# Birkhoff–James Orthogonality in the Trace Norm, with Applications to Quantum Resource Theories

Nathaniel Johnston[1,2]   Shirin Moein[1,2,3]   Rajesh Pereira[2]   and Sarah Plosker[2,3]

September 15, 2021

### Abstract

We develop numerous results that characterize when a complex Hermitian matrix is Birkhoff–James orthogonal, in the trace norm, to a (Hermitian) positive semidefinite matrix or set of positive semidefinite matrices. For example, we develop a simple-to-test criterion that determines which Hermitian matrices are Birkhoff–James orthogonal, in the trace norm, to the set of all positive semidefinite diagonal matrices. We then explore applications of our work in the theory of quantum resources. For example, we characterize exactly which quantum states have modified trace distance of coherence equal to 1 (the maximal possible value), and we establish a connection between the modified trace distance of 2-entanglement and the NPPT bound entanglement problem.

**Keywords:** Trace norm, Birkhoff–James orthogonality, quantum coherence, quantum entanglement, quantum resource theory

**MSC2010 Classification:** 15A03, 15A60, 15B57, 81P40

## 1   Introduction

Birkhoff–James orthogonality was introduced in [Jam47] to provide a definition of orthogonality in normed vector spaces that extends the usual one from inner product spaces. Necessary and sufficient conditions for Birkhoff–James orthogonality of matrices in the operator norm were given in [BŠ99], and a study of Birkhoff–James orthogonality of matrices in the Schatten $p$-norms as well as operator norm was performed in [LS02]. Birkhoff–James orthogonality of a given Hermitian matrix to every member of the subspace of real diagonal matrices, under the operator norm, was studied in [ALRV12].

In this work, we characterize when a given complex Hermitian matrix is Birkhoff–James orthogonal in the trace norm to a given (Hermitian) positive semidefinite matrix. We then explore numerous consequences of this main result, such as a simple-to-check characterization of when a given Hermitian matrix is Birkhoff–James orthogonal in the trace norm to every positive semidefinite diagonal matrix.

Our result has several natural applications in quantum information theory, where (mixed) quantum states, also known as density matrices, are positive semidefinite trace-one matrices. When working with a quantum resource theory, it is natural to ask how far a given quantum state is from a particular convex set of states of interest [Reg18], and the natural norm to use to measure distance is the trace norm [Wat18, Chapter 3]. For example, in the resource theory of quantum coherence [BCP14], which is of particular interest

---

[1]Department of Mathematics & Computer Science, Mount Allison University, Sackville, NB, Canada E4L 1E4
[2]Department of Mathematics & Statistics, University of Guelph, Guelph, ON, Canada N1G 2W1
[3]Department of Mathematics & Computer Science, Brandon University, Brandon, MB, Canada R7A 6A9

in quantum optics, quantum biology, and quantum thermodynamics [Gla63, Sud63], the aforementioned convex set of states is exactly those that are diagonal. In this case, our results concerning Birkhoff–James orthogonality to all positive semidefinite diagonal matrices provide a characterization of states that are a trace distance of 1 (the largest such distance possible) from that set.

## 1.1   Notation and Terminology

We now introduce the mathematical preliminaries that we need to discuss and prove our results concerning Birkhoff–James orthogonality. We defer a brief introduction to quantum information theory to Section 4, when we need it.

Let $M_n$ be the set of all $n \times n$ matrices with complex entries and let $M_n^+$ be the subset of them that are (Hermitian) positive semidefinite. We use bold lower case letters such as $\mathbf{v}$ and $\mathbf{w}$ to denote vectors in $\mathbb{C}^n$, with the entries of $\mathbf{v} \in \mathbb{C}^n$ denoted by $v_1, v_2, \ldots, v_n$. We denote the standard basis vectors by $\{\mathbf{e_j}\}_{j=1}^n$, and we use $A_{i,j}$ to denote the $(i,j)$-entry of a matrix $A \in M_n$. The eigenvalues of $A$ are denoted by $\lambda_j$ or by $\mu_j$, while the singular values of $A$—the non-negative square roots of the eigenvalues of $AA^*$—are denoted by $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n \geqslant 0$. If $A$ and $B$ are Hermitian matrices, we use the notation $A \succeq B$ (or $B \preceq A$) to denote $A - B$ being positive semidefinite).

The trace norm of $A$ is $\|A\|_{\mathrm{tr}} \overset{\text{def}}{=} \sum_{j=1}^n \sigma_j$ and the operator norm of $A$ is $\|A\| \overset{\text{def}}{=} \sigma_1$. If $A$ is Hermitian then its singular values are the absolute values of its eigenvalues, so $\|A\|_{\mathrm{tr}} = \sum_{j=1}^n |\lambda_j|$. When discussing norms in general, we use the notation $\|\!|\cdot|\!\|$, to avoid confusion with the operator norm. Given a Hermitian matrix $A$, we denote the eigenprojection matrices, projecting onto the direct sum of the eigenspaces corresponding to the positive, zero, and negative eigenvalues, respectively, by $P^+$, $P^0$, and $P^-$, and we note that $P^+ + P^0 + P^- = I$. For simplicity of terminology, we refer to these as the orthogonal projections onto the positive, zero, and negative eigenspaces of $A$ from now on.

As a generalization of orthogonality of vectors in a Hilbert space, we have the following notion of orthogonality in Banach spaces [Jam47]:

**Definition 1.** *Suppose $(\mathcal{X}, \|\!|\cdot|\!\|)$ is a Banach space over $\mathbb{R}$, and $A, B \in \mathcal{X}$. We say that $A$ is* Birkhoff–James orthogonal *to $B$ if*

$$\|\!|A|\!\| \leqslant \|\!|A + \lambda B|\!\| \quad \text{for all} \quad \lambda \in \mathbb{R}.$$

We note that this notion of orthogonality is homogeneous and additive, but not symmetric: if $A$ is Birkhoff–James orthogonal to $B$, then it is not necessarily the case that $B$ is Birkhoff–James orthogonal to $A$.

Here, we are interested in the case where a Hermitian matrix $A$ is Birkhoff–James orthogonal to under the trace norm $\|\cdot\|_{\mathrm{tr}}$ to a positive semidefinite matrix $B$. For this reason, from now on we simply use the terminology *Birkhoff–James orthogonal* to signify Birkhoff–James orthogonal *with respect to the trace norm*, and we note that $A$ being Birkhoff–James orthogonal to $B$ means exactly that

$$\|A\|_{\mathrm{tr}} \leqslant \|A + B\|_{\mathrm{tr}}.$$

Definition 1 of Birkhoff–James orthogonality works just fine if the field $\mathbb{R}$ is replaced by $\mathbb{C}$. For now (and only now), we briefly consider Birkhoff–James orthogonality in this sense, and recall the following result:

**Theorem 1.** *[LS02, Theorem 3.3] Let $A, B \in M_n$. Then the following are equivalent.*

  (a)  *$A$ is Birkhoff–James orthogonal to $B$ (in the trace norm).*

  (b)  *There exists a matrix $M \in M_n$ with $\|M\| \leqslant 1$ such that $\mathrm{Tr}(AM^*) = \|A\|_{\mathrm{tr}}$ and $\mathrm{Tr}(BM^*) = 0$.*

While this result in this form is stated for the complex vector space $M_n$, it applies straightforwardly to the real vector space of Hermitian matrices as well. Indeed, if $A$ is Hermitian and $\text{Tr}(AM^*)$ is real, then

$$\text{Tr}(AM^*) = \text{Tr}(AM) = \text{Tr}\left(A\left(\frac{M + M^*}{2}\right)\right),$$

and hence $M$ can be chosen to be Hermitian in this case. We thus only consider Birkhoff–James orthogonality in the *real* vector space of Hermitian matrices from now on.

It is perhaps worth noting that Theorem 1 can be proved in a different way than was done in [LS02]: via the subgradient [Wat92, Theorem 2], which is a useful tool in convex optimization. Indeed, the same is true of many of our upcoming results, but we avoid using the subgradient in this paper, opting instead for more elementary proofs.

## 1.2 Arrangement of the Paper

In Section 2, we present and prove our main result, which characterizes when a Hermitian matrix is Birkhoff–James orthogonal to a positive semidefinite one. In Section 3, we apply our main result to the special case of Birkhoff–James orthogonality to diagonal matrices. In particular, we characterize which Hermitian matrices are Birkhoff–James orthogonal to all positive semidefinite diagonal matrices, and we briefly consider the problem of Birkhoff–James orthogonality with the set of *all* (not necessarily positive semidefinite) diagonal matrices.

In Section 4, we introduce quantum resource theories and show how our results characterize which quantum states are "most resourceful" in the sense of the modified trace distance of that resource. Furthermore, in Section 4.4 we show that Birkhoff–James orthogonality of certain matrices is equivalent to a long-standard conjecture from quantum information theory, which states that a certain quantum state is "bound entangled". Finally, we close in Section 5 with some conclusions and open questions related to our work.

