THE BETA-JACOBI MATRIX MODEL, THE CS DECOMPOSITION, AND GENERALIZED SINGULAR VALUE PROBLEMS

ALAN EDELMAN AND BRIAN D. SUTTON

ABSTRACT. We provide a solution to the *β-Jacobi matrix model problem* posed by Dumitriu and the first author. The random matrix distribution introduced here, called a *matrix model*, is related to the model of Killip and Nenciu, but the development is quite different. We start by introducing a new matrix decomposition and an algorithm for computing this decomposition. Then we run the algorithm on a Haar-distributed random matrix to produce the β-Jacobi matrix model.

The Jacobi ensemble on $\mathbb{R}^n$, parameterized by $\beta > 0$, $a > -1$, and $b > -1$, is the probability distribution whose density is proportional to $\prod_i \lambda_i^{\frac{\beta}{2}(a+1)-1}(1-\lambda_i)^{\frac{\beta}{2}(b+1)-1}\prod_{i<j}|\lambda_i - \lambda_j|^\beta$. The matrix model introduced in this article is a probability distribution on structured orthogonal matrices. If J is a random matrix drawn from this distribution, then a CS decomposition can be taken,

$$J = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix} \begin{bmatrix} C & S \\ -S & C \end{bmatrix} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^T,$$

in which C and S are diagonal matrices with entries in $[0,1]$. J is designed so that the diagonal entries of C, squared, follow the law of the Jacobi ensemble.

When $\beta = 1$ (resp., $\beta = 2$), the matrix model is derived by running a numerically-inspired algorithm on a Haar-distributed random matrix from the orthogonal (resp., unitary) group. Hence, the matrix model generalizes certain features of the orthogonal and unitary groups beyond $\beta = 1$ and $\beta = 2$ to general $\beta > 0$.

Observing a connection between Haar measure on the orthogonal (resp., unitary) group and pairs of real (resp., complex) Gaussian matrices, we find a direct connection between multivariate analysis of variance (MANOVA) and the new matrix model.

1. Introduction

The three *classical ensembles* of random matrix theory are Hermite, Laguerre, and Jacobi, with the following densities:

2000 *Mathematics Subject Classification.* 15A52, 62H10, 65F15, 82B05.

Key words and phrases. random matrix, Jacobi ensemble, CS decomposition, generalized singular value decomposition.

Published in *Foundations of Computational Mathematics*. The original publication is available at www.springerlink.com: http://dx.doi.org/10.1007/s10208-006-0215-9.

Ensemble	Joint density (up to a constant factor)
Hermite	$e^{-\frac{\beta}{2}\sum_i \lambda_i^2} \prod_{i<j} \lvert\lambda_i - \lambda_j\rvert^\beta$
Laguerre	$e^{-\frac{\beta}{2}\sum_i \lambda_i} \prod_i \lambda_i^{\frac{\beta}{2}(a+1)-1} \prod_{i<j} \lvert\lambda_i - \lambda_j\rvert^\beta$
Jacobi	$\prod_i \lambda_i^{\frac{\beta}{2}(a+1)-1}(1-\lambda_i)^{\frac{\beta}{2}(b+1)-1} \prod_{i<j} \lvert\lambda_i - \lambda_j\rvert^\beta$

Traditionally, the Hermite ensemble is modeled by the *eigenvalues* of a symmetric matrix with Gaussian entries, and the Laguerre ensemble is modeled by the *singular values* of a matrix with Gaussian entries. This article begins by showing that the Jacobi ensemble arises from a *CS decomposition* problem. Specifically, the $\beta = 1$ Jacobi ensemble arises from the CS decomposition of a Haar-distributed orthogonal matrix, and the $\beta = 2$ Jacobi ensemble arises from the CS decomposition of a Haar-distributed unitary matrix. This observation completes the following table, and enables the development of a new "general β" random matrix model.

Ensemble	Random linear algebra problem
Hermite	eigenvalue decomposition
Laguerre	singular value decomposition
Jacobi	CS decomposition

For several decades, random matrix theory concentrated on three values of β in the ensemble densities. The $\beta = 1$ ensembles were shown to arise from real random matrices, the $\beta = 2$ ensembles from complex random matrices, and the $\beta = 4$ ensembles from quaternion random matrices, according to Dyson's "threefold way" [3]. In recent years, the development of a *general* β theory, extending beyond $\beta = 1, 2, 4$ to all $\beta > 0$, has gained momentum. One of the fundamental problems in developing a general β theory is to find a random matrix distribution that "models" the desired ensemble in some fashion. Dumitriu and Edelman solved the matrix model problems for the Hermite and Laguerre ensembles [1]. In the Hermite case, for example, they provided a random symmetric tridiagonal matrix for each β whose eigenvalues follow the law of the Hermite ensemble. Dumitriu and Edelman posed the development of a β-Jacobi matrix model as an open problem, which has been considered in [6, 7].

The major contribution of this article is the introduction of a *β-Jacobi matrix model*, displayed in Figure 1. This matrix model is a distribution on structured orthogonal matrices, parameterized by $\beta > 0$, $a > -1$, and $b > -1$. Its CS decomposition has entries from the Jacobi ensemble with the same parameters. The matrix model can be sampled in Matlab using Listing 1 in Appendix A, and its CS values can be computed with Listing 2.

The development of the model is in the spirit of [1], utilizing an algorithm inspired by bidiagonalization and tridiagonalization algorithms from numerical linear algebra.

$$J^\beta_{a,b} \sim \begin{bmatrix} B_{11}(\Theta,\Phi) & B_{12}(\Theta,\Phi) \\ B_{21}(\Theta,\Phi) & B_{22}(\Theta,\Phi)\end{bmatrix}$$

$$= \left[\begin{array}{cccc|cccc} c_n & -s_n c'_{n-1} & & & s_n s'_{n-1} & & & \\ & c_{n-1}s'_{n-1} & \ddots & & c_{n-1}c'_{n-1} & s_{n-1}s'_{n-2} & & \\ & & \ddots & -s_2c'_1 & & \ddots & \ddots & \\ & & & c_1 s'_1 & & & c_1c'_1 & s_1 \\ \hline -s_n & -c_nc'_{n-1} & & & c_ns'_{n-1} & & & \\ & -s_{n-1}s'_{n-1} & \ddots & & -s_{n-1}c'_{n-1} & c_{n-1}s'_{n-2} & & \\ & & \ddots & -c_2c'_1 & & \ddots & \ddots & \\ & & & -s_1s'_1 & & & -s_1c'_1 & c_1\end{array}\right]$$

$$\beta > 0,\ a,\ b > -1$$

$$\begin{array}{ll} \Theta = (\theta_n, \dots, \theta_1) \in [0, \frac{\pi}{2}]^n & \Phi = (\phi_{n-1}, \dots, \phi_1) \in [0, \frac{\pi}{2}]^{n-1} \\ c_i = \cos\theta_i & c'_i = \cos\phi_i \\ s_i = \sin\theta_i & s'_i = \sin\phi_i \\ c_i^2 \sim \mathrm{beta}(\frac{\beta}{2}(a+i), \frac{\beta}{2}(b+i)) & (c'_i)^2 \sim \mathrm{beta}(\frac{\beta}{2}i, \frac{\beta}{2}(a+b+1+i)) \end{array}$$

FIGURE 1. The β-Jacobi matrix model. $\mathrm{beta}(c,d)$ denotes a beta-distributed random variable, with p.d.f. $\frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)}x^{c-1}(1-x)^{d-1}$. The angles $\theta_1, \dots, \theta_n, \phi_1, \dots, \phi_{n-1}$ are independent. See Definition 1.8 for details.

The use of the CS decomposition breaks from previous work, which has focused on eigenvalues. Notable among the existing work is that of Killip and Nenciu [6], which provides a random matrix model whose *eigenvalues* follow the law of the Jacobi ensemble. In fact, the model of Killip and Nenciu can be obtained from ours via the transformation $2(I - 2\Omega B_{11}^T B_{11}\Omega)$, in which I is the n-by-n identity matrix, B_{11} is the upper-left n-by-n block of our model (Figure 1), and Ω is a diagonal matrix with alternating ± 1's along the diagonal. Our matrix model, based on the CS decomposition, has the following advantages:

- Our matrix model is a random orthogonal matrix, generalizing certain features of the orthogonal and unitary groups to general β.
- CS decomposition is used in place of eigenvalue decomposition, which is natural considering that the Jacobi ensemble is a distribution on $[0,1]^n$ rather than all of $\mathbb{R}^n$. (CS values lie in $[0,1]$ by definition.)

- The matrix model has both left and right CS vectors, rather than just eigenvectors.
- The development of the matrix model is illuminating, based on a numerically-inspired algorithm.
- There is an immediate connection to multivariate analysis of variance (MANOVA), based on the similarity between CS decomposition and generalized singular value decomposition.

More details on the β-Jacobi matrix model, notably asymptotics for large matrix sizes, can be found in the Ph. D. thesis of the second author [11].

1.1. **Background.** The Jacobi ensemble, a distribution on n-vectors whose p.d.f. is proportional to

$$\prod_{i=1}^{n} \lambda_i^{\frac{\beta}{2}(a+1)-1}(1-\lambda_i)^{\frac{\beta}{2}(b+1)-1} \prod_{i<j} |\lambda_i - \lambda_j|^{\beta},$$

in which β, a, and b are assignable parameters, has been studied extensively, motivated by applications in both physics and statistics.

In statistical mechanics, the ensemble arises in the context of log gases. A log gas is a system of charged particles on the real line that are subject to a logarithmic interaction potential as well as Brownian-like fluctuations. If the particles are constrained to the interval $[0,1]$ and are also subject to the external potential $\sum_{i=1}^{n}(\frac{a+1}{2}-\frac{1}{\beta})\log\lambda_i + \sum_{i=1}^{n}(\frac{b+1}{2}-\frac{1}{\beta})\log(1-\lambda_i)$, then the long term stationary distribution of the system of charges is the Jacobi ensemble [2, 4, 13].

In statistics, the ensemble arises in the context of MANOVA, starting from a pair of independent Gaussian matrices N_1, N_2. If N_1 and N_2 have independent real entries, then their generalized singular values, squared, follow the law of the Jacobi ensemble with $\beta = 1$. If they have independent complex entries, then their generalized singular values (squared) follow the law of the Jacobi ensemble with $\beta = 2$. Now we define the generalized singular value decomposition (GSVD) and make these statements precise.

