INVARIANCE OF THE DISTRIBUTIONS OF NORMALIZED GRAM MATRICES

Stephen D. Howard[1] *Songsri Sirianunpiboon*[1] *Douglas Cochran*[2]

[1]Defence Science and Technology Organisation
PO Box 1500, Edinburgh 5111, Australia

[2]School of Mathematical and Statistical Sciences
Arizona State University, Tempe AZ 85287-5706 USA

ABSTRACT

Normalized Gram matrices formed from multiple vectors of sensor data, and functions of the eigenvalues of such matrices in particular, have a long history in connection with multiple-channel detection. The determinant and various other functions of the eigenvalues of these matrices arise as detection statistics in a variety of multi-channel problems, and knowledge of their distributions under the H_0 assumption that the sensor channels are independent and contain only white gaussian noise is consequently important for determining false-alarm probabilities for multi-channel detectors. Invariance of the H_0 distribution of the eigenvalues to one data channel is significant in some applications. This paper derives the H_0 distribution of a normalized Gram matrix and, as corollaries, obtains the distribution of the determinant as well as invariance results for the matrix that carry over to its spectrum. The essential symmetry property of white gaussian noise on which these results depend is also noted.

Index Terms— Coherence, Multiple-channel detection, Gram matrix, Stiefel manifold

1. INTRODUCTION

Normalized gram matrices have a rich history in connection with multi-channel signal analysis, dating back at least as far as the use of the magnitude-squared coherence (MSC) estimate as a statistic for detecting the presence of a common but unknown signal in two noisy receiver channels [1]. The M-channel generalized coherence (GC) estimate, introduced in [2] and developed more fully in [3], is explicitly formulated in terms of the determinant of an $M \times M$ normalized gram matrix $\hat{G}$ as $1 - |\hat{G}|$ and shown to reduce to the MSC estimate when $M = 2$. Recently, functions of the eigenvalues of normalized gram matrices have been proposed as detection statistics for spectrum sensing applications, and statistics arising naturally (e.g., from Bayesian and generalized likelihood ratio formulations) in multi-measurement signal rank estimation and in detection of signals having known rank.

Even before introduction of the GC estimate, the distribution of the MSC estimate was shown not to depend on the distribution of one of the two data vectors from which is is formed, provided that the second vector is zero-mean white gaussian noise and is independent from the first [4, 5]. This result is of practical importance because it allows false alarm thresholds computed under the H_0 assumption that both channels contain white gaussian noise, and are independent, to be used when one channel has arbitrary distribution; e.g., it contains a structured signal detectable by single-channel processing. A similar result for the GC estimate was given in [6], and invariance of the individual eigenvalues of $\hat{G}$ to the statistics of one channel was shown in [7].

This paper develops a geometric perspective on the distribution of normalized Gram matrices that is used to derive the distribution of the matrix under suitable H_0 assumptions. As a direct corollary, this yields an invariance result for the distribution of the matrix that implies invariance of, for example, functions of its eigenvalues as special cases. Section 2 introduces the geometric formulation, together with necessary notation and terminology, and also establishes a few preliminary results used in the proof of the main theorem. Section 3 gives a precise statement of the invariance theorem that is the main result of the paper. Section 4 uses this result to derive the form of the determinant of the normalized Gram matrix under signal-absent hypotheses used in other literature, thus reproducing an important known result in this area using the approaches introduced here. A short conclusion closes the paper.

2. MATHEMATICAL FORMULATION

This section establishes notation and develops the geometrical framework that are used in Section 3 to prove the main result of this paper; i.e., that the distribution of a normalized Gram matrix is, under suitable null hypotheses about its constituent data vectors, invariant to the distribution of one of those vectors.