## 2 Main Result

We now explore the question of when a Hermitian matrix is Birkhoff–James orthogonal to a given positive semidefinite matrix. First, however, we need the following well-known inequality that bounds the trace of a product of Hermitian matrices (see [Bha97, Problem III.6.14], for example).

**Lemma 1.** *Suppose $H, M \in M_n$ are two Hermitian matrices with eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ and $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n$, respectively. Then*

$$\text{Tr}(HM) \leqslant \sum_{j=1}^{n} \lambda_j \mu_j,$$

*with equality if and only if there exists an orthonormal basis $\{\mathbf{v_j}\}_{j=1}^{n} \subset \mathbb{C}^n$ such that, for all $j$, $\mathbf{v_j}$ is an eigenvector of both $H$ and $M$ corresponding to $\lambda_j$ and $\mu_j$, respectively.*

We also need one more lemma before we will be able to state and prove our main result:

**Lemma 2.** *Suppose $H, M \in M_n$ are two Hermitian matrices, and let $P^+$ and $P^-$ denote the orthogonal projections onto the strictly positive and strictly negative eigenspaces of $H$, respectively. Then $2P^+ - I \preceq M \preceq I - 2P^-$ if and only if $\|M\| \leqslant 1$ and $\text{Tr}(HM) = \|H\|_{\text{tr}}$.*

*Proof.* Let $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ and $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n$ denote the eigenvalues of $H$ and $M$, respectively. Then $\|M\| \leqslant 1$ if and only if $-1 \leqslant \mu_j \leqslant 1$ for all $j$. If this condition holds then Lemma 1 tells us that

$$\text{Tr}(HM) \leqslant \sum_{j=1}^{n} \lambda_j \mu_j \leqslant \sum_{j=1}^{n} |\lambda_j| = \|H\|_{\text{tr}}.$$

We thus have $\text{Tr}(HM) = \|H\|_{\text{tr}}$ if and only if both of the previous two inequalities are equalities. The first inequality is an equality if and only if there exists an orthonormal basis $\{\mathbf{v_j}\}_{j=1}^{n} \subset \mathbb{C}^n$ such that, for all $j$, $\mathbf{v_j}$ is an eigenvector of both $H$ and $M$ corresponding to $\lambda_j$ and $\mu_j$, respectively. The second inequality is an equality if and only if $\mu_j = 1$ whenever $\lambda_j > 0$ and $\mu_j = -1$ whenever $\lambda_j < 0$. These two conditions together are equivalent to $2P^+ - I \preceq M \preceq I - 2P^-$, as claimed. $\qquad\square$

We now have enough machinery to state and prove our main result:

**Theorem 2.** *Suppose $H, B \in M_n$ are Hermitian and $B$ is positive semidefinite. Let $P^+$ and $P^-$ denote the orthogonal projections onto the strictly positive and negative eigenspaces of $H$, respectively. Then $H$ is Birkhoff–James orthogonal to $B$ in the trace norm if and only if both*

$$\text{Tr}(BP^+) \leqslant \frac{1}{2}\text{Tr}(B) \quad and \quad \text{Tr}(BP^-) \leqslant \frac{1}{2}\text{Tr}(B).$$

*Proof.* Theorem 1 and Lemma 2 together imply that $H$ is Birkhoff–James orthogonal to $B$ in the trace norm if and only if there exists a Hermitian matrix $M$ with $2P^+ - I \preceq M \preceq I - 2P^-$ and $\text{Tr}(BM) = 0$. These conditions imply that both $\text{Tr}(B(I - 2P^+)) \geqslant 0$ and $\text{Tr}(B(I - 2P^-)) \geqslant 0$, which is equivalent to $\text{Tr}(BP^+) \leqslant \frac{1}{2}\text{Tr}(B)$ and $\text{Tr}(BP^-) \leqslant \frac{1}{2}\text{Tr}(B)$ by the linearity of the trace.

Conversely, suppose we have both $\text{Tr}(BP^+) \leqslant \frac{1}{2}\text{Tr}(B)$ and $\text{Tr}(BP^-) \leqslant \frac{1}{2}\text{Tr}(B)$. If either $\text{Tr}(B(I - 2P^+)) = 0$ or $\text{Tr}(B(I - 2P^-)) = 0$, we choose $M = 2P^+ - I$ or $M = I - 2P^-$, respectively. On the other hand, if we have $\alpha := \text{Tr}(B(I - 2P^+)) > 0$ and $\beta := \text{Tr}(B(I - 2P^-)) > 0$, we choose

$$M = \frac{1}{\alpha + \beta}\big(\beta(2P^+ - I) + \alpha(I - 2P^-)\big).$$

In either case, we then have $2P^+ - I \preceq M \preceq I - 2P^-$ and $\text{Tr}(BM) = 0$, which means that $H$ is Birkhoff–James orthogonal to $B$ in the trace norm. $\qquad\square$

# 3 Birkhoff–James Orthogonality to Diagonal Matrices

We now consider the problem of determining which Hermitian matrices are Birkhoff–James orthogonal to every diagonal matrix. This problem was considered for the operator norm in [ALRV12], and was solved in small dimensions there, whereas we consider the trace norm version of it. We start with a result that determines when a Hermitian matrix is Birkhoff–James orthogonal in the trace norm to all *positive semidefinite* diagonal matrices:

**Theorem 3.** *Suppose $H \in M_n$ is Hermitian, and let $P^+$ and $P^-$ be the orthogonal projections onto its strictly positive and negative eigenspaces, respectively. Then the following are equivalent:*

  *(a) $H$ is Birkhoff–James orthogonal in the trace norm to every positive semidefinite diagonal matrix.*

  *(b) $H$ is Birkhoff–James orthogonal in the trace norm to $\mathbf{e_j}\mathbf{e_j^*}$ for all $1 \leqslant j \leqslant n$.*

  *(c) $P_{j,j}^+ \leqslant 1/2$ and $P_{j,j}^- \leqslant 1/2$ for all $1 \leqslant j \leqslant n$.*

*Proof.* The equivalence of (b) and (c) follows from choosing $B = \mathbf{e_j}\mathbf{e_j^*}$ in Theorem 2. The equivalence of (a) and (c) similarly follows from choosing $B$ to be an arbitrary diagonal positive semidefinite matrix scaled so that $\text{Tr}(B) = 1$ in Theorem 2. $\qquad\square$

We now start looking at the more difficult problem of characterizing which Hermitian matrices are Birkhoff–James orthogonal to *all* (not necessarily positive semidefinite) diagonal matrices. Theorem 3 provides a necessary condition, but it is not a sufficient one when $n \geqslant 3$, as demonstrated by the following example:

**Example 1.** *The Hermitian matrix*

$$H = \begin{bmatrix} -1 & 5 & 2 \\ 5 & -1 & 2 \\ 2 & 2 & 2 \end{bmatrix}$$

*has projections onto its negative and positive eigenspaces*

$$P^- = \frac{1}{2} \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad and \quad P^+ = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix},$$

*respectively. It follows that $P_{j,j}^- \leqslant 1/2$ and $P_{j,j}^+ \leqslant 1/2$ for all $1 \leqslant j \leqslant 3$, so Theorem 3 tells us that $H$ is Birkhoff–James orthogonal to every positive semidefinite diagonal matrix.*

*However, $H$ is not Birkhoff–James orthogonal to* every *diagonal matrix. To see this, we can compute $\|H\|_{\text{tr}} = 12$, but if $D = \text{diag}(6, 6, -6/5)$ then $\|H + D\|_{\text{tr}} = 54/5 < 12$ (and in fact, semidefinite programming can be used to show that this $D$ is optimal, so $H + D$ is Birkhoff–James orthogonal to every diagonal matrix).*

On the other hand, if $n = 2$ then the condition of Theorem 3 is both necessary *and* sufficient for Birkhoff–James orthogonality to every diagonal matrix. Our starting point towards proving this fact is the following simple corollary, which tells us that if a Hermitian matrix is invertible then Birkhoff–James orthogonality places very strong restrictions on its eigenvectors:

**Corollary 1.** *Suppose $H \in M_n$ is Hermitian and invertible, and let $P^+$ and $P^-$ be the orthogonal projections onto its positive and negative eigenspaces, respectively. The following are equivalent:*

*(a) $H$ is Birkhoff–James orthogonal in the trace norm to $\mathbf{e_j}\mathbf{e_j^*}$ for all $1 \leqslant j \leqslant n$.*

*(b) $P^+ = (I + U)/2$ and $P^- = (I - U)/2$ for some Hermitian unitary $U$ with zeros on its diagonal.*

*(c) $P_{j,j}^+ = P_{j,j}^- = 1/2$ for all $1 \leqslant j \leqslant n$.*

*Proof.* By Theorem 3, we know that $H$ is Birkhoff–James orthogonal to each $\mathbf{e_j}\mathbf{e_j^*}$ if and only if $P_{j,j}^+ \leqslant 1/2$ and $P_{j,j}^- \leqslant 1/2$ for all $1 \leqslant j \leqslant n$. Since $H$ is invertible, $P^+ + P^- = I$, so $P_{j,j}^+ + P_{j,j}^- = 1$ for all $j$. It follows that $P_{j,j}^+ \leqslant 1/2$ and $P_{j,j}^- \leqslant 1/2$ is equivalent to $P_{j,j}^+ = P_{j,j}^- = 1/2$, so (a) is equivalent to (c).