Definition 1.1 (GSVD)**.** Let A be $(n+a)$-by-n and B be $(n+b)$-by-n with complex entries. Then there exist matrices R, U_1, U_2, V, C, and S such that

$$\left[\begin{array}{c} A \\ \hline B \end{array}\right] = \left[\begin{array}{c|c} U_1 & \\ \hline & U_2 \end{array}\right] \left[\begin{array}{c} C \\ 0 \\ \hline -S \\ 0 \end{array}\right] V^* R, \tag{1}$$

in which R is n-by-n upper triangular, U_1 is $(n+a)$-by-$(n+a)$ unitary, U_2 is $(n+b)$-by-$(n+b)$ unitary, V is n-by-n unitary, and C and S are nonnegative diagonal, satisfying $C^2 + S^2 = I$. The diagonal entries of C are known as the *generalized singular values* of the pair A, B, and the factorization in (1) is a *generalized singular value decomposition* (GSVD).

There are a few observations worth mentioning. First, this definition does not define the GSVD uniquely. (C and S are unique up to reordering, but the other matrices have extra freedom when some of the generalized singular values occur multiply.) Second, if A and B have real entries, then R, U_1, U_2, and V may be taken to have real entries as well. Third, many authors refer to the cotangents $\frac{c_k}{s_k}, k = 1, \ldots, n$, instead of the cosines c_k, as generalized singular values.

One way to construct a GSVD, which may not be the most numerically accurate, is to first compute a QR decomposition of $\left[\begin{smallmatrix} A \\ B \end{smallmatrix}\right]$, and then to compute SVD's for the top and bottom blocks of Q. See [12] for details.

The Jacobi ensemble can be seen in the generalized singular values of a pair of Gaussian matrices. A real standard Gaussian random variable has p.d.f. $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. A complex standard Gaussian random variable is distributed as $\frac{1}{\sqrt{2}}(G_1 + \sqrt{-1} G_2)$, in which G_1 and G_2 are independent real standard Gaussians.

Proposition 1.2. *Let N_1 and N_2 be independent random matrices. Suppose that N_1 is $(n+a)$-by-n and N_2 is $(n+b)$-by-n, each with i.i.d. real (resp., complex) standard Gaussian entries. Then the generalized singular values, squared, of the pair N_1, N_2 follow the law of the Jacobi ensemble with parameters a, b, for $\beta = 1$ (resp., $\beta = 2$).*

Proof. The generalized singular values, squared, are equal to the eigenvalues of $N_1^* N_1 (N_1^* N_1 + N_2^* N_2)^{-1}$, which behave as the Jacobi ensemble [8]. To see this, note that $N_1^* N_1 (N_1^* N_1 + N_2^* N_2)^{-1} = N_1^* N_1 \left(\left[\begin{smallmatrix} N_1 \\ N_2 \end{smallmatrix}\right]^* \left[\begin{smallmatrix} N_1 \\ N_2 \end{smallmatrix}\right] \right)^{-1}$, so if the CSD of $\left[\begin{smallmatrix} N_1 \\ N_2 \end{smallmatrix}\right]$ is

$$\begin{bmatrix} N_1 \\ N_2 \end{bmatrix} = \left[\begin{array}{c|c} U_1 & \\ \hline & U_2 \end{array}\right] \left[\begin{array}{c} C \\ 0 \\ \hline -S \\ 0 \end{array}\right] V^* R,$$

then $N_1^* N_1 (N_1^* N_1 + N_2^* N_2)^{-1} = (R^* V) C^2 (R^* V)^{-1}$. □

The preceding proposition provides matrix models for the Jacobi ensemble in the cases $\beta = 1$ and $\beta = 2$, for integral a and b. The primary contribution of this article is a general β matrix model, which also removes the quantization on a and b.

1.2. **Results.** We show that the $\beta = 1, 2$ Jacobi ensembles arise from Haar measure on compact matrix groups, through the CS decomposition (CSD). This viewpoint is central to the development of the general β-Jacobi matrix model. The CS decomposition is perhaps less familiar than the eigenvalue and singular value decompositions (SVDs), but it has the same flavor. A proof of the following proposition can be found in [10].

Proposition 1.3. *Let X be an m-by-m unitary matrix, and let p, q be nonnegative integers such that $p \geq q$ and $p + q \leq m$. Then there exist unitary matrices U_1, U_2, V_1, and V_2, of sizes p-by-p, $(m-p)$-by-$(m-p)$, q-by-q, and $(m-q)$-by-$(m-q)$, respectively, such that*

$$X = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix} \left[\begin{array}{c|c|c} C & S & \\ & & I_{p-q} \\ \hline -S & C & \\ \hline & & I_{m-p-q} \end{array}\right] \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^*, \tag{2}$$

with C and S q-by-q nonnegative diagonal. The relationship $C^2 + S^2 = I$ is guaranteed.

Definition 1.4 (CSD)**.** Assume that in the factorization (2), the diagonal entries of C are distinct. Then the factorization is made unique by imposing that the diagonal entries of C are increasing and that the last nonzero entry in each column of $V_1 \oplus V_2$ is real positive. This factorization is known as the *CS decomposition* of X (with partition size p-by-q), and the entries of C will be called the (p-by-q) *CS values* of X.

This form of CSD is similar to the "Davis-Kahan-Stewart direct rotation form" of [10].

There is a deep connection between CSD and GSVD. Specifically, if a unitary X is partitioned into $X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$, with X_{11} of size p-by-q, then the generalized singular values of the pair X_{11}, X_{21} equal the p-by-q CS values of X. This fact is evident from the definitions. The connection between CSD and GSVD allows us to see the Jacobi ensemble in the CSD of a Haar-distributed orthogonal or unitary matrix.

Theorem 1.5. *Let n be a positive integer, let a and b be nonnegative integers, and define $m = 2n + a + b$. Let X be an m-by-m Haar-distributed orthogonal matrix, and take the CS decomposition of X with partition size $(n+a)$-by-n. Then the CS values of X, squared, follow the law of the $\beta = 1$ Jacobi ensemble with parameters a, b. If, instead, X is a Haar-distributed unitary matrix, then the CS values, squared, obey the law of the $\beta = 2$ Jacobi ensemble.*

Proof. Let $\begin{bmatrix} A \\ B \end{bmatrix}$ be an m-by-n matrix of independent standard Gaussian entries, with A $(n+a)$-by-n and B $(n+b)$-by-n. We claim that the CS values of X share the same distribution with the generalized singular values of the pair A, B. Upon showing this, the proof will follow by Proposition 1.2.

With probability 1, the generalized singular values are distinct, so we can take a QR decomposition, $\begin{bmatrix} A \\ B \end{bmatrix} = QR$, with R invertible. Next, randomize signs, $QR = (QD)(D^*R)$, using a diagonal matrix D with i.i.d. entries chosen uniformly from either $\{-1, 1\}$ (if X is real orthogonal) or the unit circle (if X is complex unitary). It is well known that QD shares the same distribution with the first n columns of X. Therefore, the CS values of X

share the same distribution with the singular values of the first $n+a$ rows of QD. But these singular values equal the generalized singular values of the pair A, B. (Note that $\left[\begin{smallmatrix} A \\ B \end{smallmatrix}\right] = (QD)(D^*R)$, generalized singular values are invariant under right multiplication by an invertible matrix, and the first $n+a$ and last $n+b$ rows of QD must have the same right singular vector matrix since $(QD)^*(QD) = I$.) □

Now we introduce the β-Jacobi matrix model, valid for all $\beta > 0$ and $a, b > -1$. The model is a distribution on orthogonal matrices with a special structure.

Definition 1.6. Given $\Theta = (\theta_n, \dots, \theta_1)$ and $\Phi = (\phi_{n-1}, \dots, \phi_1)$, we define four n-by-n bidiagonal matrices, $B_{11}(\Theta, \Phi)$, $B_{12}(\Theta, \Phi)$, $B_{21}(\Theta, \Phi)$, and $B_{22}(\Theta, \Phi)$, as follows.

$$\left[\begin{array}{c|c} B_{11}(\Theta,\Phi) & B_{12}(\Theta,\Phi) \\ \hline B_{21}(\Theta,\Phi) & B_{22}(\Theta,\Phi) \end{array}\right] =$$

(3)

$$= \left[\begin{array}{cccc|cccc}
c_n & -s_n c'_{n-1} & & & s_n s'_{n-1} & & & \\
 & c_{n-1} s'_{n-1} & \ddots & & c_{n-1} c'_{n-1} & s_{n-1} s'_{n-2} & & \\
 & & \ddots & -s_2 c'_1 & & \ddots & \ddots & \\
 & & & c_1 s'_1 & & & c_1 c'_1 & s_1 \\
\hline
-s_n & -c_n c'_{n-1} & & & c_n s'_{n-1} & & & \\
 & -s_{n-1} s'_{n-1} & \ddots & & -s_{n-1} c'_{n-1} & c_{n-1} s'_{n-2} & & \\
 & & \ddots & -c_2 c'_1 & & \ddots & \ddots & \\
 & & & -s_1 s'_1 & & & -s_1 c'_1 & c_1
\end{array}\right],$$

in which $c_i = \cos\theta_i$, $s_i = \sin\theta_i$, $c'_i = \cos\phi_i$, $s'_i = \sin\phi_i$.

To clarify, the $(n-1, n-1)$ entry of $B_{12}(\Theta, \Phi)$ is $s_2 s'_1$, and the $(n-1, n-1)$ entry of $B_{22}(\Theta, \Phi)$ is $c_2 s'_1$. Also, if $n = 1$, then the matrices are

$$\left[\begin{array}{c|c} B_{11}((\theta_1),()) & B_{12}((\theta_1),()) \\ \hline B_{21}((\theta_1),()) & B_{22}((\theta_1),()) \end{array}\right] = \left[\begin{array}{c|c} c_1 & s_1 \\ \hline -s_1 & c_1 \end{array}\right].$$

Lemma 1.7. *For any real Θ, Φ, the matrix* $\begin{bmatrix} B_{11}(\Theta,\Phi) & B_{12}(\Theta,\Phi) \\ B_{21}(\Theta,\Phi) & B_{22}(\Theta,\Phi) \end{bmatrix}$ *is orthogonal.*

The proof is left to the reader.

The β-Jacobi matrix model can now be defined. It is a random matrix of the form (3), defined by placing a distribution on Θ, Φ. Hence, by the lemma, the β-Jacobi matrix model is a random orthogonal matrix.