Consider a set of M complex vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M \in \mathbb{C}^N$ with $M \leq N$. The Gram matrix of this set of vectors, denoted by $G(\mathbf{x}_1, \cdots, \mathbf{x}_M)$, is the $M \times M$ positive semi-definite hermitian matrix whose elements are $g_{ij} = \langle \mathbf{x}_i, \mathbf{x}_j \rangle = \mathbf{x}_j^\dagger \mathbf{x}_i$, where † denotes conjugate transpose. Denoting by X the $N \times M$ matrix whose m^{th} column is $\mathbf{x}_m$, the Gram matrix can be written as $G = X^\dagger X$.

The normalized Gram matrix $\widehat{G}$ is obtained by normalizing the vectors $\mathbf{x}_j$ to unit length; i.e.,

$$\hat{g}_{ij} = \left\langle \frac{\mathbf{x}_i}{\|\mathbf{x}_i\|}, \frac{\mathbf{x}_j}{\|\mathbf{x}_j\|} \right\rangle = \frac{\langle \mathbf{x}_i, \mathbf{x}_j \rangle}{\|\mathbf{x}_i\| \|\mathbf{x}_j\|}. \tag{1}$$

The elements on the main diagonal of $\hat{G}$ are $\hat{g}_{ii} = 1$, and its determinant is

$$|\widehat{G}| = \frac{|G|}{\|\mathbf{x}_1\|^2 \cdots \|\mathbf{x}_M\|^2}$$

In [8] and [9], this structure is extended to the context of vector-valued time series. In this generalization, $X_j \in \mathbb{C}^{N \times K}$ for $j = 1, \ldots, M$. Denoting $\mathbf{X} = (X_1, X_2, \ldots, X_M)$. The Gram matrix associated with $\mathbf{X}$ is $\mathbf{G} = \mathbf{X}^\dagger \mathbf{X}$, which is an $M \times M$ block matrix

This work was supported in part by the U.S. Air Force Office of Scientific Research under Grant No. FA9550-12-1-0225.

with block elements $G_{ij} = X_i^\dagger X_j$. It is always possible to decompose X_j uniquely as

$$X_j = \widehat{X}_j R_j \qquad (2)$$

where $\widehat{X}_j$ is semi-unitary; i.e., $\widehat{X}_j^\dagger \widehat{X}_j = I_K$, the $K \times K$ identity matrix, and $R_j \in \mathbb{C}^{K\times K}$ is an upper triangular matrix with positive diagonal elements. Note that $R_j^\dagger R_j = X_j^\dagger X_j$. The block elements of the normalized Gram matrix $\widehat{\mathbf{G}}$ are

$$\widehat{G}_{ij} = (R_i^{-1})^\dagger X_i^\dagger X_j R_j^{-1}.$$

When $K = 1$, this reduces to (1). The normalized Gram matrix has the properties that $\widehat{G}_{ii} = I_K$ and

$$|\widehat{\mathbf{G}}| = \frac{|\mathbf{G}|}{|G_{11}|\cdots|G_{MM}|},$$

which is the detection statistic defined and analyzed in [8, 9].

In the decomposition (2), the columns of $\widehat{X}_i$ comprise a set of K orthonormal vectors in $\mathbb{C}^N$. Such a set is known as a K-*frame*, and the collection of all K-frames in $\mathbb{C}^N$ defines a smooth manifold $\mathcal{V}_{K,N}$ of dimension $2NK - K^2$ called the Stiefel manifold [10]. $\mathcal{V}_{K,N}$ admits an invariant measure $d\mu_{\mathcal{V}_{K,N}}(\widehat{X})$ and has volume

$$\mathrm{vol}(\mathcal{V}_{K,N}) = \int_{\mathcal{V}_{K,N}} d\mu_{\mathcal{V}_{K,N}}(\widehat{X}) = \prod_{\ell=1}^{K} \mathrm{vol}(S^{2(N-\ell)+1})$$

where

$$\mathrm{vol}(S^{m-1}) = \frac{2\pi^{m/2}}{\Gamma(m/2)}$$

is the volume of the unit $(m-1)$-sphere.

