

Sampling Uniformly From the Set of Positive Definite Matrices With Trace Constraint

Martin Mittelbach, Bho Matthiesen, *Student Member, IEEE*, and Eduard A. Jorswieck, *Senior Member, IEEE*

Abstract—We derive a parameterization of positive definite matrices using the Cholesky decomposition in combination with hyperspherical coordinates. Based on the parameterization we develop a simple and efficient method to randomly generate positive definite matrices with constant or bounded trace according to a uniform distribution. Further, we present an efficient implementation using the inversion method and either rejection sampling or transforming a beta distribution. The matrix parameterization might be of independent interest, whereas the random sampling algorithm finds applications in Monte-Carlo simulations, testing of algorithms, and performance studies. With the help of an abstract example we describe how the sampling method can be used to approximate the optimum in a difficult, e.g. non-convex, optimization problem for which no solution or efficient global optimization algorithm is known. In this paper we consider real as well as complex matrices.

Index Terms—Matrix parameterization, random covariance matrix, random matrix generation, random positive definite matrix, uniform distribution.

I. INTRODUCTION

THE purpose of this article is to derive an efficient method to sample uniformly from the set of real or complex $n \times n$ positive definite matrices with constant or bounded trace.

Motivation. Generating positive definite matrices (or equivalently covariance matrices) at random finds applications in multivariate statistics, numerical analysis, or signal processing. Examples include clustering and classification methods, testing numerical algorithms (e.g. Cholesky factorization), and studying the efficiency of data compression and decorrelation algorithms [1], [2]. This work is mainly motivated by applications in the area of multiple antenna wireless communications system design and optimization, which will be explained below. However, other areas of application in signal processing for communications are possible.

The transmit strategy in multiple antenna wireless systems is described by a linear precoding matrix or its corresponding transmit covariance matrix. It is necessarily positive (or non-negative) definite and the trace is bounded due to the transmit power constraint [3]. In single-user multiple-input multiple-output (MIMO) systems, the optimization problems are usually convex programming problems and can be solved either by

exploiting the optimality conditions or by algorithms which adapt to the concrete problem structure.

In multiuser MIMO scenarios, however, the resulting programming problems for system optimization are non-convex. Examples are the MIMO broadcast channel with linear precoding [4] and the MIMO interference channel [5], where one is interested in the achievable rate region or the optimized weighted sum rate [6]. The maximization of secrecy rates or secret key rates in MIMO systems also leads to non-convex programming problems [7].

At least a numerical approximation of the solution of these problems can be obtained by evaluating the objective function for a suitable set of positive (or nonnegative) definite matrices that satisfy the trace constraint. This approximation can serve not only as a performance measure but also as bound to be compared with suboptimal and heuristic transmit strategies. The method requires the number of considered matrices to be sufficiently large and without being able to exploit additional knowledge about the optimization problem, it is reasonable to sample the matrices according to a uniform distribution. The efficient generation of such a set of matrices is important in this context.

Related work. Extensive research has been devoted to the problem of randomly generating positive definite matrices with 1's on the diagonal, also known as correlation matrices. Previous work in this area is mainly based on combining either specified [8], [9], [10], [2] or random [1] eigenvalues with random orthogonal matrices, multiplying random triangular matrix factors [10], [1], or randomly generating so called partial correlations [11], [12]. In principle one can build on randomly generated correlation matrices to randomly generate covariance matrices. An approach of this type has been used in [13] to randomly generate positive definite matrices with bounded trace according to a uniform distribution. Even if the existing methods show useful directions to consider also the generation of complex covariance matrices, it is not straightforward to directly extend these approaches and still keep control over the resulting distribution.

Contribution and Outline. The main contribution of this work is to derive an efficient method to sample uniformly from the set of real as well as from the set of complex positive definite matrices with either constant or with bounded trace. We aim at a simple and elementary approach.

First, we derive in Section II-A a parameterization of real positive definite matrices with unit trace using the Cholesky decomposition in combination with hyperspherical coordinates. The parameterization is quite natural and might be of independent interest. It constitutes the foundation of the whole approach and allows us to develop the results in a

Manuscript received September 18, 2011; revised January 18, 2012. This research was supported in part by the Deutsche Forschungsgemeinschaft (DFG) under grant JO 801/3-1.

The authors are with the Chair for Communications Theory, Communications Laboratory, Dresden University of Technology, Dresden 01062, Germany (e-mail: martin.mittelbach@tu-dresden.de, bho.matthiesen@ieee.org, eduard.jorswieck@tu-dresden.de).

Copyright © 2012 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

modular fashion: On the one hand it can be easily extended to parameterize real matrices with variable trace. In particular the extension to complex positive definite matrices is straightforward. On the other hand it enables us to find a simple, efficient, and unified method to generate a uniform distribution on all considered sets of matrices.

In Section II-B we derive in detail the parameter distribution that leads to a uniform distribution on the set of real positive definite matrices with constant trace. We provide in Section II-C details of the implementation to generate the desired distribution using either rejection sampling or transforming a beta distribution.

Starting from the results of Section II we derive in Section III a uniform distribution on the set of real positive definite matrices with bounded trace by treating the trace as independent multiplication parameter. An implementation of the derived trace distribution using the inversion method is proposed. Note, in [13] a uniform distribution on the same set of real matrices is obtained by a special transformation of beta and Liouville distributed random variables. However, another approach is used there which is based on a completely different, not intuitively accessible parameterization. Only the real case is considered and no efficient implementation is provided.

Section IV is concerned with the derivation of a uniform distribution on the set of complex positive definite matrices with constant or bounded trace. Due to our approach the derivations are largely identical to the real-valued case so that we focus on the main steps and results. We point out the differences and omit repeating details.

The developed sampling methods lead to an efficient MATLAB-algorithm which is described by pseudo code and for reproducibility [14] it is available for download at [15]. The implementation allows to uniformly sample a number of positive definite matrices, either based on rejection sampling or on transforming a beta distribution, either with real or complex entries, either with constant or with bounded trace. The option of sampling matrices with fixed or with bounded trace is useful for some applications in connection with our motivating example.

A side effect of the performed analysis is that we obtain the volumes of the sets for which a uniform distribution is derived. Corresponding expressions are developed in the paper.

In Section V an example application is studied, where the optimization of some non-convex function over the set of real positive definite matrices with constant trace is considered. It is shown how the required number of randomly generated matrices is computed in order to obtain a solution close to the global optimum with a predefined probability. The results show clearly the advantage of the proposed approach compared to other heuristic generation methods or a deterministic method.

Notation and Preliminaries. We denote the transpose of a matrix A by A' , the conjugate transpose by A^* , the trace by $\text{tr}(A)$, and the determinant by $\det(A)$. A real or complex $n \times n$ matrix A is said to be positive definite, denoted by $A \succ 0$, if

$$x^* A x > 0 \quad (1)$$

holds for all nonzero $x \in \mathbb{C}^n$. If the strict inequality required in (1) is weakened to $x^* A x \geq 0$, then A is said to be nonnegative

definite (or positive semidefinite). We define

$$\mathcal{R}_c := \{A \in \mathbb{R}^{n \times n} : A \succ 0, \text{tr}(A) = c\}, \quad \mathcal{R} := \mathcal{R}_1, \quad (2)$$

$$\mathcal{R}_{(\underline{c}, \bar{c}]} := \{B \in \mathbb{R}^{n \times n} : B \succ 0, \underline{c} < \text{tr}(B) \leq \bar{c}\}, \quad (3)$$

i.e., the set of real $n \times n$ positive definite matrices with trace equal to the constant $c > 0$, with unit trace, and with trace in the interval $(\underline{c}, \bar{c}]$ for constants $0 \leq \underline{c} < \bar{c}$, respectively. Accordingly, we define the following sets of complex $n \times n$ positive definite matrices:

$$\mathcal{C}_c := \{A \in \mathbb{C}^{n \times n} : A \succ 0, \text{tr}(A) = c\}, \quad \mathcal{C} := \mathcal{C}_1, \quad (4)$$

$$\mathcal{C}_{(\underline{c}, \bar{c}]} := \{B \in \mathbb{C}^{n \times n} : B \succ 0, \underline{c} < \text{tr}(B) \leq \bar{c}\}. \quad (5)$$

Due to its symmetry, a matrix $B \in \mathcal{R}_{(\underline{c}, \bar{c}]}$ has $\frac{1}{2}n(n+1)$ independent entries. Therefore, the set $\mathcal{R}_{(\underline{c}, \bar{c}]}$ is a $(\frac{1}{2}n(n+1))$ -dimensional subset of the set of all real $n \times n$ matrices, denoted by $\mathbb{R}^{n \times n}$. Adding the constraint of a constant trace reduces the number of independent matrix entries by one. Thus, the set \mathcal{R}_c is a $(\frac{1}{2}n(n+1) - 1)$ -dimensional subset of $\mathbb{R}^{n \times n}$. Let ζ be a random variable with values in $\mathbb{R}^{n \times n}$ and assume ζ to be defined on a probability space with probability measure \mathbb{P} . We say that ζ has a uniform distribution on $\mathcal{R}_{(\underline{c}, \bar{c}]}$ if

$$\mathbb{P}(\zeta \in \mathcal{A}) = \frac{\text{vol}(\mathcal{A} \cap \mathcal{R}_{(\underline{c}, \bar{c}]})}{\text{vol}(\mathcal{R}_{(\underline{c}, \bar{c}]})} \quad (6)$$

holds and that ζ has a uniform distribution on \mathcal{R}_c if

$$\mathbb{P}(\zeta \in \mathcal{A}) = \frac{\text{vol}(\mathcal{A} \cap \mathcal{R}_c)}{\text{vol}(\mathcal{R}_c)} \quad (7)$$

holds for all $\mathcal{A} \subset \mathbb{R}^{n \times n}$. In (6) $\text{vol}(\mathcal{B})$ denotes the $(\frac{1}{2}n(n+1))$ -dimensional volume of an arbitrary set $\mathcal{B} \subseteq \mathcal{R}_{(\underline{c}, \bar{c}]}$ and in (7) $\text{vol}(\mathcal{B})$ denotes the $(\frac{1}{2}n(n+1) - 1)$ -dimensional volume of an arbitrary set $\mathcal{B} \subseteq \mathcal{R}_c$. A uniform distribution on the sets $\mathcal{C}_{(\underline{c}, \bar{c}]}$ and \mathcal{C}_c is defined in a similar way. Note that a uniform distribution can be defined equivalently by a constant probability density function (pdf), if a pdf exists for the considered space.

Subsequently, Γ denotes the gamma function given by $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$ for $x > 0$ as defined in [16, 8.310.1]. Further, $\mathbb{1}_{\mathcal{A}}(x)$ denotes the indicator function at point x , which is one if $x \in \mathcal{A}$ and zero otherwise. $\Re(x)$ and $\Im(x)$ represent the real and imaginary part of a complex number x . Throughout this article n is assumed to be an arbitrary but fixed integer with $n \geq 2$. Most of the quantities introduced or derived are functions of n . However, in order to keep notation simple and because it is always clear from the context, we abandon to express this dependency explicitly.

