



# CHALMERS

## Chalmers Publication Library

### **A map based estimator for inverse complex covariance matrices**

This document has been downloaded from Chalmers Publication Library (CPL). It is the author's version of a work that was accepted for publication in:

**2012 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012. Kyoto, 25 - 30 March 2012 (ISSN: 15206149)**

Citation for the published paper:

Nordenvaad, M. ; Svensson, L. (2012) "A map based estimator for inverse complex covariance matrices". 2012 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012. Kyoto, 25 - 30 March 2012 pp. 3369-3372.

<http://dx.doi.org/10.1109/ICASSP.2012.6288638>

Downloaded from: <http://publications.lib.chalmers.se/publication/166043>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

# A MAP BASED ESTIMATOR FOR INVERSE COMPLEX COVARIANCE MATRICES

Magnus L. Nordenvaad

Department of Information Technology  
Uppsala University  
SE-751 05 Uppsala, Sweden  
magnus.lundberg-nordenvaad@it.uu.se

Lennart Svensson

Department of Signals and Systems  
Chalmers University of Technology  
SE-412 96 Göteborg, Sweden  
lennart.svensson@chalmers.se

## ABSTRACT

A novel approach to estimate (inverse) complex covariance matrices is proposed. By considering the class of unitary invariant estimators, the main challenge lies in estimating the underlying eigenvalues from sampled versions. By exploiting that the distribution of the sample eigenvalues can be derived in closed form, a Maximum A Posteriori (MAP) based scheme is then derived. The performance of the derived estimator is simulated and results indicate that the proposed scheme shows performance similar to one of the best estimators known to date. The main advantage lies in that the proposed solution only requires numerical optimization over a  $P$ -dimensional space where  $P$  is the size of the covariance matrix.

## 1. INTRODUCTION

Complex covariance matrices appear as natural components in a vast variety of array and statistical signal processing applications. Meanwhile, in most of these applications, the underlying covariance matrices, or more often their inverses, are not known but rather have to be inferred, either explicitly or implicitly, from measured data. Covariance matrix estimation is hence a fundamental problem that has turned out to be particularly challenging. Although proper treatment of the covariance matrix is vital for many classical array and signal processing problems, up until recently only limited efforts can be noticed in the signal processing community while most of the results stem from the statistics literature and the real valued problem.

The main difficulty is that the natural estimator, the sample covariance matrix (SCM), has a tendency to separate estimated eigenvalues, see, e.g., [1] and references therein. This in turn lead to deteriorated conditioning and noise enhancement while used in subsequent processing. Although this behavior stems from the underlying structure of the problem, its effect has mostly been studied for Gaussian distributed data, for which the SCM corresponds to the Maximum Likelihood (ML) estimator.

Estimators that attempt to overcome the eigenvalue spread, so-called shrinkage estimators, have therefore been studied frequently in the literature. Most of the efforts focus on the real-valued problem and most are based either on the Bayesian methodology or by asymptotic arguments as the dimension of the problem increases, see e.g. [1, 2, 3] and references therein. The Bayesian approaches often suffer from the fact that the most promising solutions tend to lack analytical solutions, whereby numerical integration is often necessary. As the number of unknowns is in the order of  $P^2$  where  $P$  is the size of the covariance matrix, this often limit the applicability of these techniques, especially for large problems. Meanwhile, estimators derived from asymptotic arguments cannot guarantee appropriate performance for small problems. The purpose of this paper

is to derive a well behaved estimator that does not assume asymptotical arguments and only require numerical optimization over a  $P$ -dimensional space.

As mentioned, most of the results concerning covariance matrix estimation target the real-valued problem. Meanwhile, array and signal processing problems often treat the complex data case. Pleasingly for signal processing applications, however, it turns out that the complex-valued problem has a simpler structure that opens up for further analytical studies, and improved results. A noticeable distinction is for instance that the distribution of the SCM eigenvalues, for the complex case, can be derived in closed form. Since the SCM constitutes a sufficient statistic for the problem at hand, this result together with a unitary invariant assumption gives us all the components required in order to infer the covariance matrix from data. To the best of the authors knowledge, this approach has yet to appear in the literature and is the focus of this paper.

