An Efficient Sampling Scheme for the Eigenvalues of Dual Wishart Matrices

Ignacio Santamaria, *Senior Member, IEEE*, Víctor Elvira, *Senior Member, IEEE*

***Abstract*—Despite the numerous results in the literature about the eigenvalue distributions of Wishart matrices, the existing closed-form probability density function (pdf) expressions do not allow for efficient sampling schemes from such densities. In this letter, we present a stochastic representation for the eigenvalues of 2×2 complex central uncorrelated Wishart matrices with an arbitrary number of degrees of freedom (referred to as dual Wishart matrices). The draws from the joint pdf of the eigenvalues are generated by means of a simple transformation of a chi-squared random variable and an independent beta random variable. Moreover, this stochastic representation allows a simple derivation, alternative to those already existing in the literature, of some eigenvalue function distributions such as the condition number or the ratio of the maximum eigenvalue to the trace of the matrix. The proposed sampling scheme may be of interest in wireless communications and multivariate statistical analysis, where Wishart matrices play a central role.**

***Index Terms*—Random matrices, Wishart matrices, eigenvalue distribution, Rayleigh fading channels, multiple-input multiple-output (MIMO) communications**

I. Introduction

The statistical properties of the eigenvalues of Wishart matrices play a fundamental role in multiple applications. For instance, they characterize the performance of multiple-input multiple-output (MIMO) wireless communication systems in Rayleigh fading channels. As another example, many of the standard test statistics in multivariate analysis are functions of eigenvalues of a Wishart matrix. Given the importance of the problem, an extensive body of literature exists on the distribution of the eigenvalues of Wishart matrices that can be traced back to the classical works of James [1] and Katri [2] in the 1960s and, more recently, to Edelman's thesis [3]. In the context of MIMO communications, the distribution of the unordered eigenvalues of a Wishart matrix of the form $\mathbf{W} = \mathbf{H}^H\mathbf{H}$ (where $\mathbf{H}$ is the MIMO channel with complex Gaussian entries) was used to derive the ergodic capacity when the channel is assumed to be known at the receiver but not at the transmitter in [4]. The distribution of the largest eigenvalue of $\mathbf{W}$ characterizes the performance of beamforming MIMO systems that apply maximum-ratio transmission/combining [5], while the distribution of the smallest eigenvalue has application to antenna selection techniques [6], [7], and beamforming [8]. Many other expressions of interest concerning the distributions of the eigenvalues of Wishart matrices and their functions have appeared in the vast literature dealing with this subject [9]–[13].

Despite the large number of results about the eigenvalue distributions of Wishart matrices and their functions, the simulation from such distributions is in general not possible, since there are no methods to generate draws from the existing probability density function (pdf) closed-form expressions. Therefore, the conventional method to simulate the joint pdf of ordered eigenvalues is still to generate a Gaussian random matrix, $\mathbf{H}$, calculate the Wishart Hermitian matrix $\mathbf{W} = \mathbf{H}^H\mathbf{H}$, and finally obtain its eigenvalues. This procedure can be clearly inefficient, so it would be of interest in many problems to have simpler sampling methods for the eigenvalues of Wishart matrices, for instance for proposing more efficient methods to characterize the symbol error rate (SER) performance of MIMO detectors [14]–[16]. In this letter, we propose a stochastically equivalent representation for the eigenvalues of 2×2 complex Wishart matrices with an arbitrary number of degrees of freedom that, following [17], are termed as *dual Wishart matrices*. The proposed stochastic characterization is based on a simple transformation of a chi-squared random variable and a beta random. Hence, it leads to an efficient sampling scheme that may be of interest in applications of wireless communications and multivariate statistical analysis where Wishart matrices play a central role. Our results suggest that the proposed sampling scheme is orders of magnitude faster than the conventional simulation procedure. Furthermore, it allows a simple derivation, alternative to those already existing in the literature, of some eigenvalue function distributions such as the condition number or the ratio of the maximum eigenvalue to the trace.