Definitions and theorems employed in this article that are standard knowledge in analysis or linear algebra and which can be found in any standard textbook on the topic (see e. g. [17] and [18]) are referred to as 'well-known' and explicit references will be omitted.

II. REAL MATRICES WITH UNIT OR CONSTANT TRACE

Consider the sets \mathcal{R} and \mathcal{R}_c of real $n \times n$ positive definite matrices with unit and with constant trace, respectively, as defined in (2). In this section, we are concerned with the derivation of a uniform distribution on these sets. Observe that

once we have a uniform distribution on \mathcal{R} , we obtain a uniform distribution on \mathcal{R}_c by the linear one-to-one correspondence $A \mapsto cA$. Therefore, we can restrict the derivations in this section to matrices with unit trace.

A. Parameterization

It is well-known that A is a real positive definite matrix if and only if there exists a unique real upper triangular matrix U with positive diagonal entries such that

$$A = U'U \quad (8)$$

holds. This is called the Cholesky decomposition of A .

Let $A \in \mathcal{R}$ and consider its Cholesky decomposition, where the entries of the matrix U are named as follows

$$U = \begin{pmatrix} x_1 & x_2 & x_4 & \cdots & x_{\frac{(n-1)n}{2}+1} \\ 0 & x_3 & x_5 & \cdots & x_{\frac{(n-1)n}{2}+2} \\ 0 & 0 & x_6 & \cdots & x_{\frac{(n-1)n}{2}+3} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \cdots & x_{\frac{n(n+1)}{2}} \end{pmatrix}. \quad (9)$$

By matrix multiplication and due to the trace constraint of A we obtain

$$\text{tr}(A) = \text{tr}(U'U) = \sum_{k=1}^{\frac{1}{2}n(n+1)} x_k^2 = 1, \quad (10)$$

which is simply the Frobenius norm of the matrix U . If we rewrite the upper triangular part of the matrix U as a vector $x = (x_1, x_2, x_3, \dots, x_{\frac{1}{2}n(n+1)})$, then it follows from (10) that all valid vectors x correspond to points on the unit sphere in the $\frac{1}{2}n(n+1)$ -dimensional Euclidean space. This observation suggests to use the well-known hyperspherical coordinates to parameterize the vector x , which indeed turns out to be convenient for generating a uniform distribution on \mathcal{R} .

The use of hyperspherical coordinates for a vector $x = (x_1, x_2, \dots, x_{\frac{1}{2}n(n+1)})$ satisfying the last equality of (10) yields

$$x_k = \begin{cases} \cos \varphi_k \prod_{l=1}^{k-1} \sin \varphi_l & \text{if } k = 1, 2, \dots, \frac{1}{2}n(n+1) - 1, \\ \prod_{l=1}^{k-1} \sin \varphi_l & \text{if } k = \frac{1}{2}n(n+1), \end{cases} \quad (11)$$

with

$$\varphi_l \in \begin{cases} [0, \pi] & \text{if } l = 1, 2, \dots, \frac{1}{2}n(n+1) - 2, \\ [0, 2\pi] & \text{if } l = \frac{1}{2}n(n+1) - 1. \end{cases} \quad (12)$$

With this parameterization the components x_k corresponding to diagonal entries of the matrix U are nonnegative if and only if the intervals specified in (12) are modified to

$$\varphi_l \in \begin{cases} [0, \frac{\pi}{2}] & \text{if } l = \frac{1}{2}i(i+1), \quad i = 1, \dots, n-1, \\ [0, \pi] & \text{else.} \end{cases} \quad (13)$$

Reducing the parameter ranges this way does not reduce the range of the components x_k corresponding to off-diagonal entries of U . Further, recall that due to the properties of the Cholesky decomposition, we need to parameterize only matrices U with positive diagonal entries. This is achieved if

and only if the closed intervals in (13) are changed to open intervals. These observations all follow with (11) and obvious properties of the sine and cosine function.

We are now able to write the entries of the matrix $A \in \mathcal{R}$ as functions of the parameters φ_l , $l = 1, 2, \dots, \frac{1}{2}n(n+1) - 1$. To simplify the subsequent presentation, let us define the mapping

$$\begin{aligned} \gamma : \mathcal{R} &\rightarrow \gamma(\mathcal{R}) =: \tilde{\mathcal{R}} \subset \mathbb{R}^{\frac{1}{2}n(n+1)-1} \\ A = (a_{ij})_{i,j=1}^n &\mapsto a = (a_{11}, a_{12}, a_{22}, \dots, a_{n-1,n}), \end{aligned} \quad (14)$$

which converts the upper triangular part of the matrix A into a vector, leaving out the last diagonal entry a_{nn} . This is similar to applying the vech-operator, which is common in matrix calculus. Due to its symmetry and the unit trace, the matrix A is completely determined by the vector $a = \gamma(A)$. In fact, the function γ simply extracts the independent entries of the matrix A . Since γ is a linear one-to-one correspondence it follows that once we have a uniform distribution on $\tilde{\mathcal{R}}$, we obtain a uniform distribution on \mathcal{R} by the inverse mapping γ^{-1} . Consequently, in this section we can restrict ourselves to derive a method to sample uniformly from the set $\tilde{\mathcal{R}}$.

For this purpose let us define the mapping

$$\begin{aligned} g : \mathcal{I} &\rightarrow \tilde{\mathcal{R}} \\ \varphi &= a = \\ (\varphi_1, \varphi_2, \dots, \varphi_{\frac{n(n+1)}{2}-1}) &\mapsto (a_{11}, a_{12}, a_{22}, \dots, a_{n-1,n}) \end{aligned} \quad (15)$$

based on the hyperspherical parameterization from above, where the domain \mathcal{I} is defined by

$$\begin{aligned} \mathcal{I} &:= \mathcal{I}_1 \times \mathcal{I}_2 \times \dots \times \mathcal{I}_{\frac{n(n+1)}{2}-1}, \quad \varphi_l \in \mathcal{I}_l, \\ \mathcal{I}_l &= \begin{cases} (0, \frac{\pi}{2}) & \text{if } l = \frac{1}{2}i(i+1), \quad i = 1, 2, \dots, n-1, \\ (0, \pi) & \text{else.} \end{cases} \end{aligned} \quad (16)$$

The components a_{ij} in (15) are given through (8) and (9) by

$$a_{ij} = \sum_{m=1}^i x_{\frac{(i-1)i}{2}+m} x_{\frac{(j-1)j}{2}+m}, \quad i \leq j, \quad (17)$$

and with (11) we obtain the component functions

$$a_{ii}(\varphi) = \sum_{m=1}^i \left[\left(\prod_{l=1}^{\frac{1}{2}(i-1)i+m-1} \sin^2 \varphi_l \right) \cos^2 \varphi_{\frac{(i-1)i}{2}+m} \right] \quad (18)$$

for $i = 1, \dots, n-1$ and

$$\begin{aligned} a_{ij}(\varphi) &= \sum_{m=1}^i \left[\left(\prod_{l=1}^{\frac{1}{2}(i-1)i+m-1} \sin^2 \varphi_l \right) \left(\prod_{l=\frac{1}{2}(i-1)i+m}^{\frac{1}{2}(j-1)j+m-1} \sin \varphi_l \right) \right. \\ &\quad \left. \cdot \cos \varphi_{\frac{(i-1)i}{2}+m} \cos \varphi_{\frac{(j-1)j}{2}+m} \right] \end{aligned} \quad (19)$$

for $i < j$.

Note, that the function g is a one-to-one correspondence, which follows from the uniqueness of the matrix U in the Cholesky decomposition (8) and from the well-known fact that the hyperspherical coordinate transformation is a one-to-one mapping for the considered open parameter intervals. Thus, the composition $\gamma^{-1} \circ g$ is a one-to-one correspondence between \mathcal{I} and \mathcal{R} .

Remark. The mapping $\gamma^{-1} \circ g$ can also be employed to parameterize nonnegative definite matrices with unit trace. We only have to change the domain of the function g by substituting the open intervals in (16) by the corresponding closed intervals, i.e., substituting \mathcal{I} by its closed version, denoted by $\bar{\mathcal{I}}$. Then the parameter ranges are given by (13) and a singular matrix A is the image of a point on the boundary of $\bar{\mathcal{I}}$. The reason why we have to modify the domain of g is that the diagonal elements of the matrix U in the decomposition (8) are nonnegative in the nonnegative definite case. Further note that as a consequence, the function g is not a one-to-one correspondence anymore. The reason is twofold. On the one hand, the matrix U in the decomposition (8) is not necessarily unique for nonnegative definite matrices. On the other hand, the hyperspherical coordinate transformation is no longer a one-to-one mapping.

B. Parameter distribution

Let ξ be a real $(\frac{1}{2}n(n+1)-1)$ -dimensional random vector. Furthermore, let g be a function mapping the set of real $(\frac{1}{2}n(n+1)-1)$ -dimensional vectors into itself. We define g on the set \mathcal{I} by (15), (18) and (19) with \mathcal{I} given by (16). Outside \mathcal{I} we set g equal to the zero vector. Then $\eta := g(\xi)$ is a random vector with the same dimension as ξ . If the values of ξ are almost surely in \mathcal{I} , then the values of η are almost surely in $g(\mathcal{I}) = \tilde{\mathcal{R}}$, with $\tilde{\mathcal{R}}$ defined in (14).

Our goal is to find the distribution of ξ such that η is uniformly distributed on $\tilde{\mathcal{R}}$, i.e., the pdf of η shall be

$$f_\eta(a) = c_\eta \mathbb{1}_{\tilde{\mathcal{R}}}(a), \quad a \in \mathbb{R}^{\frac{1}{2}n(n+1)-1}, \quad (20)$$

where c_η is a constant making f_η a pdf. Due to the simplicity of f_η , we immediately obtain the pdf of ξ by applying the well-known integral transformation formula. The pdf of ξ is then given by

$$f_\xi(\varphi) = c_\eta |\det J_g(\varphi)| \mathbb{1}_{\mathcal{I}}(\varphi), \quad (21)$$

$$\varphi = (\varphi_1, \varphi_2, \dots, \varphi_{\frac{n(n+1)}{2}-1}) \in \mathbb{R}^{\frac{1}{2}n(n+1)-1},$$

where J_g denotes the Jacobian matrix of the function g . Note that indeed all conditions to apply the transformation formula are satisfied: The sets \mathcal{I} and $\tilde{\mathcal{R}}$ are open sets. The function g is a one-to-one correspondence between \mathcal{I} and $\tilde{\mathcal{R}}$, which is discussed in Section II-A and it is continuous, which can be seen from (18) and (19). Furthermore, from (23)–(25) below we see that the Jacobian determinant has no zero on \mathcal{I} .

