DISTRIBUTIONS OF PROJECTIONS OF UNIFORMLY DISTRIBUTED K-FRAMES

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

A geometrical perspective is introduced that enables unification and generalization of several results regarding the distributions of quantities that arise in connection with an important class of multiple-channel detectors. Standard models on sets of normalized vectors following from joint Gaussian assumptions in this context are relaxed to the geometrically appealing model of uniform distributions on the Stiefel manifold of K-frames in N-dimensional space. In addition to bolstering geometric insight, several prior results are subsumed and strengthened by results obtained under this formulation. Additionally, a generalization of a classical theorem of W. G. Cochran is enabled by this framework.

***Index Terms*—** Coherence, Multiple-channel detection, Stiefel manifold, Cochran's theorem

I. INTRODUCTION

A broad class of multiple-channel detectors are formulated under Gaussian assumptions on the data under the null hypothesis. Early examples include the two-channel magnitude-squared coherence detector [1], [2] and its M-channel counterpart, the generalized coherence detector [3], [4]. Inspired in part by recent application interest in spectrum sensing [5], [6], [7] and passive radar [8], [9], [10], work on multiple-channel detectors for signals having known rank [6], [11], [12], [13], multiple-channel estimators of signal rank [5], [12], [14], and multi-channel detection of spatially correlated signals [15], [16] have received considerable attention in the research literature over the past few years.

The geometrical nature of detection and estimation problems in this context is well recognized in the literature, and it has been exploited in connection with invariances of detection statistics (e.g., [17], [18], [16], [19]) and Bayesian formulations [20]. Recently, geometrical insight has been strengthened through formulations in terms of Grassmannian and Stiefel manifolds, which arise naturally when considering collections of subspaces having given dimension in a vector space of higher dimension.

This paper advances the geometrical insight that has characterized most of the work on multiple-channel detection and estimation in which normalized Gram matrices play a central role. After establishing the necessary mathematical framework in Section II, Section III proceeds to show how the Gaussian null hypothesis assumed in much of the preceding work in this vein can be relaxed to one that assumes a uniform distribution on a Stiefel manifold. This model strengthens geometric insight and leads to a derivation of the distribution of the normalized Gram matrix under the null hypothesis that subsumes several previously published results as special cases. In Section IV, this perspective leads to a generalization of a classical theorem of W. G. Cochran [21] regarding the joint distribution of quantities obtained when a random vector is projected into a collection of mutually orthogonal subspaces that partition a vector space. Some concluding remarks are given in Section V.

II. MATHEMATICAL FRAMEWORK

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 $\dagger$ 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 $G = XX^\dagger$.

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 $\widehat{G}$ are $\hat{g}_{ii} = 1$, and its determinant is

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

In [15] and [16], this structure is extended to the context of vector-valued time series. In this generalization, $X_j \in \mathbb{C}^{K \times N}$ for $j = 1, \ldots, M$. Denoting $\mathbf{X} = (X_1^\dagger, X_2^\dagger, \ldots, X_M^\dagger)^\dagger$. The Gram matrix associated with $\mathbf{X}$ is $\mathbf{G} = \mathbf{X}\mathbf{X}^\dagger$, which is an $M \times M$ block matrix with block elements $G_{ij} = X_i X_j^\dagger$. It is always possible to decompose X_j uniquely as

$$X_j = R_j \widehat{X}_j \tag{2}$$

where $\widehat{X}_j$ is semi-unitary; i.e., $\widehat{X}_j \widehat{X}_j^\dagger = I_K$, 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 $R_j R_j^\dagger = X_j X_j^\dagger$. The block elements of the normalized Gram matrix $\widehat{\mathbf{G}}$ are

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

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

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

which is the detection statistic defined and analyzed in [15], [16].