Notation. We use upper case boldface for matrices. The $n \times n$ identity matrix is denoted as $\mathbf{I}_n$. The symbol $()^H$ denotes Hermitian. The trace and determinant of a matrix $\mathbf{W}$ are denoted as $\mathrm{tr}(\mathbf{W})$ and $\det(\mathbf{W})$, respectively. The symbol $\sim$ denotes "distributed as", and $\propto$ means "proportional to". The notation used for distributions is as follows: $x \sim \mathcal{CN}(0,1)$ denotes a complex Gaussian distribution with zero mean and unit variance; $\mathbf{W} \sim \mathcal{CW}_T(\mathbf{\Sigma}, R)$ means that the $T \times T$ matrix $\mathbf{W}$ follows a complex central Wishart distribution with parameter $\mathbf{\Sigma}$ and R degrees of freedom; $x \sim \chi^2_\nu$ denotes a central chi-square with ν degrees of freedom; $x \sim \mathrm{Beta}(\alpha, \beta)$ and $x \sim \mathrm{Gamma}(\alpha, \beta)$ denote beta and gamma distributions, respectively, with parameters α and β.

I. Santamaria is with the Department of Communications Engineering, University of Cantabria, Santander, Spain (e-mail: i.santamaria@unican.es).

V. Elvira is with School of Mathematics, University of Edinburgh, United Kingdom (e-mail victor.elvira@ed.ac.uk).

The work of I. Santamaria was supported by Ministerio de Ciencia e Innovación and AEI (10.13039/501100011033), under grant PID2019-104958RB-C43 (ADELE). The work of V. E. is supported by the *Agence Nationale de la Recherche* of France under PISCES (ANR-17-CE40-0031-01), and by ARL/ARO under grant W911NF-20-1-0126.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2021.3121197, IEEE Signal Processing Letters

II. Preliminaries

Let us define a complex normal random matrix $\mathbf{H}$ of dimensions $R \times 2$ ($R \geq 2$) whose elements are independent and identically distributed (i.i.d.) complex normal random variables $h_{i,j} \sim \mathcal{CN}(0,1)$. In wireless communications, for instance, $\mathbf{H}$ represents the $R \times 2$ Rayleigh MIMO channel between the low-complexity user equipment (UE) with only $T = 2$ antennas and the base station (BS) with R antennas. This scenario is particularly important in MIMO wireless communications, where the cost and size constraints of the mobile terminals limit the number of antennas of UEs, while the BS can have a much larger number of antennas. In fact, with 4G Long Term Evolution (LTE) networks the use of two receiver antennas in the user equipment (UE) has become the norm [18].

The 2×2 Grammian $\mathbf{W} = \mathbf{H}^H\mathbf{H}$ follows a complex central Wishart distribution with R degrees of freedom $\mathbf{W} \sim \mathcal{CW}_2(\mathbf{I}_2, R)$. Its density is $f(\mathbf{W}) \propto \det(\mathbf{W})^{R-2}\exp(-\operatorname{tr}(\mathbf{W}))$. The Wishart distribution in the particular case of dual matrices ($T = 2$) was first derived by Fisher in 1915 [19], and for a general $T \geq 2$ was derived by Wishart in 1928 [20].

The joint distribution of the ordered eigenvalues $\lambda_1 \geq \lambda_2 \geq 0$ is

$$f(\lambda_1, \lambda_2) = \frac{e^{-(\lambda_1+\lambda_2)}}{\Gamma(R)\Gamma(R-1)}(\lambda_1\lambda_2)^{R-2}(\lambda_1-\lambda_2)^2. \quad (1)$$

Although the distribution in Eq. (1) admits a simple expression, to the best of the authors' knowledge there are no efficient procedures to sample from this distribution. An approximate method based on importance sampling has been proposed in [21], but it only provides an approximation of the true distribution of Eq. (1), and further, it is useful primarily for approximating the distribution of arbitrary eigenvalue functions, but not for generating samples of the joint distribution. In practice, the standard method for simulating from $f(\lambda_1, \lambda_2)$ is to generate $R \times 2$ complex normal matrices with $\mathcal{CN}(0,1)$ entries, and then directly calculate the eigenvalues of $\mathbf{W} = \mathbf{H}^H\mathbf{H}$. Although this procedure does not represent any difficulty in most cases, it can be inefficient especially when R grows. This is due to the need to generate $2R$ complex normal random variables, then calculate the 2×2 Wishart matrix, and finally obtain its eigenvalues by solving a simple quadratic equation.

