ORIGINAL RESEARCH

On the Maximum of Entropy for Orthogonal Matrices

K. T. Arasu[1] · Manil T. Mohan[2] · Akhilesh Pathak[3] · Ramya Ramachandran Janaki[4]

Received: 20 April 2023 / Accepted: 19 September 2025 / Published online: 6 October 2025

Abstract
Local and global bounds for the Shannon entropy of real orthogonal matrices have been established for various matrix orders. It has been shown that whenever a Hadamard matrix of order n exists, the entropy reaches its global maximum value of $n \ln n$. For other orders, where a Hadamard matrix does not exist, local maximal bounds are derived using orthogonal matrices associated with finite projective planes, biplanes, and triplanes. In the setting of complex inverse orthogonal matrices, explicit constructions are provided that attain this global maximum bound. Furthermore, the Rényi and Tsallis entropies for real orthogonal matrices are introduced, and both their local and global optimal bounds are determined for different values of n.

Keywords Shannon Entropy · Orthogonal Matrices · Optimization · Hadamard Matrices · Rényi Entropy · Tsallis Entropy · Kullback-Leibler Divergence

Mathematics Subject Classification 94A17 · 65K05 · 15B10 · 15B34

1 Introduction

The concept of *entropy* in Information Theory is defined as a quantitative measure of uncertainty associated with a random variable (or a probability distribution). When the uncertainty in the outcomes of an event decreases, we interpret this as gaining *information*. The notion of entropy was first introduced by Claude E. Shannon in his seminal work [24]. In signal processing and communication, one often encounters situations where probabilities must be specified despite limited information, a challenge deeply rooted in classical probability theory. Intuitively, one might expect that the greatest uncertainty arises when all possible outcomes are equally likely. This intuition was formalized by mathematicians such as Jacob Bernoulli and Pierre-Simon Laplace through the so-called "principle of indifference" (cf. [15]). Later, George Boole, John Venn, and others refined this idea into what became known as the *principle of insufficient reason* (see [22]). In the twentieth century, Information Theory advanced these classical notions by providing a precise and unambiguous framework for quantifying uncertainty. This measure, associated with a discrete probabilistic random variable, was formally named *entropy*.

Communicated by B V Rajarama Bhat

 https://doi.org/10.1007/s13226-025-00873-8

The characterization of Shannon's entropy can be described as follows. Let $\mathbb{X} = (\mathrm{X}, p)$ be a discrete probability space, where X is a set of discrete random variables taking finitely many possible values $x_1, x_2, \ldots, x_n$ with associated probabilities $p_1, p_2, \ldots, p_n$, satisfying $\sum_{i=1}^{n} p_i = 1$. The corresponding measure of uncertainty (the *Shannon entropy*) of the distribution is given by (with the convention $0 \ln 0 = 0$):

$$\mathscr{H}(\mathbb{X}) = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i}.$$

It is immediate that $\mathscr{H}(\mathbb{X}) \geq 0$. The minimum entropy value, 0, occurs when one outcome is certain and all others are impossible, that is,

$$p_i = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \quad \text{for some } j.$$

By Jensen's inequality and the convexity of the function $-\log_2$, it follows that

$$\mathscr{H}(\mathbb{X}) \leq \log_2 n.$$

Moreover, the strict convexity of $-\log_2$ ensures that the maximum entropy $\log_2 n$ is attained precisely when the distribution is uniform, that is,

$$p_i = \frac{1}{n}, \quad \text{for all } i = 1, 2, \ldots, n.$$

The following definition of Shannon entropy for an orthogonal matrix is given in [11].

Definition 1.1 *(Shannon entropy of a real orthogonal matrix)* Let M_n be an $n \times n$ orthogonal matrix with real entries a_{ij} in the $(i, j)^{\text{th}}$ position. The *Shannon entropy* of M_n is defined by

$$\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_n] \;=\; \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2 \, \ln\left(\frac{1}{a_{ij}^2}\right). \tag{1.1}$$

As usual in the definition of entropy, we adopt the conventions $0 \ln 0 = 0$ and $0 \times \infty = 0$, when needed. Clearly, $\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_n] \geq 0$ for every real orthogonal matrix M_n, and equality holds precisely when

$$a_{ij}^2 = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

that is, when M_n is a permutation of the identity matrix. Thus, 0 is the absolute minimum of the Shannon entropy for real orthogonal matrices.

In this paper, we investigate both local and global bounds on the Shannon entropy for various values of n (see also [20]). A key feature is the connection between the optimal entropy bound for an $n \times n$ orthogonal matrix and the existence of Hadamard matrices. In fact, we show that the global maximum entropy is attained precisely by Hadamard matrices (see Theorem 5.1). For orders where Hadamard matrices do not exist, we establish local maximal bounds. Section 2 provides a detailed introduction to Hadamard matrices and their properties. We formulate the problem of maximizing entropy as a constrained optimization problem in Section 3, derive necessary and sufficient conditions for the existence of local maxima, and present a globally convergent monotone curvilinear search algorithm. The numerical results of [11], which identified locally optimal entropy-maximizing matrices of orders 3 and 5, are here confirmed analytically (see Section 4). Moreover, we construct various local and global optimal orthogonal matrices of different orders that maximize Shannon entropy (see Section 5). To establish the universal entropy bound, we apply Jensen's inequality, while for the sharper bound in the case $n = 3$, we employ

 https://doi.org/10.1007/s13226-025-00873-8

the characterization of Stewart [3, 26]. Finally, in Section 6, we extend the analysis to Rényi and Tsallis entropies for real orthogonal matrices, and obtain their respective local and global optimal bounds. Connections between these optimization problems and conference matrices as well as weighing matrices have also been established.

A more general yet compact formulation of the problems considered in this work can be found in the literature (see, e.g., [1]). Furthermore, [2] explores the relationship between the entropy of orthogonal matrices and the minimum distance of orthostochastic matrices from the uniform van der Waerden matrices. However, in [1, 2], orthogonal matrices arising from finite projective planes, biplanes, and triplanes were not considered. The main contributions of this work can be summarized as follows:

(i) We establish the optimal bound on the Shannon entropy for a general $n \times n$ real orthogonal matrix and construct the corresponding optimal matrices when n is a multiple of 4 (see Section 5).
(ii) We provide explicit constructions of real orthogonal matrices of orders 3, 5, and 6 that maximize Shannon entropy (globally for $n = 3$, and locally for $n = 5, 6$; see Section 5.1). For higher small orders up to $n = 15$, we obtain local maxima of Shannon entropy using orthogonal matrices associated with finite projective planes, biplanes, and triplanes.
(iii) For complex inverse orthogonal matrices, we construct entropy-maximizing matrices for all orders (see Section 5.2).
(iv) We introduce the Rényi entropy, Kullback-Leibler divergence, Rényi divergence, and Tsallis entropy for real orthogonal matrices, and derive maximal bounds for both Rényi and Tsallis entropy (see Section 6). In addition, we classify different families of orthogonal matrices that correspond to local maxima, local minima, and saddle points of the entropy maximization problem.

Optimization problems of this type arise in a wide range of applications, including quantum communication, polynomial and combinatorial optimization, eigenvalue problems, sparse PCA, p-harmonic flows, one-bit compressive sensing, and matrix rank minimization, among others. These problems are particularly challenging because the constraints are not only non-convex but also computationally expensive to maintain during iterative methods. Moreover, the presence of orthogonality constraints often leads to multiple local maximizers, and in fact, several special cases of these problems are known to be NP-hard (see [30]). The work [31] further investigates entropy of large matrices and demonstrates its application in quantum optics, specifically in studying photon entanglement, thereby highlighting the relevance of our framework to quantum communication.

2 Preliminaries

In this section, we recall some fundamental definitions and preliminary results that will be used throughout the paper. We begin with the definition and basic properties of real Hadamard matrices.

Definition 2.1 *(Real Hadamard matrix)* A *real Hadamard matrix* H_n is an $n \times n$ matrix with entries in $\{+1, -1\}$ such that

$$\mathrm{H}_n \mathrm{H}_n^\top = n\mathrm{I}_n,$$

where $\mathrm{H}_n^\top$ denotes the transpose of H_n and I_n is the $n \times n$ identity matrix.

Lemma 2.2 ([19], [27, Theorem 4.4], [29, Lemma 1.1.4]) *If a real Hadamard matrix of order n exists, then either* $n = 1, 2$ *or* $n \equiv 0 \pmod 4$.

Conjecture 2.3 (Hadamard conjecture, [19]) *If* $n \equiv 0 \pmod 4$, *then there exists a real Hadamard matrix of order* n.

Now we introduce the definition and basic properties of real conference matrices.

Definition 2.4 *(Real conference matrix)* A *real conference matrix* of order $n > 1$ is an $n \times n$ matrix C_n with diagonal entries equal to 0 and off-diagonal entries in $\{\pm 1\}$ such that

$$\mathrm{C}_n \mathrm{C}_n^\top = (n-1)\mathrm{I}_n.$$

A conference matrix $\mathrm{C}_n = (c_{ij})_{i,j=1}^n$ is called *symmetric* if $c_{ij} = c_{ji}$ for all $1 \le i, j \le n$, and *skew-symmetric* if $c_{ij} = -c_{ji}$ for all $1 \le i, j \le n$.

Lemma 2.5 ([10, Corollary 2.2]) *Any real conference matrix of order $n > 2$ is equivalent, under multiplication of rows and columns by -1, to a symmetric conference matrix when $n \equiv 2 \pmod 4$, and to a skew-symmetric conference matrix when $n \equiv 0 \pmod 4$.*

Let us now provide the definition and properties of real weighing matrices.

Definition 2.6 *(Real weighing matrix)* A *real weighing matrix* $\mathrm{W}_{n,k}$ is an $n \times n$ matrix with entries in $\{0, \pm 1\}$ such that each row and each column contains exactly k nonzero entries, and the inner product of any two distinct rows is zero. Equivalently, $\mathrm{W}_{n,k}$ satisfies

$$\mathrm{W}_{n,k}\mathrm{W}_{n,k}^\top = k\mathrm{I}_n.$$

The integer k is called the *weight* of $\mathrm{W}_{n,k}$.

The following existence results are known for weighing matrices.

Theorem 2.7 ([12, Proposition 23])

(1) If n is odd, then a $\mathrm{W}_{n,k}$ exists only if

- *(i) k is a perfect square, and*
- *(ii) $(n-k)^2 - (n-k) + 1 \ge n$.*

(2) If $n \equiv 2 \pmod 4$, then for a $\mathrm{W}_{n,k}$ to exist,

- *(i) $k \le n-1$, and*
- *(ii) k is expressible as the sum of two integer squares.*

(3) A skew-symmetric $\mathrm{W}_{n,k}$ can exist only if k is the sum of three integer squares.

In addition, the following conjectures are widely believed to hold.

Conjecture 2.8

(1) If $n \equiv 0 \pmod 4$, then a $\mathrm{W}_{n,k}$ exists for all $1 \le k \le n$.
(2) If $n \equiv 4 \pmod 8$, then a skew-symmetric $\mathrm{W}_{n,k}$ exists if and only if

$$k \le n-1, \quad k = a^2 + b^2 + c^2 \textit{ for integers } a, b, c,$$

with the exception that $n-2$ must itself be the sum of two integer squares.
(3) If $n \equiv 0 \pmod 8$, then a skew-symmetric $\mathrm{W}_{n,k}$ exists for all $1 \le k \le n-1$.

In order to extend the concept of a plane by requiring that any two distinct lines intersect in exactly one point, the notion of a *projective plane* was introduced.

Definition 2.9 A *projective plane* of order $n \ge 2$ consists of a finite set of points and a collection of lines (each line being a set of points) such that:

(i) Every line contains exactly $n + 1$ points.
(ii) Every point lies on exactly $n + 1$ lines.
(iii) Any two distinct lines intersect in a unique point.
(iv) Any two distinct points are contained in a unique line.

Lemma 2.10 ([14, Theorem 21], [21, Lemma 4]) *In a projective plane of order n, the total number of points is* $n^2 + n + 1$*, and the total number of lines is also* $n^2 + n + 1$.

Definition 2.11 A *biplane* (or biplane geometry) is an incidence structure in which every pair of distinct points lies on exactly two lines, and any two distinct lines intersect in exactly two points.

Biplanes can be viewed as an analogue of finite projective planes, with the roles of points and lines relaxed so that pairs determine two incidences instead of one. A biplane of order n has $1 + \frac{(n+2)(n+1)}{2}$ points and the same number of lines, with each line containing exactly $n + 2$ points.

Projective planes and biplanes can also be represented using their *incidence matrices*. In this representation, the rows correspond to lines, while the columns correspond to points, resulting in a square matrix of size $(n^2 + n + 1) \times (n^2 + n + 1)$. Each entry indicates whether a given point lies on a given line. Specifically, the entries of the incidence matrix $\mathrm{A}_{n^2+n+1} = (a_{ij})_{i,j=1}^{n^2+n+1}$ are defined by

$$a_{ij} = \begin{cases} 1 & \text{if line } i \text{ contains point } j, \\ 0 & \text{otheorwise.} \end{cases}$$

The incidence matrix of a finite projective plane satisfies the following properties:

(i) Each row contains exactly $n + 1$ ones.
(ii) Each column contains exactly $n + 1$ ones.
(iii) Any two distinct rows share exactly one column in which both entries are 1.
(iv) Any two distinct columns share exactly one row in which both entries are 1.

Thus, the existence of such a matrix of size $(n^2 + n + 1) \times (n^2 + n + 1)$ is equivalent to the existence of a projective plane of order n.

Example 2.12 The matrix

$$\mathrm{A}_7 = \begin{pmatrix} 1&1&1&0&0&0&0 \\ 1&0&0&1&1&0&0 \\ 1&0&0&0&0&1&1 \\ 0&1&0&1&0&1&0 \\ 0&1&0&0&1&0&1 \\ 0&0&1&1&0&0&1 \\ 0&0&1&0&1&1&0 \end{pmatrix}$$

is the incidence matrix of the projective plane of order 2, also known as the *Fano plane*.

Remark 2.13 ([17, Section 2.3]) A λ-*plane* is an incidence structure in which every set of λ points lies on exactly λ lines, and dually, any λ lines intersect in exactly λ points. A λ-plane of order n contains $1 + \frac{(n+\lambda)(n+\lambda-1)}{\lambda}$ points and the same number of lines, with each line containing exactly $n + \lambda$ points. It has been conjectured that only finitely many λ-planes exist for $\lambda > 1$. For a detailed discussion of the connection between λ-planes and symmetric designs, see [8]. Moreover, whenever a λ-plane exists, one can associate with it a corresponding orthogonal matrix whose entries are given by

$$x = \frac{n\lambda - n(n + \lambda - 1)\sqrt{n}}{n(n^2 + (2\lambda - 1)n + \lambda^2)} \quad \text{and} \quad y = \frac{\lambda(n + (n + \lambda)\sqrt{n})}{n(n^2 + (2\lambda - 1)n + \lambda^2)}. \tag{2.1}$$

For example, the orthogonal matrix corresponding to the Fano plane is given by

$$\mathcal{I}_7 = \begin{pmatrix} x & x & y & y & y & x & y \\ y & x & x & y & y & y & x \\ x & y & x & x & y & y & y \\ y & x & y & x & x & y & y \\ y & y & x & y & x & x & y \\ y & y & y & x & y & x & x \\ x & y & y & y & x & y & x \end{pmatrix}, \quad \text{where } x = \frac{1-2\sqrt{2}}{7} \text{ and } y = \frac{2+3\sqrt{2}}{14}. \tag{2.2}$$

and the orthogonal matrix corresponding finite projective plane of order 3 is given by

$$\mathcal{I}_{13} = \begin{pmatrix} x & x & x & x & y & y & y & y & y & y & y & y & y \\ x & y & y & y & x & x & x & y & y & y & y & y & y \\ x & y & y & y & y & y & y & x & x & x & y & y & y \\ x & y & y & y & y & y & y & y & y & y & x & x & x \\ y & x & y & y & y & y & x & y & x & y & y & y & x \\ y & x & y & y & y & x & y & y & y & x & x & y & y \\ y & x & y & y & x & y & y & x & y & y & y & x & y \\ y & y & x & y & y & y & x & x & y & y & x & y & y \\ y & y & x & y & x & y & y & y & y & x & y & y & x \\ y & y & x & y & y & x & y & y & x & y & y & x & y \\ y & y & y & x & y & x & y & x & y & y & y & y & x \\ y & y & y & x & y & y & x & y & y & x & y & x & y \\ y & y & y & x & x & y & y & y & x & y & x & y & y \end{pmatrix}, \tag{2.3}$$

where $x = \frac{1-3\sqrt{3}}{13}$ and $y = \frac{3+4\sqrt{3}}{39}$. The following is an example of an orthogonal matrix corresponding to the Paley biplane of order 3.

$$\mathrm{B}_{11} = \begin{pmatrix} y & x & y & x & x & x & y & y & y & x & y \\ y & y & x & y & x & x & x & y & y & y & x \\ x & y & y & x & y & x & x & x & y & y & y \\ y & x & y & y & x & y & x & x & x & y & y \\ y & y & x & y & y & x & y & x & x & x & y \\ y & y & y & x & y & y & x & y & x & x & x \\ x & y & y & y & x & y & y & x & y & x & x \\ x & x & y & y & y & x & y & y & x & y & x \\ x & x & x & y & y & y & x & y & y & x & y \\ y & x & x & x & y & y & y & x & y & y & x \\ x & y & x & x & x & y & y & y & x & y & y \end{pmatrix}, \tag{2.4}$$

where $x = \frac{1-2\sqrt{3}}{11}$ and $y = \frac{3+5\sqrt{3}}{33}$. The following is an example of an orthogonal matrix corresponding to the triplane of order 4.

 https://doi.org/10.1007/s13226-025-00873-8

$$
\mathrm{T}_{15} = \begin{pmatrix}
x & x & x & x & x & x & x & y & y & y & y & y & y & y & y \\
x & x & x & y & y & y & y & x & x & x & x & y & y & y & y \\
x & x & x & y & y & y & y & y & y & y & y & x & x & x & x \\
x & y & y & x & x & y & y & x & x & y & y & x & x & y & y \\
x & y & y & x & x & y & y & y & y & x & x & y & y & x & x \\
x & y & y & y & y & x & x & x & x & y & y & y & y & x & x \\
x & y & y & y & y & x & x & y & y & x & x & x & x & y & y \\
y & x & y & x & y & x & y & x & y & x & y & x & y & x & y \\
y & x & y & x & y & y & x & x & y & y & x & y & x & y & x \\
y & x & y & y & x & x & y & y & x & y & x & x & y & y & x \\
y & x & y & y & x & y & x & y & x & x & y & y & x & x & y \\
y & y & x & x & y & x & y & y & x & y & x & y & x & x & y \\
y & y & x & x & y & y & x & y & x & x & y & x & y & y & x \\
y & y & x & y & x & x & y & x & y & x & y & y & x & y & x \\
y & y & x & y & x & y & x & x & y & y & x & x & y & x & y
\end{pmatrix}, \tag{2.5}
$$

where $x = -\frac{1}{5}$ and $y = \frac{3}{10}$.