Denote by $\mathcal{T}_{K,N}$ the space of matrices of the form $[A^\mathsf{T}\ \ T^\mathsf{T}]^\mathsf{T}$ where $A \in \mathbb{C}^{(N-K)\times K}$ and T is a $K \times K$ upper triangular matrix with real positive diagonal elements. Thus $R_i \in \mathcal{T}_{K,K}$ in (2).

Suppose $X \in \mathbb{C}^{N\times K}$ is a random matrix having probability density f with respect to Lebesgue measure on $\mathbb{C}^{N\times K}$. The matrix X is said to have a unitarily invariant probability distribution if

$$f(VX) = f(X) \qquad (3)$$

for all $N \times N$ unitary matrices $V \in U(N)$. Decomposing X according to (2) as $X = \widehat{X}R$, and noting that there always exists a $V \in U(N)$ that maps an arbitrary K-frame to a standard K-frame $\widehat{E}_0$, (3) implies that

$$f(X) = f(\widehat{X}R) = f(\widehat{E}_0 R) = h(R^\dagger R) = h(X^\dagger X)$$

for some function h.

An additional element preliminary to establishing the main result, Theorem 1, in the following Section is calculating the Jacobian for the transformation (2). If $X \in \mathbb{C}^{N\times K}$, then under the decomposition $X = \widehat{X}R$ Lebesgue measure in $\mathbb{C}^{N\times K}$ becomes

$$dX = |R|^{2N} d\mu_{\mathcal{T}_{K,K}}(R)\ d\mu_{\mathcal{V}_{K,N}}(\widehat{X}),$$

where

$$d\mu_{\mathcal{T}_{K,K}}(R) = \prod_{j=1}^{K} r_{jj}^{-2j+1} dr_{jj} \prod_{i<j=1}^{K} d\,\mathrm{Re}(r_{ij}) d\,\mathrm{Im}(r_{ij}).$$

In this expression, r_{ij} are the elements of R and $d\mu_{\mathcal{V}_{K,N}}(\widehat{X})$ denotes invariant measure on the Stiefel manifold.

For a random matrix $X \in \mathbb{C}^{N\times K}$ with unitarily invariant probability density $f(X) = h(X^\dagger X)$,

$$\begin{aligned} 1 &= \int_{\mathbb{C}^{N\times K}} f(X) dX \\ &= \int_{\mathcal{V}_{K,N}} \int_{\mathcal{T}_{K,K}} h(R^\dagger R)|R|^{2N} d\mu_{\mathcal{T}_{K,K}}(R)\ d\mu_{\mathcal{V}_{K,N}}(\widehat{X}). \end{aligned}$$

Thus,

$$\int_{\mathcal{T}_{K,K}} h(R^\dagger R)|R|^{2N} d\mu_{\mathcal{T}_{K,K}}(R) = \frac{1}{\mathrm{vol}(\mathcal{V}_{K,N})}. \qquad (4)$$

3. INVARIANCE OF THE DISTRIBUTION OF $\widehat{\mathbf{G}}$

With the notation and preliminary results set forth in Section 2, it is now feasible to establish the central result of this paper:

Theorem 1. *Let $X_1, X_2, \ldots, X_M$ be independently distributed random matrices in $\mathbb{C}^{N\times K}$ with X_j, $j = 2, \ldots, M$ having unitary invariant probability distributions and X_1 having an arbitrary distribution. Then the normalized Gram matrix $\widehat{\mathbf{G}}(X_1, \ldots, X_M)$ has distribution*

$$dF(\widehat{\mathbf{G}}) = \frac{\mathrm{vol}(\mathcal{V}_{KM,N})}{\mathrm{vol}(\mathcal{V}_{K,N})^M} |\widehat{\mathbf{G}}|^{N-MK} \prod_{i<j}^{M} d\widehat{G}_{ij}$$

where $d\widehat{G}_{ij} = \prod_{\ell,m=1}^{K} d\,\mathrm{Re}[\widehat{G}_{ij}]_{\ell m}\ d\,\mathrm{Im}[\widehat{G}_{ij}]_{\ell m}$.