Analysis of the statistics of these detectors under the null hypothesis involves knowledge of various properties of the probability distribution of $\widehat{X} P \widehat{X}^\dagger$ or, more generally, the joint distribution of $\{\widehat{X} P_1 \widehat{X}^\dagger, \cdots, \widehat{X} P_M \widehat{X}^\dagger\}$, where the P_j are orthonormal projectors on $\mathbb{C}^N$ such that $\sum_{j=1}^M P_j = I_M$. These quantities also arise in the analysis of the effect of compression on detection and estimation [22]. When $K = 1$ and X is a Gaussian random vector with iid component, then Cochran's theorem [21], [23] shows that the random variables $X P_j X^\dagger$ are independent and χ^2 distributed with the degrees of freedom of the distribution depending only on the dimensions of the subspaces associated with the orthogonal projections P_j. Furthermore, each of the quantities $\widehat{X} P_j \widehat{X}^\dagger$ is beta distributed. Cochran's theorem can be extended to $K > 1$. If $K > 1$ and the matrix X has iid Gaussian components, then the quantities $X P_j X^\dagger$ are independently Wishart distributed [24], [25], [26] and each of the random matrices $\widehat{X} P_j \widehat{X}^\dagger$ are matrix beta distributed.

This paper has two goals. The first is to weaken the Gaussian null hypothesis to one that assumes only that the matrix $\widehat{X}$ is uniformly distributed on the Stiefel manifold. The usual assumption that there exists a Gaussian X associated with $\widehat{X}$ is an unnecessary artifice, and deriving results directly from uniform distributions on the Stiefel manifold gives greater geometric insight. The proof that $\widehat{X} P_j \widehat{X}^\dagger$ given here that is matrix beta distributed with parameters depending only on the rank of P_j follows directly from the uniformity of the distribution of $\widehat{X}$. This proof involves a judicious choice of the matrix Y and the construction of the invariant measure on the Stiefel manifold (3).

The second goal of this paper is to give a simple geometric proof that the joint distribution of $\{\widehat{X} P_1 \widehat{X}^\dagger, \cdots, \widehat{X} P_M \widehat{X}^\dagger\}$ is matrix Dirichlet distributed with parameters depending only on the ranks of the orthogonal projectors $P_1, \cdots, P_M$. This has been shown by Tan [27] (see also [28]), using an argument involving Gaussian random matrices. Theorem 2 is a step towards a version of Cochran's theorem for uniformly distributed K-frames.

For a subspace $S \subset \mathbb{C}^N$ with $\dim S > K$, the Stiefel manifold $\mathcal{V}_K(S)$ is the space of all sets of K orthonormal vectors in S. Such a set of K orthonormal vectors is called an orthonormal K-frame or just a K-frame when no confusion arises. A K-frame can be regarded as an $N \times K$ matrix X, satisfying $X^\dagger X = I_K$, so $\mathcal{V}_K(S)$ can be regarded the space of all such matrices X. Note that here and in the remainder of this paper the notation $\widehat{\ }$ denoting orthonormality is dropped. $\mathcal{V}_K(S)$ is a smooth manifold of dimension $2(\dim S)K - K^2$ and is a sub-manifold of $\mathcal{V}_K(\mathbb{C}^N)$.

The invariant measure on $\mathcal{V}_K(S)$ can be constructed following James [29]. For each $X \in \mathcal{V}_K(S)$ choose a $K \times (\dim S - K)$ matrix Y such that $\begin{pmatrix} X \\ Y \end{pmatrix}$ is unitary and Y is a smooth function of X. The invariant measure on $\mathcal{V}_K(S)$ is constructed by taking the exterior product of the independent entries in the matrix of differential forms $\begin{pmatrix} X \\ Y \end{pmatrix} dX^\dagger$, which results in

$$\begin{aligned} d\mu_{\mathcal{V}_K(S)}(X) = &\prod_{j=1}^{\dim S - K} \prod_{i=1}^{K} \mathrm{Re}(y_j^\dagger dx_i)\, \mathrm{Im}(y_j^\dagger dx_i) \\ &\times \prod_{i<j}^{K} \mathrm{Re}(x_j^\dagger dx_i) \prod_{i \le j}^{K} \mathrm{Im}(x_j^\dagger dx_i). \end{aligned} \tag{3}$$

The matrix Y cannot be chosen to be a smooth function of X across the whole manifold $\mathcal{V}_K(S)$, but can be constructed in a set of domains whose union is the entire manifold. The measure does not depend on the particular choice of Y (see [29]). The volume of $\mathcal{V}_K(S)$ is