## 2. BACKGROUND AND NOTATION

Assume a measured  $P \times N$ -dimensional matrix with  $N \geq P$ ,

$$\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N]. \quad (1)$$

The  $P$ -dimensional vectors,  $\mathbf{z}_n, n = 1, \dots, N$ , are assumed to be independent and identically distributed (i.i.d.) samples from a zero-mean complex Gaussian distribution with covariance matrix  $\mathbf{R}$ . For the purpose of estimating the covariance matrix,  $\mathbf{R}$ , or its inverse,  $\mathbf{R}^{-1}$ ,

$$\mathbf{S} = \mathbf{Z}\mathbf{Z}^H = \sum_{n=1}^N \mathbf{z}_n \mathbf{z}_n^H \quad (2)$$

is a *minimal sufficient statistic*. Hence, all information regarding  $\mathbf{R}$  or  $\mathbf{R}^{-1}$  which is included in  $\mathbf{Z}$  is also contained in  $\mathbf{S}$  which is known to be complex Wishart distributed, see e.g., [4, 5]. Let  $\lambda_1, \dots, \lambda_P$  denote the eigenvalues of  $\mathbf{S}$  and define  $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_P]^T$  and  $\boldsymbol{\Lambda} = \text{diag}(\boldsymbol{\lambda})$ , with  $\text{diag}(\cdot)$  being the diagonal matrix with the elements of the vector argument along the main diagonal. Using the eigenvalue decomposition, the matrix  $\mathbf{S}$  can now be expressed as

$$\mathbf{S} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H, \quad (3)$$

where  $\mathbf{U}$  is a unitary matrix containing the eigenvectors of  $\mathbf{S}$ . Our primary concern in this paper is to estimate  $\mathbf{R}^{-1}$ , not  $\mathbf{R}$ . In fact, this is not a restriction of the approach but rather the choice is made out of notational convenience, as will be understood later. Furthermore, inferring the inverse is often more desirable as, due to pre-whitening,  $\mathbf{R}^{-1}$  and not  $\mathbf{R}$  usually appear as the natural component in many array and signal processing applications. Using similar notation as

above, we can decompose the underlying inverse covariance matrix  $\mathbf{R}^{-1}$  as

$$\mathbf{R}^{-1} = \mathbf{V}\mathbf{D}\mathbf{V}^H. \quad (4)$$

As previously highlighted, the main challenge concerns estimation of the eigenvalues  $d_1, \dots, d_p$  contained in  $\mathbf{D}$ . We here make the assumption that the considered estimator is unitary invariant, see [6]. With the above notation, the estimator can then be written on the form,

$$\hat{\mathbf{R}}^{-1}(\mathbf{S}) = \mathbf{U} \text{diag}(\phi(\boldsymbol{\lambda})) \mathbf{U}^H \quad (5)$$

where  $\hat{\mathbf{d}} = \phi(\boldsymbol{\lambda}) = [\phi_1(\boldsymbol{\lambda}), \dots, \phi_p(\boldsymbol{\lambda})]^T$  contains the estimated eigenvalues of the inverse covariance matrix. Note that most considered estimators, e.g., the inverse sample covariance and the inverse regularized (diagonally loaded) estimator, share this property.

### 3. EIGENVALUE ESTIMATION

From (5) it is apparent that the prime target is to find a suitable mapping/estimator  $\hat{\mathbf{d}} = \phi(\boldsymbol{\lambda})$ . For this purpose it is natural to study the distribution of  $\boldsymbol{\lambda}$  given  $\mathbf{d}$ , i.e., the (integrated) likelihood. This task is simplified by the fact that unitary invariance applies also to the distribution of  $\boldsymbol{\lambda}$ , i.e., the distribution of  $\boldsymbol{\lambda}$  depends only on  $\mathbf{d}$  and not on  $\mathbf{V}$ , as shown below.