Proof. Under the hypotheses of the theorem, $\mathbf{X} = (X_1, X_2, \ldots, X_M)$ has distribution

$$dF(\mathbf{X}) = f_1(X_1) \prod_{m=2}^{M} f_m(X_m) \prod_{j=1}^{M} dX_j, \qquad (5)$$

where the distributions $f_m(X_m) dX_m$, $j = 2, \cdots, M$ are unitarily invariant. In this expression (and subsequently), the shorthand notation $dX_j = d\,\mathrm{Re}(X_j) d\,\mathrm{Im}(X_j)$ is used. Now, $\mathbf{X}$ has a unique decomposition $\mathbf{X} = \mathbf{L}\mathbf{T}$ with $\mathbf{L} \in \mathcal{V}_{MK,N}$ and $\mathbf{T} \in \mathcal{T}_{MK,MK}$. Write $\mathbf{T} = (T_1, \ldots, T_M)$ where $T_j \in \mathcal{T}_{K,jK}$. From [11, 10],

$$\prod_{j=1}^{M} dX_j = |\mathbf{T}|^{2N} \prod_{j=1}^{M} d\mu_{\mathcal{T}_{K,jK}}(T_j)\ d\mu_{\mathcal{V}_{MK,N}}(\mathbf{L}). \qquad (6)$$

Substituting (6) into (5), the joint density of $\mathbf{T}$ and $\mathbf{L}$ is obtained as

$$\begin{aligned} &dF(\mathbf{T}, \mathbf{L}) \\ &= f_1(\mathbf{L}T_1) \prod_{m=2}^{M} f_m(\mathbf{L}T_m) |\mathbf{T}|^{2N} \prod_{j=1}^{M} d\mu_{\mathcal{T}_{K,jK}}(T_j)\ d\mu_{\mathcal{V}_{MK,N}}(\mathbf{L}). \end{aligned} \qquad (7)$$

The Gram matrix of $\mathbf{X}$ is $\mathbf{G} = \mathbf{T}^\dagger \mathbf{T}$. Consequently (see [10]),

$$\prod_{i\le j}^{M} dG_{ij} = 2^{MK} |\mathbf{T}|^{2MK} \prod_{j=1}^{M} d\mu_{\mathcal{T}_{K,jK}}(T_j). \qquad (8)$$

Substituting (8) into (7) and noting that $|\mathbf{G}| = |\mathbf{T}|^2$, the joint density of $\mathbf{G}$ and $\mathbf{L}$ is obtained as

$$\begin{aligned} dF(\mathbf{G}, \mathbf{L}) &= 2^{-MK} f_1(\mathbf{L}T_1(G_{11})) d\mu_{\mathcal{V}_{MK,N}}(\mathbf{L}) \\ &\quad \times \prod_{m=2}^{M} h_m(G_{mm}) |\mathbf{G}|^{N-MK} \prod_{i\le j}^{M} dG_{ij}, \end{aligned}$$

which depends on the assumption that the f_m for $m = 2, \cdots, M$ are unitarily invariant.

Next, the Gram matrix $\mathbf{G}$ is normalized by letting $\mathbf{G} = \mathbf{R}^\dagger \widehat{\mathbf{G}} \mathbf{R}$ where $\mathbf{R} = \operatorname{diag}(R_1, \ldots, R_M)$ with $R_j \in \mathcal{T}_{K,K}$. Note that $G_{mm} = R_m^\dagger R_m$. Hence,

$$d\mathbf{G} = d\mathbf{R}^\dagger \widehat{\mathbf{G}} \mathbf{R} + \mathbf{R}^\dagger d\widehat{\mathbf{G}} \mathbf{R} + \mathbf{R}^\dagger \widehat{\mathbf{G}} d\mathbf{R},$$

which implies that

$$(\mathbf{R}^{-1})^\dagger d\mathbf{G}\mathbf{R}^{-1} = (\mathbf{R}^{-1})^\dagger d\mathbf{R}^\dagger \widehat{\mathbf{G}} + d\widehat{\mathbf{G}} + \widehat{\mathbf{G}} d\mathbf{R}\mathbf{R}^{-1}.$$