The fact that (b) implies (c) is trivial. To see that (c) implies (b), notice that (c) implies we can write $P^+ = I/2 + X$ for some matrix $X$ with diagonal entries equal to 0. Since $P^+$ is Hermitian, so is $X$, and since $P^+$ is a projection, we have $(I/2 + X)^2 = I/2 + X$, so $X^2 = I/4$, so $U := X/2$ is unitary. Then $P^+ = (I + U)/2$, and since $P^+ + P^- = I$ we must have $P^- = (I - U)/2$. $\qquad\square$

**Theorem 4.** *A Hermitian matrix $H \in M_2$ is Birkhoff–James orthogonal to* every *diagonal matrix if and only if it is Birkhoff–James orthogonal to $\mathbf{e_1}\mathbf{e_1^*}$ and $\mathbf{e_2}\mathbf{e_2^*}$.*

*Proof.* The "only if" direction follows immediately from Theorem 3, so we only prove the "if" direction. To this end, we suppose that $\mathbf{e_1}\mathbf{e_1^*}$ and $\mathbf{e_2}\mathbf{e_2^*}$ and then we split into two cases depending on whether $H$ is invertible or not:

If $H$ is invertible then Corollary 1 tells us that $P_{j,j}^+ = P_{j,j}^- = 1/2$ for $j = 1, 2$. The only $2 \times 2$ orthogonal projections with diagonal entries equal to $1/2$ are

$$\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

If we denote the positive and negative eigenvalues of $H$ by $\lambda_+$ and $\lambda_-$, respectively, then this tells us that either

$$H = \frac{1}{2}\begin{bmatrix} \lambda_+ + \lambda_- & \lambda_+ - \lambda_- \\ \lambda_+ - \lambda_- & \lambda_+ + \lambda_- \end{bmatrix} \quad \text{or} \quad H = \frac{1}{2}\begin{bmatrix} \lambda_- + \lambda_+ & \lambda_- - \lambda_+ \\ \lambda_- - \lambda_+ & \lambda_- + \lambda_+ \end{bmatrix}.$$

Depending on which of these cases occurs, we can choose $M = \pm(\mathbf{e_1}\mathbf{e_2^*} + \mathbf{e_2}\mathbf{e_1^*})$ as appropriate to see that $\mathrm{Tr}(HM) = |\lambda_+ - \lambda_-| = \|H\|_{\mathrm{tr}}$ and, if $D$ is any diagonal matrix, then $\mathrm{Tr}(DM) = 0$, so Theorem 1 tells us that $H$ is Birkhoff–James orthogonal to $D$.

On the other hand, if $H$ is not invertible then it has the form $H = \pm\mathbf{v}\mathbf{v}^*$ for some $\mathbf{v} \in \mathbb{C}^n$ (without loss of generality, we will assume that $H = \mathbf{v}\mathbf{v}^*$). Then $P^+ = \mathbf{v}\mathbf{v}/\|\mathbf{v}\|^2$, and we again see that if $P_{j,j}^+ \leqslant 1/2$ for $j = 1, 2$, then in fact we must have $P_{j,j}^+ = 1/2$ for $j = 1, 2$. It follows that either

$$H = \frac{\|\mathbf{v}\|^2}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{or} \quad H = \frac{\|\mathbf{v}\|^2}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Depending on which of these cases occurs, we can choose $M = \pm(\mathbf{e_1}\mathbf{e_2^*} + \mathbf{e_2}\mathbf{e_1^*})$ in Theorem 1 just like before to see that $H$ is Birkhoff–James orthogonal to every diagonal matrix. $\square$

We now present several corollaries of Theorem 3 that place restrictions on which Hermitian matrices are Birkhoff–James orthogonal to the diagonal matrices $\mathbf{e_j}\mathbf{e_j^*}$ $(1 \leqslant j \leqslant n)$. To start, we note that these conditions bound the inertia of such matrices quite strongly.

**Corollary 2.** *Suppose $H \in M_n$ is Hermitian and Birkhoff–James orthogonal to $\mathbf{e_j}\mathbf{e_j^*}$ for each $1 \leqslant j \leqslant n$, and let $\mu_+$ and $\mu_-$ denote how many eigenvalues it has that are positive and negative, respectively. Then*

$$\mu_- \leqslant n/2 \quad \text{and} \quad \mu_+ \leqslant n/2.$$

*Proof.* Theorem 3 tells us that the positive and negative eigenprojections $P^+$ and $P^-$ satisfy $P_{j,j}^+ \leqslant 1/2$ and $P_{j,j}^- \leqslant 1/2$ for all $1 \leqslant j \leqslant n$. Taking the trace, we obtain $\sum_{j=1}^n P_{j,j}^+ = \mu_+ \leqslant n/2$, and similarly for $\mu_-$. $\square$

**Corollary 3.** *Suppose $H \in M_n$ is Hermitian and $n$ is odd. If $H$ is Birkhoff–James orthogonal to $\mathbf{e_j}\mathbf{e_j^*}$ for each $1 \leqslant j \leqslant n$ then it must have $0$ as an eigenvalue.*

*Proof.* Let $\mu_+$, $\mu_-$, and $\mu_0$ denote how many eigenvalues $H$ has that are positive, negative, and $0$, respectively. Suppose (for the sake of establishing a contradiction) that $\mu_0 = 0$. Then $\mu_+ + \mu_0 + \mu_- = n$ implies that $\mu_+ + \mu_- = n$. However, Corollary 2 implies, since $n$ is odd, that $\mu_+ \leqslant (n-1)/2$ and similarly for $\mu_-$, which yields $\mu_+ + \mu_- \leqslant n - 1 < n$, a contradiction. $\square$

In the even-dimensional case, matrices that are Birkhoff–James orthogonal to every diagonal matrix do not necessarily have 0 as an eigenvalue. For example, Theorem 4 tells us that the invertible matrix

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

is Birkhoff–James orthogonal to every diagonal matrix.

In addition to the 2-dimensional case that we saw in Theorem 4, there are some other extra conditions that we can add to $H$ that ensure that the properties of Theorem 3 are equivalent to $H$ being Birkhoff–James orthogonal to *every* diagonal matrix. In particular, positive semidefiniteness of $H$ lets us increase the dimension a bit:

**Theorem 5.** *Suppose $H \in M_n^+$ is positive semidefinite and at least one of the following conditions hold: (a) $n \leqslant 4$, (b) $\mathrm{rank}(H) = 1$, (c) $\mathrm{rank}(H) = n/2$, or (d) the orthogonal projection onto $\mathrm{range}(H)$ has constant diagonal. Then $H$ is Birkhoff–James orthogonal to* every *diagonal matrix if and only if it is Birkhoff–James orthogonal to $\mathbf{e_j}\mathbf{e_j^*}$ for all $1 \leqslant j \leqslant n$.*

*Proof.* Since the "only if" implication of this theorem is trivial, we only prove the "if" direction.

We start with part (b). To see that the result holds when $\mathrm{rank}(H) = 1$, notice that in this case we can write $H = \mathbf{vv^*}$. By the same argument used in the proof of [JLP18, Theorem 2], there exists a matrix $M = (\mathbf{vv^*} - \mathbf{ww^*})/\|\mathbf{v}\|^2$ with all of its diagonal entries equal to 0, $|w_j| = |v_j|$ for all $1 \leqslant j \leqslant n$, and $\mathbf{v^*w} = 0$ (this matrix $M$ was called $Y$ in the proof of [JLP18, Theorem 2]). For an arbitrary diagonal matrix $D$, we thus have $\mathrm{Tr}(DM^*) = 0$, $\|M\| \leqslant 1$, and $\mathrm{Tr}(HM^*) = \mathbf{v^*}M\mathbf{v} = 1 = \|H\|_{\mathrm{tr}}$, so Theorem 1 tells us that $H$ is Birkhoff–James orthogonal to $D$.

To see that part (d) of the theorem holds, let $c$ be the constant diagonal value (i.e., if $P$ is the orthogonal projection onto $\mathrm{range}(H)$ then $c = P_{j,j}$ for all $j$). If we let $M = P/c - I$, then by Lemma 2 we have $\|M\| \leqslant 1$ and $\mathrm{Tr}(HM) = \|H\|_{\mathrm{tr}}$. Since all diagonal entries of $M$ are zero, $\mathrm{Tr}(MD) = 0$ for all diagonal matrices. It then follows from Theorem 1 that $H$ is Birkhoff–James orthogonal to every diagonal matrix.