The calculation of the Jacobian determinant is significantly simplified by the fact that the component function a_{ij} only depends on $\varphi_1, \varphi_2, \dots, \varphi_{\frac{1}{2}(j-1)j+i}$, as can be seen from (18) and (19). It follows that the derivative of a_{ij} with respect to φ_l is zero for all $l > \frac{1}{2}(j-1)j+i$, which results in a lower triangular Jacobian matrix

$$J_g = \begin{pmatrix} \frac{\partial a_{11}}{\partial \varphi_1} & 0 & 0 & \dots & 0 \\ \frac{\partial a_{12}}{\partial \varphi_1} & \frac{\partial a_{12}}{\partial \varphi_2} & 0 & \dots & 0 \\ \frac{\partial a_{22}}{\partial \varphi_1} & \frac{\partial a_{22}}{\partial \varphi_2} & \frac{\partial a_{22}}{\partial \varphi_3} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\partial a_{n-1n}}{\partial \varphi_1} & \frac{\partial a_{n-1n}}{\partial \varphi_2} & \frac{\partial a_{n-1n}}{\partial \varphi_3} & \dots & \frac{\partial a_{n-1n}}{\partial \varphi_{\frac{n(n+1)}{2}-1}} \end{pmatrix}. \quad (22)$$

The Jacobian determinant is then simply the product of the diagonal entries of J_g , i.e.,

$$\det J_g = \frac{\partial a_{11}}{\partial \varphi_1} \frac{\partial a_{12}}{\partial \varphi_2} \dots \frac{\partial a_{n-1n}}{\partial \varphi_{\frac{n(n+1)}{2}-1}}$$

$$= \prod_{i=1}^{n-1} \prod_{j=i}^n \frac{\partial a_{ij}}{\partial \varphi_{\frac{(j-1)j}{2}+i}}$$

$$= \left(\prod_{i=1}^{n-1} \frac{\partial a_{ii}}{\partial \varphi_{\frac{i(i+1)}{2}}} \right) \left(\prod_{i=1}^{n-1} \prod_{j=i+1}^n \frac{\partial a_{ij}}{\partial \varphi_{\frac{(j-1)j}{2}+i}} \right). \quad (23)$$

With (18) and (19) we can calculate the required derivatives, which results in

$$\frac{\partial a_{ii}}{\partial \varphi_{\frac{i(i+1)}{2}}}(\varphi) =$$

$$-2 \cos \varphi_{\frac{i(i+1)}{2}} \sin \varphi_{\frac{i(i+1)}{2}} \left(\prod_{l=1}^{\frac{1}{2}i(i+1)-1} \sin^2 \varphi_l \right) \mathbb{1}_{\mathcal{I}}(\varphi) \quad (24)$$

for $i = 1, \dots, n-1$, and

$$\frac{\partial a_{ij}}{\partial \varphi_{\frac{(j-1)j}{2}+i}}(\varphi) =$$

$$- \cos \varphi_{\frac{i(i+1)}{2}} \left(\prod_{l=1}^{\frac{1}{2}i(i+1)-1} \sin^2 \varphi_l \right) \left(\prod_{l=\frac{1}{2}(j-1)j+i}^{\frac{1}{2}(j-1)j+i} \sin \varphi_l \right) \mathbb{1}_{\mathcal{I}}(\varphi) \quad (25)$$

for $i < j$. Finally taking the absolute value of the Jacobian determinant, we obtain with (21) the pdf of ξ

$$f_\xi(\varphi) =$$

$$c_\eta \left[\prod_{i=1}^{n-1} 2 \cos \varphi_{\frac{i(i+1)}{2}} \sin \varphi_{\frac{i(i+1)}{2}} \left(\prod_{l=1}^{\frac{1}{2}i(i+1)-1} \sin^2 \varphi_l \right) \right]$$

$$\cdot \left[\prod_{i=1}^{n-1} \prod_{j=i+1}^n \cos \varphi_{\frac{i(i+1)}{2}} \left(\prod_{l=1}^{\frac{1}{2}i(i+1)-1} \sin^2 \varphi_l \right) \right]$$

$$\cdot \left(\prod_{l=\frac{1}{2}i(i+1)}^{\frac{1}{2}(j-1)j+i} \sin \varphi_l \right) \mathbb{1}_{\mathcal{I}}(\varphi). \quad (26)$$

Let $\xi_l, l = 1, 2, \dots, \frac{1}{2}n(n+1)-1$, be the components of the real random vector ξ and let f_{ξ_l} be the pdf of ξ_l . Then, after some rearrangements, we obtain from (26)

$$f_\xi(\varphi) = \prod_{l=1}^{\frac{1}{2}n(n+1)-1} f_{\xi_l}(\varphi_l), \quad (27)$$

where f_{ξ_l} is given by

$$f_{\xi_l}(\varphi_l) = c_l \cos^{p_l}(\varphi_l) \sin^{q_l}(\varphi_l) \mathbb{1}_{\mathcal{I}_l}(\varphi_l), \quad \varphi_l \in \mathbb{R}, \quad (28)$$

with \mathcal{I}_l specified in (16). For the exponents p_l and q_l we have

$$p_l = \begin{cases} (n+1) - i & \text{if } l = \frac{1}{2}i(i+1), \\ i = 1, 2, \dots, n-1, \\ 0 & \text{else,} \end{cases} \quad (29)$$

and

$$q_l = n^2 - \kappa_l n - \lambda_l, \quad (30)$$

where

$$\kappa_l = i - 1, \quad \lambda_l = i + 1 + m \quad (31)$$

if $l = \frac{1}{2}i(i+1) + m$, $i = 1, 2, \dots, n-1$, $m = 0, 1, \dots, i$.

Please refer to Appendix A for a derivation of (29)–(31).

The normalizing constant c_l in (28), which makes f_{ξ_l} a pdf, can be obtained using [16, 3.621.5 and 8.384.1]

$$c_l = \begin{cases} 2 \frac{\Gamma\left(\frac{p_l+1}{2} + \frac{q_l+1}{2}\right)}{\Gamma\left(\frac{p_l+1}{2}\right)\Gamma\left(\frac{q_l+1}{2}\right)} & \text{if } l = \frac{1}{2}i(i+1), \\ & i = 1, 2, \dots, n-1, \\ \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\frac{q_l+1}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{q_l+1}{2}\right)} & \text{else.} \end{cases} \quad (32)$$

For convenience of the reader, the constants specified in (29)–(32) and the resulting pdfs f_{ξ_l} are evaluated for $n = 2, 3, 4$ in Appendix B, Table II.

From (27) follows that the random vector ξ has stochastically independent components ξ_l . Thus, to generate the desired distribution of ξ is equivalent to independently generate the distributions of ξ_l , which are specified by the pdfs in (28). This is actually the main advantage of the proposed matrix parameterization.

Remarks. (i) To obtain a uniform distribution on the set of nonnegative definite matrices with unit trace, we just have to replace \mathcal{I} in (21) by its closed version, denoted by $\bar{\mathcal{I}}$. This is equivalent to replacing the open intervals \mathcal{I}_l in (28) by the corresponding closed intervals. Even though the function g is not invertible on $\bar{\mathcal{I}}$ (see Remark in Section II-A), this result can be obtained also by the integral transformation formula, since the $(\frac{1}{2}n(n+1) - 1)$ -dimensional volume of $\bar{\mathcal{I}} \setminus \mathcal{I}$ is zero. However, note that the resulting probability of the set of singular matrices is zero.

(ii) The constant c_η in (20) is equal to the reciprocal of the $(\frac{1}{2}n(n+1) - 1)$ -dimensional volume of the set $\bar{\mathcal{R}}$. Comparing (26) with (27) and (28) yields

$$\begin{aligned} \text{vol}(\bar{\mathcal{R}}) &= c_\eta^{-1} = 2^{n-1} \prod_{l=1}^{\frac{1}{2}n(n+1)-1} c_l^{-1} \\ &= \pi^{\frac{1}{4}n(n-1)} \frac{\prod_{k=2}^n \Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{n(n+1)}{2}\right)}. \end{aligned} \quad (33)$$

For the last equality we have used (32) and a simplification given in [13, Def. 5]. This result might be of independent interest.

C. Implementation

As before, let \mathcal{R} be given by (2), γ by (14) and g by (15)–(19). Then according to the derivations in Sections II-A and II-B we obtain a uniform distribution on \mathcal{R} by applying the transformation $\gamma^{-1} \circ g$ to the real random vector $\xi = (\xi_1, \xi_2, \dots, \xi_{\frac{1}{2}n(n+1)-1})$. The components ξ_l of the vector ξ are stochastically independent, each having pdf f_{ξ_l} as specified in (28)–(32). Thus, to implement this method on a computer we have to generate random numbers drawn from $\frac{1}{2}n(n+1) - 1$ distributions independently, each with pdf f_{ξ_l} .

There are several methods to generate random numbers with a desired distribution on the basis of efficient random number

generators available in standard software like MATLAB or R. For example, we can generate a distribution by applying the inverse of its cumulative distribution function (cdf) to a uniform distribution on the interval $(0, 1)$. This is well-known as the inversion method described in [19, Ch. II.2]. With [16, 2.511] it is not difficult to obtain the cdf for f_{ξ_l} . However, for most exponents in (28) the cdf can only be inverted numerically, which is not efficient. Subsequently, we propose two efficient methods, each having its advantages.

Rejection sampling. First, we develop the well-known rejection method described in [19, Ch. II.3] or [20, Ch. 3.4.1]. To apply this method to generate a distribution with pdf f_{ξ_l} , we have to find a so called instrumental distribution, say with pdf f_{ϕ_l} , for which an efficient generation method exists and for which the inequality

$$f_{\xi_l}(t) \leq \nu_l f_{\phi_l}(t), \quad t \in \mathbb{R}, \quad (34)$$

holds, where $\nu_l > 1$ is a constant. If ν_l is close to 1 this method is efficient. In fact, the expected portion of samples that are rejected during the generation process is given by $(1 - 1/\nu_l)$.

In the considered case, the normal distribution is a suitable candidate to serve as instrumental distribution. Let ϕ_l be a random variable with normal distribution, i.e., with pdf

$$f_{\phi_l}(t) = \frac{1}{\sqrt{2\pi\sigma_l^2}} \exp\left(-\frac{(t-\mu_l)^2}{2\sigma_l^2}\right), \quad t \in \mathbb{R}, \quad (35)$$

where $\mu_l \in \mathbb{R}$ and $\sigma_l^2 > 0$ are the expectation and variance parameters, respectively. The advantage of this approach is that very efficient algorithms are available in all commonly used numerical computing environments to generate normally distributed random numbers. In addition, with proper parameters μ_l and σ_l^2 the constant ν_l can be chosen close to 1.