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

where $\mathrm{vol}(\mathcal{S}^{m-1}) = 2\pi^{m/2}/\Gamma(m/2)$ is the volume of the unit $(m-1)$-sphere. In what follows, it will be assumed that the invariant measure is normalized. Further, noting that $\mathrm{vol}(\mathcal{V}_K(S))$ depends only on S through its dimension, the notation $\mathrm{vol}(\mathcal{V}_{K,\dim S})$ will be adopted. Finally, B_K will denote the multidimensional beta function which, for integer arguments, can be written conveniently as

$$B_K(M_1, \cdots, M_L) = \frac{\prod_{j=1}^{L} \mathrm{vol}(\mathcal{V}_{K,M_j})}{2^{LK}\, \mathrm{vol}(\mathcal{V}_{K,N})} \tag{4}$$

where $M_j \ge K$ for $j = 1, \cdots, L$, $\sum_{j=1}^{L} M_j = N$, and $B_1(M_1, M_2)$ is the usual beta function.

As an example of this geometric view, suppose that $\{\hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_K\}$ is a random orthonormal K-frame, uniformly distributed on the Stiefel manifold $\mathcal{V}_{K,N}$. Let $S \subset \mathbb{C}^N$ be a subspace of dimension $L > K$ and P_S be the orthogonal projector on to V. The properties of Gram determinants imply that

$$\begin{aligned} |X P_S X^\dagger| &= |G(P_S \hat{\mathbf{x}}_1, \cdots, P_S \hat{\mathbf{x}}_K)| \\ &= |G(P_S \hat{\mathbf{x}}_1, \cdots, P_S \hat{\mathbf{x}}_{K-1})| \|P_{W_{K-1}} \hat{\mathbf{x}}_K\|^2, \end{aligned} \tag{5}$$

where $W_{K-1} = S \cap < \hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_{K-1} >^\perp$ and $P_{W_{K-1}}$ is the orthogonal projector onto this subspace. In this expression, $< \hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_{K-1} >$ denotes the subspace spanned by the vectors $\{\hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_{K-1}\}$ and $^\perp$ denotes orthogonal complement. Now $\hat{\mathbf{x}}_K$ is a uniformly distributed unit vector in $< \hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_{K-1} >^\perp$ and W_{K-1} is an L-dimensional subspace of $< \hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_{K-1} >^\perp$, except on a set of measure zero. Theorem 1, for $K = 1$ (see [30, Theorem 2]), implies that $\|P_{W_{K-1}} \hat{\mathbf{x}}_K\|^2 \sim \mathcal{B}(N-L-(K-1), L)$, and this distribution depends only on the subspace W_{K-1} only through its dimension. Thus the two factors in (5) are independently distributed. Continuing in this way yields

$$|G(P_S \hat{\mathbf{x}}_1, \cdots, P_S \hat{\mathbf{x}}_K)| = \prod_{k=1}^{K} \|P_{W_{k-1}} \hat{\mathbf{x}}_k\|^2,$$

where the factors on the right-hand side are independently beta distributed as

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

Here $W_{k-1} = S \cap < \hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_{k-1} >^\perp$ and $W_0 = S$. The main result in [16] is a consequence of the above result and some elementary properties of Gram matrices.

III. UNIFORMLY DISTRIBUTED K-FRAMES

This section begins with some background. Suppose the space of M-dimensional positive definite Hermitian matrices $\mathcal{P}_M$ is parameterized by the matrix eigenvalues $\lambda_1, \cdots, \lambda_M$ and eigenvectors $\mathbf{u}_1, \cdots, \mathbf{u}_M$,

$$G = U^\dagger \Lambda U \tag{6}$$

This parameterization is redundant as it stands [31], since multiplying U by a unitary matrix U_0 such that $U_0^\dagger \Lambda U_0$ is still diagonal gives an alternative decomposition for the same matrix. This redundancy can be removed by choosing the phases of one column of U, or of its diagonal, and by choosing an ordering for the eigenvalues. Whatever the choice, $U \in \tilde{U}(M) = U(M)/(T_M \times S_M)$, where T_M is the group of diagonal $M \times M$ unitary matrices (maximal torus) and S_M denotes the symmetric group of $M \times M$ permutation matrices. $\tilde{U}(M)$ is a smooth manifold of real dimension $M(M-1)$.