It is possible to design a more efficient simulation scheme based on the well-known Bartlett's factorization theorem [22]. Bartlett's theorem states that the $R \times 2$ complex Gaussian data matrix $\mathbf{H}$ can be factored as $\mathbf{H} = \mathbf{Q}\mathbf{R}$, where the $R \times 2$ matrix $\mathbf{Q}$ is a slice of an $R \times R$ unitary matrix (uniformly distributed on the Stiefel manifold), and $\mathbf{R}$ is a 2×2 upper triangular matrix with positive elements on its diagonal

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix}. \quad (2)$$

According to Bartlett's factorization theorem, the elements of $\mathbf{R}$ are distributed as follows: $r_{11}^2 \sim \frac{1}{2}\chi^2_{2R}$, $r_{22}^2 \sim \frac{1}{2}\chi^2_{2(R-1)}$ and $r_{12} \sim \mathcal{CN}(0,1)$. Moreover, they are independent. Note that the dual Wishart matrix is stochastically equivalent to

$$\mathbf{W} = \mathbf{H}^H\mathbf{H} = \mathbf{R}^H\mathbf{R} = \begin{bmatrix} r_{11}^2 & r_{11}r_{12} \\ r_{12}^* r_{11} & |r_{12}|^2 + r_{22}^2 \end{bmatrix}. \quad (3)$$

Hence, a more efficient procedure to sample from $f(\lambda_1, \lambda_2)$ amounts to generating 3 random variables: $r_{11}^2 \sim \frac{1}{2}\chi^2_{2R}$, $r_{22}^2 \sim \frac{1}{2}\chi^2_{2(R-1)}$ and $r_{12} \sim \mathcal{CN}(0,1)$; then form $\mathbf{W}$ as in (3) and obtain its eigenvalues.

III. Proposed stochastic characterization

The main contribution of this letter is to propose a yet more efficient sampling method that only requires the generation of two random variables and then performs a simple transformation of them. The method is based on a stochastic characterization of the eigenvalues obtained through the characteristic polynomial of the dual Wishart matrix. The eigenvalues $\lambda_1 \geq \lambda_2 \geq 0$ of $\mathbf{W}$ satisfy the characteristic polynomial

$$\det(\mathbf{W} - \lambda\mathbf{I}_2) = \lambda^2 - \lambda\operatorname{tr}(\mathbf{W}) + \det(\mathbf{W}), \quad (4)$$

whose roots are

$$\lambda_i = \frac{1}{2}\operatorname{tr}(\mathbf{W})\left(1 \pm \sqrt{1 - \frac{\det(\mathbf{W})}{\left(\frac{1}{2}\operatorname{tr}(\mathbf{W})\right)^2}}\right), \quad i = 1, 2, \quad (5)$$

where λ_1 (λ_2) corresponds to the root with positive (negative) sign. It is interesting to note that the term inside the square root, $\eta = \frac{\det(\mathbf{W})}{\left(\frac{1}{2}\operatorname{tr}(\mathbf{W})\right)^2}$, is the sphericity statistic [23], which is the generalized likelihood ratio (GLR) statistic to test whether an i.i.d. sequence of vector-valued normal random vectors are spatially white or not. The exact distribution of the sphericity-test in the complex-valued case has been derived in [24], [25]. A stochastic representation of the sphericity statistic under the null hypothesis as a product of independent beta distributions was derived in [26]. The distribution of η when $\mathbf{W} \sim \mathcal{CW}_2(\mathbf{I}_2, R)$ reduces to

$$\eta = \frac{\det(\mathbf{W})}{\left(\frac{1}{2}\operatorname{tr}(\mathbf{W})\right)^2} \sim \operatorname{Beta}\left(R-1, \frac{3}{2}\right), \quad (6)$$

and it is independent of $\operatorname{tr}(\mathbf{W}) = \operatorname{tr}(\mathbf{H}^H\mathbf{H}) = \lambda_1 + \lambda_2 \sim \frac{1}{2}\chi^2_{4R}$. Note that if $\eta \sim \operatorname{Beta}(R-1, 3/2)$, then $1 - \eta \sim \operatorname{Beta}(3/2, R-1)$.