We now introduce a different class of matrices, known as complex inverse orthogonal matrices (see [9, 28]).

Definition 2.14 *(Complex inverse orthogonal matrix)* An $n \times n$ *complex inverse orthogonal matrix* M_n is a matrix with nonzero complex entries a_{ij} such that

$$
\mathrm{M}_n \cdot \frac{1}{\mathrm{M}_n} = n\mathrm{I}_n,
$$

where $\frac{1}{\mathrm{M}_n}$ denotes the matrix whose $(i, j)^{\text{th}}$ entry is $1/a_{ij}$, and M_n has unimodular columns (i.e., each column has entries of modulus 1).

Definition 2.15 *(Complex Hadamard matrix)* A *complex Hadamard matrix* H_n is an $n \times n$ matrix with entries of modulus 1 in $\mathbb{C}$ such that

$$
\mathrm{H}_n\mathrm{H}_n^* = n\mathrm{I}_n,
$$

where H_n^* denotes the conjugate transpose of H_n.

Lemma 2.16 *For every integer $n \geq 1$, there exists a complex Hadamard matrix of order n.*

Proof An explicit example is provided by the *Fourier matrix* F_n, whose entries are given by

$$
[\mathrm{F}_n]_{ij} = e^{2\pi \mathbf{i}(i-1)(j-1)/n}, \quad i, j = 1, \ldots, n. \tag{2.6}
$$

It is straightforward to verify that $\mathrm{F}_n\mathrm{F}_n^* = n\mathrm{I}_n$, so F_n is indeed a complex Hadamard matrix. $\square$

Note that if H_n is a complex Hadamard matrix, then so is H_n^*, $\overline{\mathrm{H}}_n$ and $\mathrm{H}_n^\top$, where $\overline{\mathrm{H}}_n$ is the conjugate of H_n. If H_{n_1} and H_{n_2} are complex Hadamard matrices, then so is $\mathrm{H}_{n_1} \otimes \mathrm{H}_{n_2}$. The following are complex Hadamard matrices of orders $n = 1, 2$ and 3, respectively:

$$
\mathrm{F}_1 = [1], \;\; \mathrm{F}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \;\; \mathrm{F}_2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix}, \tag{2.7}
$$

where $\{1, \omega, \omega^2\}$ are the cube roots of unity. The complex Hadamard matrix given in (2.6) is an example of a complex inverse orthogonal matrix, since

$$\begin{aligned}\left[(\mathrm{F}_n)\left(\frac{1}{\mathrm{F}_n}\right)\right]_{ij} &= \sum_{k=1}^{n} (\mathrm{F}_n)_{ik}\left(\frac{1}{\mathrm{F}_n}\right)_{kj} = \sum_{k=1}^{n} \mathbf{e}^{2\pi\mathbf{i}(i-1)(k-1)/n}\mathbf{e}^{-2\pi\mathbf{i}(k-1)(j-1)/n}\\ &= \sum_{k=1}^{n} \mathbf{e}^{2\pi\mathbf{i}(i-j)(k-1)/n} = \begin{cases} n, & \text{if } i = j,\\ 0, & \text{if } i \neq j.\end{cases}\end{aligned}$$

For more details on the complex Hadamard matrices, see [29].

3 The Optimization Problem

In this section, we formulate an optimization problem with orthogonal matrix constraints. For notational convenience, we denote $n \times n$ matrices by symbols such as A, B, G, I, W, X, Y, Z, etc. The entries of a matrix X are denoted by X_{ij}, while $\mathrm{X}^{i,j}$ refers to a general matrix in $\mathbb{R}^{n\times n}$. The general optimization problem over real orthogonal matrices (see [30]) can be stated as

$$\max_{\mathrm{X}\in\mathbb{R}^{n\times n}} \mathcal{F}(\mathrm{X}) \quad \text{subject to} \quad \mathrm{X}^\top\mathrm{X} = \mathrm{I}, \tag{3.1}$$

where $\mathcal{F} : \mathbb{R}^{n\times n} \to \mathbb{R}$ is a differentiable function. The feasible set

$$\mathscr{M}_n := \left\{\mathrm{X} \in \mathbb{R}^{n\times n} : \mathrm{X}^\top\mathrm{X} = \mathrm{I}\right\}$$

is commonly referred to as the *Stiefel manifold*.

Given a differentiable function, $\mathcal{F}(\cdot) : \mathbb{R}^{n\times n} \to \mathbb{R}$, the gradient of $\mathcal{F}$ with respect to X is denoted by $\mathrm{G} := \mathcal{D}\mathcal{F}(\mathrm{X}) = \left(\frac{\partial\mathcal{F}(\mathrm{X})}{\partial\mathrm{X}_{ij}}\right)_{i,j=1}^{n}$. The derivative of $\mathcal{F}$ at X in the direction Z is given by

$$\mathcal{D}\mathcal{F}(\mathrm{X})[\mathrm{Z}] := \lim_{t\to 0}\frac{\mathcal{F}(\mathrm{X}+t\mathrm{Z}) - \mathcal{F}(\mathrm{X})}{t} = \langle\mathcal{D}\mathcal{F}(\mathrm{X}), \mathrm{Z}\rangle\,.$$

Here $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product between two matrices and is defined by

$$\langle \mathrm{A}, \mathrm{B}\rangle := \sum_{j,k=1}^{n} \mathrm{A}_{jk}\mathrm{B}_{jk} = \mathrm{Tr}(\mathrm{A}^\top\mathrm{B}),\ \text{ for all } \mathrm{A}, \mathrm{B} \in \mathbb{R}^{n\times n},$$

where $\mathrm{Tr}(\mathrm{A})$ is the trace of A, i.e., the sum of the diagonal elements of A. We use $\nabla\mathcal{F}$ for gradients in tangent planes. The Frobenius norm of $\mathrm{A} \in \mathbb{R}^{n\times n}$ is defined as $\|\mathrm{A}\|_{\mathrm{F}}^2 := \sum\limits_{i,j=1}^{n} \mathrm{A}_{ij}^2$. Given a feasible point X and the gradient G, we define a skew-symmetric matrix A as either

$$\mathrm{A} := \mathrm{G}\mathrm{X}^\top - \mathrm{X}\mathrm{G}^\top \text{ or } \mathrm{A} := (\mathrm{P_X}\mathrm{G})\mathrm{X}^\top - \mathrm{X}(\mathrm{P_X}\mathrm{G})^\top,\ \text{ where } \mathrm{P_X} := \left(\mathrm{I} - \frac{1}{2}\mathrm{X}\mathrm{X}^\top\right). \tag{3.2}$$

Following [30], we next present the first- and second-order optimality conditions in the form of the following lemmas. Since the constraint $\mathrm{X}^\top\mathrm{X} = \mathrm{I}$ is symmetric, the corresponding Lagrange multiplier Λ is also a symmetric matrix. The *Lagrangian function* associated with the optimization problem (3.1) is given by

$$\mathcal{L}(\mathrm{X}, \Lambda) = \mathcal{F}(\mathrm{X}) - \frac{1}{2}\mathrm{Tr}\Big(\Lambda\big(\mathrm{X}^\top\mathrm{X} - \mathrm{I}\big)\Big). \tag{3.3}$$

Lemma 3.1 (First order necessary conditions, [30, Lemma 1]) *If* X *is a local maximizer of the problem (3.1), then* X *satisfies the following first order necessary conditions:* $\mathcal{D}_{\mathrm{X}}\mathcal{L}(\mathrm{X}, \Lambda) = \mathrm{G} - \mathrm{X}\mathrm{G}^\top\mathrm{X} = \mathbf{0}$ *and* $\mathrm{X}^\top\mathrm{X} = \mathrm{I}$, *with the associated Lagrangian multiplier* $\Lambda = \mathrm{G}^\top\mathrm{X}$. *Define*

$$\nabla\mathcal{F}(\mathrm{X}) := \mathrm{G} - \mathrm{X}\mathrm{G}^\top\mathrm{X}, \ \ \textit{and}\ \mathrm{A} := \mathrm{G}\mathrm{X}^\top - \mathrm{X}\mathrm{G}^\top. \tag{3.4}$$

Then, $\nabla\mathcal{F}(\mathrm{X}) = \mathrm{A}\mathrm{X}$. *Moreover,* $\nabla\mathcal{F}(\mathrm{X}) = \mathbf{0}$, *if and only if* $\mathrm{A} = \mathbf{0}$.

Lemma 3.2 (Second order necessary conditions, [18, Theorem 12.5], [30, Lemma 2]) *Suppose that* $\mathrm{X} \in \mathscr{M}_n$ *is a local maximizer for the problem (3.1). Then* X *satisfies*

$$\mathrm{Tr}\left(\mathrm{Z}^\top\mathcal{D}\left(\mathcal{D}\mathcal{F}(\mathrm{X})\right)[\mathrm{Z}]\right) - \mathrm{Tr}\left(\Lambda\mathrm{Z}^\top\mathrm{Z}\right) \le 0,\ \ \textit{where}\ \Lambda = \mathrm{G}^\top\mathrm{X}, \tag{3.5}$$

for all

$$\mathrm{Z} \in \mathcal{T}_{\mathrm{X}}\mathscr{M} := \left\{\mathrm{Z} \in \mathbb{R}^{n\times n} : \mathrm{X}^\top\mathrm{Z} + \mathrm{Z}^\top\mathrm{X} = \mathbf{0}\right\},$$

which is the tangent space of $\mathscr{M}_n$ *at* X.

Lemma 3.3 (Second-order sufficient conditions, [18, Theorem 12.6], [30, Lemma 2]) *Let* $\mathrm{X} \in \mathscr{M}_n$ *and suppose there exists a symmetric Lagrange multiplier* Λ *such that the first-order optimality conditions are satisfied. If, in addition,*

$$\mathrm{Tr}\Big(\mathrm{Z}^\top\mathcal{D}\big(\mathcal{D}\mathcal{F}(\mathrm{X})\big)[\mathrm{Z}]\Big) - \mathrm{Tr}\big(\Lambda\mathrm{Z}^\top\mathrm{Z}\big) < 0 \tag{3.6}$$

for all $\mathrm{Z} \in \mathcal{T}_{\mathrm{X}}\mathscr{M}_n$, *then* X *is a* strict local maximizer *of the optimization problem* (3.1).

3.1 The optimization problem for the Shannon entropy

The optimization problem for the Shannon entropy of a real orthogonal matrix can be formulated as follows. Let $\mathrm{X} = (a_{ij})_{i,j=1}^n$, and consider

$$\max_{\mathrm{X}\in\mathbb{R}^{n\times n}} \sum_{i=1}^n\sum_{j=1}^n a_{ij}^2 \ln\left(\frac{1}{a_{ij}^2}\right), \quad \text{subject to} \quad \sum_{k=1}^n a_{ki}a_{kj} = \begin{cases}1, & i=j,\\ 0, & i\neq j.\end{cases} \tag{3.7}$$

Remark 3.4 The function $f(x) = -x^2\ln(x^2)$ is continuous and differentiable for all $x \in \mathbb{R}$, achieving its global maximum at $x = \pm e^{-1/2}$. Moreover, $f(0) = 0$ is its global minimum.

The derivative matrix G and the matrix A are given by

$$\mathrm{G} := \left(\frac{\partial\mathcal{F}(\mathrm{X})}{\partial\mathrm{X}_{ij}}\right)_{i,j=1}^n = \left(2a_{ij}\ln\left(\frac{1}{a_{ij}^2}\right) - 2a_{ij}\right)_{i,j=1}^n = 2(-\mathrm{X}\circ\ln(\mathrm{X}) - \mathrm{X}), \tag{3.8}$$

and

$$\mathrm{A} := \mathrm{G}\mathrm{X}^\top - \mathrm{X}\mathrm{G}^\top = 2\left(\sum_{k=1}^{n} a_{ik}a_{jk}\ln\left(\frac{a_{jk}^2}{a_{ik}^2}\right)\right)_{i,j=1}^{n}. \tag{3.9}$$

In (3.8), the symbol $\circ$ denotes the *Hadamard product* meaning $(\mathrm{A}\circ\mathrm{B})_{ij} = \mathrm{A}_{ij}\mathrm{B}_{ij}$. Suppose there exists a matrix $\mathrm{X} = (a_{ij})_{i,j=1}^{n} \in \mathrm{SO}(n,\mathbb{R})$ that maximizes the function in (3.7). Then, the first-order optimality condition from Lemma 3.1 reduces to

$$\mathrm{G}\mathrm{X}^\top = \mathrm{X}\mathrm{G}^\top,$$

which implies, for all i and j, that

$$\sum_{k=1}^{n} a_{ik}a_{jk}\ln\left(\frac{a_{jk}^2}{a_{ik}^2}\right) = 0. \tag{3.10}$$

Also for the second order necessary and sufficient conditions given in Lemmas 3.2 and 3.3, we first calculate $\mathcal{D}\left(\mathcal{DF}(\mathrm{X})\right)[\mathrm{Z}]$ as

$$\begin{aligned}\mathcal{D}\left(\mathcal{DF}(\mathrm{X})\right)[\mathrm{Z}] = \mathrm{Tr}\left(\mathrm{Z}^\top\mathcal{DF}(\mathrm{X})\right) &= \sum_{i,j=1}^{n} z_{ij}\left(\frac{\partial\left(\mathcal{DF}(\mathrm{X})\right)_{ij}}{\partial \mathrm{X}_{ij}}\right)_{i,j=1}^{n} \\ &= 2\left(z_{ij}\left(\ln\left(\frac{1}{a_{ij}^2}\right) - 3\right)\right)_{i,j=1}^{n}.\end{aligned} \tag{3.11}$$

Thus, we find $\mathrm{Tr}\left(\mathrm{Z}^\top\mathcal{D}\left(\mathcal{DF}(\mathrm{X})\right)[\mathrm{Z}]\right)$ as

$$\mathrm{Tr}\left(\mathrm{Z}^\top\mathcal{D}\left(\mathcal{DF}(\mathrm{X})\right)[\mathrm{Z}]\right) = \sum_{i,j=1}^{n} z_{ij}\left(\mathcal{D}\left(\mathcal{DF}(\mathrm{X})\right)[\mathrm{Z}]\right)_{ij} = 2\sum_{i,j=1}^{n} z_{ij}^2\left(\ln\left(\frac{1}{a_{ij}^2}\right) - 3\right). \tag{3.12}$$

For $1 \le i < j \le n$, let us define $\mathrm{A}^{i,j} \in \mathfrak{so}(n,\mathbb{R})$ as the skew-symmetric matrix having $+1$ in the (i,j)-entry and -1 in the (j,i)-entry. Now, for $\mathrm{Z}^{i,j} \in \mathcal{T}_{\mathrm{X}}\mathscr{M}_n$ with $1 \le i < j \le n$, using the orthogonality condition, we obtain

$$\mathrm{Tr}\left(\Lambda\mathrm{Z}^{i,j\top}\mathrm{Z}^{i,j}\right) = 2\sum_{k=1}^{n}\left[a_{ki}^2\ln\left(\frac{1}{a_{ki}^2}\right) + a_{kj}^2\ln\left(\frac{1}{a_{kj}^2}\right)\right] - 4. \tag{3.13}$$

From (3.12), we also have

$$\mathrm{Tr}\left(\mathrm{Z}^{i,j\top}\mathcal{D}\left(\mathcal{DF}(\mathrm{X})\right)[\mathrm{Z}^{i,j}]\right) = 2\sum_{k=1}^{n}\left[a_{ki}^2\ln\left(\frac{1}{a_{kj}^2}\right) + a_{kj}^2\ln\left(\frac{1}{a_{ki}^2}\right)\right] - 12. \tag{3.14}$$

Thus, we deduce

$$\begin{aligned}\xi^{i,j} &:= \operatorname{Tr}\left(Z^{i,j\top}\mathcal{D}\left(\mathcal{D}\mathcal{F}(X)\right)[Z^{i,j}]\right) - \operatorname{Tr}\left(\Lambda Z^{i,j\top}Z^{i,j}\right)\\ &= -8 + 2\sum_{k=1}^{n}(a_{ki}^2 - a_{kj}^2)\left[\ln\left(\frac{1}{a_{kj}^2}\right) - \ln\left(\frac{1}{a_{ki}^2}\right)\right]\\ &= -8 + 2\sum_{k=1}^{n}(a_{ki}^2 - a_{kj}^2)\ln\left(\frac{a_{ki}^2}{a_{kj}^2}\right).\end{aligned} \tag{3.15}$$

If $X = (a_{ij})_{i,j=1}^{n}$ is a strict local maximum, then a sufficient condition is that $\xi^{i,j} < 0$, for all $1 \le i < j \le n$.

Remark 3.5 It is well known that concave functions are closed under conical combinations. In other words, if $f_i(x)$ is concave for each $1 \le i \le N$ and the coefficients $c_i \ge 0$, then the function

$$f(x) = c_1 f_1(x) + \cdots + c_N f_N(x)$$

is also concave (see [7, Proposition 2.20]). In our setting, the function is

$$f(x) = x^2 \ln\left(\tfrac{1}{x^2}\right),$$

and the Shannon entropy is concave if and only if

$$|a_{ij}| \ge e^{-3/2} \approx 0.223, \quad \text{for all } 1 \le i, j \le n.$$

However, since the orthogonality constraint is non-convex, the maximization problem for $\mathscr{H}_1^{\mathbb{R}}[M_n]$ over the space of $n \times n$ real orthogonal matrices may admit multiple local maxima and minima.

4 Known Results and Analytical Study

As reported in [11], certain results on the maximization of the Shannon entropy function (see Definition 1.1) have been obtained numerically. In this section, we employ the optimization method from Section 3 to provide an analytical verification of those computational findings.

(1) For $n = 2$ and $n \equiv 0 \pmod 4$, the scaled real Hadamard matrix saturates the entropy, say for example, for $n = 2$,

$$M_2 := \frac{1}{\sqrt{2}}H_2 = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}, \tag{4.1}$$

and for $n = 4$,

$$M_4 := \frac{1}{2}H_4 = \begin{bmatrix} 1/2 & -1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & 1/2 & 1/2 \end{bmatrix}, \tag{4.2}$$

gives a *global maxima* for the problem (1.1).