Indeed, for $l = 1, 2, \dots, \frac{1}{2}n(n+1) - 1$ we set

$$\mu_l = \begin{cases} \arctan \sqrt{\frac{q_l}{p_l}} & \text{if } l = \frac{1}{2}i(i+1), \\ & i = 1, 2, \dots, n-1, \\ \frac{\pi}{2} & \text{else,} \end{cases} \quad (36)$$

and

$$\sigma_l^2 = \frac{1}{(\sqrt{p_l} + \sqrt{q_l})^2}, \quad (37)$$

where p_l and q_l are the exponents of the cosine and sine function in (28). Then we choose ν_l in (34) to be equal to

$$\nu_l = \sqrt{2\pi\sigma_l^2} \frac{c_l}{d_l}, \quad (38)$$

where c_l is the constant specified in (32) and d_l is given by

$$d_l = \begin{cases} \frac{(1+q_l/p_l)^{\frac{p_l}{2}}}{(1+p_l/q_l)^{-\frac{q_l}{2}}} & \text{if } l = \frac{1}{2}i(i+1), \\ & i = 1, 2, \dots, n-1, \\ 1 & \text{else.} \end{cases} \quad (39)$$

Observe that the different cases in (36) and (39) are equivalent to the cases $p_l = 0$ and $p_l > 0$.

Both functions f_{ξ_l} and $\nu_l f_{\phi_l}$ have a unique maximum and with μ_l as in (36) the maxima are at the same position.

Choosing ν_l as in (38) their value is also identical. Together with this property the inequality

$$\frac{d^2 \ln f_{\xi_l}(t)}{dt^2} \leq \frac{d^2 \ln f_{\phi_l}(t)}{dt^2}, \quad t \in \mathbb{R}, \quad (40)$$

implies (34), which follows from the monotonicity of the logarithm, the fundamental theorem of calculus, and the monotonicity of the integral. The smallest possible σ_l^2 such that (40) still holds is then given by (37). This can be seen from the following facts: First, the RHS of (40) is equal to $-1/\sigma_l^2$ for all $t \in \mathbb{R}$. Second, the maximum of the LHS is equal to $-(\sqrt{p_l} + \sqrt{q_l})^2$, which is obtained as the value of the LHS at the zero of its derivative. Note that the calculations have to be performed separately for the cases $p_l = 0$ and $p_l > 0$.

For illustration, f_{ξ_l} and $\nu_l f_{\phi_l}$ are compared in Appendix C, Fig. 1 for the marked parameter sets of Table II. Fig. 1 (a) and (b) show two examples for $p_l = 0$, i.e., where f_{ξ_l} is symmetric. Increasing q_l from 2 to 8 reduces the expected portion of rejected samples from about 11% to 3%. This portion further decreases as q_l increases. If $p_l > 0$ and $p_l \neq q_l$, then f_{ξ_l} is asymmetric as can be seen in Fig. 1 (c) and (d). We observe that the asymmetry increases with increasing difference between p_l and q_l . Fig. 1 (c) and (d) illustrate characteristic examples, where the expected portion of rejection lies between 8% and 10%.

Transformation of beta distribution. As an alternative, we can generate the distribution of the random variable ξ_l directly from a beta distribution by applying an appropriate transformation. The efficiency of this method relies on the efficient generation of beta distributed random numbers.

Let ψ_l be a random variable with beta distribution, i.e., with pdf

$$f_{\psi_l}(t) = c_{\psi_l} t^{\alpha_q-1} (1-t)^{\alpha_p-1} \mathbb{1}_{(0,1)}(t), \quad t \in \mathbb{R}, \quad (41)$$

where $\alpha_p, \alpha_q > 0$ are shape parameters and $c_{\psi_l} = \frac{\Gamma(\alpha_q + \alpha_p)}{\Gamma(\alpha_q)\Gamma(\alpha_p)}$ is a normalizing constant. Consider the mapping

$$\begin{aligned} \theta : (0, 1) &\rightarrow (0, \frac{\pi}{2}), \\ y &\mapsto \arcsin \sqrt{y}. \end{aligned} \quad (42)$$

It is a continuous one-to-one correspondence between two open intervals and its inverse θ^{-1} has derivative

$$\frac{d\theta^{-1}}{dy}(y) = 2 \sin y \cos y, \quad y \in (0, \frac{\pi}{2}), \quad (43)$$

which is a positive function. Thus we can calculate the pdf of the random variable $\theta(\psi_l)$ using the integral transformation formula. For the sake of completeness we extend θ outside $(0, 1)$ by 0. The pdf of $\theta(\psi_l)$ is then given by

$$f_{\theta(\psi_l)}(y) = f_{\psi_l}(\theta^{-1}(y)) \frac{d\theta^{-1}}{dy}(y) \quad (44)$$

$$= 2 c_{\psi_l} \cos^{2\alpha_p-1}(y) \sin^{2\alpha_q-1}(y) \mathbb{1}_{(0, \frac{\pi}{2})}(y), \quad y \in \mathbb{R}. \quad (45)$$

For $l = \frac{1}{2}i(i+1)$, $i = 1, 2, \dots, n-1$, the exponent p_l is positive due to (29). For this case $\theta(\psi_l)$ and ξ_l have the same distribution if we put $\alpha_p = \frac{1}{2}(p_l + 1)$ and $\alpha_q = \frac{1}{2}(q_l + 1)$, which follows immediately from comparing (45) with (16), (28), and (32).

It remains to treat the case $p_l = 0$. We observe that the random variable $\pi - \theta(\psi_l)$ has pdf

$$f_{\theta(\psi_l)}(\pi - y), \quad y \in \mathbb{R}. \quad (46)$$

If β_l is a random variable with Bernoulli distribution and if the probabilities of $\beta_l = 0$ and $\beta_l = 1$ are equal, then the random variable

$$(1 - \beta_l)\theta(\psi_l) + \beta_l(\pi - \theta(\psi_l)) \quad (47)$$

has pdf

$$\frac{1}{2}f_{\theta(\psi_l)}(y) + \frac{1}{2}f_{\theta(\psi_l)}(\pi - y), \quad y \in \mathbb{R}, \quad (48)$$

given β_l and ψ_l are stochastically independent. Combining (45) and (48) and comparing this to (16), (28), and (32), yields for the case $p_l = 0$, that the random variable in (47) has the same distribution as ξ_l , if we put $\alpha_p = \frac{1}{2}$ and $\alpha_q = \frac{1}{2}(q_l + 1)$.

The complete method to sample uniformly from the set of real $n \times n$ positive definite matrices with unit (or constant) trace is summarized in Appendix D in Algorithm 1, 2, and 3 as pseudo code listing. The algorithm illustrates the basic approach developed so far. For the sake of accessibility it is presented in a way close to the derivations. However, further improvements are possible. In particular the calculation of some constants can be omitted.

Matlab implementation. An optimized version of the algorithm has been implemented in MATLAB, which is provided for download at [15]. The method based on rejection sampling as well as the method based on transforming a beta distribution has been made available. The former algorithm works with standard MATLAB, whereas for the latter the Statistics Toolbox is required, since the beta distribution is obtained from two gamma distributions. According to [19, p.431] it is very competitive to utilize a gamma distribution, in particular for the parameter constellations required, since gamma generators are very efficient.

Both generation methods show similar performance. However, comparing the algorithms based on a MATLAB implementation can only be a rough estimate. The numerical stability of the implementation has been tested and verified for matrix dimensions up to $n = 25$. The verification was accomplished by comparing the empirical cdfs of the parameters with their theoretical cdfs. A sample size of 100 000 was used for obtaining the empirical cdfs. The theoretical cdfs were calculated by numerically integrating (28). For higher matrix dimensions the numerical stability has not been tested.

Finally we want to give a rough estimate of the computational time of the MATLAB implementation: On a desktop computer with Intel Core i5 (2x3.33 GHz) processor, 4 GB memory, and Windows 7 (64bit) operating system, it requires less than half a second to generate 5000 real 10×10 matrices with unit trace.

III. REAL MATRICES WITH BOUNDED TRACE

Consider the sets \mathcal{R} and $\mathcal{R}_{(\underline{c}, \bar{c}]}$ of real $n \times n$ positive definite matrices with unit and with bounded trace, respectively, as defined in (2) and (3). In this section we are concerned with the derivation of a uniform distribution on $\mathcal{R}_{(\underline{c}, \bar{c}]}$. We will

show the following result: The product distribution of a certain distribution on the interval $(\underline{c}, \bar{c}]$ and a uniform distribution on \mathcal{R} is transformed into a uniform distribution on $\tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]}$ by the one-to-one correspondence $(t, A) \mapsto tA$.

A. Parameterization

Let us define the mapping

$$\begin{aligned} \gamma : \mathcal{R}_{(\underline{c}, \bar{c}]} &\rightarrow \gamma(\mathcal{R}_{(\underline{c}, \bar{c}]}) =: \tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]} \subset \mathbb{R}^{\frac{1}{2}n(n+1)} \\ B = (b_{ij})_{i,j=1}^n &\mapsto b = (b_{11}, b_{12}, b_{22}, \dots, b_{nn}), \end{aligned} \quad (49)$$

which converts the upper triangular part of the matrix B , i.e., its independent entries, into a vector. We can restrict subsequent derivations to the set $\tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]}$, since γ is a linear one-to-one correspondence that preserves the uniform distribution.

We parameterize the elements of $\tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]}$ by the mapping

$$\begin{aligned} h : (\underline{c}, \bar{c}] \times \tilde{\mathcal{R}} &\rightarrow \tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]} \\ (t, a) &= (t, a_{11}, a_{12}, a_{22}, \dots, a_{n-1n}) \mapsto \\ b &= (b_{11}, b_{12}, b_{22}, \dots, b_{nn}) \\ &= t(a_{11}, a_{12}, a_{22}, \dots, a_{n-1n}, 1 - \sum_{i=1}^{n-1} a_{ii}), \end{aligned} \quad (50)$$

with $\tilde{\mathcal{R}}$ as specified in (14). That means we use the trace as multiplication parameter. The function h satisfies all conditions to be applied in connection with the integral transformation formula: The sets $(\underline{c}, \bar{c}] \times \tilde{\mathcal{R}}$ and $\tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]}$ are open sets and the $(\frac{1}{2}n(n+1))$ -dimensional volume of the sets $\{\bar{c}\} \times \tilde{\mathcal{R}}$ and $\tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]} \setminus \tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]}$ is zero. The function h is a continuous one-to-one correspondence with inverse

$$\begin{aligned} h^{-1} : \tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]} &\rightarrow (\underline{c}, \bar{c}] \times \tilde{\mathcal{R}} \\ b &\mapsto (t, a) \\ &= \left(\sum_{i=1}^n b_{ii} \right)^{-1} \left(\left(\sum_{i=1}^n b_{ii} \right)^2, b_{11}, b_{12}, b_{22}, \dots, b_{n-1n} \right). \end{aligned} \quad (51)$$

It is a simple task to calculate the Jacobian matrix of h^{-1} and after some routine column manipulations we obtain the Jacobian determinant

$$\det J_{h^{-1}}(b) = (-1)^{n-1} \left(\sum_{i=1}^n b_{ii} \right)^{-\frac{1}{2}n(n+1)+1}, \quad (52)$$

which has obviously no zero on $\tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]}$.