To summarize, an exact procedure to sample from the joint $f(\lambda_1, \lambda_2)$ in Eq. (1) is:

1) Generate $s \sim \operatorname{Beta}\left(\frac{3}{2}, R-1\right)$ and $t \sim \frac{1}{2}\chi^2_{4R}$
2) Calculate λ_1 and λ_2 as the following transformation of s and t

$$\begin{aligned} \lambda_1 &= \frac{1}{2}t\left(1 + \sqrt{s}\right), \\ \lambda_2 &= \frac{1}{2}t\left(1 - \sqrt{s}\right), \end{aligned} \quad (7)$$

The procedure does not involve matrix operations or eigenvalue decompositions, simply one square root, two additions and two multiplications.

a) The case of real central Wishart matrices: When $\mathbf{H}$ is a real $R \times 2$ normal matrix with i.i.d. entries, then $\mathbf{W} = \mathbf{H}^T\mathbf{H} \sim \mathcal{W}_2(\mathbf{I}_2, R)$ follows a real central Wishart matrix with R degrees of freedom. In this case, $\mathrm{tr}(\mathbf{W}) \sim \chi^2_{2R}$ and, as proved in [27], the distribution of the sphericity ratio is

$$\eta = \frac{\det(\mathbf{W})}{\left(\frac{1}{2}\,\mathrm{tr}(\mathbf{W})\right)^2} \sim \mathrm{Beta}\left(\frac{R-1}{2}, 1\right). \tag{8}$$

Therefore, to sample from the joint $f(\lambda_1, \lambda_2)$ in the real case we generate $s \sim \mathrm{Beta}\left(1, \frac{R-1}{2}\right)$ and $t \sim \chi^2_{2R}$, and then calculate λ_1 and λ_2 as in (7).

A. Characterization of eigenvalue functions

The proposed stochastic characterization allows for a simple derivation of the distribution of certain functions of the eigenvalues of dual Wishart matrices. Let us begin with the condition number $c = \lambda_1/\lambda_2$. In numerical analysis and matrix algebra, the condition number determines the sensitivity of solving linear systems, which are ill-posed when the condition number is very large. In MIMO communications a large condition number of $\mathbf{H}^H\mathbf{H}$ implies a near rank-deficient matrix, which in turn impacts the main figures of merit such as multiplexing gain or symbol error rate. An expression for the distribution of the condition number of $R \times 2$ complex Gaussian matrices was originally given by Edelman in [3, Chapter 7, Eq. (7.2)]. The expression was derived by integrating the joint distribution with respect to the condition number. The derivation is simple but tedious, so it is omitted in [3]. A comprehensive study on the condition number distribution of complex Wishart matrices was carried out in [17]. Theorem 2 in [17] provides a closed-form expression of the cumulative density function (cdf) of the condition number of $\mathbf{W} \sim \mathcal{CW}_2(\mathbf{I}_2, R)$, and the corresponding pdf is obtained by derivating the cdf in Corollary 1 (cf. [17, Eq. (15)]). The pdf expression, however, involves the summation of R terms.

Using the parametrization in terms of t and s, it follows that the condition number is stochastically equivalent to

$$c = \frac{1+\sqrt{s}}{1-\sqrt{s}}, \quad s \sim \mathrm{Beta}\left(3/2, R-1\right), \tag{9}$$

so it is a one-to-one mapping of a beta random variable from the interval $0 \le s \le 1$ to the interval $1 \le c \le \infty$. It is then straightforward to obtain the pdf of c as

$$f(c) = K\frac{(c-1)^2 c^{R-2}}{(c+1)^{2R}}, \quad 1 \le c \le \infty, \tag{10}$$

where the normalizing constant is

$$K = \frac{2^{(2R-1)}\Gamma\left(R+\frac{1}{2}\right)}{\sqrt{\pi}\Gamma(R-1)} = \frac{\Gamma(2R)}{\Gamma(R)\Gamma(R-1)}. \tag{11}$$

The pdf $f(c)$ is related to [3, Eq. (7.2)], which is the pdf of $c^{1/2}$. We believe that the derivation based on the new stochastic characterization is much simpler than the procedure followed in either [3] or [17].

In addition, this result opens up the possibility of applying alternative Monte Carlo schemes such as importance sampling [28]. For example, in certain applications it may be of interest to generate eigenvalues of Wishart matrices that are ill-conditioned (values of $c >> 1$) with higher probability than that given by the true pdf in Eq. (10). To do this, it is sufficient to sample from $s \sim \mathrm{Beta}\left(a, R-1\right)$ with $a > 3/2$, while maintaining the trace distribution $t \sim \frac{1}{2}\chi^2_{4R}$. The samples of s generated this way are more likely to be close to 1, thus increasing the probability of obtaining higher condition numbers.