Part (c) then holds because, in this case we must have $\mathrm{Tr}(P) = \mathrm{rank}(H) = n/2$. When $H$ is Birkhoff–James orthogonal to $\mathbf{e_j}\mathbf{e_j^*}$ for all $1 \leqslant j \leqslant n$, we must have all diagonal entries of $P$ exactly equal to $1/2$ to also satisfy both the trace condition and condition (c) of Theorem 3. Part (d) of this theorem then implies the result.

Finally, to see that part (a) of the theorem holds, recall that we already proved the $n = 2$ case in Theorem 4. If $n = 3$ then Corollary 2 tells us that $\mathrm{rank}(H) = \mu_+ \leqslant n/2 = 3/2$, so $\mathrm{rank}(H) = 1$, and we already proved the rank-1 case. If $n = 4$ then Corollary 2 tells us that $\mathrm{rank}(H) = \mu_+ \leqslant n/2 = 2$. If $\mathrm{rank}(H) = 1$ then we are again already done by part (b) of this theorem, and if $\mathrm{rank}(H) = 2$ then we are done by part (c) of this theorem. □

We note that Theorem 5 is in a sense tight: if all four of the conditions (a)–(d) fail, then the conclusion of the theorem needs not hold, as demonstrated by the next example that has $n = 5$ and $1 < \mathrm{rank}(H) = 2 < 5/2$.

**Example 2.** *Let $\alpha = \sqrt{(\sqrt{10} - 1)/3}$ and let $P \in M_5^+$ be the orthogonal projection onto the subspace*

$$\mathrm{span}\big((1, 1, 1, 0, 0), (0, 0, 1, \alpha, 1/\alpha)\big).$$

*Then the following claims are all straightforward to verify:*

1) *$P$ is positive semidefinite with rank 2;*

2) $P_{j,j} \leqslant 1/2$ *for all $j$ (with $P_{5,5} = 1/2$ exactly), so Theorem [3] tells us that it is Birkhoff–James orthogonal to every positive semidefinite diagonal matrix. However;*

3) *If $D = \mathrm{diag}(0, 0, 3, -1, 3)$ then $\|P - D/40\|_{\mathrm{tr}} \approx 1.99441235 < 2 = \|P\|_{\mathrm{tr}}$, so $P$ is not Birkhoff–James orthogonal to $D$.*

We now present a *sufficient* condition for a Hermitian matrix to be Birkhoff–James orthogonal to every diagonal matrix (in contrast with Theorem [3], which provides a necessary condition).

**Theorem 6.** *Suppose $H \in M_n$ is a Hermitian matrix. If there is a polar decomposition $H = UP$ in which the unitary $U$ is Hermitian and has all of its diagonal entries equal to $0$, then $H$ is Birkhoff–James orthogonal to every diagonal matrix. If $H$ is invertible then the converse also holds.*

*Proof.* Suppose $H = UP$ in which the unitary $U$ is a Hermitian matrix which has all of its diagonal entries equal to zero. Then $\|U\| = 1$, $\mathrm{Tr}(HU^*) = \|H\|_{\mathrm{tr}}$, and $\mathrm{Tr}(DU^*) = 0$ for every diagonal matrix $D$. It follows from Theorem [1] that $H$ is Birkhoff–James orthogonal to every diagonal matrix $D$.

To see that the converse holds when $H$ is invertible, recall from Theorem [1] that if $H$ is Birkhoff–James orthogonal to every diagonal matrix $D$ then there must exist a Hermitian matrix $M$ with diagonal entries all equal to $0$ such that $\|M\| \leqslant 1$ and $\mathrm{Tr}(HM) = \|H\|_{\mathrm{tr}}$. If we denote the eigenvalues of $H$ and $M$ by $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ and $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n$, respectively, then Lemma [1] tells us that the only way that, since $\mathrm{Tr}(HM) = \|H\|_{\mathrm{tr}} = \sum_{j=1}^{n} |\lambda_j|$, it must be the case that $M$ is diagonalized by the same orthonormal basis as $H$, and $\mu_j = 1$ whenever $\lambda_j > 0$ and $\mu_j = -1$ whenever $\lambda_j < 0$.

It follows that $M$ is unitary and $MH = |H|$, so we can choose $U = M$ as the Hermitian unitary matrix with diagonal entries equal to $0$ in the polar decomposition of $H$. $\qquad\square$

It is perhaps worth recalling that if $H$ is invertible then its polar decomposition is unique, so the necessary and sufficient condition of Theorem [6] is easy to check in this case.

## 4  Applications to Quantum Resource Theories

In quantum information theory, a *mixed quantum state* or *density matrix* is a positive semidefinite trace-one matrix, typically denoted by $\rho \in M_n$ or $\sigma \in M_n$. A *pure quantum state* is a unit vector in $\mathbb{C}^n$, which we denote using "bra-ket" notation: $|v\rangle \in \mathbb{C}^n$ is a unit column vector, while $\langle v| \stackrel{\text{def}}{=} |v\rangle^*$ is the corresponding (dual) row vector. From now on, whenever we use lowercase Greek letters like $\rho$, we are assuming that it is a quantum state normalized to have $\mathrm{Tr}(\rho) = 1$, and similarly if we use $|v\rangle$ then we are assuming it is a *unit* vector. For un-normalized matrices and vectors, we denote them like $A \in M_n^+$ and $\mathbf{v} \in \mathbb{C}^n$, just like in the earlier sections of this paper.

A quantum resource theory [CG19] consists of two things: (1) a closed convex cone of positive semidefinite matrices $C \subseteq M_n^+$, and (2) a set of linear maps that send $C$ back to itself (for our purposes, only (1) is relevant). The quantum states $\rho \in C$ (i.e., the members of $C$ with trace 1) are called "free states", and are thought of as the states that are "useless" for some given task in quantum information theory or quantum computation.

The most well-known resource theory is that of *entanglement*, where $C = C_{\mathrm{ent}}$ is the set of *separable* quantum states [GT09, HHHH09, Wer89]:

$$C_{\mathrm{ent}} \stackrel{\text{def}}{=} \left\{ \sum_j X_j \otimes Y_j : X_j \in M_m^+, Y_j \in M_n^+ \text{ for all } j \right\} \subseteq (M_m \otimes M_n)^+.$$

However, numerous other resource theories are studied as well [Reg18]. For example, another widely-used resource theory is that of *coherence* [Abe06, BCP14, Gla63, Sud63], in which the set of free states $C = C_{\text{coh}}$ consists of exactly those that are diagonal (in the standard basis):

$$C_{\text{coh}} \stackrel{\text{def}}{=} \left\{ \sum_j x_j \mathbf{e_j} \mathbf{e_j^*} : 0 \leqslant x_j \in \mathbb{R} \text{ for all } j \right\} \subseteq M_n^+.$$

In this context, the members of $C_{\text{coh}}$ are called *incoherent*.

Given a closed convex cone $C$, we can define the following *modified trace distance of $C$*, which simply measures the distance from a given quantum state to $C$:

$$D_C(\rho) \stackrel{\text{def}}{=} \min \left\{ \|\rho - X\|_{\text{tr}} : X \in C \right\}. \tag{1}$$

This quantity is a "proper" (i.e., physically relevant) measure of the resource specified by $C$ in the sense of [Reg18]. When $C = C_{\text{coh}}$, it is the *modified trace distance of coherence* [YZXT16], and when $C = C_{\text{ent}}$, it is the *modified trace distance of entanglement* [RFWG19]. For notational convenience, we omit the central level of subscripting when denoting this distance. For example, we denote the modified trace distance of coherence simply by $D_{\text{coh}}$, rather than $D_{C_{\text{coh}}}$.

As a bit of a historical note, we comment that an analogous (non-modified) *trace distance of $C$* was originally defined in the $C = C_{\text{coh}}$ case in [RPL16] via the minimization

$$\min \left\{ \|\rho - X\|_{\text{tr}} : X \in C, \text{Tr}(X) = 1 \right\}.$$

However, this quantity was subsequently shown to *not* be a "proper" measure of coherence (i.e., it lacked certain physically desirable properties that a quantification of a quantum resource should satisfy). The "modified" trace distance of $C$ was then introduced to fix this problem [YZXT16].

By applying our theorems concerning Birkhoff–James orthogonality (in particular, Theorem 2) in this setting, we get the following characterization of which mixed quantum states have $D_C(\rho) = 1$. These states are of interest because they are exactly those that are "most resourceful" or "most useful": it is clear from the definition of Equation (1) that it is always the case that $D_C(\rho) \leqslant 1$.