(2) For $n = 3$, the matrix below with rational entries gives a *global maxima* for (1.1):

$$\mathrm{K}_3 = \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}. \tag{4.3}$$

(3) For $n = 5$, the matrix below with rational entries gives a *local maxima* for (1.1):

$$\mathrm{K}_5 = \begin{bmatrix} -3/5 & 2/5 & 2/5 & 2/5 & 2/5 \\ 2/5 & -3/5 & 2/5 & 2/5 & 2/5 \\ 2/5 & 2/5 & -3/5 & 2/5 & 2/5 \\ 2/5 & 2/5 & 2/5 & -3/5 & 2/5 \\ 2/5 & 2/5 & 2/5 & 2/5 & -3/5 \end{bmatrix}. \tag{4.4}$$

Although a global maximum for the case $n = 5$ is not yet established, (4.4) appears to be the global maximizer.

We now apply the optimization techniques from the previous section to validate these results on a case-by-case basis.

(1) For $n = 2$, the matrix in (4.1) represents a stationary point, and condition (3.10) is satisfied in this case, since $a_{ij}^2 = \frac{1}{2}$ for all i, j. Moreover, the tangent space to M_2 is given by

$$\mathcal{T}_{\mathrm{M}_2}\mathscr{M}_2 = \operatorname{span}\left\{\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\right\}.$$

The derivative matrix G_2 and the Lagrange multiplier $\Lambda_2 = \mathrm{G}_2^\top \mathrm{M}_2$ are given by

$$\mathrm{G}_2 = 2(\ln 2 - 1)\mathrm{M}_2 \;\text{ and }\; \Lambda_2 = 2(\ln 2 - 1)\mathrm{I}_2.$$

Observe that the sufficient condition in (3.15) is also fulfilled, since

$$\xi_2 := \operatorname{Tr}\left(\mathrm{Z}_2^\top \mathcal{D}\left(\mathcal{D}\mathcal{F}(\mathrm{M}_2)\right)[\mathrm{Z}_2]\right) - \operatorname{Tr}\left(\Lambda_2 \mathrm{Z}_2^\top \mathrm{Z}_2\right) = -8 < 0,$$

for all $\mathrm{Z}_2 \in \mathcal{T}_{\mathrm{M}_2}\mathscr{M}_2$, it follows that (4.1) is a local maximizer of $\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_2]$, with the maximum value $2\ln 2$. Moreover, this is also a global maximum, since Hadamard matrices achieve the global maximum of the Shannon entropy (see Theorem 5.1 below). For $n = 2$, a general orthogonal matrix takes the form

$$\mathrm{M}_2 := \begin{bmatrix} \pm a & \mp\sqrt{1-a^2} \\ \pm\sqrt{1-a^2} & \pm a \end{bmatrix} \text{ or } \begin{bmatrix} \pm a & \pm\sqrt{1-a^2} \\ \pm\sqrt{1-a^2} & \mp a \end{bmatrix}, \tag{4.5}$$

$a \in [-1, 1]$. Hence, maximization problem of the Shannon entropy reduces to maximizing the one variable function

$$\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_2] = 2\left[a^2 \ln\left(\frac{1}{a^2}\right) + (1-a^2)\ln\left(\frac{1}{1-a^2}\right)\right],$$

for which $a = 0$ and $a = \pm\frac{1}{\sqrt{2}}$ are the stationary points with $a = 0$ giving a local minimum and $a = \pm\frac{1}{\sqrt{2}}$, a local maximum, justifying our previous results.

(2) The matrix in (4.3) clearly satisfies condition (3.10), and is therefore a stationary point. Furthermore, the tangent space to K_3 is given by

$$\mathcal{T}_{K_3}\mathscr{M}_3 = \operatorname{span}\left\{ \begin{pmatrix} -2 & -1 & 0 \\ 1 & 2 & 0 \\ -2 & 2 & 0 \end{pmatrix}, \begin{pmatrix} -2 & 0 & -1 \\ -2 & 0 & 2 \\ 1 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 0 & -2 & 2 \\ 0 & -2 & -1 \\ 0 & 1 & 2 \end{pmatrix} \right\}. \tag{4.6}$$

The derivative matrix G_3 and the Lagrange multiplier Λ_3 are given by

$$G_3 = 2\left(\ln\left(\frac{9}{4}\right) - 1\right) K_3 - \frac{2}{3}\ln 4 I_3 \text{ and } \Lambda_3 = 2\left(\ln\left(\frac{9}{4}\right) - 1\right) I_3 - \frac{2}{3}\ln 4 K_3.$$

Note that $\xi_3^{i,j} = \frac{4}{3}(-6 + \ln 4) < 0$, and the sufficient condition is satisfied. In fact, it is a strict local maxima for the Shannon entropy $\mathscr{H}_1^{\mathbb{R}}[K_3]$, and the maximum value is $3\ln\left(\frac{9}{4}\right) + \frac{1}{3}\ln 4 \approx 2.895 < 3.296 \approx 3\ln 3$.

(3) For the matrix in (4.3), it is straightforward to verify that condition (3.10) is satisfied, confirming it as a stationary point. In this case, the derivative matrix G_5 and the Lagrange multiplier Λ_5 are obtained as

$$G_5 = 2\left(\ln\left(\frac{25}{4}\right) - 1\right) K_5 - \frac{6}{5}\ln\left(\frac{4}{9}\right) I_5 \text{ and } \Lambda_5 = 2\left(\ln\left(\frac{25}{4}\right) - 1\right) I_5 - \frac{6}{5}\ln\left(\frac{4}{9}\right) K_5.$$

A straightforward calculation yields $\xi_5^{i,j} = \frac{4}{5}\left(-10 + \ln\left(\frac{9}{4}\right)\right) < 0$. Hence, the matrix M_5 in (4.3) is a strict local maximizer of the Shannon entropy $\mathscr{H}_1^{\mathbb{R}}[K_5]$, with local maximum value $5\ln\left(\frac{25}{4}\right) - \frac{9}{5}\ln\left(\frac{9}{4}\right) \approx 7.703 < 8.047 \approx 5\ln 5$.

5 Bounds on the Shannon Entropy of Real Orthogonal Matrices

In this section, we derive maximum bounds for the Shannon entropy across all orders. We construct the matrices K_3 and K_5 (see (4.3) and (4.4)), which serve as local maximizers of the Shannon entropy for orders 3 and 5, respectively. In fact, for order 3, K_3 is a global maximizer. We begin by establishing a universal bound for the Shannon entropy defined in (1.1) for arbitrary n.

Theorem 5.1 *For all $n \in \mathbb{N}$, the Shannon entropy satisfies $\mathscr{H}_1^{\mathbb{R}}[M_n] \leq n\ln n$, with equality if and only if $a_{ij} = \pm\frac{1}{\sqrt{n}}, \quad 1 \leq i, j \leq n$.*

Proof Consider an $n \times n$ orthogonal matrix M_n with entries a_{ij} at the $(i, j)^{\text{th}}$ position, and define $x_{ij} = a_{ij}^2$. Then, by the definition of the Shannon entropy (see Definition 1.1), we obtain

$$\mathscr{H}_1^{\mathbb{R}}[M_n] = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}\ln\left(\frac{1}{x_{ij}}\right). \tag{5.1}$$

Since the logarithmic function is concave, Jensen's inequality implies that for each i, we have

$$\ln n = \ln\left(\sum_{j=1}^{n} x_{ij}\left(\frac{1}{x_{ij}}\right)\right) \geq \sum_{j=1}^{n} x_{ij}\ln\left(\frac{1}{x_{ij}}\right). \tag{5.2}$$

From (5.2), it can be easily seen that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}\ln\left(\frac{1}{x_{ij}}\right) \leq \sum_{i=1}^{n}\ln n = n\ln n, \tag{5.3}$$

and is an upper bound for $\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_n]$ in (5.1). Therefore, we have

$$0 \leq \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_n] \leq n\ln n, \tag{5.4}$$

for all n. Because the logarithmic function is strictly concave, the inequality above is strict unless $x_{ij} = a_{ij}^2 = \frac{1}{n}$. Consequently, equality holds if and only if $a_{ij} = \pm\frac{1}{\sqrt{n}}$ for all $1 \leq i, j \leq n$. □

From Theorem 5.1, it follows that whenever Hadamard matrices exist for a given n, the upper bound is attained uniquely by $\mathrm{M}_n = \frac{1}{\sqrt{n}}\mathrm{H}_n$. Thus, optimal matrices are available for $n = 1, 2, 4, 8, 12, \ldots$. More generally, for $n = 1, 2$ or $n \equiv 0 \pmod 4$, the existence of an optimal orthogonal matrix is equivalent to the validity of the Hadamard conjecture (see [13, Chapter I], [23, Lemma 2.2, and Chapter 11]).

Remark 5.2 If the matrix M_n is only required to satisfy the condition $\sum_{j=1}^{n} a_{ij}^2 = 1$ for $1 \leq i \leq n$, then the matrix

$$\mathrm{M}_n = \frac{1}{\sqrt{n}}\begin{bmatrix} \pm 1 & \cdots & \pm 1 \\ \vdots & \ddots & \vdots \\ \pm 1 & \cdots & \pm 1 \end{bmatrix}$$

achieves this bound for all n.

5.1 The bounds for Shannon entropy when $n = 3, 5, 6$:

In cases where Hadamard matrices do not exist, the upper bound of the entropy given in (5.4) is strict and cannot be attained. For such instances, we construct real orthogonal matrices that serve as local maximizers of the Shannon entropy for $n = 3$, 5, and 6.

Case (1): $\boldsymbol{n = 3}$. A 3×3 real orthogonal matrix contains exactly three independent real variables. Consider the form

$$\mathrm{M}_3 = \begin{bmatrix} * & x & y \\ * & * & z \\ * & * & * \end{bmatrix}.$$

The entry a_{11} is determined from the condition that the sum of the squares of the elements in the first row equals 1, yielding two possible values. Once a_{11} is fixed, the second row has two free variables, a_{21} and a_{22}, subject to two orthogonality constraints: the dot product of the first and second rows is zero, and the sum of the squares of the elements in the second row equals 1. This provides at most four possible pairs (a_{21}, a_{22}). After a_{11}, a_{21}, and a_{22} are determined, the third row contains three unknowns with three orthogonality conditions, yielding at most eight solutions. Consequently, a given triplet (x, y, z) can generate up to 64 orthogonal matrices, so the overall complexity of enumerating such orthogonal matrices reduces to analyzing the possible triplets (x, y, z).

Another approach to construct a 3×3 orthogonal matrix with exactly three independent real variables is as follows. Choose three diagonal entries (with absolute values less than 1) freely, and determine the remaining six entries from the six orthogonality conditions. Define

$$\mathrm{M}_3 := \begin{bmatrix} x & * & * \\ * & y & * \\ * & * & z \end{bmatrix},$$

and set

$$\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_3] = \left[x^2 \ln\left(\frac{1}{x^2}\right) + y^2 \ln\left(\frac{1}{y^2}\right) + z^2 \ln\left(\frac{1}{z^2}\right)\right] + \sum_{i \neq j} a_{ij}^2 \ln\left(\frac{1}{a_{ij}^2}\right). \tag{5.5}$$

For any fixed triplet (x, y, z), the first bracketed term remains constant. To optimize $\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_3]$, we apply Jensen's inequality to the six off-diagonal terms a_{ij}^2 with $i \neq j$. Since

$$\sum_{i \neq j} a_{ij}^2 = 3 - (x^2 + y^2 + z^2) = \text{constant},$$

the maximum entropy is achieved (for fixed diagonal entries) when all off-diagonal values satisfy $a_{ij}^2 = \text{constant}$, provided such an orthogonal matrix exists. Thus, the maximum of $\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_3]$ occurs precisely when such a symmetric orthogonal matrix can be constructed.

Stewart [26] provided a characterization of symmetric orthogonal matrices based on Householder reflections. Specifically, any such matrix can be expressed as

$$\mathrm{M}_n(\mathrm{u}) = \mathrm{I}_n - 2\frac{\mathrm{u}\mathrm{u}^\top}{\|\mathrm{u}\|^2},$$

where u is a nonzero vector in $\mathbb{R}^n$ and $\|\cdot\|$ denotes the Euclidean norm. Moreover, any $n \times n$ orthogonal matrix can be written as a product of at most n such reflections. From the above expression, it follows directly that $\mathrm{M}_n^\top = \mathrm{M}_n$ and $\mathrm{M}_n\mathrm{M}_n^\top = \mathrm{M}_n^\top\mathrm{M}_n = \mathrm{I}_n$.

For $n = 3$, let $\mathrm{u} = (p, q, r)$ with norm $\|\mathrm{u}\|^2 = p^2 + q^2 + r^2 = d$. Then the matrix M_3 takes the form

$$\mathrm{M}_3(p, q, r) = \begin{bmatrix} 1 - 2p^2/d & -2pq/d & -2pr/d \\ -2pq/d & 1 - 2q^2/d & -2qr/d \\ -2pr/d & -2qr/d & 1 - 2r^2/d \end{bmatrix}. \tag{5.6}$$

Suppose M_3 has column vectors v_1, v_2, v_3. Define

$$f(\mathrm{v}_1) = f(x_1, y_1, z_1) = x_1^2 \ln\left(\frac{1}{x_1^2}\right) + y_1^2 \ln\left(\frac{1}{y_1^2}\right) + z_1^2 \ln\left(\frac{1}{z_1^2}\right).$$

Then the Shannon entropy is given by

$$\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_3] = f(\mathrm{v}_1) + f(\mathrm{v}_2) + f(\mathrm{v}_3).$$

To maximize $\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_3]$, it follows from the structure of M_3 in (5.6) that we must have

$$f(\mathrm{v}_1) = f(\mathrm{v}_2) = f(\mathrm{v}_3).$$

Applying this condition to the columns of $M_3(p, q, r)$ gives $p = q = r$. Thus, with $u = (p, p, p)$ and $\|u\|^2 = 3p^2 = d$, we obtain

$$M_3(p, p, p) = \begin{bmatrix} 1/3 & -2/3 & -2/3 \\ -2/3 & 1/3 & -2/3 \\ -2/3 & -2/3 & 1/3 \end{bmatrix} \text{ or } (-1) \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}, \tag{5.7}$$

Thus, (5.7) gives the optimal 3×3 orthogonal matrix. For this matrix, the optimal Shannon entropy is

$$\mathscr{H}_1^{\mathbb{R}}[M_3] \approx 2.895 \ < \ 3 \ln 3 \approx 3.296.$$

Case (2): $\boldsymbol{n = 5}$. In general, an $n \times n$ real orthogonal matrix involves n^2 variables subject to $n + \binom{n}{2}$ orthogonality constraints, leaving $\binom{n}{2}$ independent variables. For $n = 3$, the number of independent variables, $\binom{n}{2}$, coincides with the number of diagonal entries ($= n$). However, for $n > 3$, this equality no longer holds, so the diagonal-fixing approach is not applicable. In such cases, we adopt a more general strategy, extending the construction method to higher orders.

The optimal value $n \ln n$ is attained when all entries of the orthogonal matrix have absolute value $\frac{1}{\sqrt{n}}$. If such a configuration is not achievable, then the matrix must contain at least two distinct absolute values among its entries. Suppose each row of the $n \times n$ matrix contains m entries equal to a and $(n - m)$ entries equal to b, with $m \geq (n - m)$. Then, the row norm condition is given by

$$ma^2 + (n - m)b^2 = 1. \tag{5.8}$$

Let x denote the number of entries equal to b that overlap in two rows. Then, the inner product of the corresponding elements of these two rows is given by

$$xb^2 + (2m + x - n)a^2 + (2n - 2m - 2x)ab = 0. \tag{5.9}$$

From (5.9), we have

$$x\left(\frac{b}{a}\right)^2 + 2(n - m - x)\frac{b}{a} + (2m + x - n) = 0. \tag{5.10}$$

To ensure that $\frac{b}{a}$ is real, we obtain the following constraint:

$$x \leq \frac{(n - m)^2}{n}. \tag{5.11}$$

In our attempt to attain the upper bound $n \ln n$, we aim for the absolute values of a and b to be as close as possible to $\frac{1}{\sqrt{n}}$ for all n. Consequently, the condition $\left|\frac{b}{a}\right| \approx 1$ provides an approximate relation between x, m, and n from (5.10):

$$x + m \approx \frac{3n}{4}. \tag{5.12}$$

This relation confirms that when n is a multiple of 4 (the Hadamard matrix case), the bound $n \ln n$ is achievable by setting $x = 0$ and $m = \frac{3n}{4}$, which is an integer. Substituting this value of m into (5.8) and (5.9), one verifies that $|a| = |b| = \frac{1}{\sqrt{n}}$.

For $n = 3$, as noted earlier, $\frac{3n}{4} = 2.25$, for which the nearest integers give $x + m = 2$ or 3. Since $m \geq (n - m)$, we must have $m = 2$ or 3. The constraint on x (see (5.11)) yields $x = 0$ for both possible values of m. However, when $m = 3$, the two orthogonality conditions contradict each other. Thus, we fix $m = 2$, meaning that each row of a 3×3 matrix must contain two entries of one value and one entry of the other. Solving (5.8) and (5.9) in this case gives

$$a = \pm\tfrac{2}{3},\ b = \mp\tfrac{1}{3},$$

which agrees with the result obtained earlier for the case $n = 3$.

For $n = 5$, we have $\frac{3n}{4} = 3.75$, so the nearest integers are $x + m = 3$ or 4. From (5.11), for both possible values of m, the only integer solution for x is 0. Thus, $m = 3$ or 4. For $m = 3$, we obtain

$$a = \tfrac{4}{5\sqrt{2}}, \qquad b = \tfrac{-1}{5\sqrt{2}}.$$

In this case, the Shannon entropy is

$$\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_5] \approx 6.252 \ < \ 5\ln 5 \approx 8.047.$$

Two rows of such a matrix M_5 (if it exists) would have the form

$$\begin{bmatrix} 4/5\sqrt{2} & -1/5\sqrt{2} & 4/5\sqrt{2} & -1/5\sqrt{2} & 4/5\sqrt{2} \\ 4/5\sqrt{2} & 4/5\sqrt{2} & -1/5\sqrt{2} & 4/5\sqrt{2} & -1/5\sqrt{2} \end{bmatrix}.$$

However, since $x = 0$ and $(n - m) = 2$ (i.e., there is no overlap of the b's and each row contains exactly two b's), constructing a full 5×5 orthogonal matrix of this form is impossible.