The exterior product of the independent elements of the matrix on the left-have hand side is

$$|\mathbf{R}|^{-2MK} \prod_{i \le j}^{M} dG_{ij}$$

By construction, the diagonal blocks of $\widehat{\mathbf{G}}$ are I_K and those of $d\widehat{\mathbf{G}}$ are zero. Thus the exterior product of the independent elements of the matrix on the right hand side is

$$2^{MK} \prod_{i<j}^{M} d\widehat{G}_{ij} \prod_{m=1}^{M} d\mu_{\mathcal{T}_{K,K}}(R_m).$$

Therefore,

$$\prod_{i \le j}^{M} dG_{ij} = 2^{MK} |\mathbf{R}|^{2MK} \prod_{m=1}^{M} d\mu_{\mathcal{T}_{K,K}}(R_m) \prod_{i<j}^{M} d\widehat{G}_{ij}. \tag{9}$$

In addition, $|\mathbf{G}|^{N-MK} = |\mathbf{R}|^{2(N-MK)} |\widehat{\mathbf{G}}|^{N-MK}$. Substituting (9) into (1) and marginalizing with respect to $\mathbf{L}$ and $\mathbf{R}$ yields

$$\begin{aligned} dF(\widehat{\mathbf{G}}) = & \int\limits_{V_{MK,N}} \int_0^\infty f_1(L_1 R_1) \prod_{\ell=1}^{K} [R_1]_{\ell\ell}^{2(N-\ell)+1} dR_1 d\mu(\mathbf{L}) \\ & \times \prod_{m=2}^{M} \int\limits_0^\infty h_m(R_m^\dagger R_m) \prod_{\ell=1}^{K} [R_m]_{\ell\ell}^{2(N-\ell)+1} dR_m |\widehat{\mathbf{G}}|^{N-MK} \prod_{i<j}^{M} d\widehat{G}_{ij} \end{aligned} \tag{10}$$

where L_1 is a matrix comprised of the first K columns of $\mathbf{L}$.

Writing $\mathbf{L} = [L_1, \mathbf{L}']$ and using the recursive property of the invariant measure on the Stiefel manifold [10], the normalization of the density f_1 implies

$$\begin{aligned} & \int\limits_{V_{MK,N}} \int\limits_0^\infty f_1(L_1 R_1) |R_1|^{2K} d\mu_{\mathcal{T}_{K,K}}(R_1)\, d\mu_{\mathcal{V}_{MK,N}}(\mathbf{L}) \\ & = \int_{\mathcal{V}_{(M-1)K,N-K}} d\mu_{\mathcal{V}_{(M-1)K,N-K}}(\mathbf{L}') \\ & \times \int\limits_{V_{K,N}} \int\limits_0^\infty f_1(L_1 R_1) |R_1|^{2K} d\mu_{\mathcal{T}_{K,K}}(R_1)\, d\mu_{\mathcal{V}_{K,N}}(L_1) \\ & = \operatorname{vol}(\mathcal{V}_{(M-1)K,N-K}) \end{aligned}$$

Finally, using this and (4) in (10) proves Theorem 1. Note that $\operatorname{vol}(\mathcal{V}_{MK,N}) = \operatorname{vol}(\mathcal{V}_{(M-1)K,N-K}) \operatorname{vol}(\mathcal{V}_{K,N})$. □