**Theorem 7.** *Suppose $C \subseteq M_n^+$ is a closed convex cone, $\rho \in M_n^+$ is a mixed quantum state, and $P$ is the orthogonal projection onto $\text{range}(\rho)$. The following are equivalent:*

*(a) $D_C(\rho) = 1$.*

*(b) $\rho$ is Birkhoff–James orthogonal in the trace norm to every member of $C$.*

*(c) $\text{Tr}(P\sigma) \leqslant 1/2$ for every mixed quantum state $\sigma \in C$.*

*Proof.* To see that (a) is equivalent to (b), note (by definition) that $\rho$ is Birkhoff–James orthogonal in the trace norm to every member of $C$ if and only if

$$1 = \|\rho\|_{\text{tr}} \leqslant \|\rho + \lambda X\|_{\text{tr}}$$

for all $\lambda \in \mathbb{R}$ and $X \in C$. However, we can safely ignore positive values of $\lambda$ since adding the positive semidefinite matrix $\lambda X$ to the positive semidefinite matrix $\rho$ can only increase its eigenvalues (and thus its trace norm). We can also safely ignore all negative values of $\lambda$ other than $\lambda = -1$, since $C$ is a cone so we can absorb $|\lambda|$ into $X$.

It follows that (b) is equivalent to $\|\rho - X\|_{\text{tr}} \geqslant 1$ for all $X \in C$, which (by Equation (1)) is equivalent to $D_C(\rho) \geqslant 1$. Since $D_C(\rho) \leqslant 1$ always holds, this establishes the claim that (a) and (b) are equivalent.

To see that (b) and (c) are equivalent, we recall from Theorem 2 that $\rho$ is Birkhoff–James orthogonal to every $X \in C$ if and only if $\mathrm{Tr}(XP) \leqslant \mathrm{Tr}(X)/2$. Since this inequality is scale-invariant, without loss of generality we may set $\sigma = X/\mathrm{Tr}(X)$ to see that (b) holds if and only if $\mathrm{Tr}(P\sigma) \leqslant 1/2$ for every mixed quantum state $\sigma \in C$. $\qquad \square$

Since $C$ is a convex cone, condition (c) of Theorem 7 can be weakened slightly to say that $\mathrm{Tr}(P\sigma) \leqslant 1/2$ for every mixed quantum state $\sigma$ that lies on an extreme ray of $C$. For all of the cones $C$ that we will explore (i.e., the cones defining $k$-coherence or $k$-entanglement), these extreme quantum states are exactly the pure states (i.e., the rank-1 matrices). So for our purposes, the conditions of Theorem 7 are furthermore equivalent to $\langle v|P|v \rangle \leqslant 1/2$ for every pure quantum state $|v\rangle\langle v| \in C$.

## 4.1 The Modified Trace Distance of Coherence

The modified trace distance of coherence was shown to be of limited use for pure states, in that it reaches its maximum value (i.e., $D_{\mathrm{coh}}(|v\rangle\langle v|) = 1$) on nearly all pure states (and therefore does not differentiate between these states) in [JLP18]. We now provide the natural generalization of this result to the case of arbitrary-rank (i.e., not necessarily pure) quantum states:

**Theorem 8.** *Suppose $\rho \in M_n^+$ is a mixed quantum state and $P$ is the orthogonal projection onto* $\mathrm{range}(\rho)$. *Then $D_{\mathrm{coh}}(\rho) = 1$ if and only if $P_{j,j} \leqslant 1/2$ for all $1 \leqslant j \leqslant n$.*

*Proof.* This follows immediately from the equivalence of conditions (a) and (c) in Theorem 7: the only pure quantum states $|v\rangle\langle v| \in C$ come from choosing $|v\rangle = \mathbf{e_j}$ for some $1 \leqslant j \leqslant n$, so $D_{\mathrm{coh}}(\rho) = 1$ if and only if $P_{j,j} = \mathbf{e_j^*} P \mathbf{e_j} \leqslant 1/2$ for all $1 \leqslant j \leqslant n$. $\qquad \square$

Indeed, the above theorem is a direct generalization of [JLP18, Theorem 2] from the rank-1 case to the general case (in the rank-1 case $\rho = |v\rangle\langle v|$, notice that $P_{j,j} = |v_j|^2$).

In addition to the condition of Theorem 8 being easy to check, it has the interesting consequence of bounding the possible rank of any state with $D_{\mathrm{coh}}(\rho) = 1$. Not only are low-rank states usually maximally coherent (in the sense of $D_{\mathrm{coh}}$), but high-rank states are *never* maximally coherent:

**Corollary 4.** *Suppose $\rho \in M_n^+$ is a mixed quantum state with $D_{\mathrm{coh}}(\rho) = 1$. Then $\mathrm{rank}(\rho) \leqslant n/2$.*

*Proof.* If $D_{\mathrm{coh}}(\rho) = 1$ then Theorem 8 tells us that the projection $P$ onto $\mathrm{range}(\rho)$ satisfies $P_{j,j} \leqslant 1/2$ for all $j$, so $\mathrm{Tr}(P) \leqslant n/2$. Since $\mathrm{Tr}(P) = \mathrm{rank}(P) = \mathrm{rank}(\rho)$, the result follows. $\qquad \square$

In particular, the above corollary tells us that all quantum states with maximal modified trace distance of coherence in the $n = 2$ and $n = 3$ cases are necessarily pure (i.e., rank-1). This is no longer true when $n \geqslant 4$, however.

## 4.2 k-Coherence

As a generalization of coherence, for an integer $1 \leqslant k \leqslant n$, the resource theory of *k-coherence* [RBC$^+$18] makes use of the following set $C = C_{\mathrm{k\text{-}coh}}$ as its free states:

$$C_{\mathrm{k\text{-}coh}} \stackrel{\mathrm{def}}{=} \left\{ \sum_j \mathbf{v_j v_j^*} : \text{each } \mathbf{v_j} \in \mathbb{C}^n \text{ has at most } k \text{ non-zero entries} \right\} \subseteq M_n^+.$$

In the $k = 1$ case, this reduces down to exactly the set of incoherent states (i.e., $C_{\text{1-coh}} = C_{\text{coh}}$), whereas the $k = n$ case simply reduces to the set of *all* mixed states (i.e., $C_{\text{n-coh}} = M_n^+$). For intermediate values of $k$, these sets satisfy the obvious chain of inclusions

$$C_{\text{coh}} = C_{\text{1-coh}} \subsetneq C_{\text{2-coh}} \subsetneq C_{\text{3-coh}} \subsetneq \cdots \subsetneq C_{\text{(n-1)-coh}} \subsetneq C_{\text{n-coh}} = M_n^+.$$

By using the fact that the pure states in $C_{\text{k-coh}}$ are the unit vectors with at most $k$ non-zero entries, we immediately get a characterization of the mixed states that maximize the modified trace distance of $k$-coherence:

**Theorem 9.** *Suppose $\rho \in M_n^+$ is a mixed quantum state and $P$ is the orthogonal projection onto* $\text{range}(\rho)$. *Then $D_{\text{k-coh}}(\rho) = 1$ if and only if the operator norm of every $k \times k$ principal submatrix of $P$ is at most $1/2$.*

*Proof.* This follows immediately from the equivalence of conditions (a) and (c) in Theorem 7: the pure quantum states $|v\rangle\langle v| \in C$ come from choosing the unit vectors $|v\rangle \in \mathbb{C}^n$ with at most $k$ non-zero entries, so $D_{\text{coh}}(\rho) = 1$ if and only if $\langle v|P|v\rangle \leqslant 1/2$ for all such vectors. By focusing on all the $|v\rangle$ that have their non-zero entries in particular locations, we see that this is equivalent to the principal submatrix of $P$ corresponding to those locations having operator norm at most $1/2$. $\qquad\square$

Indeed, the above theorem reduces to Theorem 8 in the special case when $k = 1$, since the $1 \times 1$ principal submatrices of $P$ are exactly its diagonal entries.

The following concept will be useful in what follows. Let $\|A\|_{(k)}$ denote the maximal operator norm of a $k \times k$ principal submatrix of $A$ (for example, if $k = 1$ then $\|A\|_{(k)}$ is the maximal diagonal entry of $A$). For $k > 1$, this is a norm (but that fact won't be important for the coming argument). We will present a useful characterization of $\|P\|_{(k)}$ when $P$ is a projection. First we need the following lemma.

**Lemma 3.** *If $P, Q \in M_n^+$ are orthogonal projections then $\|PQP\| = \|QPQ\|$.*

*Proof.* Since $PQP$ and $QPQ$ are both positive semidefinite matrices, their operator norms are equal to the their largest eigenvalues. We note that it is well-known that if $A$ and $B$ are two square matrices, then $AB$ and $BA$ have the same eigenvalues. Our result now follows from the fact that $PQP = PQ^2P = (PQ)(QP)$ and $QPQ = QP^2Q = (QP)(PQ)$. $\qquad\square$

While this fact will not be used in the paper, we note that $\|PQP\|$ has geometric significance being the square of the cosine of the first canonical angle between the range subspaces of $P$ and $Q$ (see [LPP15, KMP21] for applications of canonical angles to quantum information).