Definition 1.8 (β-Jacobi matrix model)**.** The $2n$-by-$2n$ *β-Jacobi matrix model* $J^\beta_{a,b}$ is the random matrix distribution defined as follows.

$$J^\beta_{a,b} \sim \begin{bmatrix} B_{11}(\Theta,\Phi) & B_{12}(\Theta,\Phi) \\ B_{21}(\Theta,\Phi) & B_{22}(\Theta,\Phi) \end{bmatrix},$$

in which $\Theta = (\theta_n, \ldots, \theta_1)$ and $\Phi = (\phi_{n-1}, \ldots, \phi_1)$ are independently distributed as follows, with all angles between 0 and $\frac{\pi}{2}$.

$$\cos^2 \theta_i \sim \text{beta}\left(\tfrac{\beta}{2}(a+i), \tfrac{\beta}{2}(b+i)\right), \qquad i = n, n-1, \ldots, 1,$$
$$\cos^2 \phi_i \sim \text{beta}\left(\tfrac{\beta}{2}i, \tfrac{\beta}{2}(a+b+1+i)\right), \qquad i = n-1, n-2 \ldots, 1.$$

beta(c, d) refers to the beta distribution with parameters c and d, whose p.d.f. is $\frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)} x^{c-1}(1-x)^{d-1}$ for $0 < x < 1$. This matrix model is first derived in the real and complex cases ($\beta = 1, 2$) by applying unitary transformations to a Haar-distributed matrix from the orthogonal or unitary group. These unitary transformations are structured to preserve CS values. In fact, they are direct sums of Householder reflectors, chosen by an algorithm reminiscent of familiar algorithms from numerical analysis; see Section 2. This algorithmic approach is used in Section 3 to prove the following theorem in the special cases $\beta = 1, 2$ with a and b integers.

Theorem. *Let β be any positive real number, let n be a positive integer, and let $a, b > -1$. Take the n-by-n CS decomposition of the $2n$-by-$2n$ β-Jacobi matrix model $J_{a,b}^\beta$,*

$$J_{a,b}^\beta = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix} \begin{bmatrix} C & S \\ -S & C \end{bmatrix} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^T.$$

Then the diagonal entries of C, squared, follow the law of the Jacobi ensemble with parameters β, a, b.

The theorem is first proved in the classical cases $\beta = 1, 2$ as Corollary 3.7 and then proved in full generality in Section 4.

The article concludes with Corollary 5.1, restated here, relating our work to MANOVA.

Corollary. *Partitioning the $2n$-by-$2n$ β-Jacobi matrix model into n-by-n blocks, $J_{a,b}^\beta = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$, the generalized singular values, squared, of the pair B_{11}, B_{21} follow the law of the Jacobi ensemble with the same parameters.*

As far as generalized singular values are concerned, the pair B_{11}, B_{21} behaves in the same way as a pair of Gaussian matrices.

2. Bidiagonalization

2.1. **Bidiagonal block form.**

The β-Jacobi matrix model is a random matrix in *bidiagonal block form*, satisfying a certain sign pattern. Throughout this article, $+$ in a sign pattern denotes a nonnegative entry, $-$ denotes a nonpositive entry, $\times$ denotes an unconstrained entry, and blanks denote zero entries.

Definition 2.1. Let A be a real m-by-m matrix, and let $p \geq q$ be nonnegative integers such that $p + q \leq m$. A is in *bidiagonal block form* with partition size p-by-q if A has the sign pattern in Figure 2.

$$
\left[\begin{array}{cccc|ccccc|cccc}
+ & - & & & + & & & & & & & & \\
& + & - & & + & + & & & & & & & \\
& & + & \ddots & & + & + & & & & & & \\
& & & \ddots & & & \ddots & \ddots & & & & & \\
& & & & & & & \ddots & \ddots & & & & \\
\hline
- & - & & & + & & & & & & & & \\
& - & - & & - & + & & & & & & & \\
& & - & \ddots & & - & + & & & & & & \\
& & & \ddots & & & \ddots & \ddots & & & & & \\
\hline
& & & & & & & & & + & & & \\
& & & & & & & & & & + & & \\
& & & & & & & & & & & + & \\
& & & & & & & & & & & & \ddots
\end{array}\right]
$$

The rows are partitioned into blocks of sizes p, q, and $(m-p-q)$, and the columns are partitioned into blocks of sizes q, p, and $(m-p-q)$.

FIGURE 2. Bidiagonal block form.

Bidiagonal block form is most interesting in the context of unitary matrices. We shall see an analogy:

Finite computation	Infinite computation
tridiagonal form	eigenvalue decomposition
bidiagonal form	singular value decomposition
bidiagonal block form	CS decomposition

An anonymous referee alerted us to [14], in which bidiagonal block form arises from running a Lanczos-type iteration on a unitary matrix. In our work, an original algorithm based on Householder reflectors plays a vital role.

The following proposition is not used in this paper, so its proof is omitted.

Proposition 2.2. *If Y is an m-by-m orthogonal matrix in bidiagonal block form with partition size p-by-q, then there exist unique $\Theta = (\theta_q, \ldots, \theta_1)$ and $\Phi = (\phi_{q-1}, \ldots, \phi_1)$, with entries between 0 and $\frac{\pi}{2}$, such that*

$$
Y = \left[\begin{array}{c|cc|c}
B_{11}(\Theta,\Phi) & B_{12}(\Theta,\Phi) & & \\
& & I_{p-q} & \\
\hline
B_{21}(\Theta,\Phi) & B_{22}(\Theta,\Phi) & & \\
\hline
& & & I_{m-p-q}
\end{array}\right]. \tag{4}
$$

The following theorem is proved in Subsection 2.3.

Theorem 2.3. *Given m-by-m unitary X and $p \geq q \geq 0$ with $p+q \leq m$, there exist matrices U, Y, and V such that*

(1) $U^*XV = Y$.

(2) *U is unitary and block diagonal, with blocks of sizes p-by-p and $(m-p)$-by-$(m-p)$.*

$$\underset{\text{(unitary)}}{X} =$$

$$= \underbrace{\left[\begin{array}{ccc|ccc} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \hline \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \end{array}\right]}_{Y^{(0)}} \to \underbrace{\left[\begin{array}{ccc|ccc} \mathbf{+} & \times & \times & \times & \times & \times \\ \mathbf{0} & \times & \times & \times & \times & \times \\ \mathbf{0} & \times & \times & \times & \times & \times \\ \hline \mathbf{-} & \times & \times & \times & \times & \times \\ \mathbf{0} & \times & \times & \times & \times & \times \\ \mathbf{0} & \times & \times & \times & \times & \times \end{array}\right]}_{Y^{(1)}} \to \underbrace{\left[\begin{array}{ccc|ccc} + & \mathbf{-} & \mathbf{0} & \mathbf{+} & \mathbf{0} & \mathbf{0} \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ \hline - & \mathbf{-} & \mathbf{0} & \mathbf{+} & \mathbf{0} & \mathbf{0} \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \end{array}\right]}_{Y^{(2)}}$$

$$\to \underbrace{\left[\begin{array}{ccc|ccc} + & - & 0 & + & 0 & 0 \\ 0 & \mathbf{+} & \times & \mathbf{+} & \times & \times \\ 0 & \mathbf{0} & \times & \mathbf{0} & \times & \times \\ \hline - & - & 0 & + & 0 & 0 \\ 0 & \mathbf{-} & \times & \mathbf{-} & \times & \times \\ 0 & \mathbf{0} & \times & \mathbf{0} & \times & \times \end{array}\right]}_{Y^{(3)}} \to \underbrace{\left[\begin{array}{ccc|ccc} + & - & 0 & + & 0 & 0 \\ 0 & + & \mathbf{-} & + & \mathbf{+} & \mathbf{0} \\ 0 & 0 & \times & 0 & \times & \times \\ \hline - & - & 0 & + & 0 & 0 \\ 0 & - & \mathbf{-} & - & \mathbf{+} & \mathbf{0} \\ 0 & 0 & \times & 0 & \times & \times \end{array}\right]}_{Y^{(4)}} \to \underbrace{\left[\begin{array}{ccc|ccc} + & - & 0 & + & 0 & 0 \\ 0 & + & - & + & + & 0 \\ 0 & 0 & \mathbf{+} & 0 & \mathbf{+} & \times \\ \hline - & - & 0 & + & 0 & 0 \\ 0 & - & - & - & + & 0 \\ 0 & 0 & \mathbf{-} & 0 & \mathbf{-} & \times \end{array}\right]}_{Y^{(5)}}$$

$$\to \underbrace{\left[\begin{array}{ccc|ccc} + & - & 0 & + & 0 & 0 \\ 0 & + & - & + & + & 0 \\ 0 & 0 & + & 0 & + & \mathbf{+} \\ \hline - & - & 0 & + & 0 & 0 \\ 0 & - & - & - & + & 0 \\ 0 & 0 & - & 0 & - & \mathbf{+} \end{array}\right]}_{Y^{(6)}} = \underset{\text{(orthogonal, bidiagonal block form)}}{Y}$$

FIGURE 3. An example run of the algorithm. In this example, $m = 6$, $p = 3$, and $q = 3$. When $p + q < m$, a postprocessing step is required.

(3) *V is unitary and block diagonal, with blocks of sizes q-by-q and $(m-q)$-by-$(m-q)$.*
(4) *Y is an orthogonal matrix in bidiagonal block form with partition size p-by-q.*
(5) *X and Y share the same p-by-q CS values.*

2.2. **The algorithm.** We present an algorithm that transforms any unitary matrix into a matrix in bidiagonal block form. The transformation is accomplished using block diagonal unitary matrices, to preserve CS values. The algorithm serves as a constructive proof of Theorem 2.3.

The behavior of the algorithm is suggested graphically in Figure 3. In the first step, a pair of Householder reflectors works on column 1. The other columns are modified, but they are not directly observed. In the next step, a pair of Householder reflectors works on rows 1 and $p+1$. The algorithm continues, working on columns 2 and $q+1$, then rows 2 and $p+2$, then columns 3 and $q+2$, then rows 3 and $p+3$, and so on.

The algorithm is defined in Figure 4. Submatrices of Y are specified using subscripts and Matlab-style indices. The $\oplus$ operator constructs block diagonal matrices. house(z) represents a Householder reflector that maps z to the first column of an identity matrix. (If z is a "vector of length zero," then house(z) is a "0-by-0 matrix.") $\cos^{-1}$ and $\sin^{-1}$ have range $[0, \frac{\pi}{2}]$ for our purposes.