**Theorem:** *Assuming that all eigenvalues of  $\mathbf{R}^{-1}$  are distinct, and that the sample eigenvalues  $\lambda_i$  are ordered, here  $\lambda_1 > \lambda_2 > \dots > \lambda_p$ , then,*

$$f(\boldsymbol{\lambda}; \mathbf{R}^{-1}) = f(\boldsymbol{\lambda}; \mathbf{d}) = c \left( \prod_{p=1}^P d_p^N \lambda_p^{N-P} \right) \Delta(-\boldsymbol{\lambda}) \frac{\det\{\mathbf{Q}\}}{\Delta(\mathbf{d})} \quad (6)$$

where  $\mathbf{Q}_{i,j} = e^{-\lambda_i d_j}$ ,  $\Delta(-\boldsymbol{\lambda})$ ,  $\Delta(\mathbf{d})$  are the Vandermonde determinants

$$\begin{aligned} \Delta(\mathbf{d}) &= \prod_{i>j} (d_i - d_j), \\ \Delta(-\boldsymbol{\lambda}) &= \prod_{i<j} (\lambda_i - \lambda_j), \end{aligned} \quad (7)$$

and  $c$  is a constant (w.r.t. both  $\mathbf{d}$  and  $\boldsymbol{\lambda}$ ) that normalizes the density.

This result, or part thereof, is well known in statistics, statistical physics, and information theory, see for instance [7, 8]. However, to the best of the authors knowledge, it has yet not appeared in the signal processing literature. For completeness, a proof is hence provided in the Appendix. We also note that the result relatively easily, through limit calculations, can be extended to cases with repeated eigenvalues in  $\mathbf{d}$ . The details of this extension are for space concerns excluded in this paper.

Once equipped with the likelihood, several estimation procedures are available. In this paper we will focus on a *Maximum A Posteriori* (MAP) approach, for which a suitable prior is required. In order to maintain a general perspective, we would like to apply a non-informative prior. In [9] several plausible prior choices for  $\mathbf{R}^{-1}$  are proposed and analyzed. Perhaps the most favorable option is the *reference prior*, which in the eigenvalue parametrization is given as

$$\pi(\mathbf{d}, \mathbf{V}) \propto \frac{1}{\prod_{p=1}^P d_p} d\mathbf{d} d\mathbf{V}. \quad (8)$$

Here,  $d\mathbf{V}$  denotes the *invariant Haar measure* over unitary matrices [5]. A natural prior for  $\mathbf{d}$ , given the independence between  $\mathbf{d}$  and  $\mathbf{V}$

in (8), is then given as

$$\pi(\mathbf{d}) \propto \frac{1}{\prod_{p=1}^P d_p}, \quad (9)$$

yielding the posterior distribution

$$f(\mathbf{d}|\boldsymbol{\lambda}) \propto \prod_{p=1}^P d_p^{N-1} \frac{\det\{\mathbf{Q}\}}{\Delta(\mathbf{d})}. \quad (10)$$

From the construction of the problem, it is natural that the posterior in (10) is invariant to reordering of the entries in  $\mathbf{d}$ , since a corresponding reordering in terms of eigenvectors yields the same  $\mathbf{R}^{-1}$ . Meanwhile, in the final estimate for  $\mathbf{R}^{-1}$  we want the eigenvalues to be associated with the corresponding eigenvectors. As the entries in  $\boldsymbol{\lambda}$  are assumed to be in descending order and since  $\mathbf{d}$  are inverse measures, this can be obtained by using ascending ordering for  $\mathbf{d}$ , i.e.,  $0 < d_1 \leq d_2 \leq \dots \leq d_P$ . The considered eigenvalue estimates are thus given as