In particular, for the case $K = 1$, the distribution of $\widehat{G}$ is

$$dF(\widehat{G}) = \frac{\operatorname{vol}(\mathcal{V}_{M,N})}{\operatorname{vol}(V_{1,N})^M} |\widehat{G}|^{N-M} \prod_{i<j}^{M} d\hat{g}_{ij}.$$

4. DISTRIBUTION OF THE DETERMINANT OF $\widehat{\mathbf{G}}$

Determinants of normalized Gram matrices have been used as detection statistics for certain multi-channel detection problems. In fact, this statistic arises naturally in Bayesian and other tests for discrimination between diagonal covariance structure (H_0) and arbitrary non-diagonal covariance structure (H_1) [12]. Typical applications involve testing for the presence of a common but unknown signal in multiple channels of noisy sensor data [2, 8, 9].

As evident from (1), normalized Gram matrices can be regarded as Gram matrices of sets of unit vectors, or more generally, of orthonormal K-frames. The distributions of these statistics under the null hypothesis follow directly from the distributions of Gram determinants of independent uniformly distributed unit vectors, or of K-frames that are independent and uniformly distributed on the Stiefel manifold.

Among many important properties of Gram determinants are the following:

1. The Gram determinant is related to the standard inner product on the anti-symmetric tensor space $\wedge^M \mathbb{C}^N$ by
$$|G(\mathbf{x}_1, \cdots, \mathbf{x}_M)| = \langle \mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_M, \mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_M \rangle$$
2. $|G(\mathbf{x}_1, \cdots, \mathbf{x}_M)|$ is the square of the volume of the parallelotope form by $\mathbf{x}_1, \cdots, \mathbf{x}_M$ and is independent of the order of the vectors.
3. If $\mathbf{y} \in <\mathbf{x}_1, \cdots, \mathbf{x}_{M-1}>$, then
$$|G(\mathbf{x}_1, \cdots, \mathbf{x}_M + \mathbf{y})| = |G(\mathbf{x}_1, \cdots, \mathbf{x}_M)|$$
4. If $<\mathbf{x}_1, \cdots, \mathbf{x}_K>$ denotes the subspace of $\mathbb{C}^N$ spanned by $\mathbf{x}_1, \cdots, \mathbf{x}_K$ and P_{M-1} is the orthogonal projection onto its orthogonal complement $<\mathbf{x}_1, \cdots, \mathbf{x}_{M-1}>^\perp$, then
$$|G(\mathbf{x}_1, \cdots, \mathbf{x}_M)| = |G(\mathbf{x}_1, \cdots, \mathbf{x}_{M-1})| \|P_{M-1}\mathbf{x}_M\|^2$$

Properties 2-4 follow in straightforward ways from Property 1 and basic properties of determinants.

Consider Theorem 1 for the special case $K = 1$, where the normalized Gram matrix $\hat{G} = G(\hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_M)$ is formed from independent random vectors with $\hat{\mathbf{x}}_2, \cdots, \hat{\mathbf{x}}_M$ uniformly distributed on the unit sphere in $\mathbb{C}^N$ and $\hat{\mathbf{x}}_1$ having arbitrary distribution. Property 4 implies directly that

$$\begin{aligned} |G(\hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_M)| &= \prod_{m=2}^{M} \frac{|G(\hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_k)|}{|G(\hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_{k-1})|} \\ &= \prod_{m=2}^{M} \|P_{k-1}\hat{\mathbf{x}}_k\|^2 \end{aligned} \tag{11}$$

where P_k denotes the orthogonal projector onto $<\hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_K>^\perp$. Obtain the distribution of $|G(\hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_M)|$ is possible with the use of the following result, which adapts [13] to unitary invariance, rather than gaussianity assumptions:

Theorem 2. *Let $\hat{\mathbf{x}}$ be a complex unit vector uniformly distributed on the unit sphere in $\mathbb{C}^N$. Let $V \subset \mathbb{C}^N$ be a subspace of dimension K and decompose $\hat{\mathbf{x}}$ as*

$$\hat{\mathbf{x}} = \sqrt{\tau}\, \mathbf{v} + \sqrt{1-\tau}\, \mathbf{w} \tag{12}$$

where $\tau \in [0, 1]$, $\mathbf{v} \in V$ and $\mathbf{w} \in V^\perp$. Then the joint distribution of τ, $\mathbf{v}$ and $\mathbf{w}$ is

$$dF(\tau, \mathbf{v}, \mathbf{w}) = \frac{\tau^{K-1}(1-\tau)^{N-K-1}}{2\operatorname{vol}(S^{2N-1})} d\tau\, d\mu_{\mathcal{V}_{1,K}}(\mathbf{v})\, d\mu_{\mathcal{V}_{1,N-K}}(\mathbf{w})$$

Proof. The distribution of $\hat{\mathbf{x}}$ is

$$dF(\hat{\mathbf{x}}) = \frac{1}{\mathrm{vol}(S^{2N-1})} d\mu_{\mathcal{V}_{1,N}}(\hat{\mathbf{x}})$$

where $d\mu_{\mathcal{V}_{1,N}}$ is the invariant measure on the unit sphere in $\mathbb{C}^N$. The decomposition (12) implies that

$$d\hat{\mathbf{x}} = \frac{1}{2}\left(\frac{1}{\sqrt{\tau}}\mathbf{v} - \frac{1}{\sqrt{1-\tau}}\mathbf{w}\right) d\tau + \sqrt{\tau}\, d\mathbf{v} + \sqrt{1-\tau}\, d\mathbf{w}$$

Multiplying by the inverse of the $N \times N$ unitary matrix

$$U = \left(\sqrt{\tau}\,\mathbf{v} + \sqrt{1-\tau}\,\mathbf{w}, \sqrt{1-\tau}\,\mathbf{v} - \sqrt{\tau}\,\mathbf{w}, Y_V, Y_{V^\perp}\right)$$

where the columns of Y_V consist of an orthonormal set of $K-1$ vectors in V which are all orthogonal to $\mathbf{v}$, while those of $Y_{V^\perp}$ consist of an orthonormal set of $K-1$ vectors in $V^\perp$ which are all orthogonal to $\mathbf{w}$, gives

$$U^\dagger d\hat{\mathbf{x}} = \begin{pmatrix} \tau\mathbf{v}^\dagger d\mathbf{v} + (1-\tau)\mathbf{w}^\dagger d\mathbf{w} \\ 1/\left(2\sqrt{\tau}\sqrt{1-\tau}\right) d\tau + \sqrt{\tau}\sqrt{1-\tau}\left(\mathbf{v}^\dagger d\mathbf{v} - \mathbf{w}^\dagger d\mathbf{w}\right) \\ \sqrt{\tau}\, Y_V^\dagger d\mathbf{v} \\ \sqrt{1-\tau}\, Y_{V^\perp}^\dagger d\mathbf{w} \end{pmatrix}$$

Taking the exterior product of the independent elements of both sides of this equation, and noting that $\hat{\mathbf{x}}^\dagger d\hat{\mathbf{x}}$, $\mathbf{v}^\dagger d\mathbf{v}$, and $\mathbf{w}^\dagger d\mathbf{w}$ are pure imaginary, gives

$$d\mu_{\mathcal{V}_{1,N}}(\hat{\mathbf{x}}) = \frac{1}{2}\tau^{K-1}(1-\tau)^{N-K-1} d\tau\, d\mu_{\mathcal{V}_{1,K}}(\mathbf{v})\, d\mu_{\mathcal{V}_{1,N-K}}(\mathbf{w})$$

and the theorem follows. □

An immediate consequence of Theorem 2 is that is τ, the squared length of the projection of $\hat{\mathbf{x}}$ onto V, is beta-distributed; i.e., $\tau \sim \mathcal{B}(K, N-K)$