In terms of such a parameterization, the Lebesgue measure on $\mathcal{P}_M$ is [31],

$$\begin{aligned} &\prod_{j=1}^{M} dG_{jj} \prod_{i>j}^{M} d\,\mathrm{Re}(G_{ij}) d\,\mathrm{Im}(G_{ij}) \\ &= \prod_{i<j}^{M} (\lambda_i - \lambda_j)^2 \prod_{j=1}^{M} d\lambda_j \prod_{i<j}^{M} d\mu_{\tilde{U}(M)}(U), \end{aligned} \tag{7}$$

where $d\mu_{\tilde{U}(M)}(U) = \prod_{i<j}^{M} \mathrm{Re}(\mathbf{u}_i^\dagger d\mathbf{u}_j)\, \mathrm{Im}(\mathbf{u}_i^\dagger d\mathbf{u}_j)$ is the non-normalized invariant measure on $\tilde{U}(M)$. The volume of $\tilde{U}(M)$ is $\mathrm{vol}(\tilde{U}(M)) = \pi^{M(M-1)/2} / \prod_{j=1}^{M+1} \Gamma(j)$. When there is no confusion, it is convenient to write the measure on the left-hand side of (7) as dG. Note that the invariant 1-forms $\mathbf{u}_i^\dagger d\mathbf{u}_i$, $i = 1, \cdots, M$

on $\tilde{U}(M)$, which constitute the diagonal of the matrix $U^\dagger dU$, are linearly dependent on the off-diagonal elements.

Now suppose $X \in \mathbb{C}^{K\times N}$ is a K-frame, i.e., $X \in \mathcal{V}_K(\mathbb{C}^N)$, and that S is a subspace of $\mathbb{C}^N$, with $\dim S = M$. X can be uniquely expressed as $X = A + B$ with the rows of A in S and the rows of B in $S^\perp$. Applying singular value decompositions (SVDs) to A and B, X can be written non-redundantly as

$$X = U^\dagger \Lambda^{1/2} V + U^\dagger (I-\Lambda)^{1/2} W,$$

where $U \in \tilde{U}(K)$ and Λ is a diagonal matrix with elements in $[0,1]$. In terms of P, the orthogonal projection onto the subspace S,

$$(XP)(XP)^\dagger = U^\dagger \Lambda U = G,$$

the Gram matrix of the projection of the K-frame X onto S.

Theorem 1. *Let $X \in \mathbb{C}^{K\times N}$, $K < N$, be uniformly distributed on the Stiefel manifold $\mathcal{V}_K(\mathbb{C}^N)$. Let $S \subset \mathbb{C}^N$ be a subspace of dimension $M \geq K$ and decompose X as*

$$X = U^\dagger \Lambda^{1/2} V + U^\dagger (I-\Lambda)^{1/2} W \tag{8}$$

where Λ is a diagonal matrix with elements in $[0,1]$, $U \in \tilde{U}(M) = U(N)/(T_N \times S_N)$, $V \in \mathcal{V}_K(S)$, and $W \in \mathcal{V}_K(S^\perp)$. Then the joint distribution of $G = U^\dagger \Lambda U$, V and W is

$$dF(G,V,W) = \frac{1}{B_K(M,N-M)} |G|^{M-K} |I-G|^{N-M-K} dG \\ \times d\mu_{\mathcal{V}_K(S)}(V)\, d\mu_{\mathcal{V}_K(S^\perp)}(W)$$

where $dG = \prod_{j=1}^K dG_{jj} \prod_{i<j=1}^K d\,\mathrm{Re}\, G_{ij}\, d\,\mathrm{Im}\, G_{ij}$ and $d\mu_{\mathcal{V}_K(S)}$ denotes the normalized invariant measure on the Stiefel manifold $\mathcal{V}_K(S)$.