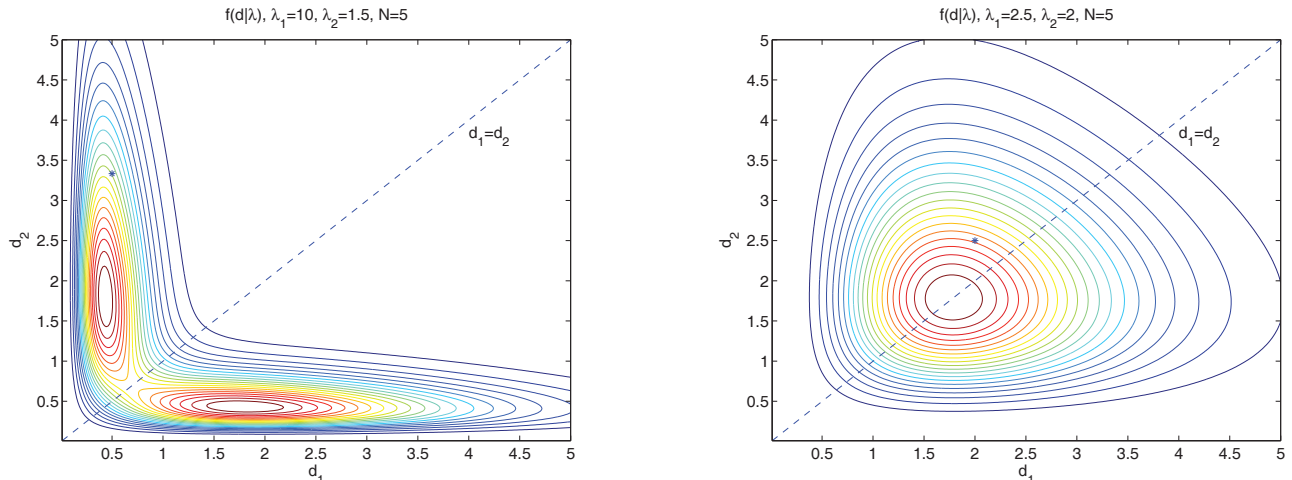
$$\hat{\mathbf{d}} = \phi(\boldsymbol{\lambda}) = \arg \max_{\mathbf{d}} \prod_{p=1}^P d_i^{N-1} \frac{\det\{\mathbf{Q}\}}{\prod_{i>j} (d_i - d_j)}, \quad (11)$$

where  $\mathbf{Q}_{i,j} = e^{-\lambda_i d_j}$ , which together with (5) yield the final estimator. At this point we are convinced that a closed form expression for the estimator in (11) does not exist. Furthermore, neither the function nor its logarithm appear to be convex. The experience, however, is that the function to be optimized in (10) is well behaved and easily amends itself to numerical optimization. To indicate the desirable characteristics of the posterior function, Fig. 1 (a)-(b) show level curves of  $f(\mathbf{d}|\boldsymbol{\lambda})$  for two  $N = 5$  and  $P = 2$  examples; one with well separated sample eigenvalues and one with fairly closely separated sample eigenvalues. In Fig. 1, the eigenvalues of the inverse SCM, i.e.,  $N/\lambda_p$ , are also shown (marked with '\*') as they could serve as a starting point for numerical optimization.

### 4. SIMULATIONS

The main challenge in covariance matrix estimation is to shrink the sample eigenvalues in an appropriate manner. It is therefore natural to investigate and compare the performance of the derived estimator under varying separation in terms of the underlying eigenvalues, contained in  $\mathbf{d}$ . For simplicity we will, in this paper, evaluate the  $P=2$ -dimensional case. The dependence on the underlying eigenvalue separation can then easily be studied by varying the conditioning  $0 < d_2/d_1 \leq 1$ . We note that  $d_2/d_1 = 1$  represents white noise while values close to zero characterize large separation.

In Fig 2, the results of a Monte Carlo experiments with 8000 repetitions are shown. For simplicity we have through out the simulations used  $d_2 = 1$ . In order to also investigate the dependence on sample support, one critically measured set-up ( $N = 2$ ) and one well sampled scenario ( $N = 4P = 8$ ) are evaluated. The performance of three different estimators are simulated. First, we consider the inverse SCM,  $\mathbf{R}^{-1} = N(\mathbf{Z}\mathbf{Z}^H)^{-1}$ , which is the ML estimator for the problem at hand. This estimator is referred to as *ISCM*. The second estimator, denoted *MAP*, is the proposed one, as defined by equations (5) and (11). As a reference estimator we have chosen the Reference prior based Bayesian estimator proposed in [1, 10]. Of course, any well behaved estimator can be used as a reference but we have chosen the one in [1, 10] as it previously have been shown to render very good characteristics for a wide variety of scenarios. The