2.3. **Analysis of the algorithm.** This section considers the correctness of the algorithm. Suppose that the algorithm is run on an m-by-m unitary matrix X with partition size p-by-q, with outputs $\Theta = (\theta_q, \ldots, \theta_1)$ and $\Phi = (\phi_{q-1}, \ldots, \phi_1)$, and let

$$J = \left[\begin{array}{c|c|c} B_{11}(\Theta,\Phi) & \begin{array}{cc} B_{12}(\Theta,\Phi) & \\ & I_{p-q} \end{array} & \\ \hline B_{21}(\Theta,\Phi) & B_{22}(\Theta,\Phi) & \\ \hline & & I_{m-p-q} \end{array}\right].$$

Also, let

$$P_{2k-1} = \begin{bmatrix} I_{q,k} & \\ & I_{m-q,k-1} \end{bmatrix} \begin{bmatrix} I_{q,k} & \\ & I_{m-q,k-1} \end{bmatrix}^T, \ k = 1, \ldots, q,$$

and

$$P_{2k} = \begin{bmatrix} I_{p,k} & \\ & I_{m-p,k} \end{bmatrix} \begin{bmatrix} I_{p,k} & \\ & I_{m-p,k} \end{bmatrix}^T, \ k = 1, \ldots, q,$$

and, for $k = 1, \ldots, q$, let $J^{(2k-1)} = JP_{2k-1}$ and $J^{(2k)} = P_{2k}J$. Note that $J^{(2k-1)}$ depends only on $\theta_q, \ldots, \theta_{q+1-k}$ and $\phi_{q-1}, \ldots, \phi_{q+1-k}$, and therefore is determined halfway through the kth iteration of the algorithm. Also, $J^{(2k)}$ depends only on $\theta_q, \ldots, \theta_{q+1-k}$ and $\phi_{q-1}, \ldots, \phi_{q-k}$, and therefore is determined by the end of the kth iteration.

The proof of correctness rests on five assertions. The code in Figure 4 indicates natural stages at which to check the assertions.

Assertion 1. This assertion concerns the entries of Y suggested by the squares in the following matrix:

$$\left[\begin{array}{cccccc|cccccc} + & - & & & & & + & & & & & \\ & + & - & & & & + & + & & & & \\ & & \square & \times & \cdots & \times & & \square & \times & \times & \cdots & \times \\ & & \square & \times & \cdots & \times & & \square & \times & \times & \cdots & \times \\ & & \vdots & \vdots & \ddots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ & & \square & \times & \cdots & \times & & \square & \times & \times & \cdots & \times \\ \hline - & - & & & & & + & & & & & \\ & - & - & & & & - & + & & & & \\ & & \square & \times & \cdots & \times & & \square & \times & \times & \cdots & \times \\ & & \square & \times & \cdots & \times & & \square & \times & \times & \cdots & \times \\ & & \vdots & \vdots & \ddots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ & & \square & \times & \cdots & \times & & \square & \times & \times & \cdots & \times \end{array}\right].$$

Let $r = q + 1 - k$. For the $k = 1$ case,

$$\left[\begin{array}{c} Y^{(0)}_{1:p,1} \\ \hline Y^{(0)}_{p+1:m,1} \end{array}\right] = \left[\begin{array}{c} z_q \\ \hline -w_q \end{array}\right].$$

Algorithm `bdb`

Input: X (m-by-m unitary) and $p \geq q \geq 0$ s.t. $p+q \leq m$
Output: Y (bidiagonal block form) and $\theta_q, \ldots, \theta_1, \phi_{q-1}, \ldots, \phi_1$

Y `:=` $Y^{(0)}$ `:=` X

`for` k `= 1 :` q

r `:=` $q+1-k$

$$z_r \; \texttt{:=} \; \begin{cases} Y_{1:p,1} & \texttt{if } k=1 \\ s'_r Y_{k:p,k} + c'_r Y_{k:p,q-1+k} & \texttt{if } k>1 \end{cases}$$

$$w_r \; \texttt{:=} \; \begin{cases} -Y_{p+1:m,1} & \texttt{if } k=1 \\ -s'_r Y_{p+k:m,k} - c'_r Y_{p+k:m,q-1+k} & \texttt{if } k>1 \end{cases}$$

$$\theta_r \; \texttt{:=} \; \begin{cases} \cos^{-1}\|z_r\| & \texttt{if } \|z_r\| < \frac{1}{\sqrt{2}} \\ \sin^{-1}\|w_r\| & \texttt{if } \|z_r\| \geq \frac{1}{\sqrt{2}} \end{cases}$$

`assert(#1)`
U_k `:=` $(I_{k-1} \oplus \text{house}(z_r) \oplus I_{k-1} \oplus \text{house}(w_r))^*$
Y `:=` $Y^{(2k-1)}$ `:=` $U_k^* Y$
`assert(#2)`

r `:=` $q-k$

$$(z'_r)^* \; \texttt{:=} \; \begin{cases} -s_{r+1} Y_{k,k+1:q} - c_{r+1} Y_{p+k,k+1:q} & \texttt{if } k<q \\ () & \texttt{if } k=q \end{cases}$$

$(w'_r)^*$ `:=` $s_{r+1} Y_{k,q+k:m} + c_{r+1} Y_{p+k,q+k:m}$

$$\texttt{if } k<q \texttt{ then } \phi_r \; \texttt{:=} \; \begin{cases} \cos^{-1}\|z'_r\| & \texttt{if } \|z'_r\| < \frac{1}{\sqrt{2}} \\ \sin^{-1}\|w'_r\| & \texttt{if } \|z'_r\| \geq \frac{1}{\sqrt{2}} \end{cases}$$

`assert(#3)`
V_k `:=` $(I_k \oplus \text{house}(z'_r) \oplus I_{k-1} \oplus \text{house}(w'_r))^*$
Y `:=` $Y^{(2k)}$ `:=` YV_k
`assert(#4)`

`end`

V_{q+1} `:=` $I_{2q} \oplus (Y_{[1+q:p \;\; p+q+1:m],2q+1:m})^*$
Y `:=` YV_{q+1}
`assert(#5)`

FIGURE 4. The algorithm for transforming a unitary matrix into bidiagonal block form while preserving CS values. A more intuitive definition for θ_r is $\tan\theta_r = \frac{\|w_r\|}{\|z_r\|}$, and a more intuitive definition for ϕ_r is $\tan\phi_r = \frac{\|w'_r\|}{\|z'_r\|}$. (It will be shown that $\|z_r\|^2 + \|w_r\|^2 = \|z'_r\|^2 + \|w'_r\|^2 = 1$.) The definitions for θ_r and ϕ_r used in the algorithm are chosen for numerical stability.

When $k > 1$,

$$\left[\frac{Y^{(2k-2)}_{k:p,k}}{Y^{(2k-2)}_{p+k:m,k}}\right] = s'_r \left[\frac{z_r}{-w_r}\right] \text{ and } \left[\frac{Y^{(2k-2)}_{k:p,q-1+k}}{Y^{(2k-2)}_{p+k:m,q-1+k}}\right] = c'_r \left[\frac{z_r}{-w_r}\right].$$

For all $k = 1, \ldots, q$, $(z_r^*, w_r^*)^*$ has unit norm, $\|z_r\| = c_r$, and $\|w_r\| = s_r$.

Assertion 2. This assertion says that halfway through the kth iteration, certain rows and columns have attained their final values.

$$Y^{(2k-1)} P_{2k-1} = J^{(2k-1)},\ k = 1, \ldots, q,$$
$$P_{2k-2} Y^{(2k-1)} = J^{(2k-2)},\ k = 2, \ldots, q.$$

Assertion 3. This assertion concerns the entries of Y suggested by the squares in the following matrix:

$$\left[\begin{array}{cccccc|cccccc}
+ & - & & & & & + & & & & & \\
 & + & - & & & & + & + & & & & \\
 & & + & \square & \cdots & \square & & + & \square & \square & \cdots & \square \\
 & & & \times & \cdots & \times & & & \times & \times & \cdots & \times \\
 & & & \vdots & \ddots & \vdots & & & \vdots & \vdots & \ddots & \vdots \\
 & & & \times & \cdots & \times & & & \times & \times & \cdots & \times \\
\hline
- & - & & & & & + & & & & & \\
 & - & - & & & & - & + & & & & \\
 & & - & \square & \cdots & \square & & - & \square & \square & \cdots & \square \\
 & & & \times & \cdots & \times & & & \times & \times & \cdots & \times \\
 & & & \vdots & \ddots & \vdots & & & \vdots & \vdots & \ddots & \vdots \\
 & & & \times & \cdots & \times & & & \times & \times & \cdots & \times
\end{array}\right].$$

Let $r = q - k$. When $k < q$,

$$\left[\ Y^{(2k-1)}_{k,k+1:q} \ \middle|\ Y^{(2k-1)}_{k,q+k:m}\ \right] = s_{r+1} \left[\ -(z'_r)^* \ \middle|\ (w'_r)^*\ \right],$$
$$\left[\ Y^{(2k-1)}_{p+k,k+1:q} \ \middle|\ Y^{(2k-1)}_{p+k,q+k:m}\ \right] = c_{r+1} \left[\ -(z'_r)^* \ \middle|\ (w'_r)^*\ \right],$$

and when $k = q$,

$$\left[\ Y^{(2q-1)}_{q,2q:m}\ \right] = s_1 \left[\ (w'_0)^*\ \right],$$
$$\left[\ Y^{(2q-1)}_{p+q,2q:m}\ \right] = c_1 \left[\ (w'_0)^*\ \right].$$

When $k < q$, $((z'_r)^*, (w'_r)^*)^*$ has unit norm, $\|z'_r\| = c'_r$, and $\|w'_r\| = s'_r$. When $k = q$, $\|w'_r\| = 1$.

Assertion 4. This assertion says that by the end of the kth iteration, certain rows and columns have attained their final values.

$$P_{2k} Y^{(2k)} = J^{(2k)},\ k = 1, \ldots, q,$$
$$Y^{(2k)} P_{2k-1} = J^{(2k-1)},\ k = 1, \ldots, q.$$

Assertion 5. The final matrix Y is in bidiagonal block form, defined by the angles $\theta_q, \ldots, \theta_1$ and $\phi_{q-1}, \ldots, \phi_1$. In other words, $Y = J$.

Lemma 2.4. *The assertions in Algorithm* `bdb` *(Figure 4) hold.*

The proof can be found in Appendix B.

Proof of Theorem 2.3. The algorithm produces matrices Y, $U = U_1 \cdot \cdots \cdot U_q$, and $V = V_1 \cdot \cdots \cdot V_q V_{q+1}$ that deliver the five conclusions of the theorem. Conclusions (1), (2), and (3) are obvious from the design of the algorithm, and (5) follows immediately from (1), (2), and (3). (4) is a consequence of Assertion 5. □

3. Real and complex random matrices

The β-Jacobi matrix model is first obtained in the real and complex cases by running the algorithm on a Haar-distributed random matrix X. Recall that the CS values of the random matrix X follow the law of the Jacobi ensemble. The algorithm produces a random orthogonal matrix Y in bidiagonal block form, with the same CS values as X. Hence, the CS values of Y follow the Jacobi ensemble as well. We show that the distribution of Y is the β-Jacobi matrix model. This proves that the CS values of the β-Jacobi matrix model follow the law of the Jacobi ensemble, as desired, in the classical real and complex cases.