For $m = 4$, we obtain

$$a = \tfrac{2}{5}, \qquad b = \tfrac{-3}{5}.$$

Here, the Shannon entropy is

$$\mathscr{H}_1^{\mathbb{R}}[\mathrm{K}_5] \approx 7.703 \ > \ 6.252 \quad \text{and} \quad \mathscr{H}_1^{\mathbb{R}}[\mathrm{K}_5] \approx 7.703 \ < \ 5\ln 5 \approx 8.047.$$

Thus, in this case, the locally optimal matrix is K_5 and is given by

$$\mathrm{K}_5 = \begin{bmatrix} -3/5 & 2/5 & 2/5 & 2/5 & 2/5 \\ 2/5 & -3/5 & 2/5 & 2/5 & 2/5 \\ 2/5 & 2/5 & -3/5 & 2/5 & 2/5 \\ 2/5 & 2/5 & 2/5 & -3/5 & 2/5 \\ 2/5 & 2/5 & 2/5 & 2/5 & -3/5 \end{bmatrix}.$$

Proposition 5.3 *The orthogonal matrix*

$$\mathrm{K}_n := \frac{1}{n}\begin{bmatrix} -(n-2) & 2 & \cdots & 2 \\ 2 & -(n-2) & \cdots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ 2 & 2 & \cdots & -(n-2) \end{bmatrix}, \tag{5.13}$$

is always a stationary point of the maximization problem (3.7). For $n \geq 3$, a sufficient condition for K_n to be a local maximum is

$$(n-4)\ln\left(\frac{n-2}{2}\right) < n.$$

If K_n *is a local maximum, then the local maximum value is* $\frac{(n-2)^2}{n}\ln\left(\frac{n^2}{(n-2)^2}\right)+\frac{4(n-1)}{n}\ln\left(\frac{n^2}{4}\right)$.

Proof For the orthogonal matrix given in (5.13), we have

$$\mathrm{A}_n=\mathrm{G}_n\mathrm{M}_n^\top-\mathrm{M}_n\mathrm{G}_n^\top=\mathbf{0}_n,$$

where

$$\mathrm{G}_n=\frac{2}{n}\begin{bmatrix}-(n-2)\left(\ln\left(\frac{n^2}{(n-2)^2}\right)-1\right) & \cdots & 2\left(\ln\left(\frac{n^2}{4}\right)-1\right)\\ \vdots & \ddots & \vdots\\ 2\left(\ln\left(\frac{n^2}{4}\right)-1\right) & \cdots & -(n-2)\left(\ln\left(\frac{n^2}{(n-2)^2}\right)-1\right)\end{bmatrix}=\mathrm{G}_n^\top.$$

It can be easily calculated that

$$\xi_n^{i,j}=-8+\frac{4}{n}(n-4)\ln\left(\frac{(n-2)^2}{4}\right),$$

$1\le i<j\le n$. Hence, $\xi_n^{i,j}<0$ for $(n-4)\ln\left(\frac{n-2}{2}\right)<n$, and it gives a sufficient condition for M_n to be a local maximum. □

Remark 5.4 From, the graph of the function $f(n)=n-(n-4)\ln\left(\frac{n-2}{2}\right)$ it is evident that M_n is expected to be a local maximum of the Shannon entropy for $n\le 11$, and a local minimum for $n\ge 12$.

(4) **Case $n=6$**: In this case as well, since a Hadamard matrix does not exist, we look for an orthogonal matrix whose entries have at most two distinct absolute values. Proceeding in a manner similar to the $n=5$ case, we obtain the following conditions on m,x,a, and b. For $n=6$, the admissible values of m are 3, 4, and 5, and we additionally have

$$x+m\approx\frac{3n}{4}=4.5.$$

The nearest integers to $x+m$ are 4 and 5. From (5.11), we know that

$$x\le\frac{(6-m)^2}{6},$$

and hence for $m=4$ or $m=5$, we have $x=0$. Let us now discuss all the cases separately.

(i) *Case* $m=4$ *and* $x=0$: Solving for a and b, we obtain

$$a=\pm\tfrac{2}{3\sqrt{2}},\qquad b=\mp\tfrac{1}{3\sqrt{2}}.$$

An orthogonal matrix with these values can be constructed as follows:

$$\mathrm{M}_6=\begin{bmatrix}-1/3\sqrt{2} & 2/3\sqrt{2} & 2/3\sqrt{2} & 1/3\sqrt{2} & -2/3\sqrt{2} & -2/3\sqrt{2}\\ 2/3\sqrt{2} & -1/3\sqrt{2} & 2/3\sqrt{2} & -2/3\sqrt{2} & 1/3\sqrt{2} & -2/3\sqrt{2}\\ 2/3\sqrt{2} & 2/3\sqrt{2} & -1/3\sqrt{2} & -2/3\sqrt{2} & -2/3\sqrt{2} & 1/3\sqrt{2}\\ -1/3\sqrt{2} & 2/3\sqrt{2} & 2/3\sqrt{2} & -1/3\sqrt{2} & 2/3\sqrt{2} & 2/3\sqrt{2}\\ 2/3\sqrt{2} & -1/3\sqrt{2} & 2/3\sqrt{2} & 2/3\sqrt{2} & -1/3\sqrt{2} & 2/3\sqrt{2}\\ 2/3\sqrt{2} & 2/3\sqrt{2} & -1/3\sqrt{2} & 2/3\sqrt{2} & 2/3\sqrt{2} & -1/3\sqrt{2}\end{bmatrix}. \tag{5.14}$$

The first-order condition stated in (3.10) can be readily verified for the matrix M_6, confirming that it is a stationary point. The corresponding derivative matrix G_6 and the Lagrange multiplier Λ_6 are given by

$$G_6 = 2\left(\ln\left(\frac{9}{2}\right) - 1\right) M_6 - \frac{2}{3}\ln 4\,(M_2 \otimes I_3) \text{ and}$$
$$\Lambda_6 = 2\left(\ln\left(\frac{9}{2}\right) - 1\right) I_6 - \frac{2}{3}\ln 4\begin{pmatrix} M_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & M_3 \end{pmatrix},$$

where M_2 and $M_3 = K_3$ are defined in (4.1) and (4.2) respectively. Note that

$$\xi_6^{i,j} = \frac{4}{3}(-6 + \ln 4) < 0,$$

for $1 \le i < j \le 6$, which makes M_6, a candidate for the local maximum for the Shannon entropy $\mathscr{H}_1^{\mathbb{R}}[M_6]$. The value of the Shannon entropy for the above matrix is

$$\mathscr{H}_1^{\mathbb{R}}[M_6] = 6\ln\left(\frac{9}{2}\right) + \frac{2}{3}\ln 4 = 2\mathscr{H}_1^{\mathbb{R}}[M_3] + 3\mathscr{H}_1^{\mathbb{R}}[M_2] \approx 9.949 < 6\ln 6 \approx 10.751.$$

Observation: In the 6×6 case, the desired matrix can be expressed as the Kronecker product of the optimal matrices obtained for the 2×2 and 3×3 cases. Specifically,

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \otimes \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix},$$

and hence $M_6 = M_2 \otimes M_3$.

(ii) *Case $m = 5$ and $x = 0$*: Solving for a and b, we obtain

$$a = \pm\tfrac{1}{3}, \qquad b = \mp\tfrac{2}{3}.$$

An orthogonal matrix with these values can be constructed as follows:

$$M_6 = \begin{bmatrix} -2/3 & 1/3 & 1/3 & 1/3 & 1/3 & 1/3 \\ 1/3 & -2/3 & 1/3 & 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & -2/3 & 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 & -2/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 & 1/3 & -2/3 & 1/3 \\ 1/3 & 1/3 & 1/3 & 1/3 & 1/3 & -2/3 \end{bmatrix}$$

By Proposition 5.3, M_6 is also a candidate for a local maximum. The corresponding Shannon entropy value for this matrix is

$$\mathscr{H}_1^{\mathbb{R}}[M_6] = 6\ln\left(\frac{9}{2}\right) + \frac{1}{3}\ln 4 \approx 9.487 < 9.949.$$

(iii) *Case $m = 3$ and $x = 0$*: Solving for a and b, we obtain

$$a = \pm\tfrac{1}{\sqrt{3}}, \qquad b = 0,$$

with the condition that a and b must be nonnegative. Such matrices, however, may correspond to saddle points.

(iv) *Case* $m = 3$ *and* $x = 1$: Solving for a and b, we obtain

$$a^2 = \tfrac{1}{17-8\sqrt{3}}, \qquad b^2 = \tfrac{7-4\sqrt{3}}{17-8\sqrt{3}}.$$

In this case (even if such a matrix exists), the Shannon entropy satisfies

$$\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_6] \approx 8.112 < 9.487 < 9.949.$$

Hence, for $n = 6$, we conclude that the matrix $\mathrm{M}_6 := \mathrm{M}_2 \otimes \mathrm{M}_3$ attains the maximal value of the Shannon entropy. However, it remains inconclusive whether this represents a global maximum.

The Kronecker product. It is straightforward to verify that the Kronecker product of two orthogonal matrices is itself orthogonal. Specifically, let $\mathrm{M}_1 = \mathrm{M}_{n_1}$ and $\mathrm{M}_2 = \mathrm{M}_{n_2}$ be orthogonal matrices with $n_1, n_2 \geq 2$. Then $\mathrm{M}_{n_1}\mathrm{M}_{n_1}^\top = \mathrm{I}_{n_1},\ \mathrm{M}_{n_2}\mathrm{M}_{n_2}^\top = \mathrm{I}_{n_2}$. The Kronecker product $\mathrm{M} = \mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2}$, and

$$\begin{aligned}\mathrm{M}\mathrm{M}^\top &= (\mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2})(\mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2})^\top = (\mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2})(\mathrm{M}_{n_1}^\top \otimes \mathrm{M}_{n_2}^\top)\\ &= (\mathrm{M}_{n_1}\mathrm{M}_{n_1}^\top) \otimes (\mathrm{M}_{n_2}\mathrm{M}_{n_2}^\top) = \mathrm{I}_{n_1} \otimes \mathrm{I}_{n_2} = \mathrm{I}_{n_1 n_2}.\end{aligned}$$

Similarly, $\mathrm{M}^\top\mathrm{M} = \mathrm{I}_{n_1 n_2}$, therefore the matrix $\mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2}$ is an orthogonal matrix. Then, we have the following lemma:

Lemma 5.5 *Let* M_{n_1} *be an* $n_1 \times n_1$ *orthogonal matrix and* M_{n_2} *be an* $n_2 \times n_2$ *orthogonal matrix, then* $\mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2}$ *is also orthogonal matrix of order* $n_1 n_2 \times n_1 n_2$.

In general, the Kronecker product of two matrices is not commutative; however the value of entropy of a Kronecker product is independent of the order, i.e.,

$$\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2}] = \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{n_2} \otimes \mathrm{M}_{n_1}]. \tag{5.15}$$

Proposition 5.6 *Let the orthogonal matrices* M_{n_1} *and* M_{n_2} *be stationary points of the maximization problem (3.1) of orders* n_1 *and* n_2 *respectively. Then* $\mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2}$ *is a stationary point of the corresponding maximization problem of order* $n_1 n_2$.

If $\mathrm{M}_{n_1} = \frac{1}{\sqrt{n_1}}\mathrm{H}_{n_1}$ *and* M_{n_2} *is a local maximum of the problem (3.1) of order* n_2, *then the matrix* $\mathrm{T}_{n_1 n_2} := \mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2}$ *satisfies the sufficient condition* (3.15) *for local maximum of the the problem (3.1) of order* $n_1 n_2$.

Proof Let the orthogonal matrices $\mathrm{M}_{n_1} = (a_{ij})_{i,j=1}^{n_1}$ and $\mathrm{M}_{n_2} = (b_{ij})_{i,j=1}^{n_2}$ be stationary points of the minimization problem (3.1) of orders n_1 and n_2 respectively. Then, from (3.1), we know that

$$\sum_{k=1}^{n_1} a_{ik}a_{jk} \ln\left(\frac{a_{jk}^2}{a_{ik}^2}\right) = 0 \ \text{ and } \ \sum_{l=1}^{n_2} b_{fl}b_{gl} \ln\left(\frac{b_{gl}^2}{b_{fl}^2}\right) = 0, \tag{5.16}$$

for all $1 \leq i, j \leq n_1$, $1 \leq f, g \leq n_2$. Our aim is to show that

$$\sum_{h=1}^{n_1 n_2} c_{ih}c_{jh} \ln\left(\frac{c_{jh}^2}{c_{ih}^2}\right) = 0, \tag{5.17}$$

for all $1 \leq i, j \leq n_1 n_2$. For $1 \leq i \leq n_1 n_2$ and $1 \leq j \leq n_2$, we have

$$\sum_{h=1}^{n_1 n_2} c_{ih} c_{jh} \ln\left(\frac{c_{jh}^2}{c_{ih}^2}\right) = \left(\sum_{k=1}^{n} a_{ik}^2\right)\left(\sum_{h=1}^{m} b_{ih} b_{jh} \ln\left(\frac{b_{jh}^2}{b_{ih}^2}\right)\right) = 0, \tag{5.18}$$

by using (5.16). Now, for $1 \leq i \leq n_1 n_2$ and $n_2 + 1 \leq j \leq 2n_2$, we have

$$\sum_{h=1}^{nm} c_{ih} c_{jh} \ln\left(\frac{c_{jh}^2}{c_{ih}^2}\right) = \left(\sum_{l=1}^{n_1} b_{il}^2\right)\left(\sum_{h=1}^{n_2} a_{ih} a_{gh} \ln\left(\frac{a_{gh}^2}{a_{ih}^2}\right)\right) = 0, \tag{5.19}$$

where $g = 1, \cdots, n_1$, by using (5.16). The other values of i and j can be obtained in a similar way and hence $\mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2}$ is a stationary point of the problem (3.1).

From Theorem 5.1, we know that $\frac{1}{\sqrt{n_1}}\mathrm{H}_{n_1}$ is a global maximum of the optimization problem (3.1) of order n_1. If the matrix $\mathrm{M}_{n_2} = (a_{ij})_{i,j=1}^{n_2}$ is a strict local maximum of the optimization problem (3.1) of order n_2, then we know that $\sum\limits_{k=1}^{n_1} a_{ik} a_{jk} \ln\left(\frac{a_{jk}^2}{a_{ik}^2}\right) = 0$, for all $1 \leq i, j \leq n_2$, and

$$\xi_{\mathrm{M}_{n_2}}^{i,j} = -8 + 2\sum_{k=1}^{n_2}(a_{ki}^2 - a_{kj}^2) \ln\left(\frac{a_{ki}^2}{a_{kj}^2}\right) < 0,$$

for $1 \leq i < j \leq n_2$. Now, we consider the matrix $\mathrm{T}_{n_1 n_2} := \frac{1}{\sqrt{n_1}}\mathrm{H}_{n_1} \otimes \mathrm{M}_{n_2} = (c_{ij})_{i,j=1}^{n_1 n_2}$. Since the entries of H_{n_1} are ± 1, one can easily get $\xi_{\mathrm{T}_{n_1 n_2}}^{i,j} := -8 + 2\sum\limits_{k=1}^{n_1 n_2}(c_{ki}^2 - c_{kj}^2) \ln\left(\frac{c_{ki}^2}{c_{kj}^2}\right)$ is

$$\text{either } -8 + 2\sum_{k=1}^{n_2}(a_{ki}^2 - a_{kj}^2) \ln\left(\frac{a_{ki}^2}{a_{kj}^2}\right) < 0, \quad \text{or } -8 < 0,$$

for $1 \leq i < j \leq n_1 n_2$. Thus the matrix $\mathrm{T}_{n_1 n_2} := \mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2}$ satisfies the sufficient condition (3.15) for local maximum of the the problem (3.1) of order $n_1 n_2$. □

Proposition 5.7 *Let the orthogonal matrices* $\mathrm{M}_{n_1} = (a_{ij})_{ij=1}^{n_1}$ *and* $\mathrm{M}_{n_2} = (b_{ij})_{i,j=1}^{n_2}$ *be stationary points of the maximization problem (3.1) for orders* n_1 *and* n_2 *respectively. Then, we have*

$$\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2}] = n_2 \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{n_1}] + n_1 \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{n_2}]. \tag{5.20}$$

Proof If $\mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2} = (c_{ij})_{i,j=1}^{n_1 n_2}$, then we have

$$\begin{aligned}
\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{n_1} \otimes \mathrm{M}_{n_2}] &= \sum_{i,j=1}^{n_1 n_2} c_{ij}^2 f\left(\frac{1}{c_{ij}^2}\right) = \sum_{k,l=1}^{n_1} \sum_{i,j=1}^{n_2} a_{kl}^2 b_{ij}^2 \ln\left(\frac{1}{a_{kl}^2 b_{ij}^2}\right) \\
&= \sum_{k,l=1}^{n_1} \sum_{i,j=1}^{n_2} a_{kl}^2 b_{ij}^2 \left(\ln\left(\frac{1}{a_{kl}^2}\right) + \ln\left(\frac{1}{b_{ij}^2}\right)\right) \\
&= \left(\sum_{k,l=1}^{n_1} a_{kl}^2 \ln\left(\frac{1}{a_{kl}^2}\right)\right)\left(\sum_{i,j=1}^{n_2} b_{ij}^2\right) + \left(\sum_{k,l=1}^{n_1} a_{kl}^2\right)\left(\sum_{i,j=1}^{n_2} b_{ij}^2 \ln\left(\frac{1}{b_{ij}^2}\right)\right) \\
&= n_2 \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{n_1}] + n_1 \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{n_2}],
\end{aligned} \tag{5.21}$$

which completes the proof. □

(4) **Case $n = 7$**: From Proposition 5.3, it follows that the matrix K_7 is a candidate for a local maximum of the optimization problem (3.1) of order 7, with a local maximum value of $10.994 < 13.621 = 7 \ln 7$. Previous studies on almost Hadamard matrices [4–6, 17] indicate that the orthogonal matrix associated with the Fano plane also corresponds to a local maximum of the Shannon entropy. This matrix is given in (2.2). Furthermore, in Proposition 5.8, we show that the matrix $\mathcal{I}_7$ is another candidate for a local maximum of (3.1) of order 7. If $\mathcal{I}_7$ indeed represents a local maximum, its corresponding Shannon entropy value is $12.840 < 13.621 \approx 7 \ln 7$.