**Fig. 1.** Characteristics of the posterior function,  $f(\mathbf{d}|\boldsymbol{\lambda})$ , for  $P = 2$  and  $N = 5$ . ‘- -’ indicate the constraints to insure proper ordering, where plausible values lie on or above the dashed line. ‘\*’ indicates the the ML estimate, i.e.,  $N/\lambda_p$  which could serve as a starting point for numerical optimization. From left to right (a)-(b). (a) Well separated sample eigenvalues,  $\lambda_1 = 10$  and  $\lambda_2 = 1.5$ . (b) Closely separated sample eigenvalues  $\lambda_1 = 2.5$  and  $\lambda_2 = 2$ .

major drawback in utilizing it is that implementing the considered estimator requires integration over a  $P^2$ -dimensional space, rendering it difficult to apply for large dimensional problems. In Fig 2, we refer to this estimator as *ref*.

The estimators are evaluated using two performance measures. The first measure, shown in Fig 2 (a) is given by the average error in Frobenius norm,

$$d_F(\mathbf{R}^{-1}, \hat{\mathbf{R}}^{-1}) = \|\mathbf{R}^{-1} - \hat{\mathbf{R}}^{-1}\|_F.$$

Although it is usually considered to be a bad measure of error between two positive-definite matrices, see [11], it is included here for comparison purposes as many authors decide to evaluate covariance matrix estimators using this measure. A more appealing performance measure can be obtained by studying the geometry of the space of positive-definite matrices. Then, the natural distance (geodesic) can be derived as,

$$d_N(\mathbf{R}^{-1}, \hat{\mathbf{R}}^{-1}) = \sqrt{\sum_{p=1}^P (\ln(\nu_p))^2},$$

where  $\nu_p$  make up the eigenvalues of  $\mathbf{R}\hat{\mathbf{R}}^{-1}$ , see for instance [11, 12]. Among the many appealing properties of this measure are that the distance to a non positive definite matrix is infinity and that  $d_N(\mathbf{R}^{-1}, \hat{\mathbf{R}}^{-1}) = d_N(\mathbf{R}, \hat{\mathbf{R}})$ , see [12]. The performance of the estimators in terms of natural distance is shown in Fig 2 (b).

From Fig 2 we note some pleasing characteristics of the proposed estimator. First of all, it dominates the inverse SCM uniformly both in terms of the amount of measured data and in terms of the underlying eigenvalue spread. This is natural for any well designed estimator. More pleasing is that the performance of the proposed estimator overall is similar to that of the reference prior based Bayesian estimator which is known to be very good. We note that the proposed estimator has a tendency to perform better for  $d_2/d_1$  close to one whereas its performance is slightly worse for problems with large underlying spread. This indicates that the proposed estimator, similar to [6], has a tendency to over-shrink the eigenvalues. In fact, this

property can be suspected from the properties of Fig 1 (b) as sample eigenvalues close to each other have a tendency to end up in the diagonal, representing equal eigenvalues. Two ways to overcome this would be to modify the chosen prior and/or to consider estimators based on other criteria than MAP. These investigations are left for future research. From Fig 2(a) we also note some of the limitations of evaluating properties using the Frobenius norm. First of all, the above described properties are not that evident in Fig 2(a). Also, we note a limitation in that the Frobenius norm is very sensitive to ill-conditioning, as the  $N = 2$  simulation indicates. Even though the result is based on 8000 repetitions, the corresponding curve is very inconsistent.