**Lemma 4.** *Let $P \in M_n$ be a projection and let $1 \leqslant k \leqslant n$. Then we have the equality*

$$\|P\|_{(k)} = \max \left\{ \sum_{j=1}^{k} |v_j^\downarrow|^2 : |v\rangle \in \text{range}(P) \right\},$$

*where $v_j^\downarrow$ is the entry of $|v\rangle$ with the j-th largest absolute value.*

*Proof.* Let $i_1, i_2, ..., i_k$ be the rows of the $k \times k$ principal submatrix of $P$ with the largest operator norm. Let $Q$ be the projection onto $\text{span}\{\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_k}\}$. Then $\|P\|_{(k)} = \|QPQ\| = \|PQP\|$. It is well known that the operator norm of a Hermitian matrix is equal to its numerical radius (i.e $\|H\| = \max_{\|x\|=1} |\langle x|H|x\rangle|$) and we apply this fact to $PQP$. Let $|v\rangle$ be the unit vector which maximizes $\langle x|PQP|x\rangle$. If $|v\rangle \notin \text{range}(P)$, we can replace $|x\rangle = |v\rangle$ with $|x\rangle = P|v\rangle/\|P|v\rangle\|$ giving us an even larger value for $\langle x|PQP|x\rangle$; hence $|v\rangle \in \text{range}(P)$. Therefore $\|P\|_{(k)} = \langle v|PQP|v\rangle = \langle v|Q|v\rangle = \sum_{j=1}^{k} |v_{i_j}|^2 = \sum_{j=1}^{k} |v_j^\downarrow|^2$. The final

equality follows from the fact that if $\{v_{i_j}\}_{j=1}^k$ were not the $k$ largest entries of $|v\rangle$ in absolute value, this would contradict our assumption that $i_1, i_2, \ldots, i_k$ are the rows of the $k \times k$ principal submatrix of $P$ with the largest operator norm. $\qquad \square$

We use the above lemmas to get a rank bound that generalizes Corollary 4 to this setting of $k$-coherence:

**Corollary 5.** *Suppose $\rho \in M_n^+$ is a mixed quantum state with $D_{\text{k-coh}}(\rho) = 1$. Then*

$$\text{rank}(\rho) \leqslant \frac{n(n+1-2k)}{2(n-k)}.$$

*Proof.* Theorem 9 tells us that if $D_{\text{k-coh}}(\rho) = 1$ then $\|P\|_{(k)} \leqslant 1/2$. On the other hand, we also have some lower bounds on $\|P\|_{(k)}$: it must be the case that $\|P\|_{(1)} \geqslant \text{Tr}(P)/n$, since otherwise the sum of the diagonal entries of $P$ would be strictly less than $n(\text{Tr}(P)/n) = \text{Tr}(P)$, which is a contradiction.

For larger values of $k$, we claim that $\|P\|_{(k)} \geqslant \|P\|_{(1)} + (1 - \|P\|_{(1)})(k-1)/(n-1)$. If we let $|w\rangle$ be a particular $|v\rangle$ attaining this supremum for $\|P\|_{(1)}$ given in Lemma 4, then we have

$$\sum_{j=2}^n \left|w_j^{\downarrow}\right|^2 = \sum_{j=1}^n \left|w_j^{\downarrow}\right|^2 - \|P\|_{(1)} = 1 - \|P\|_{(1)}.$$

Since the $\left|w_j^{\downarrow}\right|$'s are sorted in non-increasing order, this implies

$$\sum_{j=2}^k \left|w_j^{\downarrow}\right|^2 \geqslant \frac{k-1}{n-1}(1 - \|P\|_{(1)}).$$

Finally, this gives us

$$\|P\|_{(k)} \geqslant \sum_{j=1}^k \left|w_j^{\downarrow}\right|^2 = \|P\|_{(1)} + \sum_{j=2}^k \left|w_j^{\downarrow}\right|^2 \geqslant \|P\|_{(1)} + \frac{k-1}{n-1}(1 - \|P\|_{(1)}). \tag{2}$$

If we substitute the inequality $\|P\|_{(1)} \geqslant \text{Tr}(P)/n = \text{rank}(\rho)/n$ into Inequality (2), and then use the fact that $1/2 \geqslant \|P\|_{(k)}$ and solve for $\text{rank}(\rho)$, we get exactly the inequality given in the statement of the corollary. $\qquad \square$

In the special case when $k = 1$, Corollary 5 says that $D_{\text{coh}}(\rho) = 1$ implies $\text{rank}(\rho) \leqslant n/2$, agreeing with Corollary 4. At the other extreme, if $k > n/2$ then Corollary 5 says that $D_{\text{k-coh}}(\rho) = 1$ implies $\text{rank}(\rho) \leqslant 0$, which of course is impossible. It follows that if $D_{\text{k-coh}}(\rho) = 1$ then $k \leqslant n/2$. Furthermore, in the $k = n/2$ case if $D_{\text{k-coh}}(\rho) = 1$ then $\text{rank}(\rho) \leqslant 1$, so $\rho$ must be pure.

We end this section by noting that in the numerical linear algebra literature, there is a concept called the factor width of a matrix first defined in [BCPT05]. This paper deals entirely with real matrices, so we give the natural extension of the definition to complex matrices.

**Definition 2.** *[BCPT05, Definition 1] The factor width of a complex positive semidefinite matrix $A$ is the smallest integer $k$ for which there exists a complex (rectangular) matrix $V$ where $A = VV^*$ and each column of $V$ contains at most $k$ non-zero entries.*

It is easy to see that a density matrix is $k$-coherent if and only if it has factor width at most $k$. In [BCPT05], there is a characterization of matrices which have factor rank at most two. We first need the following definitions:

**Definition 3.** *[Var76] Let $A \in M_n$, then its comparison matrix (denoted $M(A)$) is an $n$ by $n$ matrix with $M(A)_{i,i} = |A_{i,i}|$ for all $i$ and $M(A)_{i,j} = -|A_{i,j}|$ whenever $i \neq j$.*

**Definition 4.** *A square matrix is said to be an H-matrix if its comparison matrix is positive semidefinite.*

We can now restate the main result of [BCPT05] in terms of 2-coherent states.

**Theorem 10.** *[BCPT05, Theorem 9] A density matrix is $2$-coherent if and only if it is an H-matrix.*

## 4.3   Entanglement and k-Entanglement

When we apply Theorem 7 to the resource theory of entanglement (i.e., when we take $C = C_{\text{ent}}$), we see that the modified trace distance of entanglement satisfies $D_{\text{ent}}(\rho) = 1$ if and only if $\langle v|P|v \rangle \leqslant 1/2$ for every pure *product* state $|v\rangle = |x\rangle \otimes |y\rangle \in \mathbb{C}^m \otimes \mathbb{C}^n$, where $P$ is the orthogonal projection onto range($\rho$). In the terminology of [JK10, JK11], this is equivalent to the statement that $\|P\|_{S(1)} \leqslant 1/2$, where $\|\cdot\|_{S(1)}$ is the norm defined by

$$\|X\|_{S(1)} \stackrel{\text{def}}{=} \max \left\{ (\langle x| \otimes \langle y|) X(|x\rangle \otimes |y\rangle) : |x\rangle \in \mathbb{C}^m, |y\rangle \in \mathbb{C}^n \right\}. \tag{3}$$

A natural generalization of entanglement is the resource theory of *k-entanglement*, in which the set $C = C_{\text{k-ent}}$ of free states consists of states with *Schmidt number* [TH00] at most $k$ (where $1 \leqslant k \leqslant \min\{m, n\}$ is an integer):

$$C_{\text{k-ent}} \stackrel{\text{def}}{=} \left\{ \sum_j \mathbf{v_j}\mathbf{v_j^*} : \text{there exist } \{\mathbf{x_{i,j}}\} \subset \mathbb{C}^m, \{\mathbf{y_{i,j}}\} \subset \mathbb{C}^n \text{ such that } \mathbf{v_j} = \sum_{i=1}^{k} \mathbf{x_{i,j}} \otimes \mathbf{y_{i,j}} \text{ for all } j \right\}.$$

By the spectral decomposition theorem, if $k = 1$ then we have $C_{\text{1-ent}} = C_{\text{ent}}$, whereas if $k = \min\{m, n\}$ then we have $C_{\min\{m,n\}\text{-ent}} = (M_m \otimes M_n)^+$. For intermediate values of $k$, these sets satisfy a chain of inclusions that is analogous to that of *k-coherence*:

$$C_{\text{ent}} = C_{\text{1-ent}} \subsetneqq C_{\text{2-ent}} \subsetneqq C_{\text{3-ent}} \subsetneqq \cdots \subsetneqq C_{(\min\{m,n\}-1)\text{-ent}} \subsetneqq C_{\min\{m,n\}\text{-ent}} = (M_m \otimes M_n)^+.$$