Proposition 5.8 *Let a finite projective plane of order n exists. Then the orthogonal matrix corresponding to the finite projective plane satisfies the sufficient condition* (3.15) *for the local maximum for the optimization problem* (3.1) *of order $n^2 + n + 1$ for $n = 2, \ldots, 7$. If it is a local maximum, then the local maximum value of the Shannon entropy is given by*

$$\begin{aligned}
&\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{n^2+n+1}] \\
&= \frac{2}{n^2+n+1}\left[(n+1)(n\sqrt{n}-1)^2 \ln\left(\frac{n\sqrt{n}+1}{n-1}\right) + n(n+1+\sqrt{n})^2 \ln\left((n+1)\sqrt{n}-n\right)\right].
\end{aligned} \tag{5.22}$$

Proof Let us assume that a finite projective plane of order n exists. Then there exists an orthogonal matrix $\mathrm{M}_{n^2+n+1} = (a_{ij})_{i,j=1}^{n^2+n+1}$ with the entries given in (2.1) with $\lambda = 1$. It should be noted that any two distinct rows contains $1 \begin{pmatrix} x \\ x \end{pmatrix}$, $n \begin{pmatrix} x \\ y \end{pmatrix}$, $n \begin{pmatrix} y \\ x \end{pmatrix}$ and $n(n-1) \begin{pmatrix} y \\ y \end{pmatrix}$. Thus the first order necessary condition for optimality is satisfied, since for $i \neq j$

$$\begin{aligned}
(\mathrm{A})_{i,j} = (\mathrm{G}\mathrm{X}^\top - \mathrm{X}\mathrm{G}^\top)_{i,j} &= 2 \sum_{k=1}^{n^2+n+1} a_{ik} a_{jk} \ln\left(\frac{a_{jk}^2}{a_{ik}^2}\right) \\
&= 2nxy \ln\left(\frac{x^2}{y^2}\right) + 2nyx \ln\left(\frac{y^2}{x^2}\right) = 0.
\end{aligned}$$

Thus the matrix M_{n^2+n+1} is a stationary point of the optimization problem (3.1). In order to get the sufficient condition (3.15), we consider

$$\begin{aligned}\xi^{i,j} &= -8 + 2\sum_{k=1}^{n^2+n+1}(a_{ki}^2 - a_{kj}^2)\ln\left(\frac{a_{ki}^2}{a_{kj}^2}\right)\\ &= -8 + 2n\left[(x^2-y^2)\ln\left(\frac{x^2}{y^2}\right) + (y^2-x^2)\ln\left(\frac{y^2}{x^2}\right)\right] = -8 + 4n(x^2-y^2)\ln\left(\frac{x^2}{y^2}\right)\\ &= -8\left[1 + \frac{1+n-n^2+2\sqrt{n}}{n^2+n+1}\ln(n-\sqrt{n})\right], \end{aligned} \tag{5.23}$$

for $i < j$. But it can be easily seen that $f(n) = 1 + \frac{1+n-n^2+2\sqrt{n}}{n^2+n+1}\ln(n-\sqrt{n}) > 0$, for all $n \leq 7$. Thus the matrix M_{n^2+n+1} is a candidate for local maximum of the Shannon entropy for $n = 2, 3, 4, 5, 6$ and 7. If it is a local maximum, then the local maximum value given in (5.22) can easily be obtained by using the values of x and y in (2.1). □

Proposition 5.9 *Suppose a biplane of order n exists. Then the orthogonal matrix associated with this biplane satisfies the sufficient condition* (3.15) *for being a local maximum of the optimization problem* (3.1) *of order* $\frac{(n+1)(n+2)}{2} + 1$, *for* $n = 2, \ldots, 13$. *If the matrix indeed corresponds to a local maximum, the resulting value of the Shannon entropy is given by*

$$\begin{aligned}\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{\frac{(n+1)(n+2)}{2}+1}] = \frac{1}{n^2+3n+4}\Bigg[&(n+2)((n+1)\sqrt{n}-2)^2\ln\left(\frac{(n+1)\sqrt{n}+2}{n-1}\right)\\ &+2(n+1)(n+2+\sqrt{n})^2\ln\left(\frac{(n+2)\sqrt{n}-n}{2}\right)\Bigg].\end{aligned} \tag{5.24}$$

Proof Let us assume that a biplane of order n exists. Then there exists an orthogonal matrix $\mathrm{M}_{\frac{(n+1)(n+2)}{2}+1} = (a_{ij})_{i,j=1}^{\frac{(n+1)(n+2)}{2}+1}$ with the entries given in (2.1) with $\lambda = 2$. We know that any two distinct rows contains $2\begin{pmatrix}x\\x\end{pmatrix}$, $n\begin{pmatrix}x\\y\end{pmatrix}$, $n\begin{pmatrix}y\\x\end{pmatrix}$ and $\frac{n(n-1)}{2}\begin{pmatrix}y\\y\end{pmatrix}$. Thus, the first order necessary condition for optimality is easily satisfied, since for $i \neq j$

$$\begin{aligned}(\mathrm{A})_{i,j} = (\mathrm{GX}^\top - \mathrm{XG}^\top)_{i,j} &= 2\sum_{k=1}^{\frac{(n+1)(n+2)}{2}+1} a_{ik}a_{jk}\ln\left(\frac{a_{jk}^2}{a_{ik}^2}\right)\\ &= 2nxy\ln\left(\frac{x^2}{y^2}\right) + 2nyx\ln\left(\frac{y^2}{x^2}\right) = 0.\end{aligned}$$

Hence, the matrix $\mathrm{M}_{\frac{(n+1)(n+2)}{2}+1}$ is a stationary point of the optimization problem (3.1) of order $\frac{(n+1)(n+2)}{2} + 1$. Now, one can show that

$$\begin{aligned}\xi^{i,j} &= -8 + 2\sum_{k=1}^{\frac{(n+1)(n+2)}{2}+1}(a_{ki}^2 - a_{kj}^2)\ln\left(\frac{a_{ki}^2}{a_{kj}^2}\right) = -8 + 4n(x^2-y^2)\ln\left(\frac{x^2}{y^2}\right)\\ &= -8\left[1 + \frac{4+n-n^2+4\sqrt{n}}{n^2+3n+4}\ln\left(\frac{n-\sqrt{n}}{2}\right)\right],\end{aligned} \tag{5.25}$$

for $i < j$. Therefore, the sufficient condition in (3.15) is satisfied whenever

$$g(n) = 1 + \frac{4 + n - n^2 + 4\sqrt{n}}{n^2 + 3n + 4} \ln\left(\frac{n - \sqrt{n}}{2}\right) > 0,$$

which holds for all $n \leq 12$. Consequently, the matrix $\mathrm{M}_{\frac{(n+1)(n+2)}{2}+1}$ is a candidate for a local maximum of the Shannon entropy for $n = 2, \ldots, 13$. Finally, if it indeed corresponds to a local maximum, its value can be readily computed using the expressions for x and y in (2.1), as given in (5.24). □

Analogous to Propositions 5.8 and 5.9, it can be shown that the λ-plane, for $\lambda \geq 1$, is also a candidate for a local maximum of the Shannon entropy for certain values of λ.

Proposition 5.10 *Suppose a λ-plane of order n exists. Then the orthogonal matrix associated with this λ-plane satisfies the sufficient condition* (3.15) *for being a local maximum of the optimization problem* (3.1) *of order* $\frac{(n+\lambda)(n+\lambda-1)}{\lambda} + 1$, *for values of λ and n that satisfy*

$$1 + \frac{\lambda^2 + n - n^2 + 2\sqrt{n}\lambda}{n^2 + (2\lambda - 1)n + \lambda^2} \ln\left(\frac{n - \sqrt{n}}{\lambda}\right) > 0. \tag{5.26}$$

(5) **Case $n = 9$**: In the case $n = 9$, the matrix $\mathrm{M}_9 = \mathrm{K}_3 \otimes \mathrm{K}_3$ can be regarded as a candidate for a local maximum of the optimization problem (3.1) of order 9 . If it indeed represents a local maximum, the corresponding value of the Shannon entropy is $17.369 < 19.775 \approx 9 \ln 9$.

(6) **Case $n = 10$**: Using Propositions 5.3 and 5.6, one can easily see that

$$\mathrm{M}_{10} = \frac{1}{\sqrt{2}} \mathrm{H}_2 \otimes \mathrm{K}_5 \tag{5.27}$$

is a candidate for local maximum of the Shannon entropy of order 10. If it is a local maximum, then he local maximum value is given by $22.338 < 23.026 \approx 10 \ln 10$.

(7) **Case $n = 11$**: By Proposition 5.9, the orthogonal matrix corresponding to the Paley biplane of order 3 is a candidate for a local maximum of the optimization problem (3.1) of order 11. If it indeed represents a local maximum, the corresponding Shannon entropy is $25.403 < 26.379 \approx 11 \ln 11$. An example of such a matrix is provided in (2.4).

(8) **Case $n = 13$**: By Proposition 5.9, the orthogonal matrix associated with the finite projective plane of order 3 is a candidate for a local maximum of the Shannon entropy for order 13. If it indeed corresponds to a local maximum, the value of the Shannon entropy is less than $33.344 \approx 13 \ln 13$. An example of such a matrix is $\mathcal{I}_{13}$, given in (2.3).

(9) **Case $n = 14$**: Using Propositions 5.6 and 5.9, we know that

$$\mathrm{M}_{14} = \frac{1}{\sqrt{2}} \mathrm{H}_2 \otimes \mathcal{I}_7 \tag{5.28}$$

is candidate for a local maximum of the optimization problem (3.1) of order 14. If it is a local maximum, then we use Proposition 5.7 to find out the local maximum value of the Shannon entropy of order 14 as $35.384 < 36.947 \approx 14 \ln(14)$.

(10) **Case $n = 15$**: For $n = 15$, one can show that the matrix $\mathrm{M}_{15} = \mathrm{K}_3 \otimes \mathrm{K}_5$ is a candidate for local maximum of the Shannon entropy of order 15. If it is a local maximum, then the local maximum value is given by $37.584 < 15 \ln 15 \approx 40.621$. But the matrix T_{15} given in (2.5) is also a candidate for local maximum and if it is a local maximum, then the local maximum value $39.525 < 15 \ln 15 \approx 40.621$.

Next, we consider certain special orthogonal matrices that are stationary points of the optimization problem (3.1) and contain zero as an entry.

Let M_{n_1} and M_{n_2} be two orthogonal matrices, then it can be easily seen that their direct sum $\mathrm{M}_{n_1} \oplus \mathrm{M}_{n_2}$ is also an orthogonal matrix.

Proposition 5.11 *Let the orthogonal matrices* M_{n_1} *and* M_{n_2} *be stationary points of the maximization problem (3.7) of orders* n_1 *and* n_2*, respectively. Then* $\mathrm{M}_{n_1} \oplus \mathrm{M}_{n_2}$ *is a stationary point of the corresponding maximization problem of order* $n_1 + n_2$.

Proof Let $\mathrm{M}_{n_1} = (a_{ij})_{i,j=1}^{n}$, $\mathrm{M}_{n_2} = (b_{ij})_{i,j=1}^{n}$ and $\mathrm{M}_{n_1} \oplus \mathrm{M}_{n_2} = (c_{ij})_{i,j=1}^{n+m}$. By the definition of $\oplus$, we know that $c_{i,j} = \begin{cases} a_{ij} \text{ for } 1 \le i, j \le n_1, \\ b_{ij} \text{ for } n_1 + 1 \le i, j \le n_1 + n_2, \\ 0, \text{ otherwise.} \end{cases}$ Since M_{n_1} and M_{n_2} be stationary points of the maximization problem (3.7), from (3.10), we have

$$\eta_{ij} := \sum_{k=1}^{n} a_{ik} a_{jk} \ln\left(\frac{a_{jk}^2}{a_{ik}^2}\right) = 0 \text{ and } \theta_{gl} := \sum_{k=1}^{n} b_{gk} b_{lk} \left(\frac{b_{jk}^2}{b_{ik}^2}\right) = 0, \tag{5.29}$$

for all $1 \le i, j \le n_1$ and $1 \le g, l \le n_2$. We know that

$$\zeta_{ij} := \sum_{k=1}^{n_1+n_2} c_{ik} c_{jk} \ln\left(\frac{c_{jk}^2}{c_{ik}^2}\right) = \sum_{k=1}^{n_1} c_{ik} c_{jk} \ln\left(\frac{c_{jk}^2}{c_{ik}^2}\right) + \sum_{k=n_1+1}^{n_1+n_2} c_{ik} c_{jk} \ln\left(\frac{c_{jk}^2}{c_{ik}^2}\right) =: \zeta_{ij}^1 + \zeta_{ij}^2. \tag{5.30}$$

For $1 \le i, j \le n_1$, we get $\zeta_{ij}^1 = \eta_{ij} = 0$, $\zeta_{ij}^2 = 0$ and so $\zeta_{ij} = 0$. For $n_1 + 1 \le i, j \le n_1 + n_2$, we have $\zeta_{ij}^1 = 0$, $\zeta_{ij}^2 = \theta_{ij}$ and hence $\zeta_{ij} = 0$. For the cases, $1 \le i \le n_2$, $n_1 + 1 \le j \le n_1 + n_2$ and $n_1 + 1 \le i \le n_1 + n_2$, $1 \le j \le n_1$, it can be easily seen that $\zeta_{ij} = 0$, since at least one of the multiplying entry in ζ_{ij} is zero. Hence, $\mathrm{M}_{n_1} \oplus \mathrm{M}_{n_2}$ of order $n_1 + n_2$ is a stationary point of the maximization problem (3.7). □

Remark 5.12 It can be easily shown that

$$\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{n_1} \oplus \mathrm{M}_{n_2}] = \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{n_1}] + \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{n_2}].$$

Proposition 5.13 *If a conference matrix* C_n *exists, then the orthogonal matrix* $\mathrm{Q}_n := \frac{1}{\sqrt{n-1}}\mathrm{C}_n$ *constitutes a stationary point of the maximization problem (3.1).*

Proof Let n be even, and assume that a conference matrix $\mathrm{C}_n = (c_{ij})_{i,j=1}^{n}$ exists. Define $\mathrm{Q}_n := \frac{1}{\sqrt{n-1}}\mathrm{C}_n$. Since the entries of C_n are either 0 or ± 1, the derivative matrix is $\mathrm{G}_n = 2(\ln(n-1) - 1)\,\mathrm{Q}_n$. It follows that $\mathrm{A}_n = \mathrm{G}_n \mathrm{Q}_n^\top - \mathrm{Q}_n \mathrm{G}_n^\top = \mathbf{0}_n$. Therefore, the orthogonal matrix Q_n is a stationary point of the maximization problem (3.1). □

Remark 5.14 Note that $\mathscr{H}_1^{\mathbb{R}}[\mathrm{Q}_n] = n\ln(n-1)$. However, since C_n contains n zero entries, the sufficient condition (3.15) for a local maximum cannot be verified in this case.

Proposition 5.15 *If a weighing matrix* $\mathrm{W}_{n,k}$ *exists for some* $1 \le k \le n$*, then the orthogonal matrix* $\mathrm{R}_n := \frac{1}{\sqrt{k}}\mathrm{W}_{n,k}$ *is a stationary point to the maximization problem (3.1).*

Proof Let $\mathrm{R}_n := \frac{1}{\sqrt{k}}\mathrm{W}_{n,k}$ exist for some $1 \le k \le n$. The derivative matrix is $\mathrm{G}_n = 2(\ln(k) - 1)\,\mathrm{R}_n$, and we have $\mathrm{A}_n = \mathbf{0}_n$. Therefore, R_n is a stationary point of the maximization problem (3.1). □

Remark 5.16 It should be noted that $\mathscr{H}_1^{\mathbb{R}}[\mathrm{R}_n] = n\ln k$.

For $n \equiv 2 (\mathrm{mod}\ 4)$, then one can add a $\sqrt{n-2}\mathrm{I}_2$ diagonally to a Hadamard matrix of order $n-2$ and fix the rest of the entries 0 to obtain a new matrix S_n such that $\mathrm{S}_n\mathrm{S}_n^\top = (n-2)\mathrm{I}_n = \mathrm{S}_n^\top \mathrm{S}_n$. Therefore, $\frac{1}{\sqrt{n-2}}\mathrm{S}_n$ is an orthogonal matrix whose Shannon entropy is given by

$$\mathscr{H}_1^{\mathbb{R}}\left[\frac{1}{\sqrt{n-2}}\mathrm{S}_n\right] = (n-2)\ln(n-2).$$

The next result shows that $\frac{1}{\sqrt{n-2}}\mathrm{S}_n$ is also a stationary point to the maximization problem (3.1).

Proposition 5.17 *The orthogonal* $\frac{1}{\sqrt{n-2}}\mathrm{S}_n$ *is a stationary point to the maximization problem (3.1).*

Proof Let us define

$$\mathrm{B}_n = \frac{1}{\sqrt{n-2}} S_n.$$

By construction, B_n has two blocks:

- an $(n-2) \times (n-2)$ Hadamard block with entries $\pm\frac{1}{\sqrt{n-2}}$,
- a 2×2 identity block in the lower-right corner.

If $b_{ij} = \pm\frac{1}{\sqrt{n-2}}$, then

$$b_{ij}^2 = \tfrac{1}{n-2}, \qquad \ln\!\Big(\tfrac{1}{b_{ij}^2}\Big) = \ln(n-2).$$

Hence for G_n defined in (3.8), we have

$$\mathrm{G}_{ij} = 2\Big(\pm\tfrac{1}{\sqrt{n-2}}\Big)\big(\ln(n-2) - 1\big).$$

Thus the entire $(n-2) \times (n-2)$ block equals

$$\frac{2(\ln(n-2)-1)}{\sqrt{n-2}}\,\mathrm{H}_{n-2},$$

where H_{n-2} is the Hadamard matrix (with entries ± 1). For the two diagonal entries $b_{ii} = 1$ $(i = n-1, n)$,

$$\mathrm{G}_{ii} = 2 \cdot 1 \cdot \ln(1) - 2 \cdot 1 = -2.$$

Off-diagonal entries in the last two rows/columns vanish. Collecting the blocks, we obtain

$$\mathrm{G}_n = \begin{bmatrix} \frac{2(\ln(n-2)-1)}{\sqrt{n-2}}\,\mathrm{H}_{n-2} & 0 \\ 0 & -2\mathrm{I}_2 \end{bmatrix}.$$

For $\mathrm{A}_n = \mathrm{G}_n B_n^\top - B_n \mathrm{G}_n^\top$ (3.9), we immediately have $\mathrm{A}_n = \mathbf{0}_n$ and hence B_n is a stationary point to the maximization problem (3.1). □

Remark 5.18 For $n = 6$, we have $\mathscr{H}_1^{\mathbb{R}}[\frac{1}{2}\mathrm{S}_6] = 4\ln 4 < 6\ln 6 = \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_6]$ for the matrix M_6 given in (5.14). For $n = 10$, we obtain $\mathscr{H}_1^{\mathbb{R}}[\frac{1}{2\sqrt{2}}\mathrm{S}_{10}] = 8\ln 8 \approx 16.636 < 22.338 \approx \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{10}]$ for the matrix M_{10} given in (5.27). Similarly, for $n = 14$, we infer $\mathscr{H}_1^{\mathbb{R}}[\frac{1}{2\sqrt{3}}\mathrm{S}_{14}] = 12\ln 12 \approx 29.819 < 35.384 \approx \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{14}]$ for the matrix M_{14} given in (5.28). For $n = 18$, we infer $\mathscr{H}_1^{\mathbb{R}}[\frac{1}{4}\mathrm{S}_{18}] = 16\ln 16 \approx 37.361 < 50.998 \approx 18\ln 17 = \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_{18}]$ for the matrix $M_{18} = \frac{1}{\sqrt{17}}C_{18}$. Therefore, $\frac{1}{\sqrt{n-2}}S_n$ has Shannon entropy lesser than the matrices constructed before.Here's a polished rephrasing of your passage: In general, if a symmetric conference matrix of order n exists, then $n \equiv 2 \pmod 4$ ([27, Theorem 4.11]). In this case, it follows that

$$\mathscr{H}_1^{\mathbb{R}}\left[\frac{1}{\sqrt{n-2}}S_n\right] = (n-2)\ln(n-2) < n\ln(n-1) = \mathscr{H}_1^{\mathbb{R}}\left[\frac{1}{\sqrt{n-1}}C_n\right].$$

Remark 5.19 The study of local extrema of the Shannon entropy for various orthogonal matrices is purely mathematical. As noted above, the local extrema arising from symmetric design-type constructions achieve higher entropy values than the simple perturbations of a Hadamard matrix described in Proposition 5.17.