## 5. CONCLUSIONS

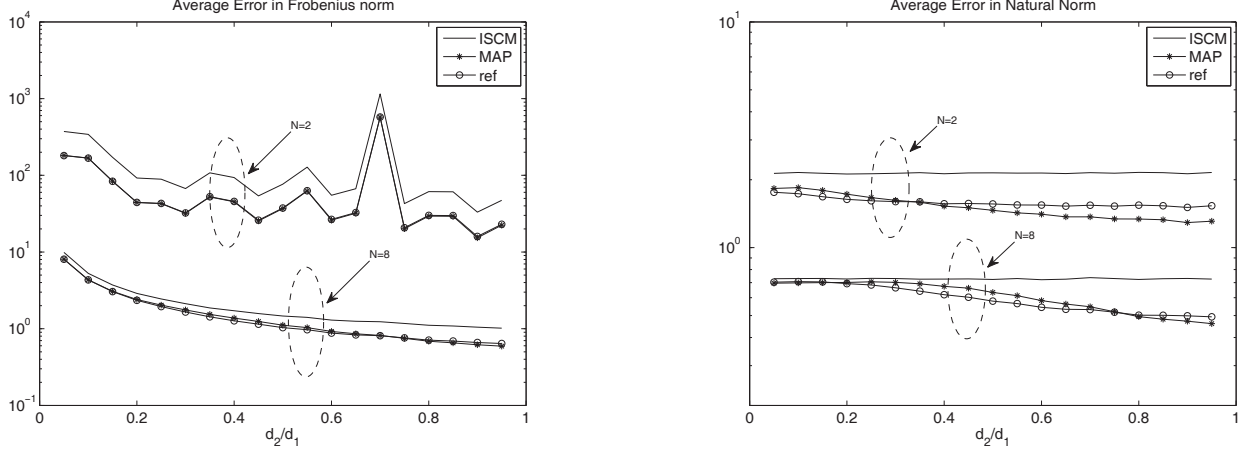
The paper proposed a new methodology to estimate inverse complex covariance matrices. As a side result, the distribution concerning the eigenvalues of Complex Wishart distributed matrices was also highlighted. A simple simulation example indicated that the proposed estimator yields performance similar to that of one of the best estimators known to date. The main advantages of the proposed estimator is that it does not rely on asymptotic assumptions nor that numerical and high-dimensional integration is required. Although the estimator is not given in closed form, the solution only requires  $P$ -dimensional optimization (where  $P$  is the size of the covariance matrix) over a function that appears to be well behaved.

## 6. APPENDIX

It is well known that  $\mathbf{S}$  in (2) is complex Wishart distributed,

$$f(\mathbf{S}; \mathbf{R}^{-1}) = c_1 |\mathbf{S}|^{N-p} |\mathbf{R}^{-1}|^N \text{etr}\{-\mathbf{R}^{-1}\mathbf{S}\} d\mathbf{S}.$$

Here,  $\text{etr}\{\cdot\} = \exp\{\text{trace}(\cdot)\}$  and  $c_1$  is a constant that normalizes the density, see for instance [4, 5]. We note that the density is given with respect to the elements of  $\mathbf{S}$ . In order to find the eigenvalue distribution, we need to switch to the parametrization in (3) and integrate out



**Fig. 2.** Simulation results for three different estimators: Maximum likelihood (ISCM), The proposed approach given by (5) and (11) (MAP), The Bayesian solution using the reference prior [1, 10] (ref). From left to right (a)-(b): (a) Average error in terms of Frobenius norm. (b) Average error in terms of natural norm.

the eigenvectors contained in  $\mathbf{U}$ . Switching parametrization requires the corresponding Jacobian, which is given by

$$d\mathbf{S} \propto \prod_{i < j} (\lambda_i - \lambda_j)^2 d\boldsymbol{\lambda} d\mathbf{U} = \Delta^2(-\boldsymbol{\lambda}) d\boldsymbol{\lambda} d\mathbf{U}$$

where  $\Delta(-\boldsymbol{\lambda})$  is the Vandermonde determinant defined in (7) and  $d\mathbf{U}$  is the invariant Haar measure over unitary matrices, see for instance [5]. The joint distribution of the eigenvalues and eigenvectors can therefore be expressed as

$$f(\boldsymbol{\lambda}, \mathbf{U}; \mathbf{R}^{-1}) \propto \quad (12)$$

$$\propto |\mathbf{R}^{-1}|^N \Delta^2(-\boldsymbol{\lambda}) \prod_{p=1}^P \lambda_p^{N-P} \text{etr}\{-\mathbf{R}^{-1} \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^H\} d\boldsymbol{\lambda} d\mathbf{U}.$$