B. Parameter distribution

Let η be the random vector considered in Section II-B, which has a uniform distribution on \mathcal{R} . Let τ be a real random variable, where its values represent a positive matrix trace. Assume τ is stochastically independent of η and has pdf f_τ . Then the joint pdf of τ and η is given by

$$\begin{aligned} f_{(\tau, \eta)}(t, a) &= f_\tau(t) \frac{1}{\text{vol}(\mathcal{R})} \mathbb{1}_{\tilde{\mathcal{R}}}(a), \\ (t, a) &\in \mathbb{R} \times \mathbb{R}^{\frac{1}{2}n(n+1)-1}, \end{aligned} \quad (53)$$

where $\text{vol}(\tilde{\mathcal{R}})$ is given in (33). If we extend the domain of the function h in (50) to $\mathbb{R}^{\frac{1}{2}n(n+1)}$ and set its values equal to the zero vector outside $(\underline{c}, \bar{c}] \times \tilde{\mathcal{R}}$, then we can define the real $(\frac{1}{2}n(n+1))$ -dimensional random vector $\zeta := h(\tau, \eta)$.

Applying the integral transformation formula we obtain the pdf f_ζ of ζ from the pdf of (τ, η) . Using (51)–(53) yields

$$\begin{aligned} f_\zeta(b) &= f_{(\tau, \eta)}(h^{-1}(b)) |\det J_{h^{-1}}(b)| \\ &= \frac{1}{\text{vol}(\tilde{\mathcal{R}})} \frac{f_\tau \left(\sum_{i=1}^n b_{ii} \right)}{\left(\sum_{i=1}^n b_{ii} \right)^{\frac{1}{2}n(n+1)-1}} \mathbb{1}_{\tilde{\mathcal{R}}}(a(b)) \end{aligned} \quad (54)$$

for all $b \in \tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]}$ and $f_\zeta(b) = 0$ otherwise. If we define f_τ by

$$f_\tau(t) = c_\tau t^{\frac{1}{2}n(n+1)-1} \mathbb{1}_{(\underline{c}, \bar{c}]}(t), \quad t \in \mathbb{R}, \quad (55)$$

then from (54) immediately follows

$$f_\zeta(b) = c_\tau \frac{1}{\text{vol}(\tilde{\mathcal{R}})} \mathbb{1}_{\tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]}}(b), \quad b \in \mathbb{R}^{\frac{1}{2}n(n+1)}. \quad (56)$$

Thus, for f_τ as in (55) and τ and η being stochastically independent the pdf f_ζ is constant on $\tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]}$ and zero otherwise, i.e., ζ has the desired uniform distribution on $\tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]}$.

Remark. It might be of interest to calculate the $(\frac{1}{2}n(n+1))$ -dimensional volume of the set $\tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]}$. From (56) we obtain

$$\begin{aligned} \text{vol}(\tilde{\mathcal{R}}_{(\underline{c}, \bar{c}]}) &= c_\tau^{-1} \text{vol}(\tilde{\mathcal{R}}) \\ &= \frac{\bar{c}^{\frac{1}{2}n(n+1)} - \underline{c}^{\frac{1}{2}n(n+1)}}{\frac{1}{2}n(n+1)} \text{vol}(\tilde{\mathcal{R}}), \end{aligned} \quad (57)$$

where $\text{vol}(\tilde{\mathcal{R}})$ is given in (33) and the last equality follows from the fact that the constant c_τ normalizes f_τ to be a pdf.

C. Implementation

As derived in the previous Sections III-A and III-B we obtain a uniform distribution on $\mathcal{R}_{(\underline{c}, \bar{c}]}$ by applying the transformation $\gamma^{-1} \circ h$ to the real random vector (τ, η) . The functions γ and h are defined in (49) and (50), the random vector η is uniformly distributed on $\tilde{\mathcal{R}}$, and the random variable τ has pdf f_τ given in (55). In Section II-C we already derived a method to generate the uniform distribution of η . Therefore and because τ and η are stochastically independent, we only have to find a method to generate the distribution of τ . This task is accomplished easily using the inversion method described in [19, Ch. II.2].

Let F_τ be the cdf of τ , which we obtain from (55) by a simple integration. The inverse of F_τ on the interval $(\underline{c}, \bar{c}]$ is then given by

$$\begin{aligned} F_\tau^{-1}(u) &= \left[\left(\bar{c}^{\frac{1}{2}n(n+1)} - \underline{c}^{\frac{1}{2}n(n+1)} \right) u \right. \\ &\quad \left. + \underline{c}^{\frac{1}{2}n(n+1)} \right]^{\frac{2}{n(n+1)}}, \quad u \in (0, 1]. \end{aligned} \quad (58)$$

Applying F_τ^{-1} to a uniform distribution on the interval $(0, 1]$ yields the desired distribution with pdf f_τ .

The complete method to sample uniformly from the set of real $n \times n$ positive definite matrices with bounded trace is summarized in Appendix D, Algorithm 4 as pseudo code listing. This algorithm is included in the MATLAB implementation, which is provided for download at [15].

IV. COMPLEX MATRICES

Consider the sets \mathcal{C} , \mathcal{C}_c , and $\mathcal{C}_{[\underline{c}, \bar{c}]}$ of complex $n \times n$ positive definite matrices with unit, constant, and bounded trace, respectively, as defined in (4) and (5). In this section we are concerned with the derivation of a uniform distribution on these sets. The approach is largely identical to the real-valued case, described in detail in Sections II and III. Therefore, it is sufficient to summarize only the main steps and results. However, the complex case is important too and it is necessary to point out the fine differences of the parameterization and of the parameter distribution.

First, we describe the generation of a uniform distribution on \mathcal{C} , which can be transformed into a uniform distribution on \mathcal{C}_c by the linear one-to-one correspondence $A \mapsto cA$. Then we construct a uniform distribution on $\mathcal{C}_{[\underline{c}, \bar{c}]}$ from the product distribution of a certain distribution on the interval $[\underline{c}, \bar{c}]$ and a uniform distribution on \mathcal{C} using the one-to-one correspondence $(t, A) \mapsto tA$.

A. Unit or constant trace

Parameterization. First, we give a suitable parameterization of the matrices contained in \mathcal{C} following the derivations in Section II-A.

It is well-known that A is a complex positive definite matrix if and only if there exists a unique complex upper triangular matrix U with real positive diagonal entries such that

$$A = U^*U \quad (59)$$

holds. This is called the Cholesky decomposition of A .

Let $A \in \mathcal{C}$ and consider its Cholesky decomposition, where the entries of the matrix U are named as follows

$$U = \begin{pmatrix} x_1 & x_2 + jx_3 & x_5 + jx_6 & \cdots & x_{(n-1)^2+1} + jx_{(n-1)^2+2} \\ 0 & x_4 & x_7 + jx_8 & \cdots & x_{(n-1)^2+3} + jx_{(n-1)^2+4} \\ 0 & 0 & x_9 & \cdots & x_{(n-1)^2+5} + jx_{(n-1)^2+6} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \cdots & x_{n^2} \end{pmatrix}. \quad (60)$$

Since

$$\text{tr}(A) = \text{tr}(U^*U) = \sum_{k=1}^{n^2} x_k^2 = 1 \quad (61)$$

holds by matrix multiplication and due to the unit trace of A , we can parameterize the components x_k , $k = 1, 2, \dots, n^2$, using hyperspherical coordinates

$$x_k = \begin{cases} \cos \varphi_k \prod_{l=1}^{k-1} \sin \varphi_l & \text{if } k = 1, 2, \dots, n^2 - 1, \\ \prod_{l=1}^{k-1} \sin \varphi_l & \text{if } k = n^2. \end{cases} \quad (62)$$

Following the same arguments as in Section II-A, the ranges \mathcal{I}_l of the parameters φ_l , $l = 1, 2, \dots, n^2 - 1$, are given by

$$\mathcal{I}_l = \begin{cases} (0, \frac{\pi}{2}) & \text{if } l = i^2, i = 1, 2, \dots, n-1, \\ (0, \pi) & \text{else.} \end{cases} \quad (63)$$

Parameter distribution. We consider the mapping

$$\begin{aligned} \gamma : \mathcal{C} &\rightarrow \gamma(\mathcal{C}) =: \tilde{\mathcal{C}} \subset \mathbb{R}^{n^2-1} \\ A &= (a_{ij})_{i,j=1}^n \mapsto \\ &(a_{11}, \Re(a_{12}), \Im(a_{12}), a_{22}, \dots, \Re(a_{n-1n}), \Im(a_{n-1n})), \end{aligned} \quad (64)$$

which converts the upper triangular part of matrix A into a real vector, leaving out the last diagonal entry a_{nn} . Note that the diagonal of a complex positive definite matrix has always real positive entries. In fact, the function γ simply extracts the independent entries of the matrix A . It is a linear one-to-one correspondence that preserves a uniform distribution, allowing us to restrict the derivations to the set $\tilde{\mathcal{C}}$.

Furthermore, we consider the mapping

$$g : \mathcal{I}_1 \times \mathcal{I}_2 \times \dots \times \mathcal{I}_{n^2-1} =: \mathcal{I} \rightarrow \tilde{\mathcal{C}} \\ (\varphi_1, \varphi_2, \dots, \varphi_{n^2-1}) \mapsto \quad (65)$$

$$(a_{11}, \Re(a_{12}), \Im(a_{12}), a_{22}, \dots, \Re(a_{n-1n}), \Im(a_{n-1n})),$$

defined by (59), (60), (62), and (63). It can be easily verified that the relevant properties of g to derive a uniform distribution on $\tilde{\mathcal{C}}$ are identical to those of the mapping defined in (15)–(19). In particular, its Jacobian matrix is a lower triangular matrix so that its determinant is the product of the diagonal entries. Thus, the parameter distribution can be obtained in a similar manner as for the real case. See Section II-B for details.

Let ξ_l , $l = 1, 2, \dots, n^2 - 1$, be stochastically independent real random variables whose values represent the parameters φ_l , $l = 1, 2, \dots, n^2 - 1$. Let the pdf f_{ξ_l} of the random variable ξ_l be given by

$$f_{\xi_l}(\varphi_l) = c_l \cos^{p_l}(\varphi_l) \sin^{q_l}(\varphi_l) \mathbb{1}_{\mathcal{I}_l}(\varphi_l), \quad \varphi_l \in \mathbb{R}, \quad (66)$$

with \mathcal{I}_l as specified in (63). Further, let the exponents p_l and q_l be given by

$$p_l = \begin{cases} 2(n-i) + 1 & \text{if } l = i^2, i = 1, 2, \dots, n-1, \\ 0 & \text{else,} \end{cases} \quad (67)$$

and

$$q_l = n^2 + \kappa_l n - \lambda_l, \quad (68)$$

where

$$\kappa_l = n - i - 1, \quad \lambda_l = (i-1)n + 1 + m \quad (69)$$

if $l = i^2 + m$, $i = 1, 2, \dots, n-1$, $m = 0, 1, \dots, 2i$.