Another useful function of the eigenvalues is the ratio $\Lambda = \frac{\lambda_1}{\lambda_1+\lambda_2}$ which, as proved by Besson and Scharf in [29], is the generalized-likelihood ratio statistic for the problem of detecting a signal whose spatial signature is known to lie in a one-dimensional subspace in the presence of white Gaussian noise of unknown variance, also known as the matched direction detector. Using our result, the distribution of this statistic is

$$\Lambda = \frac{1+\sqrt{s}}{2}, \quad s \sim \mathrm{Beta}\left(3/2, R-1\right), \tag{12}$$

and its density is

$$f(\Lambda) = 4K\left(\Lambda - \frac{1}{2}\right)^2 (1-\Lambda)^{R-2}\Lambda^{R-2}, \quad \frac{1}{2} \le \Lambda \le 1, \tag{13}$$

which is the result in [29, Eq. (25)].

Interestingly, the proposed stochastic characterization shows that the trace $t = \lambda_1 + \lambda_2$ and the condition number $c = \lambda_1/\lambda_2$ are independent random variables. This suggests yet another parametrization of the joint in (1) that can be factored as the product of the marginals of t and c. The transformation is

$$\begin{aligned} \lambda_1 &= \frac{ct}{1+c}, \\ \lambda_2 &= \frac{t}{1+c}, \end{aligned} \tag{14}$$

whose Jacobian determinant is $J = t/(1+c)^2$. Then, the joint distribution of t and c is

$$f(c,t) = \frac{t^{2R-1}e^{-t}}{\Gamma(R)\Gamma(R-1)}\frac{(c-1)^2 c^{R-2}}{(c+1)^{2R}}, \tag{15}$$

which is the product of the marginal of c, given in (10), and the marginal of the trace $f(t) = t^{2R-1}e^{-t}/\Gamma(2R)$. Therefore, c and t are independent random variables for dual Wishart matrices.

Finally, the distribution of $d = \lambda_1 - \lambda_2$ can also be obtained by resorting to the proposed stochastic characterization. We have that $d \sim t\sqrt{s}$, so after some algebra the distribution of d is found to be

$$f(d) = \sum_{k=0}^{R-2} h_k d^{2(k+1)}\Gamma(2(R-2-k)+1, d), \tag{16}$$

where $\Gamma(a,x) = \int_x^\infty t^{a-1}e^{-t}dt$ denotes the upper incomplete gamma function, and the mixture coefficients are

$$h_k = \frac{4\binom{R-2}{k}(-1)^k\Gamma\left(R+\frac{1}{2}\right)}{\sqrt{\pi}\Gamma(2R)\Gamma(R-1)}. \tag{17}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LSP.2021.3121197, IEEE Signal Processing Letters

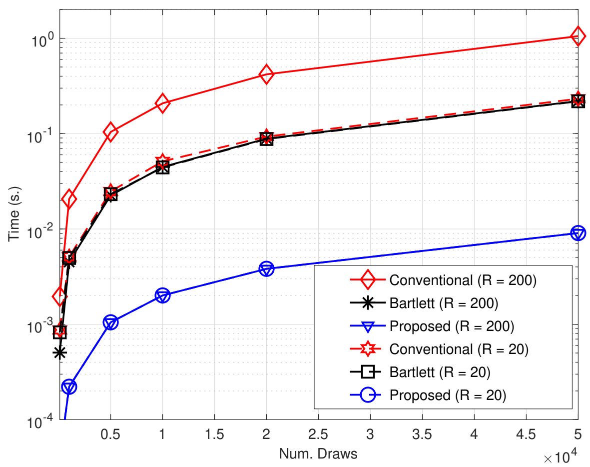

Fig. 1. Simulation time to generate a certain number of draws using the conventional method (red line), Bartlett's method (black line), and the proposed method (blue line) for $R = 20$ (dashed) and $R = 200$ (solid).