5.2 Construction of the entropy optimal complex inverse orthogonal matrix.

Let us now introduce the Shannon entropy for a complex inverse orthogonal matrix.

Definition 5.20 *(Shannon entropy of a complex inverse orthogonal matrix)* For an $n \times n$ inverse orthogonal matrix M_n with complex entries a_{ij} at the $(i, j)^{\text{th}}$ position, the *Shannon entropy* is defined as

$$\mathscr{H}_1^{\mathbb{C}}[M_n] = \sum_{i=1}^{n}\sum_{j=1}^{n} \left|a_{ij}\right|^2 \ln\left(\frac{1}{\left|a_{ij}\right|^2}\right), \tag{5.31}$$

where $\left|a_{ij}\right|$ denotes the absolute value of the complex number a_{ij}.

By applying Jensen's inequality, in the same manner as for real orthogonal matrices (see Theorem 5.1), we obtain analogous bounds on the entropy of a complex inverse orthogonal matrix M_n:

$$0 \le \mathscr{H}_1^{\mathbb{C}}[M_n] \le n\ln n. \tag{5.32}$$

This bound holds for all n, with equality if and only if $\left|a_{ij}\right|^2 = \frac{1}{n}$. Equivalently, in polar form, $a_{ij} = \frac{1}{\sqrt{n}}e^{i\theta}$, where $e^{i\theta}$ is an n^{th} root of unity.

(1) For $n = 2$, we have the following optimal matrix which achieves the bound $2\ln 2$.

$$M_2 = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \frac{1}{\sqrt{2}}F_2.$$

(2) For $n = 3$, we construct the optimal matrix as follows:

$$M_3 = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & \omega/\sqrt{3} & \omega^2/\sqrt{3} \\ 1/\sqrt{3} & \omega^2/\sqrt{3} & \omega/\sqrt{3} \end{bmatrix} = \frac{1}{\sqrt{3}}F_3,$$

where ω is the cube root of unity and for the matrix above, $\mathscr{H}_1^{\mathbb{C}}[M_3] = 3\ln 3$.

(3) For a general n, the optimal value $n\ln n$ is attained by the $n \times n$ scaled Vandermonde matrix V_n, defined as

$$M_n = \left(\frac{1}{\sqrt{n}}\right)V_n, \tag{5.33}$$

where V_n is given by the following matrix:

$$\mathrm{V}_n = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \alpha & \alpha^2 & \cdots & \alpha^{n-1} \\ 1 & \alpha^2 & \alpha^4 & \cdots & \alpha^{2n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & \alpha^{n-1} & \alpha^{2n-2} & \cdots & \alpha^{(n-1)^2} \end{bmatrix},$$

where α is the n^{th} root of unity ($\alpha \neq 1$). It is worth noting that the global maximum is attained by $\mathrm{M}_n = \frac{1}{\sqrt{n}} \mathrm{F}_n$, where F_n denotes the complex Hadamard matrix defined in (2.6). Thus, the optimal entropy is achievable for all n, and for the matrix in (5.33), we have $\mathscr{H}_1^{\mathbb{C}}[\mathrm{M}_n] = n \ln n$.

Remark 5.21 Let $\mathrm{M}_n = (a_{ij})_{i,j=1}^n$ be a unitary matrix, i.e., $\mathrm{M}_n\mathrm{M}_n^* = \mathrm{I}_n = \mathrm{M}_n^*\mathrm{M}_n$. The matrix given in (5.33) is also a global maximum for the Shannon entropy $\mathscr{H}_1^{\mathbb{C}}[\mathrm{M}_n] = \sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2 \ln\left(\frac{1}{|a_{ij}|^2}\right)$, for the unitary matrices.

6 Rényi, Tsallis and Sharma-Mittal Entropies

In this section, we introduce the Rényi, Tsallis, and Sharma-Mittal entropies for real orthogonal matrices and determine their global and local maxima for different matrix orders. The optimization problems associated with these entropies are closely connected to the concepts of almost Hadamard and p-almost Hadamard matrices (see [4–6, 16, 17], etc.).

6.1 Rényi entropy for a real orthogonal matrix

We now define the Rényi entropy for real orthogonal matrices and present its main properties.

Definition 6.1 *(Rényi entropy for a real orthogonal matrix)* For an $n \times n$ orthogonal matrix M_n with real entries a_{ij}, the definition of the *Rényi entropy* is given by

$$\mathscr{H}_\alpha^{\mathbb{R}}[\mathrm{M}_n] = \frac{n}{1-\alpha} \ln \left\{ \frac{1}{n} \sum_{i,j=1}^{n} \left(a_{ij}^2\right)^\alpha \right\}, \tag{6.1}$$

for all $0 \leq \alpha \leq \infty$ and $\alpha \neq 1$.

For $\alpha = 0$ in (6.1), we get the *Hartley or max-entropy* defined by

$$\mathscr{H}_0^{\mathbb{R}}[\mathrm{M}_n] = n \ln n = n \ln(\mathrm{o}(\mathrm{M}_n)),$$

where $\mathrm{o}(\mathrm{M}_n)$ is the order of the matrix M_n. As $\alpha \to 1$ in (6.1), by using the *L'Hospital's rule*, we have

$$\begin{aligned}\lim_{\alpha\to 1} \mathscr{H}_\alpha^{\mathbb{R}}[\mathrm{M}_n] &= \lim_{\alpha\to 1} \frac{n}{1-\alpha} \ln \left\{ \frac{1}{n} \sum_{i,j=1}^{n} \left(a_{ij}^2\right)^\alpha \right\} \\ &= -n \lim_{\alpha\to 1} \frac{1}{\frac{1}{n}\sum_{i,j=1}^{n}\left(a_{ij}^2\right)^\alpha} \frac{1}{n} \sum_{i,j=1}^{n} \left(a_{ij}^2\right)^\alpha \ln(a_{ij}^2) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2 \ln\left(\frac{1}{a_{ij}^2}\right) = \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_n],\end{aligned}$$

which is the *Shannon entropy*. For $\alpha = +\infty$ in (6.1), we get the *min-entropy* defined by

$$\mathscr{H}_{\infty}^{\mathbb{R}}[\mathrm{M}_n] = n \min_{1 \le i, j \le n} \left(-\ln a_{ij}^2 \right) = -n \left(\max_{1 \le i, j \le n} \ln a_{ij}^2 \right) = -n \ln \max_{1 \le i, j \le n} a_{ij}^2.$$

Definition 6.2 *(Collision entropy or quadratic Rényi entropy)* For an $n \times n$ orthogonal matrix M_n with real entries a_{ij} at the $(i, j)^{\text{th}}$ position, the definition of the *Collision entropy or quadratic Rényi entropy* is given by

$$\mathscr{H}_{2}^{\mathbb{R}}[\mathrm{M}_n] = -n \ln \left\{ \frac{1}{n} \sum_{i,j=1}^{n} a_{ij}^4 \right\}. \tag{6.2}$$

If a real Hadamard matrix H_n exists, then by Theorem 5.1, the matrix $\mathrm{M}_n = \frac{1}{\sqrt{n}} \mathrm{H}_n$ achieves the maximum of the Shannon entropy and is therefore a global maximizer. It can similarly be shown that $\mathrm{M}_n = \frac{1}{\sqrt{n}} \mathrm{H}_n$ also maximizes the Rényi entropy, making it a global maximum in this case as well. Consequently, the maximum value of the Rényi entropy is given by

$$\mathscr{H}_{\alpha}^{\mathbb{R}}[\mathrm{M}_n] = \frac{n}{1-\alpha} \ln \left\{ \frac{1}{n} \sum_{i,j=1}^{n} \left(\frac{1}{n} \right)^{\alpha} \right\} = \frac{n}{1-\alpha} \ln\{n^{(1-\alpha)}\} = n \ln n,$$

for all $\alpha \neq 1$.

Any 2×2 orthogonal matrix can be written as the form given in (4.5) and hence the Shannon and Rényi entropies become

$$\mathscr{H}[\mathrm{M}_2] = 2 \left[a^2 \ln \left(\frac{1}{a^2} \right) + (1 - a^2) \ln \left(\frac{1}{1 - a^2} \right) \right], \text{ and}$$
$$\mathscr{H}_{\alpha}[\mathrm{M}_2] = \frac{2}{1-\alpha} \ln \left[a^{2\alpha} + (1 - a^2)^{\alpha} \right].$$

6.2 Kullback-Leibler divergence and Rényi divergence for real orthogonal matrices

A notable property of the Rényi entropy is that it is a non-increasing function of α for any real orthogonal matrix M_n. Let us take the derivative of (6.1) with respect to α to obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\alpha} \left\{ \mathscr{H}_{\alpha}^{\mathbb{R}}[\mathrm{M}_n] \right\} &= \frac{n}{(1-\alpha)^2} \left[(1-\alpha) \frac{1}{\sum_{i,j=1}^{n} (a_{ij}^2)^{\alpha}} \sum_{i,j=1}^{n} (a_{ij}^2)^{\alpha} \ln(a_{ij}^2) + \ln \left(\frac{1}{n} \sum_{i,j=1}^{n} (a_{ij}^2)^{\alpha} \right) \right] \\
&= \frac{n}{(1-\alpha)^2} \left[\sum_{i,j=1}^{n} \frac{(a_{ij}^2)^{\alpha}}{\sum_{k,l=1}^{n} (a_{kl}^2)^{\alpha}} \ln \left(\left(a_{ij}^2 \right)^{1-\alpha} \right) \right. \\
&\quad \left. - \ln \left(\frac{1}{n} \sum_{i,j=1}^{n} \frac{(a_{ij}^2)^{\alpha}}{\frac{1}{n} \sum_{k,l=1}^{n} (a_{kl}^2)^{\alpha}} (a_{ij}^2)^{1-\alpha} \right) \right] \\
&= \frac{n}{(1-\alpha)^2} \left[\frac{1}{n} \sum_{i,j=1}^{n} y_{ij}^2 \ln \left(\left(a_{ij}^2 \right)^{1-\alpha} \right) - \ln \left(\frac{1}{n} \sum_{i,j=1}^{n} y_{ij}^2 (a_{ij}^2)^{1-\alpha} \right) \right],
\end{aligned} \tag{6.3}$$

where

$$y_{ij}^2 := \frac{(a_{ij}^2)^\alpha}{\frac{1}{n}\sum_{k,l=1}^n (a_{kl}^2)^\alpha} \text{ so that } \frac{1}{n}\sum_{k,l=1}^n (a_{kl}^2)^\alpha = \frac{(a_{ij}^2)^\alpha}{y_{ij}^2}. \tag{6.4}$$

Note also that $\frac{1}{n}\sum\limits_{i,j=1}^n y_{ij}^2 = 1$ and $\frac{1}{n}\sum\limits_{i,j=1}^n y_{ij}^2 (a_{ij}^2)^{1-\alpha} = \frac{1}{\frac{1}{n}\sum\limits_{k,l=1}^n (a_{lk}^2)^\alpha}$. From (6.3), we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\alpha}\left\{\mathscr{H}_\alpha^{\mathbb{R}}[\mathrm{M}_n]\right\} &= \frac{n}{(1-\alpha)^2}\left[\frac{1}{n}\sum_{i,j=1}^n y_{ij}^2 \ln\left(\left(a_{ij}^2\right)^{1-\alpha}\right) + \ln\left(\frac{(a_{ij}^2)^\alpha}{y_{ij}^2}\right)\right] \\ &= -\frac{1}{(1-\alpha)^2}\sum_{i,j=1}^n y_{ij}^2 \ln\left(\frac{y_{ij}^2}{a_{ij}^2}\right),\end{aligned} \tag{6.5}$$

and $\frac{1}{(1-\alpha)^2}\sum\limits_{i,j=1}^n y_{ij}^2 \ln\left(\frac{y_{ij}^2}{a_{ij}^2}\right)$ is proportional to *Kullback-Leibler divergence* or *relative entropy* (which is always non-negative) for real orthogonal matrices. This immediately gives a comparison between different kinds of entropies as

$$n\ln n = \mathscr{H}_0^{\mathbb{R}}[\mathrm{M}_n] \geq \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_n] \geq \mathscr{H}_2^{\mathbb{R}}[\mathrm{M}_n] \geq \mathscr{H}_\infty^{\mathbb{R}}[\mathrm{M}_n].$$

Definition 6.3 For two real orthogonal matrices $\mathrm{M}_n = (a_{ij})_{i,j=1}^n$ and $\mathrm{K}_n = (b_{ij})_{i,j=1}^n$, the *Kullback-Leibler divergence* is defined by

$$\mathbf{D}_{\mathrm{KL}}\left[\mathrm{M}_n \middle\| \mathrm{K}_n\right] = \sum_{i,j=1}^n a_{ij}^2 \ln\left(\frac{a_{ij}^2}{b_{ij}^2}\right). \tag{6.6}$$

The non-negativity of the Kullback-Leibler divergence can be established using the property of the natural logarithm that $\ln x \leq x-1$ for all $x>0$, with equality holding if and only if $x=1$. Thus, for the relative entropy, we have

$$-\sum_{i,j=1}^n a_{ij}^2 \ln\left(\frac{b_{ij}^2}{a_{ij}^2}\right) \geq -\sum_{i,j=1}^n a_{ij}^2\left(\frac{b_{ij}^2}{a_{ij}^2}-1\right) = -\sum_{i,j=1}^n b_{ij}^2 + \sum_{i,j=1}^n a_{ij}^2 = -n+n = 0, \tag{6.7}$$

and $\mathbf{D}_{\mathrm{KL}}\left[\mathrm{M}_n \middle\| \mathrm{K}_n\right] = 0$ if and only if $|a_{ij}| = |b_{ij}|$ for all i and j. Hence from (6.5), we have $\frac{\mathrm{d}}{\mathrm{d}\alpha}\left\{\mathscr{H}_\alpha^{\mathbb{R}}[\mathrm{M}_n]\right\} \leq 0$, for all α. Moreover, from (6.6), we infer

$$\mathbf{D}_{\mathrm{KL}}\left[\mathrm{M}_n \middle\| \mathrm{K}_n\right] = \sum_{i,j=1}^n a_{ij}^2 \ln(a_{ij}^2) - \sum_{i,j=1}^n a_{ij}^2 \ln(b_{ij}^2) = -\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_n] + \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_n, \mathrm{K}_n], \tag{6.8}$$

where $\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_n, \mathrm{K}_n]$ is the *cross-entropy*. Thus, the Shannon entropy $\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_n]$ sets a minimum for the cross-entropy $\mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_n, \mathrm{K}_n]$ if and only if $|a_{ij}| = |b_{ij}|$ for all i and j.

Definition 6.4 For two real orthogonal matrices $\mathrm{M}_n = (a_{ij})_{i,j=1}^n$ and $\mathrm{K}_n = (b_{ij})_{i,j=1}^n$, the *Rényi divergence* is defined by

$$\mathbf{D}_\alpha^{\mathrm{R}}\left[\mathrm{M}_n \middle\| \mathrm{K}_n\right] = \frac{n}{\alpha-1} \ln \left(\frac{1}{n} \sum_{i,j=1}^n a_{ij}^2 \left(\frac{a_{ij}^2}{b_{ij}^2} \right)^{\alpha-1} \right), \tag{6.9}$$

for $\alpha \neq 1$.

By using L'Hospital's rule, it can be shown that

$$\lim_{\alpha\to 1} \mathbf{D}_\alpha^{\mathrm{R}}\left[\mathrm{M}_n \middle\| \mathrm{K}_n\right] = \sum_{i,j=1}^n a_{ij}^2 \ln \left(\frac{a_{ij}^2}{b_{ij}^2} \right) = \mathbf{D}_{\mathrm{KL}}\left[\mathrm{M}_n \middle\| \mathrm{K}_n\right].$$

If the real Hadamard matrix K_n exists, then the Rényi entropy can be expressed in terms of the Rényi divergence as

$$\mathscr{H}_\alpha^{\mathbb{R}}[\mathrm{M}_n] = \mathscr{H}_\alpha^{\mathbb{R}}[\mathrm{K}_n] - \mathbf{D}_\alpha^{\mathrm{R}}\left[\mathrm{M}_n \middle\| \mathrm{K}_n\right].$$

6.3 The optimization of Rényi entropy

For an orthogonal matrix $\mathrm{M}_n = (a_{ij})_{i,j=1}^n$, the maximization problem of Rényi entropy can be formulated as

$$\max_{\mathrm{M}_n\in\mathbb{R}^{n\times n}} \left[\frac{n}{1-\alpha} \ln \left\{ \frac{1}{n} \sum_{i,j=1}^n \left(a_{ij}^2\right)^\alpha \right\} \right], \text{ such that } \sum_{k=1}^n a_{ki}a_{kj} = \begin{cases} 1 & \text{if } i=j, \\ 0 & \text{if } i\neq j. \end{cases} \tag{6.10}$$

But for $\alpha > 1$, we know that

$$\max_{\mathrm{M}_n\in\mathbb{R}^{n\times n}} \left[\frac{n}{1-\alpha} \ln \left\{ \frac{1}{n} \sum_{i,j=1}^n \left(a_{ij}^2\right)^\alpha \right\} \right] = \frac{n}{1-\alpha} \min_{\mathrm{M}_n\in\mathbb{R}^{n\times n}} \ln \left\{ \frac{1}{n} \sum_{i,j=1}^n \left(a_{ij}^2\right)^\alpha \right\},$$

and minimizing a monotonically increasing function of a variable is equivalent to minimizing the variable itself. Thus, it is enough to consider the problem

$$\min_{\mathrm{M}_n\in\mathbb{R}^{n\times n}} \sum_{i,j=1}^n \left(a_{ij}^2\right)^\alpha, \text{ such that } \sum_{k=1}^n a_{ki}a_{kj} = \begin{cases} 1 & \text{if } i=j, \\ 0 & \text{if } i\neq j. \end{cases} \tag{6.11}$$

If $\mathrm{M}_n = (a_{ij})_{i,j=1}^n$ is a stationary point of the problem (6.11), then we have

$$\mathrm{A}_n = 2(\alpha+1) \left(\sum_{k=1}^n a_{ik}a_{jk} \left(\left(a_{ik}^2\right)^\alpha - \left(a_{jk}^2\right)^\alpha \right) \right)_{i,j=1}^n = \mathbf{0}. \tag{6.12}$$

Also if M_n minimizes the function given in (6.11), then we obtain

$$\xi_n^{i,j} = 2\alpha \sum_{k=1}^n \left\{ (2\alpha-1) \left[a_{ki}^2 \left(a_{kj}^2\right)^{\alpha-1} + a_{kj}^2 \left(a_{ki}^2\right)^{\alpha-1} \right] - \left[\left(a_{ki}^2\right)^\alpha + \left(a_{kj}^2\right)^\alpha \right] \right\} > 0, \tag{6.13}$$

for all $1 \leq i < j \leq n$. The following theorem establishes the global minimum of the Rényi entropy for certain values of n (see [16] for a related result concerning p-almost Hadamard matrices).