The normalizing constant c_l in (66) is necessarily given by

$$c_l = \begin{cases} 2 \frac{\Gamma(\frac{p_l+1}{2} + \frac{q_l+1}{2})}{\Gamma(\frac{p_l+1}{2})\Gamma(\frac{q_l+1}{2})} & \text{if } l = i^2, \\ \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{q_l+1}{2} + \frac{1}{2})}{\Gamma(\frac{q_l+1}{2})} & \text{else.} \end{cases} \quad (70)$$

Then with calculations similar to those in Section II-B, it can be derived that applying the transformation $\gamma^{-1} \circ g$ to the random vector $(\xi_1, \xi_2, \dots, \xi_{n^2-1})$ yields a uniform distribution on \mathcal{C} . For convenience of the reader, the constants specified in (67)–(70) and the resulting pdfs f_{ξ_l} are evaluated in Appendix B, Table III for $n = 2, 3, 4$.

Implementation. The parameter distribution derived in this subsection has the same form as for the corresponding real case considered in Section II. Thus, the algorithm to sample uniformly from the set of complex $n \times n$ positive definite matrices with unit (or constant) trace is basically identical to one developed in Section II-C. Only some constants have to be adapted accordingly.

The MATLAB implementation provided for download at [15] includes also the algorithm to generate complex matrices based on the results presented above. The numerical stability of the complex part of the implementation has been tested and verified for matrix dimensions up to $n = 25$ using the same method as for the real case. For higher matrix dimensions the numerical stability has not been tested.

To give a rough estimate of the computational time of the MATLAB implementation we have generated 5000 complex 10×10 matrices with unit trace. Under the conditions specified at the end of Section II-C this requires less than a second.

Remark. It might be of interest to evaluate the $(n^2 - 1)$ -dimensional volume of the set $\tilde{\mathcal{C}}$. Calculations similar to those in Remark (ii) of Section II-B yield together with (70)

$$\text{vol}(\tilde{\mathcal{C}}) = 2^{n-1} \prod_{l=1}^{n^2-1} c_l^{-1} = \pi^{\frac{1}{2}n(n-1)} \frac{\prod_{k=2}^n (k-1)!}{(n^2-1)!}. \quad (71)$$

B. Bounded trace

Starting from a uniform distribution on \mathcal{C} to generate a uniform distribution on $\mathcal{C}_{(\underline{c}, \bar{c})}$ we have to follow the same steps as in Section III. Except for some natural modifications the calculations are identical to the corresponding real case. Therefore we formulate only the final results and comment on the differences.

Let η be a real $(n^2 - 1)$ -dimensional random vector with a uniform distribution on $\tilde{\mathcal{C}}$, where $\tilde{\mathcal{C}}$ is defined in (64). Let τ be a real random variable, stochastically independent of η and with pdf f_τ , where

$$f_\tau(t) = c_\tau t^{n^2-1} \mathbb{1}_{(\underline{c}, \bar{c})}(t), \quad t \in \mathbb{R}. \quad (72)$$

Further consider the mappings

$$\begin{aligned} \gamma : \mathcal{C}_{(\underline{c}, \bar{c})} &\rightarrow \gamma(\mathcal{C}_{(\underline{c}, \bar{c})}) =: \tilde{\mathcal{C}}_{(\underline{c}, \bar{c})} \subset \mathbb{R}^{n^2} \\ (b_{ij})_{i,j=1}^n &\mapsto b = (b_{11}, \Re(b_{12}), \Im(b_{12}), b_{22}, \dots, b_{nn}), \end{aligned} \quad (73)$$

and

$$\begin{aligned} h : (\underline{c}, \bar{c}) \times \tilde{\mathcal{C}} &\rightarrow \tilde{\mathcal{C}}_{(\underline{c}, \bar{c})} \\ (t, a_{11}, \Re(a_{12}), \Im(a_{12}), a_{22}, \dots, \Re(a_{n-1n}), \Im(a_{n-1n})) &\mapsto \\ t(a_{11}, \Re(a_{12}), \Im(a_{12}), a_{22}, \dots & \\ \dots, \Re(a_{n-1n}), \Im(a_{n-1n}), 1 - \sum_{i=1}^{n-1} a_{ii}). & \end{aligned} \quad (74)$$

Then applying the transformation $\gamma^{-1} \circ h$ to the random vector (τ, η) yields a uniform distribution on $\mathcal{C}_{(\underline{c}, \bar{c})}$.

The result is mainly based on the stochastic independence of τ and η and the special form of the Jacobian determinant of the inverse mapping h^{-1} , which is given by

$$\det J_{h^{-1}}(b) = (-1)^{n-1} \left(\sum_{i=1}^n b_{ii} \right)^{-n^2+1}. \quad (75)$$

To implement the method, we use the algorithm of the previous Section IV-A to generate the uniform distribution of the random vector η . The inversion method is suitable to generate the distribution of the random variable τ . We just have to apply the function F_τ^{-1} with

$$F_\tau^{-1}(u) = \left[(\bar{c}^{n^2} - \underline{c}^{n^2}) u + \underline{c}^{n^2} \right]^{\frac{1}{n^2}}, \quad u \in (0, 1], \quad (76)$$

to a uniform distribution on the interval $(0, 1]$. The complete algorithm to sample uniformly from the set of complex $n \times n$ positive definite matrices with bounded trace is included in the MATLAB implementation provided for download at [15].

Remark. The n^2 -dimensional volume of the set $\tilde{\mathcal{C}}_{(\underline{c}, \bar{c})}$ is given by

$$\text{vol}(\tilde{\mathcal{C}}_{(\underline{c}, \bar{c})}) = c_\tau^{-1} \text{vol}(\tilde{\mathcal{C}}) = \frac{(\bar{c}^{n^2} - \underline{c}^{n^2})}{n^2} \text{vol}(\tilde{\mathcal{C}}), \quad (77)$$

where $\text{vol}(\tilde{\mathcal{C}})$ is given in (71) and the last equality follows from the fact that the constant c_τ normalizes f_τ to be a pdf.

V. EXAMPLE OF APPLICATION

Let f be a continuous real-valued function on the set of real $n \times n$ matrices and let $\bar{\mathcal{R}}_c$ be the set of real $n \times n$ nonnegative definite matrices with trace equal to c . Assume we want to optimize (minimize or maximize) the function f over the set $\bar{\mathcal{R}}_c$. Further, assume the optimization problem cannot be solved analytically and the properties of the function f are such that there is no standard algorithm available to solve it numerically. Then, one way to 'estimate' the optimal value of f in a certain sense is to use the random sampling algorithm developed in the previous sections. In the following, we describe and evaluate this approach. The purpose of the section is to explain the basic principle in a more abstract way rather than analyzing a detailed example.

First, observe that the optimization problem is equivalent to optimizing the function $\tilde{f} = f \circ \gamma^{-1}$ over the set $\bar{\mathcal{R}}_c^\sim := \gamma(\bar{\mathcal{R}}_c) \subset \mathbb{R}^{\frac{1}{2}n(n+1)-1}$, where the function γ is defined as in (14) with \mathcal{R} replaced by $\bar{\mathcal{R}}_c$. In fact, this simply means we reformulate the optimization problem in terms of the independent matrix entries. Note that we have $\bar{\mathcal{R}}_c^\sim = cg(\bar{\mathcal{I}})$, where the function g is defined as in (15)–(19) with \mathcal{I} replaced by its closed version $\bar{\mathcal{I}}$.

Let $a_{\text{opt}} \in \bar{\mathcal{R}}_c^\sim$ be a vector optimizing the function \tilde{f} and let \mathcal{E}_{opt} be a neighborhood of a_{opt} . Assume that we employ the uniform sampling algorithm from Section II-C to randomly select a vector from the set $\bar{\mathcal{R}}_c^\sim$. Due to the uniform distribution, the probability, say ρ , that the generated vector lies in \mathcal{E}_{opt} is given by

$$\rho = \frac{r \text{vol}(\mathcal{E}_{\text{opt}})}{\text{vol}(\bar{\mathcal{R}}_c^\sim)} \quad \text{with} \quad r = \frac{\text{vol}(\mathcal{E}_{\text{opt}} \cap \bar{\mathcal{R}}_c^\sim)}{\text{vol}(\mathcal{E}_{\text{opt}})}, \quad (78)$$

where $\text{vol}(\mathcal{A})$ denotes the $(\frac{1}{2}n(n+1)-1)$ -dimensional volume of the set $\mathcal{A} \subset \mathbb{R}^{\frac{1}{2}n(n+1)-1}$. The factor $r \in [0, 1]$ represents the portion of \mathcal{E}_{opt} , which intersects with $\bar{\mathcal{R}}_c^\sim$. If we use the sampling algorithm to generate M vectors independently, then the probability, say δ , that at least one of these vectors lies in \mathcal{E}_{opt} is given by

$$\delta = 1 - (1 - \rho)^M. \quad (79)$$

If we solve this equation for M and round up to the next integer, we obtain

$$\lceil M \rceil = \lceil \log_{(1-\rho)}(1 - \delta) \rceil, \quad (80)$$

which gives us the minimum number of vectors we have to generate such that the probability, that at least one of the generated vectors lies in \mathcal{E}_{opt} , is at least δ .

Using the integral transformation formula, we obtain

$$\text{vol}(\overline{\mathcal{R}}_c^\sim) = c^{\frac{1}{2}n(n+1)-1} \text{vol}(\tilde{\mathcal{R}}), \quad (81)$$

where $\text{vol}(\tilde{\mathcal{R}})$ is given in (33).

To explicitly calculate (78), let us consider an example, where we specify the neighborhood \mathcal{E}_{opt} of the vector a_{opt} to be a ball with respect to the Euclidean distance with center a_{opt} and radius $\epsilon > 0$, i.e.,

$$\mathcal{E}_{\text{opt}} = \{a \in \mathbb{R}^{\frac{1}{2}n(n+1)-1} : \|a - a_{\text{opt}}\|_2 \leq \epsilon\}. \quad (82)$$

That means, we define \mathcal{E}_{opt} such that the Euclidean distance between a vector in \mathcal{E}_{opt} and the vector a_{opt} is at most ϵ . If \tilde{f} is Lipschitz continuous and has the Lipschitz constant K with respect to the Euclidean distance, then the following inequality holds for all $a \in \mathcal{E}_{\text{opt}}$

$$|\tilde{f}(a) - \tilde{f}(a_{\text{opt}})| \leq K\epsilon. \quad (83)$$

Note that the norm in (82) can be any other norm depending on what is suitable for the considered application.

For \mathcal{E}_{opt} being the ball defined in (82) we obtain with [16, 4.632.2]

$$\text{vol}(\mathcal{E}_{\text{opt}}) = \frac{\pi^{\frac{1}{2}(\frac{1}{2}n(n+1)-1)}}{\Gamma\left(\frac{1}{2}(\frac{1}{2}n(n+1)+1)\right)} \epsilon^{\frac{1}{2}n(n+1)-1}. \quad (84)$$

Using (81), (33), and (84) to evaluate (78) yields

$$\rho = \pi^{\frac{n}{2}-1} \frac{\Gamma\left(\frac{n(n+1)}{4}\right)}{\prod_{k=2}^n \Gamma\left(\frac{k+1}{2}\right)} \left(\frac{2\epsilon}{c}\right)^{\frac{1}{2}n(n+1)-1} r, \quad (85)$$

where we have used [16, 8.335.1] to simplify the expressions.