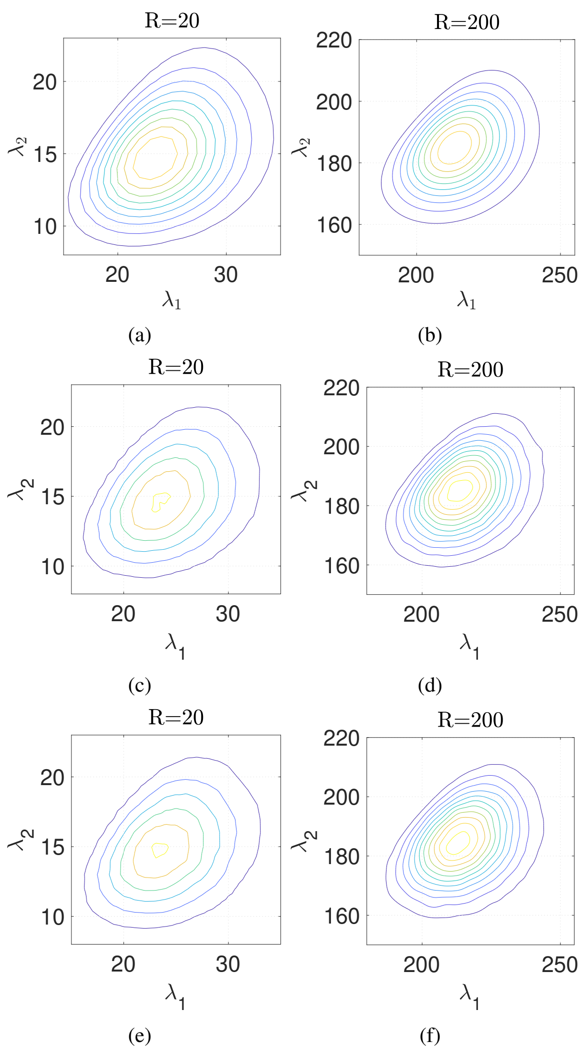

Fig. 2. Contour plot of the theoretical bi-variate pdf of λ_1 and λ_2) in Eq. (1) (top), and a kernel density estimator of the $200,000$ pairs of eigenvalues simulated by means of the conventional sampling method (middle) and the proposed stochastic characterization (bottom), for $R = 2$ (left) and $R = 20$ (right).

TABLE I
COMPARISON SAMPLING METHODS

	# r.v.'s	real operations		
		#products	#sums	#sqrt
Conventional method	$2R\ \mathcal{CN}(0,1)$	$O(R)$	$O(R)$	1
Bartlett's method	2 chi-square 1 $\mathcal{CN}(0,1)$	10	8	1
Proposed method	1 chi-square 1 beta	2	2	1

IV. RESULTS

Table I summarizes the computational complexity, in terms of the number of random variables generated and the real floating-point operations, of the different parametrizations for the simulation of one bi-variate sample of (λ_1, λ_2). The computational complexity of the conventional method grows with R. The method based on the proposed stochastic representation is by far the simplest.

For a more detailed comparison between the different simulation methods, Fig. 1 shows the execution time required to generate a certain number of realizations of the eigenvalues (λ_1, λ_2). The simulation has been performed in Matlab (R2020b) running on a personal computer with CPU at 3.60 GHz and 128 GB of RAM. Clearly, the results may vary depending on the programming language or the computer environment, so they are only intended to give an approximate idea of the simulation speedup provided by the proposed method. Clearly, as seen in Fig. 1, the proposed method is orders of magnitude faster than both the conventional method and Bartlett's method. This is partly because the proposed method does not need any loop and can generate the desired number of chi-squared and beta samples very efficiently. In addition, the simulation time of the conventional method increases with the degrees of freedom, R, while the simulation time of the proposed method, as well as Bartlett's method, do not depend on R.

In Fig. 2, we display the bi-variate pdfs for $R = 2$ (first column) and $R = 20$ (second column), for the theoretical joint pdf of (λ_1, λ_2) (top), and a kernel density estimator of the $200,000$ pairs of eigenvalues simulated by means of the conventional sampling method (middle) and the proposed stochastic characterization (bottom).

V. CONCLUSION

In this letter, we have shown that, to generate realizations of the eigenvalues of dual Wishart matrices with an arbitrary number of degrees of freedom, it is sufficient to simulate a beta random variable and an independent chi-square random variable, and then apply a simple transformation. The procedure is computationally simpler than the standard sampling method, especially for the generation of a large number of draws. Future work will consider extensions to other cases, e.g., non-central Wishart distributions or Wishart matrices of dimension greater than two.

References

[1] 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.

[2] C. G. Khatri, "Distribution of the largest or the smallest characteristic root under null hypothesis concerning complex multivariate normal populations," *Ann. Math. Statist.*, vol. 35, no. 4, pp. 1807–1810, 1964.