Let n be a positive integer, let a and b be nonnegative integers, and define $m = 2n + a + b$. Let G be either $O(m)$ or $U(m)$, i.e., either the orthogonal group or the unitary group of m-by-m matrices, and let X be a random matrix from G whose distribution is Haar measure. Running the algorithm on X with partition size $(n + a)$-by-n produces a sequence of intermediate matrices $Y^{(1)}, Y^{(2)}, \ldots, Y^{(2n)}$. Each $Y^{(i)}$ is itself a random matrix, and we are interested in its distribution.

We shall show that the distribution of $Y^{(i)}$ is defined by invariance properties. Let $\mathcal{V}_0 = \mathcal{U}_0 = G$, and for $i = 1, 2, \ldots, 2n$, let $\mathcal{V}_i = \{V \in G : J^{(i)}V = J^{(i)}\}$ and $\mathcal{U}_i = \{U \in G : U^* J^{(i)} = J^{(i)}\}$. ($J^{(i)}$ is defined in Subsection 2.3.) Claim: The distribution of $Y^{(i)}$ is uniquely determined by

1. $Y^{(2k-1)} P_{2k-1} = J^{(2k-1)}$ if $i = 2k - 1$ is odd, or $P_{2k} Y^{(2k)} = J^{(2k)}$ if $i = 2k$ is even, and
2. the distribution of $Y^{(i)}$ is both $\mathcal{V}_i$- and $\mathcal{U}_i$-invariant,

in which $\mathcal{V}_i$*-invariance* means that $Y^{(i)} V \overset{d}{=} Y^{(i)}$ for all $V \in \mathcal{V}_i$ and $\mathcal{U}_i$-*invariance* means that $U^* Y^{(i)} \overset{d}{=} Y^{(i)}$ for all $U \in \mathcal{U}_i$.

According to the following lemma, $\mathcal{V}_{2k-1}$ and $\mathcal{U}_{2k}$ are simple to describe. $\mathcal{U}_{2k-1}$ and $\mathcal{V}_{2k}$ are more complicated.

Lemma 3.1. *For* $k = 1, \ldots, n$, $\mathcal{V}_{2k-1}$ *consists of all matrices* $V \in G$ *that preserve columns* $1, \ldots, k$ *and* $n + 1, \ldots, n - 1 + k$ *of an arbitrary matrix upon right-multiplication, i.e., all matrices* V *such that* $P_{2k-1} V = P_{2k-1}$. $\mathcal{U}_{2k}$ *consists of all matrices* $U \in G$ *such that* $U P_{2k} = P_{2k}$.

Proof. If $P_{2k-1}V = P_{2k-1}$, then $J^{(2k-1)}V = JP_{2k-1}V = JP_{2k-1} = J^{(2k-1)}$, so $V \in \mathcal{V}_{2k-1}$. Conversely, if $V \in \mathcal{V}_{2k-1}$, i.e., $J^{(2k-1)}V = J^{(2k-1)}$, then $JP_{2k-1}V = JP_{2k-1}$, which implies $P_{2k-1}V = P_{2k-1}$, since J is invertible (in fact, orthogonal). If $U^*P_{2k} = P_{2k}$, then $U^*J^{(2k)} = U^*P_{2k}J = P_{2k}J = J^{(2k)}$, so $U \in \mathcal{U}_{2k}$. Conversely, if $U \in \mathcal{U}_{2k}$, i.e., $U^*J^{(2k)} = J^{(2k)}$, then $U^*P_{2k}J = P_{2k}J$, which implies $U^*P_{2k} = P_{2k}$. □

Lemma 3.2. *For $k = 1, \ldots, n$, $\mathcal{V}_{2k-2} \supset \mathcal{V}_{2k-1}$ and $\mathcal{U}_{2k-1} \supset \mathcal{U}_{2k}$.*

Proof. By the previous lemma, $\mathcal{V}_{2k-1}$ consists of all matrices V such that $P_{2k-1}V = P_{2k-1}$, i.e., all matrices V that fix columns $1, \ldots, k$ and $n+1, \ldots, n-1+k$ of an arbitrary matrix upon right-multiplication. Because the only nonzero columns of $J^{(2k-2)}$ are necessarily these columns, $J^{(2k-2)}V = J^{(2k-2)}$, so $\mathcal{V}_{2k-2} \supset \mathcal{V}_{2k-1}$. On the other hand, $\mathcal{U}_{2k}$ consists of all matrices U such that $U^*P_{2k} = P_{2k}$. Because of the zero-nonzero pattern of $J^{(2k-1)}$, $U^*J^{(2k-1)} = J^{(2k-1)}$, and therefore $\mathcal{U}_{2k-1} \supset \mathcal{U}_{2k}$. □

Lemma 3.3. *Suppose that A is a random matrix whose distribution satisfies*

(1) $AP_i = J^{(i)}$ *if i is odd, or*
(2) $P_iA = J^{(i)}$ *if i is even.*

Then the distribution of A is $\mathcal{U}_i$-invariant if and only if it is $\mathcal{V}_i$-invariant.

Proof. We prove only the case when i is odd. The case when i is even is very similar.

A can be broken into two terms, $A = AP_i + A(I - P_i) = J^{(i)} + A(I - P_i)$. Let $\hat{U}$ and $\hat{V}$ be m-by-$(m-i)$ matrices whose columns form orthonormal bases for the orthogonal complements of the column space and row space, respectively, of $J^{(i)}$. Claim: the following statements are equivalent.

(1) $A(I - P_i) \overset{d}{=} \hat{U}Q\hat{V}^*$, in which Q is a Haar-distributed $(m-i)$-by-$(m-i)$ random matrix.
(2) A is $\mathcal{U}_i$-invariant.
(3) A is $\mathcal{V}_i$-invariant.

The proofs of (1)⇔(2) and (1)⇔(3) are straightforward. □

Lemma 3.4. *The distribution of $Y^{(i)}$ is $\mathcal{U}_i$- and $\mathcal{V}_i$-invariant, for $i = 0, 1, 2, \ldots, 2n$.*

Proof. The proof uses induction. $C(M)$ will denote the column space of any given matrix M, and $R(M)$ will denote the row space. $W^\perp$ will denote the orthogonal complement of any given subspace $W \subset \mathbb{R}^m$.

Base case: $Y^{(0)} = X$ is Haar-distributed by definition, and $\mathcal{U}_0 = \mathcal{V}_0 = G$.

Induction step: Assume that the distribution of $Y^{(i-1)}$ is $\mathcal{U}_{i-1}$-invariant and $\mathcal{V}_{i-1}$-invariant. By Lemma 3.3, it suffices to prove that the distribution of $Y^{(i)}$ is *either* $\mathcal{U}_i$-invariant or $\mathcal{V}_i$-invariant. When $i = 2k-1$ is odd, we show that the distribution is $\mathcal{V}_{2k-1}$-invariant, and when $i = 2k$ is even, we show that the distribution is $\mathcal{U}_{2k}$-invariant.

Suppose that $i = 2k-1$. By induction, the distribution of $Y^{(2k-2)}$ is $\mathcal{V}_{2k-2}$-invariant, and since $\mathcal{V}_{2k-2} \supset \mathcal{V}_{2k-1}$, it is also $\mathcal{V}_{2k-1}$-invariant. Hence, columns $k+1,\dots,n$ and $n+k,\dots,m$ of $Y^{(2k-2)}$ form a random orthonormal basis for $C(Y^{(2k-2)}P_{2k-1})^{\perp}$. Applying the Householder reflectors to transform $Y^{(2k-2)}$ into $Y^{(2k-1)}$ changes the distribution of columns $k+1,\dots,n$ and $n+k,\dots,m$ from a random orthonormal basis for $C(Y^{(2k-2)}P_{2k-1})^{\perp}$ to a random orthonormal basis for $C(Y^{(2k-1)}P_{2k-1})^{\perp}$. Still, the distribution of $Y^{(2k-1)}$ is $\mathcal{V}_{2k-1}$-invariant.

Now suppose that $i = 2k$. By induction, the distribution of $Y^{(2k-1)}$ is $\mathcal{U}_{2k-1}$-invariant, and since $\mathcal{U}_{2k-1} \supset \mathcal{U}_{2k}$, it is also $\mathcal{U}_{2k}$-invariant. Hence, rows $k+1,\dots,n+a$ and $n+a+k+1,\dots,m$ of $Y^{(2k-1)}$ form a random orthonormal basis for $R(P_{2k}Y^{(2k-1)})^{\perp}$. Applying the Householder reflectors to transform $Y^{(2k-1)}$ into $Y^{(2k)}$ changes the distribution of rows $k+1,\dots,n+a$ and $n+a+k+1,\dots,m$ from a random orthonormal basis for $R(P_{2k}Y^{(2k-1)})^{\perp}$ to a random orthonormal basis for $R(P_{2k}Y^{(2k)})^{\perp}$. Still, the distribution of $Y^{(2k)}$ is $\mathcal{U}_{2k}$-invariant. □

Lemma 3.5. *If q is a random vector, uniformly distributed on the real $(c+d-1)$-sphere, then the squared norm of the vector formed from the leading c entries from q has distribution* $\mathrm{beta}(c/2, d/2)$. *If, instead, q is uniformly distributed on the* complex $(c+d-1)$*-sphere, then the squared norm is* $\mathrm{beta}(c,d)$.

Theorem 3.6. *Let n be a positive integer, let a and b be nonnegative integers, and define $m = 2n+a+b$. Suppose that X follows Haar measure on G (either $O(m)$ or $U(m)$), and run the algorithm on X using partition size $(n+a)$-by-n. Then the output Y is a random orthogonal matrix in bidiagonal block form,*

$$Y = \left[\begin{array}{c|cc|c} B_{11} & B_{12} & & \\ & & I_a & \\ \hline B_{21} & B_{22} & & \\ \hline & & & I_b \end{array}\right],$$

distributed in such a way that

$$\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \stackrel{d}{=} J^{\beta}_{a,b}.$$

In words, the distribution of $\left[\begin{smallmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{smallmatrix}\right]$ is the $2n$-by-$2n$ β-Jacobi matrix model with parameters a and b, where $\beta = 1$ if $G = O(m)$, or $\beta = 2$ if $G = U(m)$.

Proof. In light of Lemma 2.4, the only thing left to prove is that the angles $\theta_n,\dots,\theta_1$ and $\phi_{n-1},\dots,\phi_1$ have the claimed distributions.