Proof. The exterior derivative of (8) is

$$\begin{aligned} dX &= dU^\dagger \left(\Lambda^{1/2}V + (I-\Lambda)^{1/2}W\right) \\ &\quad + \frac{1}{2}U^\dagger \left(\Lambda^{-1/2} d\Lambda V - (I-\Lambda)^{-1/2} d\Lambda W\right) \\ &\quad + U^\dagger \left(\Lambda^{1/2} dV + (I-\Lambda)^{1/2} dW\right). \end{aligned}$$

Construct a matrix Y_V with columns consisting of an orthonormal set of $\dim S - K$ vectors in S, all of which are orthogonal to the rows of V. Similarly construct a matrix Y_W consisting of an orthonormal set of $(\dim S^\perp - K)$ vectors in $S^\perp$, all of which are orthogonal to the rows of W. Multiplying $dX^\dagger$ by the unitary matrix

$$Q = \begin{pmatrix} X \\ -U^\dagger(I-\Lambda)^{1/2}V + U^\dagger \Lambda^{1/2} W \\ Y_V \\ Y_W \end{pmatrix}$$

gives the matrix of invariant 1-forms

$$QdX^\dagger = \begin{pmatrix} U^\dagger(-UdU^\dagger + \Lambda^{1/2} V dV^\dagger \Lambda^{1/2} \\ + (I-\Lambda)^{1/2} W dW^\dagger (I-\Lambda)^{1/2})U \\ U^\dagger(-\frac{1}{2}(1-\Lambda)^{-1/2}\Lambda^{-1/2} d\Lambda \\ -(I-\Lambda)^{1/2} V dV^\dagger \Lambda^{1/2} \\ + \Lambda^{1/2} W dW^\dagger (I-\Lambda)^{1/2})U \\ Y_V dV^\dagger \Lambda^{1/2} U \\ Y_W dW^\dagger (I-\Lambda)^{1/2} U \end{pmatrix}. \tag{9}$$

Noting that $VdV^\dagger$, $WdW^\dagger$ and $UdU^\dagger$ are skew-Hermitian and that the diagonal of $UdU^\dagger$ is dependent on the off diagonal elements, the exterior product of the $(ij)^{\text{th}}$ element of the top block of (9) with the $(ij)^{\text{th}}$ and the complex conjugate of the $(ji)^{\text{th}}$ elements in the second block is

$$\begin{aligned} &\left(-\mathbf{u}_i d\mathbf{u}_j^\dagger + \sqrt{\lambda_i}\sqrt{\lambda_j}\mathbf{v}_i d\mathbf{v}_j^\dagger + \sqrt{1-\lambda_i}\sqrt{1-\lambda_j}\mathbf{w}_i d\mathbf{w}_j^\dagger\right) \\ &\wedge \left(-\sqrt{1-\lambda_i}\sqrt{\lambda_j}\mathbf{v}_i d\mathbf{v}_j^\dagger + \sqrt{\lambda_i}\sqrt{1-\lambda_j}\mathbf{w}_i d\mathbf{w}_j^\dagger\right) \\ &\wedge \left(\sqrt{1-\lambda_j}\sqrt{\lambda_i}\mathbf{v}_i d\mathbf{v}_j^\dagger - \sqrt{\lambda_j}\sqrt{1-\lambda_i}\mathbf{w}_i d\mathbf{w}_j^\dagger\right) \\ &= (\lambda_i - \lambda_j)\mathbf{u}_i d\mathbf{u}_j^\dagger \wedge \mathbf{v}_i d\mathbf{v}_j^\dagger \wedge \mathbf{w}_i d\mathbf{w}_j^\dagger. \end{aligned}$$

where $\wedge$ denotes the exterior product. The exterior product of the real and imaginary components of the i^{th} diagonal of the top two blocks of (9) give

$$-\frac{1}{2} d\lambda_i \wedge \mathbf{v}_i d\mathbf{v}_i^\dagger \wedge \mathbf{w}_i d\mathbf{w}_i^\dagger.$$

Therefore,

$$\begin{aligned} &dF(\Lambda,U,V,W) \\ &= \frac{1}{B_K(M,N-M)} \prod_{j=1}^K \lambda_i^{M-K}(1-\lambda_i)^{N-M-K} \prod_{i<j}(\lambda_i-\lambda_j)^2 \\ &\quad \times \left(\prod_{j=1}^K d\lambda_j\right) d\mu_{\tilde{U}(N)}(U) d\mu_{\mathcal{V}_K(S)}(V)\, d\mu_{\mathcal{V}_K(S^\perp)}(W). \end{aligned}$$