[3] A. Edelman, *Eigenvalues and condition numbers of random matrices*. Cambridge, MA, USA: Ph.D dissertation, MIT, 1989.

[4] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[5] M. Kang and M. S. Alouini, "Largest eigenvalue of complex Wishart matrices and performance analysis of MIMO MRC systems," *IEEE Journal Selected Areas in Comm.*, vol. 21, no. 3, pp. 418–426, Apr. 2003.

[6] R. W. Heath and D. J. Love, "Multimode antenna selection for spatial multiplexing systems with linear receivers," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3042–3056, Aug. 2005.

[7] R. W. Heath and A. J. Paulraj, "Switching between diversity and multiplexing in MIMO systems," *IEEE Trans. Commun.*, vol. 53, no. 6, pp. 962–968, Jun. 2005.

[8] S. Jin, M. R. McKay, X. Gao, and I. B. Collings, "MIMO multichannel beamforming: SER and outage using new eigenvalue distributions of complex noncentral Wishart matrices," *IEEE Trans. Commun.*, vol. 56, no. 3, pp. 424–434, Mar. 2008.

[9] A. Zanella, M. Chiani, and M. Z. Win, "On the marginal distribution of the eigenvalues of Wishart matrices," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1050–1060, April 2009.

[10] A. Zanella and M. Chiani, "The pdf of the l-th largest eigenvalue of central Wishart matrices and its application to the performance analysis of MIMO systems," in *Global Communications Conference (GLOBECOM)*, New Orleans (LA), USA, Nov. 2008.

[11] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[12] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.

[13] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR power offset in multiantenna ommunication," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.

[14] V. Elvira and I. Santamaria, "Multiple importance sampling for efficient symbol error rate estimation," *IEEE Signal Proc. Letters*, vol. 26, no. 3, pp. 420–424, 2019.

[15] ——, "Multiple importance sampling for symbol error rate estimation of maximum-likelihood detectors in MIMO channels," *IEEE Trans. Signal Process.*, vol. 69, pp. 1200–1212, 2021.

[16] N. Kim, Y. Lee, and H. Park, "Performance analysis of MIMO systems with linear MMSE receiver," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4474–4478, Nov. 2008.

[17] 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, 2010.

[18] C. Cox, *An introduction to LTE : LTE, LTE-advanced, SAE and 4G mobile communications*. John Wiley, 2012.

[19] R. A. Fisher, "Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population," *Biometrika*, vol. 10, no. 4, pp. 507–521, May 1915.

[20] J. Wishart, "The generalised product moment distribution in samples from a normal multivariate population," *Biometrika*, vol. 20, no. 1/2, pp. 32–52, 1928.

[21] S. Huzurbazar and R. W. Butler, "Importance sampling for p-value computations in multivariate tests," *Journal of Computational and Graphical Statistics*, vol. 7, no. 3, pp. 342–355, 1998.

[22] M. S. Bartlett, "On the theory of statistical regression," *Proc. Roy. Soc. Edinb.*, vol. 53, pp. 260–283, 1933.

[23] J. Mauchly, "Significance test for sphericity of a normal n-variate distribution," *Ann. Math. Statist.*, vol. 11, p. 204–209, 1940.

[24] D. K. Nagar, S. K. Jain, and A. K. Gupta, "Distribution of LRC for testing sphericity of a complex multivariate Gaussian model," *Int. J. Math. & Math. Sci.*, vol. 8, no. 3, pp. 555–562, 1985.

[25] Y. A. N. Spencer and A. Gorokhov, "Bounds on maximum likelihood ratio — Part I: Application to antenna array detection-estimation with perfect wavefront coherence," *IEEE Trans. Signal Process.*, vol. 52, no. 6, p. 1524–1536, Jun. 2014.

[26] N. Klausner, M. R. Azimi-Sadjadi, and L. L. Scharf, "Detection of spatially-correlated time series from a network of sensor arrays," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1396–1407, March 2014.

[27] M. S. Srivastava and C. G. Khatri, *An Introduction to Multivariate Statistics*. North Holland, 1979.

[28] V. Elvira and L. Martino, "Advances in importance sampling," *Wiley StatsRef: Statistics Reference Online*, pp. 1–22, 2020.

[29] O. Besson and L. L. Scharf, "CFAR matched direction detector," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2840–2844, July 2006.