First assume that $G = O(m)$. Because the distribution of $Y^{(2k-2)}$ is $\mathcal{U}_{2k-2}$-invariant, the unit norm vector $(z^*_{n+1-k}, w^*_{n+1-k})^*$ is uniformly distributed on the real $((n+a-k+1)+(n+b-k+1)-1)$-sphere. Hence,

$$\cos^2\theta_{n+1-k} = \|z_{n+1-k}\|^2 \sim \mathrm{beta}\left(\tfrac{1}{2}(a+n+1-k), \tfrac{1}{2}(b+n+1-k)\right).$$

Because the distribution of $Y^{(2k-1)}$ is $\mathcal{V}_{2k-1}$-invariant, the unit norm vector $((z'_{n-k})^*, (w'_{n-k})^*)^*$is uniformly distributed on the real $((n-k)+(n+a+b-k+1)-1)$-sphere. Hence,

$$\cos^2\phi_{n-k} = \|z'_{n-k}\|^2 \sim \mathrm{beta}(\tfrac{1}{2}(n-k), \tfrac{1}{2}(a+b+n+1-k)).$$

Furthermore, all of the angles are independent, because the distribution of $(I-P_{2k-2})Y^{(2k-2)}$ is independent from $\theta_n, \dots, \theta_{n+2-k}$ and $\phi_{n-1}, \dots, \phi_{n+1-k}$, and the distribution of $Y^{(2k-1)}(I-P_{2k-1})$ is independent from $\theta_n, \dots, \theta_{n+1-k}$ and $\phi_{n-1}, \dots, \phi_{n+1-k}$.

When $G = U(m)$, the proof is exactly the same, except that (1) complex spheres replace real spheres, and (2) $\beta = 2$ replaces $\beta = 1$ in

$$\cos^2\theta_{n+1-k} \sim \mathrm{beta}\left(\tfrac{\beta}{2}(a+n+1-k), \tfrac{\beta}{2}(b+n+1-k)\right),$$
$$\cos^2\phi_{n-k} \sim \mathrm{beta}\left(\tfrac{\beta}{2}(n-k), \tfrac{\beta}{2}(a+b+1+n-k)\right).$$

□

Combined with Theorem 1.5, this proves the main theorem in the classical cases:

Corollary 3.7. *When $\beta = 1$ or $\beta = 2$ and a and b are integers, the CS values of the β-Jacobi matrix model follow the law of the Jacobi ensemble with the same parameters.*

4. General β matrix models: Beyond real and complex

The main theorem has been proved in the classical real and complex cases. It remains to consider *all* $\beta > 0$ and $a, b > -1$.

Theorem 4.1. *Let β be any positive real number, let n be a positive integer, and let $a, b > -1$. Take the CS decomposition of the $2n$-by-$2n$ β-Jacobi matrix model $J^\beta_{a,b}$, using partition size n-by-n,*

$$J^\beta_{a,b} = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix} \begin{bmatrix} C & S \\ -S & C \end{bmatrix} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^*.$$

The diagonal entries of C, squared, follow the law of the β-Jacobi ensemble with parameters a, b. Also, the first row of V_1, up to sign, is distributed as a vector of i.i.d. χ_β random variables, normalized to unit length.

The proof is at the end of this section.

The β-Jacobi matrix model is a distribution on $2n$-by-$2n$ orthogonal matrices, but the CS values are completely determined by the upper-left n-by-n block. In fact, the CS values are precisely the singular values of this matrix. Their distribution will be obtained by changing variables.

Given $\Theta = (\theta_n, \dots, \theta_1)$ and $\Phi = (\phi_{n-1}, \dots, \phi_1)$, let $c_i = \cos\theta_i$, $s_i = \sin\theta_i$, $c'_i = \cos\phi_i$, and $s'_i = \sin\phi_i$. Also, let $\sigma_1 > \cdots > \sigma_n$ be the singular values of $B_{11}(\Theta, \Phi)$, and, for $i = 1, \dots, n-1$, let v_i be the first entry of the right singular vector of $B_{11}(\Theta, \Phi)$ corresponding to σ_i, constrained to be nonnegative.

Lemma 4.2. *The* $2n-1$ *parameters* $\sigma_1, \ldots, \sigma_n, v_1, \ldots, v_{n-1}$ *defined above uniquely determine a matrix of the form* $\begin{bmatrix} B_{11}(\Theta,\Phi) & B_{12}(\Theta,\Phi) \\ B_{21}(\Theta,\Phi) & B_{22}(\Theta,\Phi) \end{bmatrix}$. *The Jacobian for the change of variables between* $(c_n, \ldots, c_1, c'_{n-1}, \ldots, c'_1)$ *and* $(\sigma_1, \ldots, \sigma_n, v_1, \ldots, v_{n-1})$ *can be expressed as*

$$\prod_{i=1}^{n} \left(c_i^{\beta(i-1)+1} s_i^{\beta(i-1)} dc_i \right) \prod_{i=1}^{n-1} \left((c'_i)^{\beta i-1} (s'_i)^{\beta(i-1)+2} dc'_i \right)$$
$$= \left(\prod_{i<j} (\sigma_i^2 - \sigma_j^2)^\beta \prod_{i=1}^{n} \sigma_i d\sigma_i \right) \left(\prod_{i=1}^{n-1} v_i^{\beta-1} dv_i \right)$$

for any $\beta > 0$.

Proof. $\sigma_1^2, \ldots, \sigma_n^2$ are the eigenvalues, and $v_1, \ldots, v_{n-1}$ are the first $n-1$ entries of the first row of the eigenvector matrix, of a uniquely determined positive definite tridiagonal matrix T. (See Lemma 2.5 of [1].) There is a unique Cholesky factorization $T = B^T B$ with B upper bidiagonal with the sign pattern required by a $B_{11}(\Theta, \Phi)$ matrix. The $2n-1$ angles $\theta_n, \ldots, \theta_1, \phi_{n-1}, \ldots, \phi_1$ are uniquely determined by the entries of B. Once the angles are determined, the $2n$-by-$2n$ matrix in bidiagonal block form is determined. (Note that $c_n, \ldots, c_1, c'_{n-1}, \ldots, c'_1$ may not uniquely determine $\sigma_1, \ldots, \sigma_n, v_1, \ldots, v_{n-1}$ if there is a repeated singular value, but this occurs on a set of measure zero, so the change of variables is still sensible.)

Denote the entries of $B_{11}(\Theta, \Phi)$ by

$$B_{11}(\Theta, \Phi) = \begin{bmatrix} x_n & -y_{n-1} & & & \\ & x_{n-1} & -y_{n-2} & & \\ & & x_{n-2} & \ddots & \\ & & & \ddots & -y_1 \\ & & & & x_1 \end{bmatrix}.$$

It follows from Lemmas 2.7, 2.9, and 2.11 of [1] that

$$\left(\prod_{i<j} (\sigma_i^2 - \sigma_j^2)^\beta \prod_{i=1}^{n} \sigma_i \prod_{i=1}^{n} d\sigma_i \right) \left(\prod_{i=1}^{n} v_i^{\beta-1} \prod_{i=1}^{n-1} dv_i \right)$$
$$= \left(\prod_{i=1}^{n} x_i^{\beta(i-1)+1} \prod_{i=1}^{n} dx_i \right) \left(\prod_{i=1}^{n-1} y_i^{\beta i-1} \prod_{i=1}^{n-1} dy_i \right).$$

The Jacobian matrix $\frac{\partial(x_n,\ldots,x_1,y_{n-1},\ldots,y_1)}{\partial(c_n,\ldots,c_1,c'_n,\ldots,c'_1)}$ is

$$\left[\begin{array}{c|c|c}
\multicolumn{2}{c|}{\operatorname{diag}(1, s'_{n-1}, \ldots, s'_1)} & \begin{array}{ccc} 0 & \cdots & 0 \\ \hline \multicolumn{3}{c}{\operatorname{diag}(-c_{n-1}\frac{c'_{n-1}}{s'_{n-1}}, \ldots, -c_1\frac{c'_1}{s'_1})} \end{array} \\ \hline
\operatorname{diag}(\frac{-c_n}{s_n}c'_{n-1}, \ldots, \frac{-c_2}{s_2}c'_1) & \begin{array}{c} 0 \\ \vdots \\ 0 \end{array} & \operatorname{diag}(s_n, \ldots, s_{n-1})
\end{array}\right].$$

Using the Schur complement formula for the determinant [5], we find that the Jacobian is the product of the determinant of the top-left block with the determinant of a lower triangular matrix whose diagonal is given by the bottom-right block. Hence,

$$\prod_{i=1}^{n} dx_i \prod_{i=1}^{n-1} dy_i = \prod_{i=2}^{n} s_i dc_i \prod_{i=1}^{n-1} s'_i dc'_i.$$

Changing variables again using this Jacobian and evaluating x_i, y_i gives

$$\left(\prod_{i=1}^{n} x_i^{\beta(i-1)+1} \prod_{i=1}^{n} dx_i\right)\left(\prod_{i=1}^{n-1} y_i^{\beta i-1} \prod_{i=1}^{n-1} dy_i\right)$$
$$= \prod_{i=1}^{n} c_i^{\beta(i-1)+1} \prod_{i=1}^{n} s_i^{\beta(i-1)} \prod_{i=1}^{n-1} (c'_i)^{\beta i-1} \prod_{i=1}^{n-1} (s'_i)^{\beta(i-1)+2} \prod_{i=1}^{n} dc_i \prod_{i=1}^{n-1} dc'_i.$$

□

Proof of Theorem 4.1. The differential for the β-Jacobi matrix model is

$$dJ = \text{const} \times \prod_{i=1}^{n}\left(c_i^{\beta(a+i)-1} s_i^{\beta(b+i)-2} dc_i\right) \prod_{i=1}^{n-1}\left((c'_i)^{\beta i-1}(s'_i)^{\beta(a+b+1+i)-2} dc'_i\right).$$

Changing variables using the lemma gives

$$dJ = \text{const} \times \prod_{i=1}^{n} c_i^{\beta(a+1)-2} \prod_{i=1}^{n} s_i^{\beta(b+1)-2} \prod_{i=1}^{n-1} (s'_i)^{\beta(a+b+2)-4} \times$$
$$\times \left(\prod_{i<j} (\sigma_i^2 - \sigma_j^2)^\beta \prod_{i=1}^{n} \sigma_i d\sigma_i\right)\left(\prod_{i=1}^{n-1} v_i^{\beta-1} dv_i\right)$$
$$= \text{const} \times \prod_{i=1}^{n} x_i^{\beta(a+1)-2} \prod_{i=1}^{n} w_i^{\beta(b+1)-2} \left(\prod_{i<j} (\sigma_i^2 - \sigma_j^2)^\beta \prod_{i=1}^{n} \sigma_i d\sigma_i\right)\left(\prod_{i=1}^{n-1} v_i^{\beta-1} dv_i\right),$$

in which $x_n, \ldots, x_1$ are the diagonal entries of B_{11} and $w_n, \ldots, w_1$ are the diagonal entries of B_{21}. Now notice that $\prod_{i=1}^{n} x_i$ is the determinant of B_{11}, that $\prod_{i=1}^{n} w_i$ is the determinant of B_{21}, and that $B_{21}^T B_{21} = I - B_{11}^T B_{11}$, so that

$$\begin{aligned} dJ &= \text{const} \times \det(B_{11}^T B_{11})^{\frac{\beta}{2}(a+1)-1} \det(I - B_{11}^T B_{11})^{\frac{\beta}{2}(b+1)-1} \times \\ &\quad \times \left(\prod_{i<j} (\sigma_i^2 - \sigma_j^2)^\beta \prod_{i=1}^{n} \sigma_i d\sigma_i \right) \left(\prod_{i=1}^{n-1} v_i^{\beta-1} dv_i \right) \\ &= \text{const} \times \left(\prod_{i=1}^{n} \lambda_i^{\frac{\beta}{2}(a+1)-1} (1-\lambda_i)^{\frac{\beta}{2}(b+1)-1} \prod_{i<j} (\lambda_i - \lambda_j)^\beta \prod d\lambda_i \right) \left(\prod_{i=1}^{n-1} v_i^{\beta-1} \prod_{i=1}^{n-1} dv_i \right), \end{aligned}$$

in which $\lambda_i = \sigma_i^2$. □

5. Multivariate analysis of variance

Pairs of Gaussian matrices are often important in multivariate analysis of variance (MANOVA). The following corollary may be useful in this context.