In terms of the Gram matrix $G = U^\dagger \Lambda U$, (7) implies this can be written as

$$dF(G,V,W) = \frac{1}{B_K(M,N-M)} |G|^{M-K} |I-G|^{N-M-K} dG \\ \times d\mu_{\mathcal{V}_K(S)}(V)\, d\mu_{\mathcal{V}_K(S^\perp)}(W).$$

□

An immediate consequence of Theorem 1 is that G, the Gram matrix of the projection of X onto S, is matrix beta-distributed; i.e., $G \sim \mathcal{B}_K(M, N-M)$. Explicitly,

$$dF(G) = \frac{1}{B_K(M,N-M)} |G|^{M-K} |1-G|^{N-M-K} dG. \tag{10}$$

Note that the beta distribution (10) depends only on the dimension of the subspace S.

IV. PROJECTIONS ONTO AN ORTHOGONAL DECOMPOSITION OF $\mathbb{C}^N$

Consider an orthogonal decomposition of $\mathbb{C}^N$ into mutually orthogonal subspaces $\{S_1, \cdots, S_M\}$; i.e.,

$$\mathbb{C}^N = \bigoplus_{m=1}^M S_m.$$

This section considers the joint distribution of the Gram matrices of the projected components of a uniformly distributed K-frame with respect to such an orthogonal decomposition.

Suppose $X \in \mathbb{C}^{K\times N}$ is a K-frame; i.e., $X \in \mathcal{V}_K(\mathbb{C}^N)$. X can be uniquely decomposed as

$$X = \sum_{j=1}^M A_j$$

with the rows of A_j in S_j. Applying an SVD to A_j for each $j = 1, \cdots, M$, X can be written as

$$X = \sum_{j=1}^M U_j^\dagger \Lambda_j^{1/2} V_j$$

where $U_1, \cdots, U_M \in \tilde{U}(K)$, $\Lambda_1, \cdots, \Lambda_M$ are non-negative and diagonal, and $V_j \in \mathcal{V}_K(S_m)$ for $j = 1, \cdots, M$. Since X is a K-frame, $XX^\dagger = I_K$ and consequently $\sum_{j=1}^M G_j = I_K$, where $G_j = U_j^\dagger \Lambda_j U_j$ is the Gram matrix of the projection of X onto the subspace S_j. Denote the standard open M-simplex of $K \times K$ non-negative definite matrices by

$$\Delta_K = \left\{ (G_1, \cdots, G_M) | G_1, \cdots, G_M > 0 \text{ and } \sum_{m=1}^M G_m = I_K \right\}.$$

The following Theorem gives the joint distribution of the Gram matrices G_j.

Theorem 2. *Let $X \in \mathbb{C}^{K \times N}$, $K < N$, be uniformly distributed on the Stiefel manifold $\mathcal{V}_K(\mathbb{C}^N)$. Let $\{S_1, \cdots, S_M\}$ be an orthogonal decomposition of $\mathbb{C}^N$ and decompose X as*

$$X = \sum_{j=1}^M U_j^\dagger \Lambda_j^{1/2} V_j \tag{11}$$

where Λ_j are diagonal matrices with elements in [0,1], $U_j \in \tilde{U}(M) = U(N)/(T_N \times S_N)$, $V_j \in \mathcal{V}_K(S_j)$ for $j = 1, \cdots, M$. Then the joint distribution of $G_j = U_j^\dagger \Lambda_j U_j$, V_j, for $j = 1, \cdots, M$ is

$$\begin{aligned} &dF(G_1, \cdots, G_{M-1}, V_1, \cdots, V_M) \\ &= \frac{|I - \sum_{j=1}^{M-1} G_j|^{\dim S_M - K}}{\mathcal{B}_K(\dim S_1, \cdots, \dim S_M)} \left(\prod_{j=1}^{M-1} |G_j|^{\dim S_j - K} \right) \\ &\quad \times \prod_{j=1}^{M-1} dG_j \prod_{j=1}^{M} d\mu_{\mathcal{V}_K(S_j)}(V_j) \end{aligned}$$

for $(G_1, \cdots, G_{M-1}, I - \sum_{j=1}^{M-1} G_j) \in \Delta_K$, where $dG_j = \prod_{\ell=1}^K d[G_j]_{\ell\ell} \prod_{i<\ell=1}^K d\,\mathrm{Re}[G_j]_{i\ell}\, d\,\mathrm{Im}[G_j]_{i\ell}$ and $d\mu_{\mathcal{V}_K(S)}$ denotes the normalized invariant measure on the Stiefel manifold $\mathcal{V}_K(S)$.