In this setting, the natural generalization of the norm defined in Equation (3) is

$$\|X\|_{S(k)} \stackrel{\text{def}}{=} \max \left\{ \langle v|X|v\rangle : \text{there exist } \{\mathbf{x_i}\} \subset \mathbb{C}^m, \{\mathbf{y_i}\} \subset \mathbb{C}^n \text{ such that } |v\rangle = \sum_{i=1}^{k} \mathbf{x_i} \otimes \mathbf{y_i} \right\}.$$

This quantity is also a norm, and an almost identical argument to the one from the start of this section demonstrates the following theorem:

**Theorem 11.** *Suppose $\rho \in (M_m \otimes M_n)^+$ is a mixed quantum state and $P$ is the orthogonal projection onto* range($\rho$). *Then $D_{\text{k-ent}}(\rho) = 1$ if and only if $\|P\|_{S(k)} \leqslant 1/2$.*

Unlike the case of coherence, the condition of Theorem 11 is difficult to check even in the $k = 1$ case. Indeed, computation of $\|\cdot\|_{S(1)}$ is NP-hard [Gha10, Gur03] (which is expected, since computation of $D_{\text{ent}}$ is also NP-hard). Despite this, numerous bounds on $\|P\|_{S(k)}$ are known, which let us derive rank bounds like the following:

**Corollary 6.** *Suppose $\rho \in (M_m \otimes M_n)^+$ is a mixed quantum state with $D_{\text{k-ent}}(\rho) = 1$. Then*

$$\text{rank}(\rho) \leqslant \frac{mn(\min\{m, n\} - 2k + 1)}{2(\min\{m, n\} - k)} \quad \text{and}$$

$$\text{rank}(\rho) \leqslant \frac{(n + m + 2 - 4k)^2 - (n - m)^2}{4}.$$

*Proof.* We use [JK10, Theorem 4.15], which says that every orthogonal projection $P$ satisfies the pair of inequalities

$$\|P\|_{S(k)} \geqslant \frac{(k-1)mn + (\min\{m,n\} - k)\mathrm{rank}(P)}{mn(\min\{m,n\} - 1)} \quad \text{and}$$

$$\|P\|_{S(k)} \geqslant \min\left\{1, \frac{k}{\left\lceil \frac{1}{2}\left(n + m - \sqrt{(n-m)^2 + 4\mathrm{rank}(P) - 4}\right)\right\rceil}\right\}.$$

If $D_{\text{k-ent}}(\rho) = 1$ then Theorem 11 tells us that $1/2 \geqslant \|P\|_{S(k)}$, which gives the claimed result after a bit of routine (but somewhat ugly) algebra. $\qquad\square$

In particular, the first inequality of Corollary 6 tells us that if $k > \min\{m,n\}/2$ then $\mathrm{rank}(\rho) \leqslant 0$, which impossible. It follows that there only exist quantum states $\rho$ with $D_{\text{k-ent}}(\rho) = 1$ when $k \leqslant \min\{m,n\}/2$. Furthermore, in the extreme case when $k = \min\{m,n\}/2$, the second inequality of Corollary 6 says that any state with $D_{\text{k-ent}}(\rho) = 1$ must have $\mathrm{rank}(\rho) \leqslant |n - m + 1|$ (so if we also have $m = n$ then $\rho$ must be pure). At the other extreme, if $k = 1$ then the first inequality of Corollary 6 is the stronger one, and it says that if $D_{\text{ent}}(\rho) = 1$ then $\mathrm{rank}(\rho) \leqslant mn/2$.

## 4.4 An Application to the NPPT Bound Entanglement Problem

Given a quantum state $\rho \in (M_m \otimes M_n)^+$, one often wants to know if it can be *distilled*: transformed into (an arbitrarily good approximation of) the *pure maximally entangled* state

$$|\phi_+\rangle \overset{\text{def}}{=} \frac{1}{\sqrt{\min\{m,n\}}} \left( \sum_{j=1}^{\min\{m,n\}} \mathbf{e_j} \otimes \mathbf{e_j} \right)$$

via local operations [HHH97]. In some cases where $\rho$ is undistillable, multiple tensor copies of it are distillable [Wat04], so we say that $\rho$ is *r-copy (un)distillable* if $\rho^{\otimes r} \in (M_m^{\otimes r} \otimes M_n^{\otimes r})^+$ is (un)distillable.

It is straightforward to show that if $\rho$ is separable (i.e., if $\rho \in C_{\text{ent}}$) then it is $r$-copy undistillable for all $r \geqslant 1$. More surprisingly, it has also been shown that there is a class of entangled (i.e., not separable) states called *PPT states* that are $r$-copy undistillable for all $r \geqslant 1$ [HHH98]. Entangled states with this undistillability property are called *bound entangled*.

The *NPPT bound entanglement* problem asks whether or not there are any other (i.e., non-PPT) states that are $r$-copy undistillable for all $r \geqslant 1$, and it is one of the central open questions in the theory of quantum entanglement [HRŻ20]. It has been shown that it suffices to answer this problem for a single-parameter family of states called *Werner states* [HH99], and it has furthermore been conjectured that the specific Werner state

$$\rho_{2/n} \overset{\text{def}}{=} \left(\frac{1}{n^2 - 2}\right) I - \frac{2}{n(n^2 - 2)} \left( \sum_{i,j=1}^{n} \mathbf{e_i}\mathbf{e_j^*} \otimes \mathbf{e_j}\mathbf{e_i^*} \right) \in (M_n \otimes M_n)^+ \tag{4}$$

is bound entangled for all $n \geqslant 4$ [DSS$^+$00]. Indeed, it is straightforward to check that $\rho_{2/n}$ is not PPT and that it is (1-copy) undistillable. However, even proving the simplest case of 2-copy undistillability of this state when $n = 4$ remains elusive, despite significant effort [CHSZ21, PPHH10].

It was shown in [JK10] that $\rho_{2/n}$ is $r$-copy undistillable if and only if the orthogonal projections defined recursively by

$$\begin{aligned}
P_1 &\overset{\text{def}}{=} |\phi_+\rangle\langle\phi_+| \in (M_n \otimes M_n)^+ \quad \text{and} \\
P_r &\overset{\text{def}}{=} (I - P_1) \otimes P_{r-1} + P_1 \otimes (I - P_{r-1}) \in (M_n^{\otimes r} \otimes M_n^{\otimes r})^+ \quad \forall r \geqslant 2
\end{aligned} \tag{5}$$

satisfy $\|P_r\|_{S(2)} \leqslant 1/2$. Here, it is understood that the partitioning of the $2r$ tensor factors used to compute $\|P_r\|_{S(2)}$ is the one in which the first factors of $P_1$ always "stay together", as do the second factors of $P_1$. That is, if we label the two tensor factors of $P_1$ as $P_1 \in A \otimes B$ then we have $P_r \in (A^{\otimes r}) \otimes (B^{\otimes r})$, with the $S(2)$-norm being computed across the central tensor cut, *not* $P_r \in (A \otimes B) \otimes \cdots \otimes (A \otimes B)$.

This connection with the $S(2)$-norm leads to the following characterization of the bound entanglement problem in terms of Birkhoff–James orthogonality and the modified trace distance of 2-entanglement:

**Theorem 12.** *Let $\rho_{2/n} \in (M_n \otimes M_n)^+$ be the Werner state defined in Equation* (4) *and let $P_r$ be the orthogonal projection defined in Equation* (5)*. The following are equivalent:*

(a) *$\rho_{2/n}$ is $r$-copy undistillable.*

(b) *$P_r$ is Birkhoff–James orthogonal to every member of $C_{\text{2-ent}}$.*

(c) *There exists a quantum state $\sigma \in (M_n^{\otimes r} \otimes M_n^{\otimes r})^+$ with* $\text{range}(\sigma) = \text{range}(P_r)$ *and* $D_{\text{2-ent}}(\sigma) = 1$.

*Proof.* By Theorem 11, we know that condition (c) happens if and only if $\|P_r\|_{S(2)} \leqslant 1/2$, and we already discussed the fact that $\|P_r\|_{S(2)} \leqslant 1/2$ is equivalent to $\rho_{2/n}$ being $r$-copy undistillable [JK10]. This shows that (c) is equivalent to (a).