Corollary 5.1. *Let n be a positive integer, and let a and b be nonnegative integers. Suppose that N_1 ($(n+a)$-by-n) and N_2 ($(n+b)$-by-n) are independent random matrices, each with i.i.d. real standard Gaussian entries. Partitioning the β-Jacobi matrix model with $\beta = 1$ into n-by-n blocks,*

$$J_{a,b}^1 = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

we find

$$\text{gsvd}(N_1, N_2) \stackrel{d}{=} \text{gsvd}(B_{11}, B_{21}),$$

in which gsvd$(\cdot,\cdot)$ *gives generalized singular values in decreasing order.*

The analogous result holds in the complex case. Specifically, if $\tilde{N}_1$ and $\tilde{N}_2$ have complex Gaussian entries and the β-Jacobi matrix model with $\beta = 2$ is partitioned as

$$J_{a,b}^2 = \begin{bmatrix} \tilde{B}_{11} & \tilde{B}_{12} \\ \tilde{B}_{21} & \tilde{B}_{22} \end{bmatrix},$$

then

$$\text{gsvd}(\tilde{N}_1, \tilde{N}_2) \stackrel{d}{=} \text{gsvd}(\tilde{B}_{11}, \tilde{B}_{21}).$$

Proof. The same argument works in the real and complex cases. Because the columns of $J_{a,b}^\beta$ are orthonormal, the generalized singular values of B_{11}, B_{21} are just the CS values of $J_{a,b}^\beta$ (see [9]), whose squares follow the Jacobi law. The proof follows by Proposition 1.2. □

Appendix A. Matlab code for sampling the β-Jacobi matrix model

Listing 1. Matlab code for sampling the $2n$-by-$2n$ β-Jacobi matrix model.

```
%JACOBIMATRIXMODEL Sample the beta-Jacobi matrix model.
%  J = JACOBIMATRIXMODEL(BETA,N,A,B) is a 2N-by-2N random matrix whose
%  CS values, squared, follow the law of the BETA-Jacobi ensemble with
%  parameters A and B.  (The partition size in the CS decomposition is
%  N-by-N.)
%
%  This function requires the Statistics Toolbox.
%
%  Alan Edelman and Brian Sutton

function J=jacobimatrixmodel(beta,n,a,b)

if beta==inf
  c=sqrt((a+(1:n))./(a+b+2*(1:n)))';
  s=sqrt((b+(1:n))./(a+b+2*(1:n)))';
  cprime=sqrt((1:n-1)./(a+b+1+2*(1:n-1)))';
  sprime=sqrt((a+b+1+(1:n-1))./(a+b+1+2*(1:n-1)))';
else
  csq=betarnd(beta/2*(a+(1:n)),beta/2*(b+(1:n)))';
  c=sqrt(csq)
  s=sqrt(1-csq);
  cprimesq=betarnd(beta/2*(1:n-1),beta/2*(a+b+1+(1:n-1)))';
  cprime=sqrt(cprimesq);
  sprime=sqrt(1-cprimesq);
end

if n==1
  J=[ c s ; -s c ];return;
end

d0=c(end:-1:1).*[1;sprime(end:-1:1)];
d1=-s(end:-1:2).*cprime(end:-1:1);
B11=spdiags([d0 [nan;d1]],[0 1],n,n);

d0=s(end:-1:1).*[sprime(end:-1:1);1];
dm1=c(end-1:-1:1).*cprime(end:-1:1);
B12=spdiags([[dm1;nan] d0],[-1 0],n,n);

d0=-s(end:-1:1).*[1;sprime(end:-1:1)];
d1=-c(end:-1:2).*cprime(end:-1:1);
B21=spdiags([d0 [nan;d1]],[0 1],n,n);

d0=c(end:-1:1).*[sprime(end:-1:1);1];
dm1=-s(end-1:-1:1).*cprime(end:-1:1);
B22=spdiags([[dm1;nan] d0],[-1 0],n,n);

J=[ B11 B12; B21 B22 ];
```

Listing 2. Matlab code for sampling the n-by-1 β-Jacobi ensemble. The implementation first samples the β-Jacobi matrix model and then computes the CS values, squared. Note that only the top-left section of the matrix is required for computing the CS values, but the entire matrix would be required to compute the left and right CS vectors. (CS vectors are analogous to singular vectors).

```
%JACOBIENSEMBLE Sample the beta-Jacobi ensemble.
%  LAMBDA = JACOBIENSEMBLE(BETA,N,A,B) is a random N-by-1 vector that
%  follows the law of the BETA-Jacobi ensemble with parameters A and B.
%
%  The implementation first samples the BETA-Jacobi matrix model, and
%  then computes the squares of the CS values of the matrix model.
%
%  This function requires the Statistics Toolbox.
%
%  Alan Edelman and Brian Sutton

function lambda=jacobiensemble(beta,n,a,b)

% The CS values of J are the singular values of its top-left block,
% and the squares of the CS values follow the law of the beta-Jacobi
% ensemble.
J=jacobimatrixmodel(beta,n,a,b);
topleftblock=J(1:n,1:n);
lambda=svd(full(topleftblock)).^2;
```

Appendix B. Proof of the algorithm's correctness

Proof of Lemma 2.4. Assertion 1 ($k = 1$): $z_q = Y^{(0)}_{1:p,1}$ and $w_q = -Y^{(0)}_{p+1:m,1}$ by definition. The definition of θ_q and the equation $\|z_q\|^2+\|w_q\|^2 = 1$ (which is true because the first column of the unitary matrix $Y^{(0)}$ has unit norm) guarantees that $\|z_q\| = c_q$ and $\|w_q\| = s_q$.

Assertion 2 ($k = 1$): U_1 is chosen so that $Y^{(1)}_{1:p,1} = \|Y^{(0)}_{1:p,1}\|I_{p,1}$ and $Y^{(1)}_{p+1:m,1} = -\|Y^{(0)}_{p+1:m,1}\|I_{m-p,1}$. By Assertion 1, we have $\|Y^{(0)}_{1:p,1}\| = \|z_q\| = c_q$ and $\|Y^{(0)}_{p+1:m,1}\| = \|w_q\| = s_q$. Hence, the first column of $Y^{(1)}$ is

$$(c_q, 0, \ldots, 0, -s_q, 0, \ldots, 0)^T,$$

so $Y^{(1)}P_1 = J^{(1)}$.

Assertion 3 ($k = 1$): Assume $q > 1$. The case $q = 1$ is left to the reader. The key observation to prove is that $Y^{(1)}_{1,2:m}$, $Y^{(1)}_{p+1,2:m}$, and $(-(z'_{q-1})^*, (w'_{q-1})^*) = s_qY^{(1)}_{1,2:m} + c_qY^{(1)}_{p+1,2:m}$ all have the same direction, with the last having unit norm. By Assertion 2, $Y^{(1)}P_1 = J^{(1)}$, so the first row of $Y^{(1)}$ is $(c_q, Y^{(1)}_{1,2:m})$ and the $(p + 1)$th row is $(-s_q, Y^{(1)}_{p+1,2:m})$. Because each of these rows has unit norm, $\|Y^{(1)}_{1,2:m}\|$ must equal s_q and $\|Y^{(1)}_{p+1,2:m}\|$ must equal c_q. Because the two rows are orthogonal, $\langle Y^{(1)}_{1,2:m}, Y^{(1)}_{p+1,2:m}\rangle$ must equal c_qs_q. Therefore, $\langle Y^{(1)}_{1,2:m}, Y^{(1)}_{p+1,2:m}\rangle = \|Y^{(1)}_{1,2:m}\|\|Y^{(1)}_{p+1,2:m}\|$, i.e., $Y^{(1)}_{1,2:m}$ and $Y^{(1)}_{p+1,2:m}$ have the same direction. It follows that the unit norm vector $s_qY^{(1)}_{1,2:m} + c_qY^{(1)}_{p+1,2:m} = (-(z'_{q-1})^*, (w'_{q-1})^*)$ also has the same direction. Therefore, $\|z'_{q-1}\|^2+\|w'_{q-1}\|^2 =$

1, and from this fact, it is clear that $\|z'_{q-1}\| = \frac{\|Y^{(1)}_{1,2:q}\|}{\|Y^{(1)}_{1,2:m}\|} = \frac{\|Y^{(1)}_{p+1,2:q}\|}{\|Y^{(1)}_{p+1,2:m}\|} = c'_{q-1}$ and $\|w'_{q-1}\| = \frac{\|Y^{(1)}_{1,q+1:m}\|}{\|Y^{(1)}_{1,2:m}\|} = \frac{\|Y^{(1)}_{p+1,q+1:m}\|}{\|Y^{(1)}_{p+1,2:m}\|} = s'_{q-1}$. Finally, $-(z'_{q-1})^*$ and $Y^{(1)}_{1,2:q}$ have the same direction and $\|Y^{(1)}_{1,2:q}\| = \|Y^{(1)}_{1,2:m}\|\|z'_{q-1}\| = s_q\|z'_{q-1}\|$, so $Y^{(1)}_{1,2:q} = -s_q(z'_{q-1})^*$. Analogous arguments show that $Y^{(1)}_{p+1,2:q} = -c_q(z'_{q-1})^*$, $Y^{(1)}_{1,q+1:m} = s_q(w'_{q-1})^*$, and $Y^{(1)}_{p+1,q+1:m} = c_q(w'_{q-1})^*$.