Theorem 6.5 *For $\alpha > 1$, the global minimum of problem (6.11) is $\frac{1}{n^{\alpha-2}}$, and this minimum is attained if and only if $a_{ij}^2 = \frac{1}{n}$, for all $1 \leq i, j \leq n$.*

In particular, if a real Hadamard matrix of order n exists, then $\mathrm{P}_n = \frac{1}{\sqrt{n}}\mathrm{H}_n$ serves as a global minimizer of problem (6.11).

Proof If a real Hadamard matrix H_n of order n exists, then it is clear that $\mathrm{P}_n = \frac{1}{\sqrt{n}}\mathrm{H}_n$ is a stationary point, since $a_{ij}^2 = \frac{1}{n}$ and the condition (6.12) is easily satisfied. From (6.13), we infer

$$\xi_n^{i,j} = 8\alpha(\alpha-1)\frac{1}{n^{\alpha-1}} > 0, \tag{6.14}$$

for all $\alpha > 1$ and $1 \leq i < j \leq n$, making M_n a candidate for local minimum.

Let us use Jensen's inequality to obtain

$$\mathcal{F}(\mathrm{M}_n) = \sum_{i,j=1}^{n} a_{ij}^{2\alpha} = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}^2 (a_{ij}^2)^{\alpha-1}\right) \geq \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}^4\right)^{\alpha-1}, \tag{6.15}$$

since $x^{\alpha-1}$ is a convex function for all $\alpha > 1$. We consider the sets:

$$\mathcal{X} := \left\{\mathrm{X} = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n\times n} : \sum_{j=1}^{n} a_{ij}^2 = 1,\ \text{for } i = 1,\dots,n\right\},$$
$$\mathcal{Y} := \left\{\mathrm{X} = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n\times n} : \sum_{j=1}^{n} a_{ij}^2 = 1,\ \text{for } i = 1,\dots,n, \sum_{k=1}^{n} a_{ik}a_{jk} = 0,\ \text{for } i \neq j\right\}.$$

It is clear that the set $\mathcal{Y} \subsetneq \mathcal{X}$. Now for each fixed $1 \leq i \leq n$, we consider the problem:

$$\min_{\mathrm{X}\in\mathbb{R}^{n\times n}} \sum_{j=1}^{n} |a_{ij}|^4 \ \text{ subject to } \ \sum_{j=1}^{n} a_{ij}^2 = 1. \tag{6.16}$$

In other words, we are minimizing over the set $\mathcal{X}$. By applying the method of Lagrange multipliers, one finds that a stationary point of problem (6.16) is given by a matrix $\mathrm{M}_n = (a_{ij})_{i,j=1}^n \in \mathcal{A}$, where $|a_{ij}|^2 = \frac{1}{n}$, for all $1 \leq j \leq n$. This point is a local minimum, since the Hessian matrix is

$$\mathcal{H} = 8\,\mathrm{diag}(|a_{i1}|^2, \dots, |a_{in}|^2) = \frac{8}{n}\mathrm{I}_n,$$

which is positive definite. Moreover, it is also a global minimum, as both the cost functional and the constraint are convex. Since $\mathcal{Y} \subsetneq \mathcal{X}$ and $\mathrm{M}_n \in \mathcal{Y}$, it follows that a global minimum for this problem also exists within $\mathcal{Y}$. Therefore, for any matrix $\mathrm{M}_n = (a_{ij})_{i,j=1}^n \in \mathcal{Y}$, we have

$$\mathcal{F}(\mathrm{M}_n) = \sum_{i,j=1}^{n} |a_{ij}|^{2\alpha} \geq \sum_{i=1}^{n}\left(\sum_{j=1}^{n}\frac{1}{n^2}\right)^{\alpha-1} = \frac{1}{n^{\alpha-2}} = \mathcal{F}(\mathrm{P}_n).$$

Therefore, P_n is a global minimum for the Rényi entropy. ◻

Proposition 6.6 *If a real conference matrix* C_n *of order* n *exists, then the matrix* $\mathrm{Q}_n := \frac{1}{\sqrt{n-1}}\mathrm{C}_n$ *satisfies the sufficient condition* (6.13) *for local minimum for the problem (6.11) for* $n \geq 3$.

Proof Let n be even and a real conference matrix C_n of order n exists and let $\mathrm{Q}_n = \frac{1}{\sqrt{n-1}}\mathrm{C}_n$. Since the entries of C_n are $0, \pm 1$ and 0 occurs only along the diagonal, it is immediate that (6.12) is satisfied, making Q_n a stationary point. Note that Q_n is a global maximum for $n = 2$ and for $n \geq 3$, we have

$$\xi_n^{i,j} = \frac{4\alpha}{(n-1)^\alpha}\left[(2\alpha - 1)n - 4\alpha + 3\right] > 0 \text{ for } n \geq 3, \text{ and } \xi_n^{i,j} < 0 \text{ for } n = 2.$$

Note that $\xi_n^{i,j} > 0$ only if $n > \frac{4\alpha-3}{2\alpha-1}$ and $\lim_{\alpha\to\infty} \frac{4\alpha-3}{2\alpha-1} = 2$. For this case, if $\mathrm{Q}_n := \frac{1}{\sqrt{n-1}}\mathrm{C}_n$ is a local minimum, then the local minimum to the problem (6.11) is given by $\frac{n}{(n-1)^{\alpha-1}}$. ◻

Remark 6.7 If a real conference matrix C_n exists, then the matrix Q_n is a candidate for a local maximum of the Rényi entropy. In that case, the local maximum value is

$$\mathscr{H}_\alpha^{\mathbb{R}}[\mathrm{C}_n] = \frac{n}{1-\alpha} \ln\left\{\frac{1}{n}\sum_{i,j=1}^{n}\left(\frac{1}{n-1}\right)^\alpha\right\} = \frac{n}{1-\alpha}\ln\left((n-1)^{1-\alpha}\right) = n\ln(n-1),$$

for $\alpha \neq 1$. By analogy, one can also conclude that Q_n is a candidate for a local maximum of the Shannon entropy when $n \geq 3$.

Proposition 6.8 *Suppose a weighing matrix* $\mathrm{W}_{n,k}$ *exists for some* $1 \leq k \leq n$. *Then the associated orthogonal matrix*

$$\mathrm{R}_n := \frac{1}{\sqrt{k}}\mathrm{W}_{n,k}, \quad \mathrm{W}_{n,k} = (w_{ij})_{i,j=1}^n,$$

is a stationary point of the minimization problem (6.11).

Furthermore, for

$$\beta_{ij} = \sum_{l=1}^{n} w_{li}^2 w_{lj}^2, \quad \underline{\beta} = \min_{1\leq i<j\leq n} \beta_{ij}, \quad \overline{\beta} = \max_{1\leq i<j\leq n} \beta_{ij},$$

the following hold:

- *The sufficient condition* (6.13) *for a local minimum is satisfied if* $\underline{\beta} > \frac{k}{2\alpha-1}$.
- *The sufficient condition for a local maximum is satisfied if* $\overline{\beta} < \frac{k}{2\alpha-1}$.
- *The matrix* R_n *is a saddle point if there exist indices* i, j *such that simultaneously* $\beta_{ij} \geq \frac{k}{2\alpha-1}$ *and* $\beta_{ij} \leq \frac{k}{2\alpha-1}$.

Proof If a weighing matrix $\mathrm{W}_{n,k}$ exists, then by definition its entries are in $\{0, \pm 1\}$. Consequently, the scaled matrix $\mathrm{R}_n = \frac{1}{\sqrt{k}}\mathrm{W}_{n,k}$ is orthogonal and hence a stationary point of the minimization problem (6.11).

A direct computation shows that

$$\xi_n^{i,j} = \frac{4\alpha}{k^{\alpha-1}}\left(\frac{2\alpha-1}{k}\beta_{ij} - 1\right), \quad 1 \leq i < j \leq n,$$

from which the stated conditions for local minima, maxima, and saddle points follow. ◻

Proposition 6.9 *The matrix given in (5.13) is a stationary point of the minimization problem (6.11) and it satisfies the sufficient condition* (6.13) *for a local minimum if*

$$(2\alpha-1)\left[(n-2)^2 2^{2\alpha-2}+2^2(n-2)^{2\alpha-2}+(n-2)2^{2\alpha}\right]>(n-2)^{2\alpha}+2^{2\alpha}(n-1),$$

for all $\alpha>1$.

Proof For the orthogonal matrix given in (5.13), the matrix

$$\mathrm{A}_n=\mathrm{G}_n\mathrm{K}_n^\top-\mathrm{K}_n\mathrm{G}_n^\top=\mathbf{0}_n,$$

where $\mathrm{G}_n=\frac{2\alpha}{n^{2\alpha-1}}\begin{pmatrix}-(n-2)^{2\alpha-1} & 2^{2\alpha-1} & \cdots & 2^{2\alpha-1}\\ 2^{2\alpha-1} & -(n-2)^{2\alpha-1} & \cdots & 2^{2\alpha-1}\\ \vdots & \vdots & \ddots & \vdots\\ 2^{2\alpha-1} & 2^{2\alpha-1} & \cdots & -(n-2)^{2\alpha-1}\end{pmatrix}=\mathrm{G}_n^\top$ and $\mathrm{K}_n^\top=\mathrm{K}_n$ and $\mathrm{G}_n\mathrm{K}_n=\mathrm{K}_n\mathrm{G}_n$.
Thus, K_n is always a stationary point of the minimization problem (6.11). It can be easily seen that the sufficient condition (6.13) is satisfied if

$$\begin{aligned}\xi_n^{i,j}=\frac{4\alpha}{n^{2\alpha}}\Big\{&(2\alpha-1)\left[(n-2)^2 2^{2\alpha-2}+2^2(n-2)^{2\alpha-2}+(n-2)2^{2\alpha}\right]\\ &-(n-2)^{2\alpha}-2^{2\alpha}(n-1)\Big\}>0.\end{aligned}$$

If the matrix given in (5.13) is a local minimum, then the value is given by $\frac{(n-2)^{2\alpha}+(n-1)2^{2\alpha}}{n^{2\alpha-1}}$. □

Proposition 6.10 *Let a finite projective plane of order n exists. Then the orthogonal matrix corresponding to the finite projective plane satisfies the sufficient condition* (6.13) *for a local minimum for the optimization problem* (6.10) *of order* n^2+n+1 *for n and* α*, provided*

$$(n-\sqrt{n})^{2\alpha-2}<\frac{(2\alpha-1)(n-\sqrt{n})^2-1}{(n-\sqrt{n})^2-(2\alpha-1)}. \tag{6.17}$$

If it is a local minimum, then the local minimum value is given by

$$\mathscr{H}_\alpha^{\mathbb{R}}[\mathrm{M}_{n^2+n+1}]=\frac{1}{(n^2+n+1)^{2\alpha-1}}\left[(n+1)(n\sqrt{n}-1)^{2\alpha}+n^{2-\alpha}((n+1)+\sqrt{n})^{2\alpha}\right].$$

Proof We first assume that a finite projective plane of order n exists. Then there exists an orthogonal matrix $\mathrm{M}_{n^2+n+1}=(a_{ij})_{i,j=1}^{n^2+n+1}$ with the entries x and y given in (2.1) for $\lambda=1$. We also know that any two distinct rows contains $1\begin{pmatrix}x\\x\end{pmatrix}$, $n\begin{pmatrix}x\\y\end{pmatrix}$, $n\begin{pmatrix}y\\x\end{pmatrix}$ and $n(n-1)\begin{pmatrix}y\\y\end{pmatrix}$. Thus the first order necessary condition for optimality is satisfied, since for $i\neq j$

$$\begin{aligned}(\mathrm{A})_{i,j}=(\mathrm{GX}^\top-\mathrm{XG}^\top)_{i,j}&=\sum_{k=1}^{n^2+n+1}a_{ik}a_{jk}\left(\left(a_{ik}^2\right)^\alpha-\left(a_{jk}^2\right)^\alpha\right)\\ &=nxy(x^{2\alpha}-y^{2\alpha})+nyx(y^{2\alpha}-x^{2\alpha})=0.\end{aligned}$$

Thus the matrix M_{n^2+n+1} is a stationary point of the optimization problem (6.11). In order to obtain the sufficient condition (6.13), we consider

$$\begin{aligned}\xi^{i,j} &= 2\alpha\sum_{k=1}^{n}\left\{(2\alpha-1)\left[a_{ki}^2\left(a_{kj}^2\right)^{\alpha-1}+a_{kj}^2\left(a_{ki}^2\right)^{\alpha-1}\right]-\left[\left(a_{ki}^2\right)^{\alpha}+\left(a_{kj}^2\right)^{\alpha}\right]\right\}\\ &= 4\alpha n\left\{y^{2\alpha-2}[(2\alpha-1)x^2-y^2]-x^{2\alpha-2}[x^2-(2\alpha-1)y^2]\right\},\end{aligned} \tag{6.18}$$

for $i<j$, where x and y are defined in (2.1). We also know that the matrix M_{n^2+n+1} is a candidate for a local minimum of the optimization problem (6.11) if

$$y^{2\alpha-2}[(2\alpha-1)x^2-y^2]<x^{2\alpha-2}[x^2-(2\alpha-1)y^2] \text{ or } \left(\frac{y}{x}\right)^{2\alpha-2}<\frac{(2\alpha-1)\left(\frac{y}{x}\right)^2-1}{\left(\frac{y}{x}\right)^2-(2\alpha-1)}. \tag{6.19}$$

Substituting the values of x and y from (2.1) into (6.19), we obtain that the matrix M_{n^2+n+1} fulfills the sufficient condition (6.13) for being a local minimum of the optimization problem (6.11), provided that condition (6.17) is satisfied. □

Proposition 6.11 *Let a biplane of order n exists. Then the orthogonal matrix corresponding to the biplane satisfies the sufficient condition for a local minimum for the optimization problem* (6.10) *of order* $\frac{(n+1)(n+2)}{2}+1$ *for n and* α*, provided*

$$\left(\frac{n-\sqrt{n}}{2}\right)^{2\alpha-2}<\frac{(2\alpha-1)(n-\sqrt{n})^2-4}{(n-\sqrt{n})^2-4(2\alpha-1)}. \tag{6.20}$$

If it is a local minimum, then the local minimum value is given by

$$\begin{aligned}\mathscr{H}_{\alpha}^{\mathbb{R}}[\mathrm{M}_{\frac{(n+1)(n+2)}{2}+1}] = \frac{1}{2(n^2+3n+4)^{2\alpha-1}}\Bigg[&(n+2)((n+1)\sqrt{n}-2)^{2\alpha}\\ &+(n+1)\left(\frac{2}{n}\right)^{2\alpha-1}(n+(n+2)\sqrt{n})^{2\alpha}\Bigg].\end{aligned}$$

Proof Let us assume that a biplane of order n exists. Then, we have the existence of an orthogonal matrix $\mathrm{M}_{\frac{(n+1)(n+2)}{2}+1}=(a_{ij})_{i,j=1}^{\frac{(n+1)(n+2)}{2}+1}$ with the entries given in (2.1) for $\lambda=2$. We also know that any two distinct rows contain $2\begin{pmatrix}x\\x\end{pmatrix}, n\begin{pmatrix}x\\y\end{pmatrix}, n\begin{pmatrix}y\\x\end{pmatrix}$ and $\frac{n(n-1)}{2}\begin{pmatrix}y\\y\end{pmatrix}$. Thus, the first order necessary condition for optimality is easily satisfied, since for $i\neq j$

$$\begin{aligned}(\mathrm{A})_{i,j}=(\mathrm{G}\mathrm{X}^\top-\mathrm{X}\mathrm{G}^\top)_{i,j} &= \sum_{k=1}^{\frac{(n+1)(n+2)}{2}+1}a_{ik}a_{jk}\left(\left(a_{ik}^2\right)^{\alpha}-\left(a_{jk}^2\right)^{\alpha}\right)\\ &= nxy(x^{2\alpha}-y^{2\alpha})+nyx(y^{2\alpha}-x^{2\alpha})=0..\end{aligned}$$

It follows directly that the matrix $\mathrm{M}_{\frac{(n+1)(n+2)}{2}+1}$ is a stationary point of the optimization problem (3.1) of order $\frac{(n+1)(n+2)}{2}+1$. As shown in the proof of Proposition 6.10, the sufficient condition (6.13) holds whenever (6.19) is satisfied with x and y as defined in (2.1). Substituting these values of x and y, we conclude that $\mathrm{M}_{\frac{(n+1)(n+2)}{2}+1}$ is a candidate for a local minimum of the problem (6.11), provided condition (6.20) is met. □

Table 1 $1 \leq n \leq 20$.

n	Orthogonal matrix	Collision Entropy	n	Orthogonal Matrix	Collision Entropy
1	H_1	0	11	B_{11}	24.676
2	$\frac{1}{\sqrt{2}}H_2$	1.386	12	$\frac{1}{2\sqrt{3}}H_{12}$	29.819
3	K_3	2.694	13	$\mathcal{I}_{13}$	32.638
4	$\frac{1}{2}H_4$	5.545	14	$\frac{1}{13}C_{14}$	35.909
5	K_5	7.305	15	$T_{15}^{\dagger}$	38.655
6	$\frac{1}{\sqrt{5}}C_6^{\ddagger}$	9.657	16	H_{16}	44.361
7	$\mathcal{I}_7^{\dagger}$	12.318	17	$\frac{1}{3}W_{17,9}$	37.353
8	$\frac{1}{2\sqrt{2}}H_8$	16.636	18	$\frac{1}{17}C_{18}$	50.999
9	$K_3 \otimes K_3$	16.164	19	$\frac{1}{17}C_{18} \oplus H_1^{*}$	50.999
10	$\frac{1}{3}C_{10}$	21.972	20	$\frac{1}{2\sqrt{5}}H_{20}$	59.915

* This orthogonal matrix represents a saddle point rather than a local maximum (see Proposition 5.11). The weighing matrix $W_{19,9}$ exists for order 19, but it too corresponds to a saddle point, with a value of 41.747. However, if a quadruple plane of order 5 exists, the corresponding incidence matrix with entries $x = \frac{5+9\sqrt{5}}{95}$ and $y = \frac{1-2\sqrt{5}}{19}$ achieves a local maximum with a value of 53.83. †For order 7, a weighing matrix $W_{7,4}$ exists and $\frac{1}{2}W_{7,4}$ is also a local maximum. But the collision entropy value is 9.704. For order 15, the matrix $K_3 \otimes K_5$ is also a local maximum with the Collision entropy value 35.385. Note that a weighing matrix $W_{15,9}$ exist and $\frac{1}{3}W_{15,9}$ is also a local maximum. But the corresponding Collision entropy value is 32.958. ‡In this case, the matrices $\frac{1}{\sqrt{2}}H_2 \otimes K_3$, $\frac{1}{2}W_{6,4}$ and K_6 in (5.13) are also a local maxima of the collision entropy with values 9.547, 8.318 and 8.099, respectively. Remember that for Shannon entropy of order 6, the matrix $\frac{1}{\sqrt{2}}H_2 \otimes K_3$ gives the maximum value 9.949

Similar to Propositions 6.10 and 6.11, we can prove that the λ-plane, $\lambda \geq 1$, is also a candidate for a local minimum for the optimization problem (6.11) for certain values of λ.