To consider a numerical example, let $\frac{\epsilon}{c} = 0.1$ and $r = 1$, i.e., $\mathcal{E}_{\text{opt}} \subset \overline{\mathcal{R}}_c^\sim$. In Table I the resulting values of $\lceil M \rceil$ are listed for $n = 2, 3, 4$. Let us pick the marked parameter

$\lceil M \rceil$		2	3	4
δ	0.95	74	2339	91421
	0.99	113	3596	140537

TABLE I
NUMBER $\lceil M \rceil$ FOR BALL-SHAPED NEIGHBORHOOD \mathcal{E}_{opt} ,
 $r = 1$, $\frac{\epsilon}{c} = 0.1$, AND $n = 2, 3, 4$.

constellation in Table I to explain the results: Assume the task is to optimize the real-valued function f over the set of real 3×3 nonnegative definite matrices with unit trace. We equivalently consider optimizing the function \tilde{f} over the set $\overline{\mathcal{R}}_1^\sim$ as described at the beginning of this section and we assume \tilde{f} has Lipschitz constant K . Let us independently generate 2339 vectors of the set $\overline{\mathcal{R}}_1^\sim$ using the uniform sampling algorithm developed in Section II-C. Then with a probability of at least 95% the Euclidean distance between at least one of these vectors and a vector optimizing \tilde{f} is at most 0.1. Taking now the optimum of the function \tilde{f} on the sampled set of vectors yields a value, which differs from the true optimum by at most $0.1K$ with a probability of at least 95%. The values of Table I show that this method can be very efficient.

The condition in our numerical example is that the Euclidean distance between an optimal vector and the boundary of $\overline{\mathcal{R}}_1^\sim$ is at least 0.1, i.e. $r = 1$. Generally, we do not know if $\mathcal{E}_{\text{opt}} \subset \overline{\mathcal{R}}_c^\sim$ holds. Given the boundary of $\overline{\mathcal{R}}_c^\sim$ is sufficiently smooth and the radius of \mathcal{E}_{opt} is sufficiently small, then $r = \frac{1}{2}$ is a suitable assumption to estimate the probability ρ .

The above considerations can be adapted in an obvious way to optimization problems with respect to real nonnegative definite matrices with bounded trace or complex nonnegative definite matrices with constant or bounded trace using the results from Sections III and IV.

Remark. One might ask, what is the advantage of randomly selecting vectors from the set $\overline{\mathcal{R}}_c^\sim$ or, equivalently, matrices from the set $\overline{\mathcal{R}}_c$. Alternatively, we could use the parameterization derived in Section II-A and select matrices from $\overline{\mathcal{R}}_c$ by stepping through the set of parameters deterministically according to an equidistant grid. Since the set of parameter values is an $(\frac{1}{2}n(n+1)-1)$ -fold Cartesian product of intervals, this can be accomplished easily. However, this way we step through the corresponding set of matrices in a highly irregular manner. In contrast, with the proposed random method, we select the matrices according to a uniform distribution. Without being able to exploit any additional knowledge about the optimization problem, sampling the constraint set uniformly seems to be the better approach.

VI. CONCLUSIONS

In many application areas, some non-convex function is optimized over the set of positive definite matrices with trace constraint. A numerical approximation of the optimal solution can be obtained by evaluating the objective function for a set of uniformly distributed matrices with carefully chosen cardinality. In this paper, we have derived an efficient method to randomly generate matrices according to a uniform distribution on the constraint set for the real as well as the complex case. We expect that there are many applications of the derived algorithm to compute an approximation of the optimal solution.

APPENDIX A

DERIVATION OF EXPONENTS IN PARAMETER PDFS (REAL CASE)

To derive the exponents p_l and q_l as given in (29)–(31) we first rewrite (26) as

$$f_\xi(\varphi) = c_\eta 2^{n-1} \left[\prod_{i=1}^{n-1} \left(\cos \varphi \frac{i(i+1)}{2} \right)^{n-i+1} \right] \cdot \left[\prod_{i=2}^{n-1} \prod_{l=1}^{\frac{1}{2}i(i+1)-1} \left(\sin^2 \varphi_l \right)^{n-i+1} \right] \cdot \left[\prod_{i=1}^{n-1} \sin \varphi \frac{i(i+1)}{2} \prod_{j=i+1}^n \left(\prod_{l=\frac{1}{2}i(i+1)}^{\frac{1}{2}(j-1)j+i} \sin \varphi_l \right) \right] \mathbb{1}_{\mathcal{I}}(\varphi), \quad (86)$$

where we have separated the \cos , \sin^2 and \sin terms. From this version, we obtain for the exponents p_l and q_l in (28) the following sum representation:

$$p_l = \sum_{i=1}^{n-1} (n-i+1) \mathbb{1}_{\left\{\frac{i(i+1)}{2}\right\}}(l), \quad (87)$$

$$q_l = \underbrace{\sum_{i=2}^{n-1} 2(n-i+1) \mathbb{1}_{\{1,2,\dots,\frac{i(i+1)}{2}-1\}}(l)}_{q_l^{[1]}} + \underbrace{\sum_{i=1}^{n-1} \left[\mathbb{1}_{\left\{\frac{i(i+1)}{2}\right\}}(l) + \sum_{j=i+1}^n \mathbb{1}_{\left\{\frac{i(i+1)}{2}, \frac{i(i+1)}{2}+1, \dots, \frac{(j-1)j}{2}+i\right\}}(l) \right]}_{q_l^{[2]}}. \quad (88)$$

Equation (87) is obviously equivalent to (29). Writing out the sums of indicator functions in (88) and then collecting terms yield

$$q_l^{[1]} = 2 \sum_{k=2}^{n-i} k = n^2 - (2i-1)n + i(i-1) - 2,$$

$$q_l^{[2]} = i(n-i) - m + 1 \quad (89)$$

if $l = \frac{1}{2}i(i+1) + m$, $i = 1, 2, \dots, n-1$, $m = 0, 1, \dots, i$.

Adding $q_l^{[1]}$ and $q_l^{[2]}$ in (89) yields the final form of q_l as given in (30) and (31).

APPENDIX B

EVALUATION OF CONSTANTS AND PARAMETER PDFS

In Table II the parameter pdfs for the real case given in (28) and the involved constants specified in (29)–(32) are evaluated for $n = 2, 3, 4$. In Table III the parameter pdfs for the complex case given in (66) and the involved constants specified in (67)–(70) are evaluated for $n = 2, 3, 4$. The indicator functions are omitted in both cases. They are equal to $\mathbb{1}_{(0,\pi/2)}$ if the pdf has the form $c_l \cos^{p_l} \sin^{q_l}$ and they are equal to $\mathbb{1}_{(0,\pi)}$ if the pdf has the form $c_l \sin^{q_l}$.

APPENDIX C

PARAMETER PDFS VS. SCALED NORMAL PDFS

In Fig. 1 the parameter pdfs defined in (28) are compared to scaled normal pdfs with expectation and variance parameters given by (36) and (37) and scaling factor ν_l given by (38). The values for p_l , q_l and c_l are taken from Table II for the marked parameter sets ($n = 4$, $l = 1, 4, 6, 8$).

APPENDIX D

ALGORITHMS

(REAL CASE)

In Algorithm 1, 2, and 3 the pseudo code of the basic method to sample uniformly from the set of real $n \times n$ positive definite matrices with constant trace is listed.

In Algorithm 4 the pseudo code of the extended algorithm to sample uniformly from the set of real $n \times n$ positive definite matrices with bounded trace is listed.

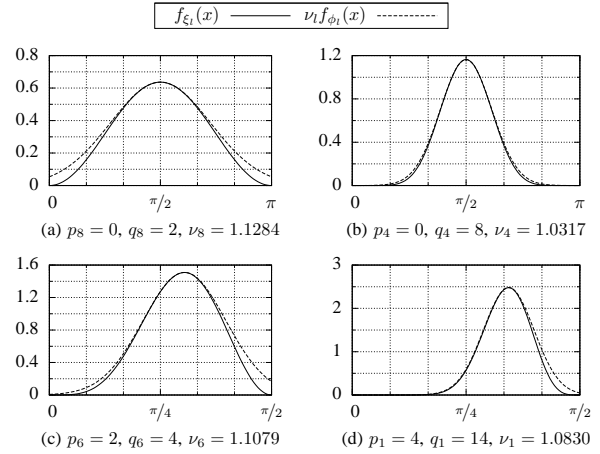


Fig. 1. Parameter pdfs vs. corresponding scaled normal pdfs for marked parameter sets from Table II ($n = 4$, $l = 1, 4, 6, 8$).

Algorithm 1 Generate M matrices from the set \mathcal{R}_c defined in (2) according to a uniform distribution.

```

1: for  $i = 1$  to  $M$  do ▷ generate  $M$  matrices  $A_1, A_2, \dots, A_M \in \mathcal{R}_c$ 
2:   for  $l = 1$  to  $\frac{1}{2}n(n+1) - 1$  do
3:     Generate  $\varphi_l$  with Algorithm 2 or 3.
4:   end for
5:   With  $\varphi_1, \dots, \varphi_{\frac{n(n+1)}{2}-1}$  compute  $x = (x_1, \dots, x_{\frac{n(n+1)}{2}})$  from (11).
6:   With vector  $x$  generate matrix  $U$  from (9).
7:    $A_i \leftarrow c U^T U$ 
8: end for

```

Algorithm 2 Generate a sample φ_l according to a distribution with pdf f_{ξ_l} defined in (28) (rejection method).

```

1: Compute  $p_l, q_l, c_l, \mathcal{I}_l$  from (29)–(32), (16). ▷ setup constants
2: With  $p_l, q_l, c_l$  compute  $\mu_l, \sigma_l^2, \nu_l$  from (36)–(39).
3: repeat ▷ rejection method
4:    $u \leftarrow \text{uni}(0, 1)$  ▷ sample uniformly from  $(0, 1)$ 
5:    $\varphi_l \leftarrow \text{normal}(\mu_l, \sigma_l^2)$  ▷ sample from normal distribution
6: until  $u \nu_l f_{\phi_l}(\varphi_l) < f_{\xi_l}(\varphi_l)$  ▷  $f_{\xi_l}$  and  $f_{\phi_l}$  as in (28), (35)

```

Algorithm 3 Generate a sample φ_l according to a distribution with pdf f_{ξ_l} defined in (28) (beta distribution method).

```

1: Compute  $p_l, q_l$  from (29)–(31). ▷ setup constants
2:  $y \leftarrow \text{beta}\left(\frac{q_l+1}{2}, \frac{p_l+1}{2}\right)$  ▷ sample from beta distribution
3: if  $p_l = 0$  then
4:    $b \leftarrow \text{bernoulli}\left(\frac{1}{2}\right)$  ▷ sample from Bernoulli distribution
5: else
6:    $b \leftarrow 0$ 
7: end if
8:  $\varphi_l \leftarrow (1-b)y + b(\pi - y)$ 