$$dF(\tau) = \frac{1}{B(K, N-K)} \tau^{K-1}(1-\tau)^{N-K-1} d\tau. \quad (13)$$

Note that the beta function $B(K, N-K)$ satisfies the identity

$$B(K, N-K) = \frac{\mathrm{vol}(S^{2K-1})\,\mathrm{vol}(S^{2(N-K)-1})}{2\,\mathrm{vol}(S^{2N-1})}$$

and the beta distribution (13) depends only on the dimension of the subspace V. Hence each factor in (11) is beta-distributed; i.e.,

$$\|P_{k-1}\hat{\mathbf{x}}_k\|^2 \sim \mathcal{B}(k-1, N-k+1)$$

Further, since each $\hat{\mathbf{x}}_k$ is independently distributed and the distribution of $\|P_{k-1}\hat{\mathbf{x}}_k\|^2$ depends on P_{k-1} only through its rank, then the factors $\|P_{k-1}\hat{\mathbf{x}}_k\|^2$, $k = 2, \cdots, M$ are independently distributed. This result from [2], that $|\hat{G}|$ is distributed as the product of independent beta distributed random variables, thus follows as a special case of the more general framework developed in this paper.

5. CONCLUSIONS

This paper has introduced a geometrical perspective to enable analysis of the distribution of normalized Gram matrices and used this approach to derive the distribution of such a matrix under a suitable null hypothesis. This derivation makes explicit the essential symmetry in the null hypothesis. It also yields as a corollary an invariance result that subsumes several results regarding invariances of (functions of) the eigenvalues of such matrices that have been published over the past four decades. The H_0 distribution of the determinant of a normalized Gram matrix is obtained using this approach, reproducing an important known result using the approach introduced here.

6. REFERENCES

[1] G. C. Carter and A. H. Nuttall, "Statistics of the estimate of coherence," *Proceedings of the IEEE*, vol. 60, pp. 465–466, April 1972.

[2] H. Gish and D. Cochran, "Generalized coherence," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, April 1988, pp. 2745–2748.

[3] D. Cochran, H. Gish, and D. Sinno, "A geometric approach to multiple-channel signal detection," *IEEE Transactions on Signal Processing*, vol. 43, no. 9, pp. 2049–2057, 1995.

[4] A. H. Nuttall, "Invariance of distribution of coherence estimate to second-channel statistics," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 2, pp. 120–122, 1981.

[5] H. Gish and D. Cochran, "Invariance of the magnitude-squared coherence estimate with respect to second-channel statistics," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 12, pp. 1774–1776, 1987.

[6] A. Clausen and D. Cochran, "An invariance property of the generalized coherence estimate," *IEEE Transactions on Signal Processing*, vol. 45, no. 4, pp. 1065–1067, 1997.

[7] K. Beaudet and D. Cochran, "Multiple-channel detection in active sensing," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2013, pp. 3910–3914.

[8] D. Ramírez, J. Vía, I. Santamaría, and L. L. Scharf, "Detection of spatially correlated Gaussian time series," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5006–5015, 2010.

[9] N. Klausner, M. Azimi-Sadjadi, L. Scharf, and D. Cochran, "Multiple-channel detection in active sensing," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2013, pp. 3919–3923.

[10] A. T. James, "Normal multivariate analysis and the orthogonal group," *Annals of Mathematical Statistics*, vol. 25, no. 1, pp. 40–75, 1954.

[11] M. S. Sriastava, "On the complex Wishart distribution," *Annals of Mathematical Statistics*, vol. 36, no. 1, pp. 313–315, 1965.

[12] S. Sirianunpiboon, S. D. Howard, and D. Cochran, "A Bayesian derivation of the generalized coherence detectors," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, March 2012, pp. 3253–3256.

[13] W. G. Cochran, "The distribution of quadratic forms in a normal system, with applications to the analysis of co- variance," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 30, no. 2, pp. 178–191, 1934.