Now integrating this w.r.t the eigenvectors yields the desired density. Furthermore, since the Haar measure is unitary invariant, we can simplify this as

$$f(\boldsymbol{\lambda}; \mathbf{R}^{-1}) = f(\boldsymbol{\Lambda}; \mathbf{d}) \propto \quad (13)$$

$$\propto \left( \prod_{p=1}^P \lambda_p^{N-P} d_p^N \right) \Delta^2(-\boldsymbol{\lambda}) \int \text{etr}\{-\mathbf{D} \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^H\} d\mathbf{U}.$$

where  $\mathbf{D}$  is a diagonal matrix containing the eigenvalues of  $\mathbf{R}^{-1}$ . We note that the challenging part consists of evaluating the integral above, known as the Harish-Chandra-Itzykson-Zuber integral. Its solution is given by

$$\int \text{etr}\{-\mathbf{D} \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^H\} d\mathbf{U} = \left( \prod_{p=1}^{P-1} p! \right) \frac{\det\{\mathbf{Q}\}}{\Delta(-\boldsymbol{\lambda}) \Delta(\mathbf{d})}, \quad (14)$$

where  $\mathbf{Q}_{i,j} = e^{-\lambda_i d_j}$ , and  $\Delta(\cdot)$  is the Vandermonde determinant, see [13, 7]. An interesting comment is that the solution to this integral does not have a counterpart in the real case, i.e., the integration over orthogonal matrices. Hence, the approach taken in this paper can not straightforwardly be applied to the real data case.

Inserting the results in (14) into (13) we then obtain the desired result. Note that the effect of ordering of the eigenvalues can be incorporated into the constant that normalizes the density.

## 7. REFERENCES

- [1] R. Yang and J. Berger, "Estimation of a covariance matrix using the reference prior," *Ann. Statist.*, vol. 22, no. 3, 1994.
- [2] X. Mestre, "Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, Nov 2008.
- [3] Y. Chen, A. Wiesel, Y.C. Eldar, and A.O. Hero, "Shrinkage algorithms for mmse covariance estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 10, 2010.
- [4] N. R. Goodman, "Statistical analysis based on a certain multivariate complex gaussian distribution," *Ann. Math. Statist.*, vol. 34, Mar. 1963.
- [5] A. Edelman, *Eigenvalues and condition numbers of random matrices*, Ph.D. thesis, MIT, 1989.
- [6] L. R. Haff, "The variational form of certain bayes estimators," *Ann. Statist.*, vol. 19, 1991.
- [7] C. Itzykson and J.B. Zuber, "Planar approximation 2," *J. Math. Phys.*, vol. 21, 1980.
- [8] K.K.K. Kumar, *Communicating Over Outage-Limited Multiple-Antenna And Cooperative Wireless Channels*, Ph.D. thesis, Univ. South. Calif., 2009.
- [9] L. Svensson and M. Lundberg, "On posterior distributions for signals in gaussian noise with unknown covariance matrix," *IEEE Trans. Signal Process.*, vol. 53, no. 5, 2005.
- [10] L. Svensson and M.L. Nordenvaad, "The reference prior for complex covariance matrices with efficient implementation strategies," *IEEE Trans. Signal Process.*, vol. 58, no. 1, 2010.
- [11] S. T. Smith, "Covariance, subspace, and intrinsic cramer-vrao bounds," *IEEE Trans. Signal Process.*, vol. 53, no. 5, May 2005.
- [12] M. Moakher, "A differential geometric approach to the geometric mean of symmetric positive-definite matrices," *SIAM Jour. on Mat. Anal. and appl.*, vol. 26, no. 3, 2005.
- [13] Harish Chandra, "Differential operators on a semi-simple lie algebra," *Amer. J. Math.*, vol. 79, 1957.