```

Algorithm 4 Generate M matrices from the set $\mathcal{R}_{(c,\bar{c})}$ defined in (3) according to a uniform distribution.

```

1: With Algorithm 1 generate  $M$  matrices  $A_1, A_2, \dots, A_M \in \mathcal{R}$ .
2: for  $i = 1$  to  $M$  do ▷ generate  $M$  matrices  $B_1, B_2, \dots, B_M \in \mathcal{R}_{(c,\bar{c})}$ 
3:    $u \leftarrow \text{uni}(0, 1)$  ▷ sample uniformly from  $(0, 1)$ 
4:    $t \leftarrow F_{\tau}^{-1}(u)$  ▷  $F_{\tau}^{-1}$  as in (58)
5:    $B_i \leftarrow t A_i$ 
6: end for

```

TABLE II
CONSTANTS AND PARAMETER PDFS (REAL CASE) FOR $n = 2, 3, 4$.

n	l	p_l	κ_l	λ_l	q_l	c_l	$f_{\xi_l}(\varphi_l)/c_l$
2	1	2	0	2	2	$\frac{16}{\pi}$	$\sin^2 \varphi_1 \cos^2 \varphi_1$
	2	0	0	3	1	$\frac{1}{2}$	$\sin \varphi_2$
3	1	3	0	2	7	40	$\sin^7 \varphi_1 \cos^3 \varphi_1$
	2	0	0	3	6	$\frac{16}{5\pi}$	$\sin^6 \varphi_2$
	3	2	1	3	3	$\frac{15}{2}$	$\sin^3 \varphi_3 \cos^2 \varphi_3$
	4	0	1	4	2	$\frac{2}{\pi}$	$\sin^2 \varphi_4$
	5	0	1	5	1	$\frac{1}{2}$	$\sin \varphi_5$
4	1	4	0	2	14	$\frac{131072}{143\pi}$	$\sin^{14} \varphi_1 \cos^4 \varphi_1$
	2	0	0	3	13	$\frac{3003}{2048}$	$\sin^{13} \varphi_2$
	3	3	1	3	9	60	$\sin^9 \varphi_3 \cos^3 \varphi_3$
	4	0	1	4	8	$\frac{128}{35\pi}$	$\sin^8 \varphi_4$
	5	0	1	5	7	$\frac{35}{32}$	$\sin^7 \varphi_5$
	6	2	2	4	4	$\frac{32}{\pi}$	$\sin^4 \varphi_6 \cos^2 \varphi_6$
	7	0	2	5	3	$\frac{3}{4}$	$\sin^3 \varphi_7$
	8	0	2	6	2	$\frac{2}{\pi}$	$\sin^2 \varphi_8$
	9	0	2	7	1	$\frac{1}{2}$	$\sin \varphi_9$

An efficient MATLAB implementation of all algorithms is available for download at [15], which also includes the algorithms to sample uniformly from the set of complex $n \times n$ positive definite matrices with constant or bounded trace.

ACKNOWLEDGMENT

Transforming a beta distribution to generate a distribution with pdf f_{ξ_l} as defined in (28) was proposed by one of the anonymous reviewers. The authors gratefully acknowledge improving the paper by contributing this method.

REFERENCES

- [1] R. B. Holmes, "On random correlation matrices," *SIAM J. Matrix Anal. Appl.*, vol. 12, no. 2, pp. 239–272, 1991.
- [2] P. I. Davies and N. J. Higham, "Numerically Stable Generation of Correlation Matrices and Their Factors," *BIT Numerical Mathematics*, vol. 40, no. 4, pp. 640–651, 2000.
- [3] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity Limits of MIMO Channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, 2003.
- [4] M. Joham, W. Utschick, and J. A. Nossek, "Linear Transmit Processing in MIMO Communications systems," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2700–2712, 2005.
- [5] S. A. Jafar and M. J. Fakhereddin, "Degrees of freedom for the MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2637–2642, 2007.
- [6] F. Negro, S. P. Shenoy, I. Ghauri, and D. T. M. Slock, "Weighted sum rate maximization in the MIMO Interference Channel," in *Proc. of PIMRC*, 2010, pp. 684–689.
- [7] Y. Liang, H. V. Poor, and S. Shamai (Shitz), *Information Theoretic Security*, ser. Foundations and Trends in Communications and Information Theory, S. Verdú, Ed. Now publishers, 2009, vol. 5, no. 4–5.
- [8] C. P. Chalmers, "Generation of correlation matrices with a given eigenstructure," *J. Statist. Comput. Simul.*, vol. 4, no. 2, pp. 133–139, 1975.

TABLE III
CONSTANTS AND PARAMETER PDFS (COMPLEX CASE) FOR $n = 2, 3, 4$.

n	l	p_l	κ_l	λ_l	q_l	c_l	$f_{\xi_l}(\varphi_l)/c_l$
2	1	3	0	1	3	12	$\sin^3 \varphi_1 \cos^3 \varphi_1$
	2	0	0	2	2	$\frac{2}{\pi}$	$\sin^2 \varphi_2$
	3	0	0	3	1	$\frac{1}{2}$	$\sin \varphi_3$
3	1	5	1	1	11	336	$\sin^{11} \varphi_1 \cos^5 \varphi_1$
	2	0	1	2	10	$\frac{256}{63\pi}$	$\sin^{10} \varphi_2$
	3	0	1	3	9	$\frac{315}{256}$	$\sin^9 \varphi_3$
	4	3	0	4	5	24	$\sin^5 \varphi_4 \cos^3 \varphi_4$
	5	0	0	5	4	$\frac{8}{3\pi}$	$\sin^4 \varphi_5$
	6	0	0	6	3	$\frac{3}{4}$	$\sin^3 \varphi_6$
	7	0	0	7	2	$\frac{2}{\pi}$	$\sin^2 \varphi_7$
	8	0	0	8	1	$\frac{1}{2}$	$\sin \varphi_8$
4	1	7	2	1	23	10920	$\sin^{23} \varphi_1 \cos^7 \varphi_1$
	2	0	2	2	22	$\frac{524288}{88179\pi}$	$\sin^{22} \varphi_2$
	3	0	2	3	21	$\frac{969969}{524288}$	$\sin^{21} \varphi_3$
	4	5	1	5	15	720	$\sin^{15} \varphi_4 \cos^5 \varphi_4$
	5	0	1	6	14	$\frac{2048}{429\pi}$	$\sin^{14} \varphi_5$
	6	0	1	7	13	$\frac{3003}{2048}$	$\sin^{13} \varphi_6$
	7	0	1	8	12	$\frac{1024}{231\pi}$	$\sin^{12} \varphi_7$
	8	0	1	9	11	$\frac{693}{512}$	$\sin^{11} \varphi_8$
	9	3	0	9	7	40	$\sin^7 \varphi_9 \cos^3 \varphi_9$
	10	0	0	10	6	$\frac{16}{5\pi}$	$\sin^6 \varphi_{10}$
	11	0	0	11	5	$\frac{15}{16}$	$\sin^5 \varphi_{11}$
	12	0	0	12	4	$\frac{8}{3\pi}$	$\sin^4 \varphi_{12}$
	13	0	0	13	3	$\frac{3}{4}$	$\sin^3 \varphi_{13}$
	14	0	0	14	2	$\frac{2}{\pi}$	$\sin^2 \varphi_{14}$
	15	0	0	15	1	$\frac{1}{2}$	$\sin \varphi_{15}$

- [9] R. B. Bendel and M. R. Mickey, "Population correlation matrices for sampling experiments," *Commun. Stat. B: Simul. Comput.*, vol. 7, no. 2, pp. 163–182, 1978.
- [10] G. Marsaglia and I. Olkin, "Generating correlation matrices," *SIAM J. Sci. and Stat. Comput.*, vol. 5, no. 2, pp. 470–475, 1984.
- [11] H. Joe, "Generating random correlation matrices based on partial correlations," *J. Multivariate Anal.*, vol. 97, no. 10, pp. 2177–2189, 2006.
- [12] D. Lewandowski, D. Kurowicka, and H. Joe, "Generating random correlation matrices based on vines and extended onion method," *J. Multivariate Anal.*, vol. 100, no. 9, pp. 1989–2001, 2009.
- [13] E. Veleva, "Uniformly distributed positive definite matrices with bounded trace," *Journal of the Technical University at Plovdiv "Fundamental Sciences and Applications"*, vol. 14, pp. 455–460, 2009.
- [14] J. Kovacevic, "How to encourage and publish reproducible research," in *Proc. of IEEE ICASSP*, 2007.
- [15] B. Matthiesen, M. Mittelbach, and E. A. Jorswieck. (2011) MATLAB implementation of the sampling algorithm. [Online]. Available: <http://www.ifn.et.tu-dresden.de/~wwwnt/sources/rndpdm/>
- [16] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed., A. Jeffrey and D. Zwillinger, Eds. Elsevier, 2007.
- [17] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.
- [18] W. Rudin, *Real and Complex Analysis*, 3rd ed. Mc-Graw Hill, 1987.
- [19] L. Devroye, *Non-Uniform Random Variate Generation*. Springer, 1986. [Online]. Available: <http://cg.scs.carleton.ca/~luc/rnbookindex.html>
- [20] D. E. Knuth, *The Art of Computer Programming: Seminumerical Algorithms*, 3rd ed. Addison-Wesley, 1998, vol. II.



Martin Mittelbach was born in Leinefelde, Germany, in 1978. He received the Diplom-Ingenieur degree in electrical engineering from the Dresden University of Technology (TUD), Germany, in 2003. Since June 2003, he has been a Research Assistant with the Communications Laboratory, Department of Electrical Engineering and Information Technology, TUD, where he is currently pursuing his doctoral degree. His research interests are within the fields of applied probability and information theory.



Bho Matthiesen (S'12) was born in Wyk/Föhr, Germany, in 1986. He is currently working towards his Diplom-Ingenieur (M.S.) degree in electrical engineering at Dresden University of Technology (TUD), Germany.



Eduard A. Jorswieck (S'01–M'05–SM'08) received his Diplom-Ingenieur (M.S.) degree and Doktor-Ingenieur (Ph.D.) degree, both in electrical engineering and computer science from the Berlin University of Technology (TUB), Germany, in 2000 and 2004, respectively.

He was with the Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institute (HHI) Berlin, Germany, from 2001 to 2006. In 2006, he joined the Signal Processing Department, Royal Institute of Technology (KTH), as a Postdoctoral Researcher and became Assistant Professor in 2007. Since February 2008, he has been the head of the Chair for Communications Theory and Full Professor at Dresden University of Technology (TUD), Germany. His main research interests are in the area of signal processing for communications and networks, applied information theory, and communications theory.

Dr. Jorswieck is elected member of the IEEE SPCOM Technical Committee. Since 2011 he has been Associate Editor for IEEE Transactions on Signal Processing. Since 2008 he has been Associate Editor for IEEE Signal Processing Letters. In 2006 Eduard received the IEEE Signal Processing Society Best Paper Award.