF F_4 given in (16) can be written in terms of F_1 and F_2 as

$$F_4\left(\frac{x}{\sqrt{2}}\right) = \cosh(\alpha(x))\sqrt{F_2(x)} \tag{33}$$

with

$$\alpha(x) = -\frac{1}{2}\int_x^\infty q(w)\ dw = \ln\left(\frac{F_1(x)}{\sqrt{F_2(x)}}\right) \tag{34}$$

from (14). The PDF f_4 can then be obtained by differentiation from (33) as

$$f_4\left(\frac{x}{\sqrt{2}}\right) = \frac{\sinh(\alpha(x))\sqrt{2F_2(x)}f_1(x)}{F_1(x)} + \frac{\exp(-\alpha(x))f_2(x)}{\sqrt{2F_2(x)}}. \tag{35}$$

It is required in (34) and (35) that $F_1(x) > 0$ and $F_2(x) > 0$. $F_4(x) = 0$ and $f_4(x) = 0$ wherever $F_1(x) = 0$ or $F_2(x) = 0$. By substituting the Gamma approximations g_β and G_β ($\beta = 1$ and 2) developed in Section III into f_β and F_β in (33) to (35), the indirect Gamma approximation $\tilde{\Gamma}_4$ is obtained. Using the double precision values obtained from [14] as reference (over $x \in [-10, 10]$ and $\Delta x = 0.0625$), the goodness-of-fit statistics (see Section III-C) are calculated as $W_4^2 = 1.1455 \times 10^{-5}$ and $K_4 = 5.4584 \times 10^{-3}$.

B. Direct Gamma approximation

The direct Gamma approximation Γ_4 was obtained using the method described in Section III-A and the numeric values from [14]. The resultant parameter values are given in Table V. The PDFs and CDFs of TW_4 from [14], Γ_4 and $\tilde{\Gamma}_4$ are displayed in Figure 9. It is evident from Figure 9 and the goodness-of-fit statistics given in the previous section and Table V that $\tilde{\Gamma}_4$ is a less accurate approximation than Γ_4. This can be expected since the approximation $\tilde{\Gamma}_4$ is based on approximations to TW_1 and TW_2.

TABLE V
PARAMETER VALUES AND RELATED RESULTS FOR THE DIRECT GAMMA APPROXIMATION TO TW_4.

Parameter	Symbol	$\beta = 4$
Shape	k	105.7442
Scale	θ	0.0700
Location	x_0	-9.7038
Mean	$k\theta + x_0$	-2.3017
Variance	$k\theta^2$	0.5181
Skewness	$2/\sqrt{k}$	0.1945
SSD	ϵ_β^2	1.0623×10^{-4}
SCvM statistic	W_β^2	1.9356×10^{-6}
Kolmogorov statistic	K_β	1.8025×10^{-3}

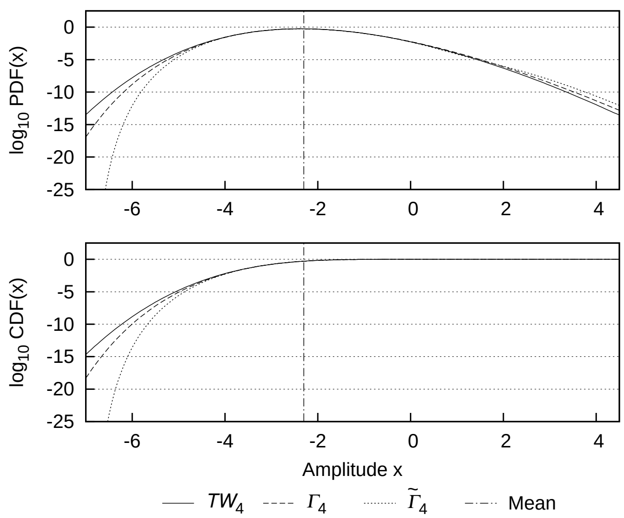

Fig. 9. Numeric and approximated PDFs and CDFs for $\beta = 4$. TW_4 refers to the numeric solution obtained from [14]. Γ_4 and $\tilde{\Gamma}_4$ refer respectively to the direct and indirect Gamma approximations to TW_4.

REFERENCES

1 Bejan, A.: ‘Largest eigenvalues and sample covariance matrices’. MSc dissertation, University of Warwick, UK, 2005
2 Zanella, A., Chiani, M., and Win, M.Z.: ‘On the marginal distribution of the eigenvalues of Wishart matrices’, IEEE Trans. Commun., 2009, 57, (4), pp. 1050–1060
3 Penna, F., Garello, R., and Spirito, M.A.: ‘Cooperative spectrum sensing based on the limiting eigenvalue ratio distribution in Wishart Matrices’, IEEE Commun. Lett., 2009, 13, (7), pp. 507–509
4 Xu, W., and Kaveh, M.: ‘Analysis of the performance and sensitivity of eigendecomposition-based detectors’, IEEE Trans. Signal Process., 1995, 43, (6), pp. 1413–1426
5 Edelman, A., and Rao, N. R.: ‘Random matrix theory’, Acta Numerica, 2005, 14, pp. 233–297
6 Telatar, I.E.: ‘Capacity of multi-antenna Gaussian channels’, Eur. Trans. Telecomm., 1999, 10, (6), pp. 585–595
7 Kwan, R., Leung, C., and Ho, P.: ‘Distribution of ordered eigenvalues of Wishart matrices’, Electronics Lett., 2007, 43, (5), pp. 31–32
8 Johnstone, I.M.: ‘Approximate null distribution of the largest root in multivariate analysis’, Ann. Appl. Stat., 2009, 3, (4), pp. 1616–1633
9 Johnstone, I.M.: ‘On the distribution of the largest eigenvalue in principal component analysis’, Ann. Statist., 2001, 29, (2), pp. 295–327
10 Tracy, C., and Widom, H.: ‘On orthogonal and symplectic matrix ensembles’, Commun. Math. Phys., 1996, 177, (3), pp. 727–754
11 Tracy, C., and Widom, H.: ’The distribution of the largest eigenvalue in the Gaussian ensembles’, in Van Diejen, J.F., Vinet, L. (Eds.): ’Calogero-Moser-Sutherland Models: CRM Series in Mathematical Physics 4’ (Springer-Verlag, 2000), pp. 461-472
12 Dieng, M.: ‘Distribution functions for edge eigenvalues in orthogonal and symplectic ensembles: Painlevé representations’. PhD thesis, University of California, Davis, USA, 2005
13 Prähofer, M., and Spohn, H.: ‘Exact scaling functions for one-dimensional stationary KPZ growth’, J. Stat. Phys., 2004, 115, (1/2), pp. 255–279
14 Bornemann, F.: ‘On the numerical evaluation of distributions in random matrix theory: a review’, Markov Processes Relat. Fields, 2010, 16, pp. 803–866
15 Wei, W., and Tirkkonen, O.: ‘Analysis of scaled largest eigenvalue based detection for spectrum sensing’. Proc. IEEE Int. Conf. Comms., Kyoto, Japan, June 2011, pp. 295–327
16 http://www-m5.ma.tum.de/KPZ, accessed April 2012
17 James, F.E.: ‘Statistical methods in experimental physics’ (World Scientific, 2006, 2nd edn. 2010)
18 Papoulis, A., and Pillai, S.U.: ‘Probability, random variables and stochastic processes’ (McGraw-Hill, 2002, 4th edn.)