Proof. The proof proceeds by induction on the number M of subspaces in the orthogonal decomposition. For $M = 2$, the result is given by Theorem 1. Suppose that (2) is true for $M - 1$, for the orthogonal decomposition

$$\mathbb{C}^N = \left(\oplus_{j=1}^{M-2} S_j \right) \oplus S'_{M-1}.$$

If S'_{M-1} is orthogonally decomposed as

$$S'_{M-1} = S_{M-1} \oplus S_M,$$

then Theorem 1 can be used to decompose the normalized invariant measure on $\mathcal{V}_K(S'_{M-1})$ as

$$\begin{aligned} &d\mu_{\mathcal{V}_K(S'_{M-1})}(V'_{M-1}) \\ &= \frac{|H|^{\dim S_{M-1} - K} |I - H|^{\dim S_M - K}}{B_K(\dim S_{M-1}, \dim S_M)} dH \\ &\quad \times d\mu_{\mathcal{V}_K(S_{M-1})}(V_{M-1}) d\mu_{\mathcal{V}_K(S_M)}(V_M). \end{aligned}$$

With the change of variable $H \mapsto G_{M-1}$ with

$$H = \big(1 - \sum_{j=1}^{M-2} G_j\big)^{-1/2} G_{M-1} \big(1 - \sum_{j=1}^{M-2} G_j\big)^{-1/2}$$

and noting that

$$dH = \big|1 - \sum_{j=1}^{M-2} G_j\big|^{-K} dG_{M-1}$$

the result extends from the case of $M - 1$ orthogonal subspaces to that of M orthogonal subspaces. □

An immediate consequence of Theorem 2 is that $(G_1, \cdots, G_{M-1})$, the set of Gram matrices of the projections of X onto the components of the orthogonal decomposition $\{S_1, \cdots, S_M\}$ of $\mathbb{C}^N$, is matrix Dirichlet-distributed; i.e., $(G_1, \cdots, G_{M-1}) \sim \mathcal{D}_K(\dim S_1, \cdots, \dim S_M)$ with

$$\begin{aligned} &dF(G_1, \cdots, G_{M-1}) \\ &= \frac{|I - \sum_{j=1}^{M-1} G_j|^{\dim S_M - K}}{\mathcal{B}_K(\dim S_1, \cdots, \dim S_M)} \left(\prod_{j=1}^{M-1} |G_j|^{\dim S_j - K} \right) \prod_{j=1}^{M-1} dG_j \end{aligned}$$

Note that this distribution depends only on the dimensions of the subspaces S_j in the orthogonal decomposition of $\mathbb{C}^N$ and not on which particular subspaces are chosen.

V. CONCLUSIONS

An important class of multiple-channel detectors and estimators are formulated in terms of Gram matrices containing normalized data vectors from the channels. Standard null-hypothesis models on the sets of unit vectors comprising these matrices arise from Gaussian assumptions on the data prior to normalization. This paper has relaxed this traditional model to the geometrically appealing case of uniform distributions on the Stiefel manifold. This not only fosters further geometric insight, but it also leads to a derivation of results regarding the distributions of quantities associated with this class of multiple-channel detectors and estimators that subsume and generalize previously known results. A generalization of a classical theorem of W. G. Cochran that is enabled by this framework was also presented. Finally, it is noted that the quantities treated here also arise in analysis of the effects of signal compression on detection and estimation.

ACKNOWLEDGMENT

This work was supported in part by the U.S. Air Force Office of Scientific Research under Grant Nos. FA9550-12-1-0225 and FA9550-12-1-0418.

VI. 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] 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.

[3] 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.