Assertion 4 ($k=1$): V_1 is chosen (1) to leave the first column of Y unchanged, i.e., to ensure $Y^{(2)}P_1 = Y^{(1)}P_1 = J^{(1)}$, and (2) to make $Y^{(2)}_{1,2:q} = -\|Y^{(1)}_{1,2:q}\|I_{1,q-1}$, $Y^{(2)}_{p+1,2:q} = -\|Y^{(1)}_{p+1,2:q}\|I_{1,q-1}$, $Y^{(2)}_{1,q+1:m} = \|Y^{(1)}_{1,q+1:m}\|I_{1,m-q}$, and $Y^{(2)}_{p+1,q+1:m} = \|Y^{(1)}_{p+1,q+1:m}\|I_{1,m-q}$. By Assertion 3, we have $\|Y^{(1)}_{1,2:q}\| = s_q\|z'_{q-1}\| = s_qc'_{q-1}$, $\|Y^{(1)}_{p+1,2:q}\| = c_q\|z'_{q-1}\| = c_qc'_{q-1}$, $\|Y^{(1)}_{1,q+1:m}\| = s_q\|w'_{q-1}\| = s_qs'_{q-1}$, and $\|Y^{(1)}_{p+1,q+1:m}\| = c_q\|w'_{q-1}\| = c_qs'_{q-1}$. Hence, rows 1 and $p+1$ of $Y^{(2)}$ are

$$(c_q, -s_qc'_{q-1}, 0, \ldots, 0, s_qs'_{q-1}, 0, 0, \ldots, 0)$$

and

$$(-s_q, -c_qc'_{q-1}, 0, \ldots, 0, c_qs'_{q-1}, 0, 0, \ldots, 0),$$

respectively, so $P_2Y^{(2)} = J^{(2)}$.

Assertion 1 ($k = 2, \ldots, q$): The key observation to prove is that $Y^{(2k-2)}_{[k:p\ p+k:m],k}$, $Y^{(2k-2)}_{[k:p\ p+k:m],q-1+k}$, and $(z_r^*, -w_r^*)^* = s'_rY^{(2k-2)}_{[k:p\ p+k:m],k} + c'_rY^{(2k-2)}_{[k:p\ p+k:m],q-1+k}$ all have the same direction, with the last having unit norm. By Assertion 4, $P_{2k-2}Y^{(2k-2)} = J^{(2k-2)}$. Because $Y^{(2k-2)}$ is unitary, columns k and $q-1+k$ have unit norm, which guarantees that $\|Y^{(2k-2)}_{[k:p\ p+k:m],k}\|$ equals s'_r and $\|Y^{(2k-2)}_{[k:p\ p+k:m],q-1+k}\|$ equals c'_r. Because the two columns are orthogonal, $\langle Y^{(2k-2)}_{[k:p\ p+k:m],k}, Y^{(2k-2)}_{[k+p\ p+k:m],q-1+k}\rangle$ must equal $c'_rs'_r$. Together, these properties guarantee that $Y^{(2k-2)}_{[k:p\ p+k:m],k}$ and $Y^{(2k-2)}_{[k:p\ p+k:m],q-1+k}$ have the same direction. It follows that the unit norm vector $s'_rY^{(2k-2)}_{[k:p\ p+k:m],k} + c'_rY^{(2k-2)}_{[k:p\ p+k:m],q-1+k} = (z_r^*, -w_r^*)^*$ also has the same direction. Therefore, $\|z_r\|^2 + \|w_r\|^2 = 1$, and from this fact, it is clear that $\|z_r\| = \frac{\|Y^{(2k-2)}_{k:p,k}\|}{\|Y^{(2k-2)}_{[k:p\ p+k:m],k}\|} = \frac{\|Y^{(2k-2)}_{k:p,q-1+k}\|}{\|Y^{(2k-2)}_{[k:p\ p+k:m],q-1+k}\|} = c_r$ and $\|w_r\| = \frac{\|Y^{(2k-2)}_{p+k:m,k}\|}{\|Y^{(2k-2)}_{[k:p\ p+k:m],k}\|} = \frac{\|Y^{(2k-2)}_{p+k:m,q-1+k}\|}{\|Y^{(2k-2)}_{[k:p\ p+k:m],q-1+k}\|} = s_r$. Finally, z_r and $Y^{(2k-2)}_{k:p,k}$ have the same direction and $\|Y^{(2k-2)}_{k:p,k}\| = \|Y^{(2k-2)}_{[k:p\ p+k:m],k}\|\|z_r\| = s'_r\|z_r\|$, so $Y^{(2k-2)}_{k:p,k} = s'_rz_r$. Analogous arguments prove $Y^{(2k-2}_{k:p,q-1+k} = c'_rz_r$, $Y^{(2k-2)}_{p+k:m,k} = -s'_rw_r$, and $Y^{(2k-2)}_{p+k:m,q-1+k} = -c'_rw_r$.

Assertion 2 ($k = 2, \dots, q$):U_k is chosen (1) to leave rows $1, \dots, k-1$ and $p+1, \dots, p-1+k$ of Y unchanged, i.e., to ensure $P_{2k}Y^{(2k-1)} = P_{2k}Y^{(2k-2)} = J^{(2k-2)}$, and (2) to make $Y^{(2k-1)}_{k:p,k} = \|Y^{(2k-2)}_{k:p,k}\| I_{p-k+1,1}$, $Y^{(2k-1)}_{k:p,q-1+k} = \|Y^{(2k-2)}_{k:p,q-1+k}\| I_{p-k+1,1}$, $Y^{(2k-1)}_{p+k:m,k} = \|Y^{(2k-2)}_{p+k:m,k}\| I_{m-(p+k)+1,1}$, and $Y^{(2k-1)}_{p+k:m,q-1+k} = \|Y^{(2k-2)}_{p+k:m,q-1+k}\| I_{m-(p+k)+1,1}$. By Assertion 1, we have $\|Y^{(2k-2)}_{k:p,k}\| = c_r s'_r$, $\|Y^{(2k-2)}_{k:p,q-1+k}\| = c_r c'_r$, $\|Y^{(2k-2)}_{p+k:m,k}\| = -s'_r s_r$, and $\|Y^{(2k-2)}_{p+k:m,q-1+k}\| = -c'_r s_r$. Hence, columns k and $q-1+k$ of $Y^{(2k-1)}$ are

$$(0, \dots, 0, -s_{r+1}c'_r, c_r s'_r, 0, \dots, 0, -c_{r+1}c'_r, -s_r s'_r, 0, \dots, 0)^T$$

and

$$(0, \dots, 0, s_{r+1}s'_r, c_r c'_r, 0, \dots, 0, c_{r+1}s'_r, -s_r c'_r, 0, 0, \dots, 0)^T,$$

respectively, so $Y^{(2k-1)}P_{2k-1} = J^{(2k-1)}$.

Assertions 3 and 4 ($k = 2, \dots, q$): The proofs of Assertions 3 and 4 for $k = 2, \dots, q-1$ are nearly identical to the proofs for $k = 1$, and are left to the reader. The case $k = q$ requires somewhat more care, but this case is also left to the reader.

Assertion 5. Because Assertion 4 holds for $k = q$, we know $P_{2q}Y^{(2q)} = J^{(2q)} = P_{2q}J$. In fact, $Y^{(2q)}$ has the following form:

$$Y^{(2q)} = \left[\begin{array}{c|cccc|ccc}
B_{11}(\Theta,\Phi) & B_{12}(\Theta,\Phi) & & & & & & \\
 & & * & \cdots & * & * & \cdots & * \\
 & & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
 & & * & \cdots & * & * & \cdots & * \\
\hline
B_{21}(\Theta,\Phi) & B_{22}(\Theta,\Phi) & & & & & & \\
\hline
 & & * & \cdots & * & * & \cdots & * \\
 & & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
 & & * & \cdots & * & * & \cdots & *
\end{array}\right].$$

(Rows 1 through q and $p+1$ through $p+q$ are as indicated because $P_{2q}Y^{(2q)} = P_{2q}J$. The remaining entries in columns 1 through $2q$ must equal zero in order for those columns to have unit norm, since the first $2q$ columns of $P_{2q}J$ are already orthonormal.) Because $Y^{(2q)}$ is unitary, the submatrix that has not yet been considered, $Y^{(2q)}_{[q+1:p\ p+q+1:m],2q:m}$, must be unitary itself. Multiplying by its Hermitian adjoint—the last step in the algorithm—produces the desired sparsity pattern, proving Assertion 5. □

Acknowledgement. The authors are grateful to Ioana Dumitriu for her inspiration, to Raj Rao Nadakuditi for valuable discussions, and to C. C. Paige for a very careful reading. The authors also thank an anonymous referee for reference [14].

References

[1] Ioana Dumitriu and Alan Edelman. Matrix models for beta ensembles. *J. Math. Phys.*, 43(11):5830–5847, 2002.

[2] Freeman J. Dyson. Statistical theory of the energy levels of complex systems. I. *J. Mathematical Phys.*, 3:140–156, 1962.

[3] Freeman J. Dyson. The threefold way. Algebraic structure of symmetry groups and ensembles in quantum mechanics. *J. Mathematical Phys.*, 3:1199–1215, 1962.

[4] Peter J. Forrester. *Log-Gases and Random Matrices*. In preparation.

[5] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.

[6] Rowan Killip and Irina Nenciu. Matrix models for circular ensembles. *Int. Math. Res. Not.*, (50):2665–2701, 2004.

[7] Ross A. Lippert. A matrix model for the β-Jacobi ensemble. *J. Math. Phys.*, 44(10):4807–4816, 2003.

[8] Robb J. Muirhead. *Aspects of multivariate statistical theory*. John Wiley & Sons Inc., New York, 1982. Wiley Series in Probability and Mathematical Statistics.

[9] C. C. Paige and M. A. Saunders. Towards a generalized singular value decomposition. *SIAM J. Numer. Anal.*, 18(3):398–405, 1981.

[10] C. C. Paige and M. Wei. History and generality of the CS decomposition. *Linear Algebra Appl.*, 208/209:303–326, 1994.

[11] Brian D. Sutton. *The stochastic operator approach to random matrix theory*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA 02139, June 2005.

[12] Charles F. Van Loan. Generalizing the singular value decomposition. *SIAM J. Numer. Anal.*, 13(1):76–83, 1976.

[13] Ming Chen Wang and G. E. Uhlenbeck. On the theory of the Brownian motion. II. *Rev. Modern Phys.*, 17:323–342, 1945.

[14] David S. Watkins. Some perspectives on the eigenvalue problem. *SIAM Rev.*, 35(3):430–471, 1993.

Alan Edelman, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139

E-mail address: `edelman@math.mit.edu`

Brian D. Sutton, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139

Current address: Department of Mathematics, Randolph-Macon College, Ashland, VA 23005

E-mail address: `bsutton@rmc.edu`