Proposition 6.12 *Let a λ-plane of order n exists. Then the orthogonal matrix corresponding to the λ-plane satisfies the sufficient condition* (6.13) *for a local minimum for the optimization problem* (6.11) *of order* $1 + \frac{(n+\lambda)(n+\lambda-1)}{\lambda}$, *for λ, α and n, provided*

$$\left(\frac{n-\sqrt{n}}{\lambda}\right)^{2\alpha-2} < \frac{\lambda^2 - (2\alpha-1)(n-\sqrt{n})^2}{(2\alpha-1)\lambda^2 - (n-\sqrt{n})^2}. \tag{6.21}$$

As in the case of Shannon entropy, if M_{n_1} and M_{n_2} are orthogonal matrices of orders n_1 and n_2, respectively, which are stationary points of the minimization problem (6.11). Then the orthogonal matrices $M_{n_1} \otimes M_{n_2}$ and $M_{n_1} \oplus M_{n_2}$ are also stationary points of the minimization problem (6.11) of orders $n_1 n_2$ and $n_1 + n_2$ respectively. If $M_{n_1} = \frac{1}{\sqrt{n_1}}H_{n_1}$ and M_{n_2} is a local maximum for the Rény entropy of order n_2, then $M_{n_1} \otimes M_{n_2}$ is a candidate for local maximum for the Rény entropy of order $n_1 n_2$. Also, we have

$$\begin{aligned}\mathscr{H}_\alpha^{\mathbb{R}}[M_{n_1} \otimes M_{n_2}] &= n_2 \mathscr{H}_\alpha^{\mathbb{R}}[M_{n_1}] + n_1 \mathscr{H}_\alpha^{\mathbb{R}}[M_{n_1}],\\ \mathscr{H}_1^{\mathbb{R}}[M_{n_1} \oplus M_{n_2}] &= \mathscr{H}_\alpha^{\mathbb{R}}[M_{n_1}] + \mathscr{H}_\alpha^{\mathbb{R}}[M_{n_2}].\end{aligned}$$

The following table presents the maximal values (not necessarily global in all cases) of the *Collision entropy* ($\alpha = 2$ in (6.1)) for $1 \leq n \leq 20$. The matrices K_3 and K_5 are defined in (4.3) and (4.4), respectively, while $\mathcal{I}_7$, B_{11}, and $\mathcal{I}_{13}$ are specified in (2.2), (2.3), and (2.4), respectively.(See Table 1)

Remark 6.13 It should be noted that, when n is even and if a conference matrix C_n exists, then the orthogonal matrix $Q_n := \frac{1}{\sqrt{n-1}}C_n$ is getting closer to the optimal value $n \ln n$ for large values of n. In the odd case, for large n, the existence of weighing matrices also provide such results.

6.4 Tsallis entropy for real orthogonal matrices

We now introduce the Tsallis entropy for real orthogonal matrices and determine its global and local maxima for different orders.

Definition 6.14 *(Tsallis entropy for a real orthogonal matrix)* For an $n \times n$ orthogonal matrix M_n with real entries a_{ij} at the $(i, j)^{\text{th}}$ position, the definition of the *Tsallis entropy* is given by

$$\mathscr{S}_q^{\mathbb{R}}[\mathrm{M}_n] = \frac{n - \sum\limits_{i,j=1}^{n} (a_{ij}^2)^q}{q-1} = \frac{n}{q-1}\left[1 - \exp\left(\frac{1-q}{n}\mathscr{H}_q^{\mathbb{R}}[\mathrm{M}_n]\right)\right], \tag{6.22}$$

for all $0 < q \leq \infty$ (called the entropic-index) and $q \neq 1$.

As $q \to 1$, Tsallis entropy becomes Shannon entropy, since by using the L'Hospital's rule, we have

$$\begin{aligned}\lim_{q\to 1} \mathscr{S}_q^{\mathbb{R}}[\mathrm{M}_n] &= \lim_{q\to 1} \frac{n - \sum\limits_{i,j=1}^{n} (a_{ij}^2)^q}{q-1} = \lim_{q\to 1} -\sum_{i,j=1}^{n} (a_{ij}^2)^q \ln\left(a_{ij}^2\right)\\ &= \sum_{i,j=1}^{n} (a_{ij})^2 \ln\left(\frac{1}{a_{ij}^2}\right) = \mathscr{H}_1^{\mathbb{R}}[\mathrm{M}_n].\end{aligned}$$

If a real Hadamard matrix H_n exists, then the matrix $\frac{1}{\sqrt{n}}\mathrm{H}_n$ attains the maximum Shannon entropy. Moreover, it can be shown that a real Hadamard matrix also maximizes the Tsallis entropy. Consequently, the maximum value of the Tsallis entropy is given by

$$\mathscr{S}_q^{\mathbb{R}}[\mathrm{M}_n] = \frac{n - \sum\limits_{i,j=1}^{n} \left(\frac{1}{n}\right)^q}{q-1} = \frac{n - n^{2-q}}{q-1},$$

for all $q \neq 1$.

The optimization problem for the Tsallis entropy can be formulated in a manner similar to that of the Shannon and Rényi entropies. From the form of the Tsallis entropy given in (6.22), it is evident that its maximization problem is equivalent to that in (6.11). Therefore, the Rényi and Tsallis entropies share the same stationary points for all $\alpha > 1$ and $q > 1$, as well as the same local and global maxima.

Definition 6.15 For two real orthogonal matrices $\mathrm{M}_n = (a_{ij})_{i,j=1}^n$ and $\mathrm{K}_n = (b_{ij})_{i,j=1}^n$, the *Tsallis divergence* is defined by

$$\mathbf{D}_\alpha^{\mathrm{T}}\left[\mathrm{M}_n \| \mathrm{K}_n\right] = \frac{1}{\alpha-1}\sum_{i,j=1}^{n} a_{ij}^2 \left[\left(\frac{a_{ij}^2}{b_{ij}^2}\right)^{\alpha-1} - 1\right] = \frac{n}{\alpha-1}\left[\exp\left(\frac{\alpha-1}{n}\mathbf{D}_\alpha^{\mathrm{R}}\right) - 1\right], \tag{6.23}$$

for $\alpha \neq 1$.

Once again, by using L'Hospital's rule, it can be shown that

$$\lim_{\alpha\to 1} \mathbf{D}_\alpha^{\mathrm{T}}\left[\mathrm{M}_n \| \mathrm{K}_n\right] = \sum_{i,j=1}^{n} a_{ij}^2 \ln\left(\frac{a_{ij}^2}{b_{ij}^2}\right) = \mathbf{D}_{\mathrm{KL}}\left[\mathrm{M}_n \| \mathrm{K}_n\right].$$

6.5 Sharma-Mittal entropy for orthogonal matrices [25]

Let us now introduce the Sharma-Mittal entropy and examine its local and global maxima.

Definition 6.16 For an $n \times n$ orthogonal matrix $\mathrm{M}_n = (a_{ij})_{i,j=1}^n$, the definition of the *Sharma-Mittal entropy* is given by

$$\mathscr{T}_{\alpha,\beta}[\mathrm{M}_n] = \frac{n}{\beta - 1}\left[1 - \left(\frac{1}{n}\sum_{i,j=1}^{n}\left(a_{ij}^2\right)^{\alpha}\right)^{\frac{1-\beta}{1-\alpha}}\right], \tag{6.24}$$

for all $0 < \alpha, \beta \leq \infty$ and $\alpha, \beta \neq 1$.

Using L'Hospital's rule, it can be easily seen that

$$\mathscr{T}_{\alpha,\alpha}[\mathrm{M}_n] = \frac{n}{\alpha - 1}\left(1 - \frac{1}{n}\sum_{i,j=1}^{n}\left(a_{ij}^2\right)^{\alpha}\right) = \mathscr{S}_{\alpha}[\mathrm{M}_n],$$

$$\lim_{\beta\to 1} \mathscr{T}_{\alpha,\beta}[\mathrm{M}_n] = \frac{n}{1-\alpha}\ln\left(\frac{1}{n}\sum_{i,j=1}^{n}\left(a_{ij}^2\right)^{\alpha}\right) = \mathscr{H}_{\alpha}[\mathrm{M}_n], \text{ and}$$

$$\lim_{\alpha\to 1}\lim_{\beta\to 1} \mathscr{T}_{\alpha,\beta}[\mathrm{M}_n] = \sum_{i,j=1}^{n} a_{ij}^2 \ln\left(\frac{1}{a_{ij}^2}\right) = \mathscr{H}_1[\mathrm{M}_n].$$

The maximization problem of the Sharma-Mittal entropy is given by

$$\mathscr{J}[\mathrm{M}_n] = \max_{\mathrm{X}\in\mathbb{R}^{n\times n}}\left\{\frac{n}{\beta - 1}\left[1 - \left(\frac{1}{n}\sum_{i,j=1}^{n}\left(a_{ij}^2\right)^{\alpha}\right)^{\frac{1-\beta}{1-\alpha}}\right]\right\}, \text{ such that } \mathrm{X}\mathrm{X}^{\top} = \mathrm{I}_n. \tag{6.25}$$

But for $\alpha, \beta > 1$, we also know that

$$\max_{\mathrm{X}\in\mathbb{R}^{n\times n}}\left\{\frac{n}{\beta - 1}\left[1 - \left(\frac{1}{n}\sum_{i,j=1}^{n}\left(a_{ij}^2\right)^{\alpha}\right)^{\frac{1-\beta}{1-\alpha}}\right]\right\} = \frac{n}{\beta - 1}\left[1 - \min_{\mathrm{X}\in\mathbb{R}^{n\times n}}\left(\frac{1}{n}\sum_{i,j=1}^{n}\left(a_{ij}^2\right)^{\alpha}\right)^{\frac{1-\beta}{1-\alpha}}\right],$$

and since x^{γ} is an increasing function for any $\gamma > 0$. Thus, as in the case of Tsallis entropy, it is enough consider the problem given in (6.11). If a Hadamard matrix H_n exists, then the matrix $\mathrm{P}_n := \frac{1}{\sqrt{n}}\mathrm{H}_n$ is a global maximum for the Sharma-Mittal entropy and the global maximum value is given by $\frac{n}{\beta-1}\left(1 - n^{\frac{(1-\beta)}{(1-\alpha)^2}}\right)$.

Remark 6.17 The Rényi, Hartley, collision, min-, relative, Tsallis, and Sharma-Mittal entropies for complex inverse orthogonal matrices can be defined in an analogous manner. Since a complex Hadamard matrix F_n exists for all orders n, the optimal values for a complex inverse orthogonal matrix are as follows: for the Rényi entropy, the maximum is $n \ln n$; for the Tsallis entropy, the optimal value is $\frac{n - n^{2-q}}{q-1}$ for all $q \neq 1$; and for the Sharma-Mittal entropy, the optimal value is $\frac{n}{\beta-1}\left(1 - n^{\frac{1-\beta}{(1-\alpha)^2}}\right)$ for all $\alpha, \beta \neq 1$.

Acknowledgements The third author, Akhilesh Pathak, tragically passed away in October 2013. The authors dedicate this paper to Akhilesh, honoring his unquenchable spirit and remarkable creative insights. The first author's research was partially supported by grants from the AFOSR and NSF. The second author gratefully acknowledges the Air Force Office of Scientific Research (AFOSR) and the National Research Council (NRC) for the Research Associateship Award, as well as the Air Force Institute of Technology (AFIT) and the Indian Institute of Technology Roorkee (IIT Roorkee) for providing a stimulating scientific environment and essential resources. Finally, the authors sincerely thank the reviewer for their valuable comments and suggestions, which greatly enhanced the quality of this paper.

Declarations

Conflicts of Interest The authors declare no competing interests.

References

1. Arasu, K. T. and Mohan, M. T.: Optimization problems with orthogonal matrix constraints, *Numerical Algebra, Control and Optimization (NACO)*, **8**(4) (2018), 413–440.
2. Arasu, K. T. and Mohan, M. T.: Entropy of orthogonal matrices and minimum distance orthostochastic matrices from the uniform van der Waerden matrices, *Discrete Optim.*, **31** (2019), 115–144.
3. Anderson, T. W., Olkin, I, Underhill, L. G.: *Generation of Random Orthogonal Matrices*, Technical Report No. **6**, Stanford University, California, 1985.
4. Banica, T., Collins, B. and Schlenker, J.-M.:, On orthogonal matrices maximizing the 1-norm, *Indiana University Mathematics Journal*, **59** (2010) 839–856.
5. Banica, T. Nechita, I. and .Zyczkowski, K.:, Almost Hadamard matrices: general theory and examples, *Open Systems & Information Dynamics*, **19**(04) (2012) 1250024 (26p).
6. Banica, T. and Nechita, I.:, Almost Hadamard matrices: the case of arbitrary exponents, *Discrete Applied Mathematics*, **161**(16) (2013), 2367–2379.
7. Clarke, F.: Functional Analysis, Calculus of Variations and Optimal Control, Graduate Texts in Mathematics **264**, Springer, New York, 2013.
8. Colbourn, C. J. and Dinitz, J. H.: *Handbook of Combinatorial Designs*, Chapman and Hall CRC, 2006.
9. Craigen, R. Equivalence Classes of Inverse Orthogonal and Unit Hadamard Matrices, *Bull. Austral. Math. Soc.* **44** (1991), 109–115.
10. Delsarte, P., Goethals, J. M. and Seidel, J. J.: Orthogonal matrices with zero diagonal II, Canadian Journal of Mathematics. Vol. **XXXIII** (5) (1971), 816–832.
11. Gadiyar, H. G., Sangeeta Maini, K. M., Padma, R. and Sharatchandra, H S.: Entropy and Hadamard matrices, *Journal of Physics A: Mathematical and General*, **36**(7) (2003), 109–112.
12. Geramita, A. V., Geramita, J. M. and Seberry, J.: Orthogonal designs, *Linear and Multilinear Algebra*, **3** (1976), 381–206.
13. Horadam, K. J.: Hadamard matrices and Their Applications, Princeton University Press, 2007.
14. Kårhström, J.: *On Projective Planes*, Bachelor's Thesis, Mid Sweden University, 2002.
15. Keynes, J. M.: *IV - The Principle of Indifference*. A Treatise on Probability, **4**, Macmillan and Company, 44–69, 1921.
16. Mohan, M. T. On some p-almost Hadamard matrices, *Operators and Matrices*, **13**(1) (2019), 253–281.
17. Mohan, M. T. p-almost Hadamard matrices and λ-planes, *J. Algebr. Comb.*, **55** (2022), 89–108.
18. Nocedal, J. and Wright, S. J.: *Numerical Optimization. Springer Series in Operations Research and Financial Engineering*, bf 2nd edn. Springer, New York, 2006.
19. Paley, R. E. A. C. : On orthogonal matrices, Journal of Mathematics and Physics, **12** (1933), 311–320.
20. Pathak, A.: Entropy Optimal Orthogonal Matrices, Master of Science Thesis, Wright State University, 2012.
21. Perrott, X.: Existence of Projective Planes, https://arxiv.org/abs/1603.05333.
22. Pruss, A. R.: *The Principle of Sufficient Reason A Reassessment*, Cambridge University Press, NY, 2006.
23. Seberry, J. and Yamada, M.: Hadamard matrices, sequences and block designs, in "Contemporary Design Theory, A Collection of Surveys", J.H. Dinitz and D.R. Stinson, Eds., John Wiley and Sons, New York, 1992
24. Shannon, C. E.: A Mathematical Theory of Communication, *The Bell System Technical Journal*, **27** (3) (1948), 379–423.
25. Sharma, B. D. and Mittal, D. P.: New non-additive measures of entropy for discrete probability distributions, *Journal of Mathematical Sciences*, **10** (1975), 28–40.
26. Stewart, G. W.: The efficient generation of random orthogonal matrices with an application to condition estimators, *SIAM Journal on Numerical Analysis*, **17**(3) (1980), 403–409.
27. Stinson, D. R.: *Combinatorial Designs: Constructions and Analysis*, Springer-Verlag New York, Inc., 2004.
28. Sylvester, J. J.: Thoughts on inverse orthogonal matrices, simultaneous sign-successions, and tessellated pavements in two or more colors, with applications to Newton's rule, ornamental tile-work, and the theory of numbers, *Phil. Mag.* **34** (1867), 461–475.

29. Szöllősi, F.: Construction, Classification and parametrization of complex Hadamard matrices, PhD Thesis, 2011, arXiv:abs/1110.5590.
30. Wen, Z. and Yin, W.: A feasible method for optimization with orthogonality constraints, Math. Program., Ser. A **142** (2013), 397–434.
31. Wihler, T. P., Bessire, B. and Stefanov, A.: Computing the entropy of a large matrix, *Journal of Physics A: Mathematical and Theoretical*, **47**, 24, (2014), 245201.

Authors and Affiliations

K. T. Arasu[1] · Manil T. Mohan[2] · Akhilesh Pathak[3] · Ramya Ramachandran Janaki[4]

✉ Manil T. Mohan
maniltmohan@ma.iitr.ac.in ; maniltmohan@gmail.com

K. T. Arasu
karasu@riversideresearch.org

Ramya Ramachandran Janaki
ramya.janaki@gmail.com

1 Trusted & Resilient Systems, Riverside Research, 2640 Hibiscus Way, 45431, OH, Beavercreek, U.S.A.

2 Department of Mathematics, Indian Institute of Technology Roorkee - IIT Roorkee, Haridwar Highway, 247667 Roorkee, Uttarakhand, India

3 Department of Mathematics & Statistics, Wright State University, 3640 Colonel Glenn Highway, 45435, OH Dayton, U.S.A.

4 Department of Biomedical and Human Factors Engineering, Wright State University, 3640 Colonel Glenn Highway, 45435, OH Dayton, U.S.A.