The fact that $\|P_r\|_{S(2)} \leqslant 1/2$ is also equivalent to (b) (and thus (a) and (c) are equivalent to (b)) follows from the equivalence of conditions (b) and (c) of Theorem 7 when $C = C_{\text{2-ent}}$. $\qquad\square$

# 5 Conclusions and Outlook

The notion of Birkhoff–James orthogonality is an extension of the usual notion of orthogonality. Here, we consider Birkhoff–James orthogonality in the setting of $n \times n$ complex-valued matrices together with the trace norm. Our main result characterizes when a Hermitian matrix $H$ is Birkhoff–James orthogonal in the trace norm to a given positive semidefinite matrix $B$, using several trace conditions involving the product of $B$ and the positive and negative eigenprojection matrices of $H$. These eigenprojection matrices are key to our study, and we use them to further characterize when $H$ is Birkhoff–James orthogonal in the trace norm to the set of all positive semidefinite matrices, and when $H$ is Birkhoff–James orthogonal in the trace norm to the set of all (positive) diagonal matrices, allowing all (not just positive) diagonal matrices in certain cases. Further descriptions can be given when $H$ is assumed to be invertible.

Our results apply to quantum resource theory; we show that the modified trace distance between any quantum state (positive semidefinite, trace-one matrix) $\rho$ and a given closed convex cone $C$ of positive $n \times n$ matrices is one (the maximum possible value) precisely when $\rho$ is Birkhoff–James orthogonal in the trace norm to every member of $C$, which is also equivalent to a trace condition involving the product of every state in $C$ and the orthogonal projection onto the range of $\rho$. We generalize a result in the literature characterizing when the modified trace distance of coherence of a given quantum state is maximal, from pure states (quantum states that are rank-one projections) to arbitrary quantum states, and show that the rank of such states is at most $n/2$. We also characterize the quantum states that have maximal modified trace distance of $k$-coherence, with a similar bound on the rank.

We apply our results to the resource theory of entanglement, including $k$-entanglement, characterizing when a given quantum state has maximal modified trace distance of $k$-entanglement by way of a norm based on inner products with quantum states with Schmidt number at most $k$, and provide bounds on the rank of $\rho$. Finally, through our use of the aforementioned norm, we make an intriguing connection to the NPPT bound entanglement problem, characterizing when a quantum state is $r$-copy undistillable via Birkhoff–James orthogonality.

Our results highlight the clear utility of Birkhoff–James orthogonality. In terms of linear algebraic considerations, one open problem would be to extend the work from Hermitian to normal matrices, which we believe is possible but we avoided the added complexity of the problem as we were interested in the applications to quantum resource theory, which involves positive semidefinite matrices. The astute reader will have noticed that the two matrices in the proof of Theorem 4 are diagonalizable by complex Hadamard matrices, and this connection could further be explored. In terms of quantum resource theory considerations, we note that 2-coherence is connected to what are known as $H$-matrices, although we did not explore that connection here. We hope that our alternate characterizations prove useful, in particular to the ongoing solution to the NPPT bound entanglement problem.

## Acknowledgements

## References

[Abe06]    J. Aberg. Quantifying superposition. E-print: arXiv:quant-ph/0612146, 2006.

[ALRV12]   E. Andruchow, G. Larotonda, L. Recht, and A. Varela. A characterization of minimal Hermitian matrices. *Linear Algebra and its Applications*, 436(7):2366–2374, 2012.

[BCP14]    T. Baumgratz, M. Cramer, and M. B. Plenio. Quantifying coherence. *Physical Review Letters*, 113:140401, 2014.

[BCPT05]   E.G. Boman, D. Chen, O. Parekh, and S. Toledo. On factor width and symmetric H-matrices. *Linear algebra and its applications*, 405:239–248, 2005.

[Bha97]    R. Bhatia. *Matrix analysis*. Springer, 1997.

[BŠ99]     R. Bhatia and P. Šemrl. Orthogonality of matrices and some distance problems. *Linear algebra and its applications*, 287(1-3):77–85, 1999.

[CG19]     E. Chitambar and G. Gour. Quantum resource theories. *Reviews of Modern Physics*, 91:025001, 2019.

[CHSZ21]   L. Chen, H. He, X. Shi, and L.-J. Zhao. Proving the distillability problem of two-copy $4 \times 4$ Werner states for monomial matrices. *Quantum Information Processing*, 20:157, 2021.

[DSS$^+$00]  D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and A. V. Thapliyal. Evidence for bound entangled states with negative partial transpose. *Physical Review A*, 61:062312, 2000.

[Gha10]    S. Gharibian. Strong NP-hardness of the quantum separability problem. *Quantum Information and Computation*, 10:343–360, 2010.

[Gla63]    R. J. Glauber. Coherent and incoherent states of the radiation field. *Physical Review*, 131:2766, 1963.

[GT09]     O. Gühne and G. Tóth. Entanglement detection. *Phys. Rep.*, 474:1–75, 2009.

[Gur03]    L. Gurvits. Classical deterministic complexity of Edmonds' problem and quantum entanglement. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, pages 10–19, 2003.

[HH99]     M. Horodecki and P. Horodecki. Reduction criterion of separability and limits for a class of distillation protocols. *Physical Review A*, 59:4206–4216, 1999.

[HHH97]    M. Horodecki, P. Horodecki, and R. Horodecki. Inseparable two spin-1/2 density matrices can be distilled to a singlet form. *Physical Review Letters*, 78:574–577, 1997.

[HHH98]    M. Horodecki, P. Horodecki, and R. Horodecki. Mixed-state entanglement and distillation: Is there a "bound" entanglement in nature? *Physical Review Letters*, 80:5239–5242, 1998.

[HHHH09]   R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, 2009.

[HRŻ20]    P. Horodecki, Ł. Rudnicki, and K. Życzkowski. *Five open problems in quantum information*. E-print: arXiv:2002.03233 [quant-ph], 2020.

[Jam47]    R. C. James. Orthogonality and linear functionals in normed linear spaces. *Transactions of the American Mathematical Society*, 61(2):265–292, 1947.

[JK10]     N. Johnston and D. W. Kribs. A family of norms with applications in quantum information theory. *Journal of Mathematical Physics*, 51:082202, 2010.

[JK11]     N. Johnston and D. W. Kribs. A family of norms with applications in quantum information theory II. *Quantum Information and Computation*, 11:104–123, 2011.

[JLP18]    N. Johnston, C.-K. Li, and S. Plosker. The modified trace distance of coherence is constant on most pure states. *Journal of Physics A: Mathematical and Theoretical*, 51(41):414010, 2018.

[KMP21]    D. W. Kribs, D. Mammarella, and R. Pereira. Isoclinic subspaces and quantum error correction. *Oper. Matrices*, 15(2):571–580, 2021.

[LPP15]    J. Li, R. Pereira, and S. Plosker. Some geometric interpretations of quantum fidelity. *Linear Algebra and its Applications*, 487:158–171, 2015.

[LS02]     C.-K. Li and H. Schneider. Orthogonality of matrices. *Linear algebra and its applications*, 347(1-3):115–122, 2002.

[PPHH10]   L. Pankowski, M. Piani, M. Horodecki, and P. Horodecki. A few steps more towards NPT bound entanglement. *IEEE Transactions on Information Theory*, 56:4085–4100, 2010.

[RBC$^+$18]  M. Ringbauer, T. R. Bromley, M. Cianciaruso, L. Lami, W. Y. S. Lau, G. Adesso, A. G. White, A. Fedrizzi, and M. Piani. Certification and quantification of multilevel quantum coherence. *Physical Review X*, 8:041007, 2018.

[Reg18]    B. Regula. Convex geometry of quantum resource quantification. *Journal of Physics A: Mathematical and Theoretical*, 51(4):045303, 2018.

[RFWG19]   B. Regula, K. Fang, X. Wang, and M. Gu. One-shot entanglement distillation beyond local operations and classical communication. *New Journal of Physics*, 21(10):103017, 2019.

[RPL16]  S. Rana, P. Parashar, and M. Lewenstein. Trace-distance measure of coherence. *Physical Review A*, 93(1):012110, 2016.

[Sud63]  E. C. G. Sudarshan. Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams. *Physical Review Letters*, 10:277, 1963.

[TH00]  B. M. Terhal and P. Horodecki. Schmidt number for density matrices. *Physical Review A*, 61:040301(R), 2000.

[Var76]  R. S. Varga. On recurring theorems on diagonal dominance. *Linear Algebra and its Applications*, 13(1-2):1–9, 1976.

[Wat92]  G. A. Watson. Characterization of the subdifferential of some matrix norms. *Linear Algebra and its Applications*, 170(0):33–45, 1992.

[Wat04]  J. Watrous. Many copies may be required for entanglement distillation. *Physical Review Letters*, 93:010502, 2004.

[Wat18]  J. Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.

[Wer89]  R. F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Physical Review A*, 40:4277–4281, 1989.

[YZXT16]  X.-D. Yu, D.-J. Zhang, G. F. Xu, and D. M. Tong. Alternative framework for quantifying coherence. *Physical Review A*, 94(6):060302(R), 2016.