[4] 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.

[5] G. Vazquez-Vilar, D. Ramírez, R. López-Valcarce, J. Vía, and I. Santamaría, "Spatial rank estimation in cognitive radio networks with uncalibrated multiple antennas," in *Proceedings of the International Conference on Cognitive Radio and Advanced Spectrum Management*, October 2011.

[6] D. Ramírez, G. Vazquez-Vilar, R. López-Valcarce, J. Vía, and I. Santamaría, "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 3764–3775, 2011.

[7] D. Ramírez, J. Vía, and I. Santamaría, "The locally most powerful test for multiantenna spectrum sensing with uncalibrated

receivers," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2012.

[8] K. S. Bialkowski, I. V. L. Clarkson, and S. D., "Generalized canonical correlation for passive multistatic radar detection," in *Proceedings of the IEEE Statistical Signal Processing Workshop*, 2011, pp. 417–420.

[9] S. D. Howard and S. Sirianunpiboon, "Passive radar detection using multiple transmitters," in *Proceedings of the 47th Asilomar Conference on Signals, Systems, and Computers*, November 2013.

[10] D. E. Hack, L. K. Patton, B. Himed, and M. A. Saville, "Centralized passive MIMO radar detection without direct-path reference signals," *IEEE Transactions on Signal Processing*, vol. 62, no. 11, pp. 3013–3023, 2014.

[11] D. Ramírez, J. Iscar, J. Vía, I. Santamaría, and L. L. Scharf, "The locally most powerful invariant test for detecting a rank-P Gaussian signal in white noise," in *Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop*, June 2012.

[12] S. Sirianunpiboon, S. D. Howard, and D. Cochran, "Multiple-channel detection of signals having known rank," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2013, pp. 6536–6540.

[13] D. E. Hack, C. W. Rossler, and L. K. Patton, "Multichannel detection of an unknown rank-N signal using uncalibrated receivers," *IEEE Signal Processing Letters*, vol. 21, no. 8, pp. 998–1002, 2014.

[14] S. Sirianunpiboon, S. D. Howard, and D. Cochran, "Maximum a posteriori estimation of signal rank," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2014, pp. 5676–5680.

[15] 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.

[16] N. Klausner, M. Azimi-Sadjadi, L. Scharf, and D. Cochran, "Space-time coherence and its exact null distribution," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2013, pp. 3919–3923.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] W. G. Cochran, "The distribution of quadratic forms in a normal system, with applications to the analysis of covariance," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 30, pp. 178–191, 4 1934.

[22] P. Pakrooh, L. L. Scharf, A. Pezeshki, and Y. Chi, "Analysis of Fisher information and the Cramér-Rao bound for nonlinear parameter estimation after compressed sensing," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2013.

[23] R. B. Bapat, *Linear Algebra and Linear Models*. Springer, 2000.

[24] R. Muirhead, *Aspects of Multivariate Statistical Theory*, ser. Wiley Series in Probability and Statistics. Wiley, 1982.

[25] C. S. Wong, J. Masaro, and T. Wang, "Multivariate versions of Cochran's theorems," *Journal of Multivariate Analysis*, vol. 39, no. 1, pp. 154 – 174, 1991.

[26] T. Mathew and K. Nordström, "Wishart and chi-square distributions associated with matrix quadratic forms," *Journal of Multivariate Analysis*, vol. 61, no. 1, pp. 129 – 143, 1997.

[27] W. Y. Tan, "Some distribution theory associated with complex Gaussian distribution," *Tamkang Journal*, vol. 7, pp. 263–301, 1968.

[28] A. K. Gupta, D. K. Nagar, and E. Bedoya, "Properties of the complex matrix variate Dirichlet distribution," *Scientiae Mathematicae Japonicae Online*, vol. e-2007, no. 20, pp. 211–220, 2007.

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

[30] S. Howard, S. Sirianunpiboon, and D. Cochran, "Invariance of the distributions of normalized gram matrices," in *Proceedings of the IEEE Statistical Signal Processing Workshop*, June 2014, pp. 352–355.

[31] A. Edelman and N. Raj Rao, "Random matrix theory," *Acta Numerica*, pp. 1 – 65, 2005.