



The Distance from a Rank $n - 1$ Projection to the Nilpotent Operators on \mathbb{C}^n

Zachary Cramer

Abstract. Building on MacDonald's formula for the distance from a rank-one projection to the set of nilpotents in $\mathbb{M}_n(\mathbb{C})$, we prove that the distance from a rank $n - 1$ projection to the set of nilpotents in $\mathbb{M}_n(\mathbb{C})$ is $\frac{1}{2} \sec\left(\frac{\pi}{n-1+2}\right)$. For each $n \geq 2$, we construct examples of pairs (Q, T) where Q is a projection of rank $n - 1$ and $T \in \mathbb{M}_n(\mathbb{C})$ is a nilpotent of minimal distance to Q . Furthermore, we prove that any two such pairs are unitarily equivalent. We end by discussing possible extensions of these results in the case of projections of intermediate ranks.

1 Introduction

Let \mathcal{H} be a complex Hilbert space of (possibly infinite) dimension n , and let $\mathcal{B}(\mathcal{H})$ denote the algebra of bounded linear operators acting on \mathcal{H} . Consider the sets

$$\mathcal{P}(\mathcal{H}) = \{P \in \mathcal{B}(\mathcal{H}) : P = P^2 = P^*\} \setminus \{0\},$$
$$\mathcal{N}(\mathcal{H}) = \{N \in \mathcal{B}(\mathcal{H}) : N^j = 0 \text{ for some } j \in \mathbb{N}\}$$

consisting of all non-zero orthogonal projections on \mathcal{H} and all nilpotent operators on \mathcal{H} , respectively. We are interested in the problem of understanding the distance between these two sets, measured in the usual operator norm on $\mathcal{B}(\mathcal{H})$. This quantity will be denoted by δ_n :

$$\delta_n := \text{dist}(\mathcal{P}(\mathcal{H}), \mathcal{N}(\mathcal{H})) = \inf \{\|P - N\| : P \in \mathcal{P}(\mathcal{H}), N \in \mathcal{N}(\mathcal{H})\}.$$

The problem of computing δ_n is by no means new to the world of operator theory. In 1972, Hedlund [3] proved that $\delta_2 = 1/\sqrt{2}$, and that $1/4 \leq \delta_n \leq 1$ for all $n \geq 3$. This lower bound was increased to $1/2$ by Herrero [4] shortly thereafter. At that time Herrero, also showed that $\delta_n = 1/2$ whenever n is infinite, thus reducing the problem to the case in which $\mathcal{H} = \mathbb{C}^n$ for some $n \in \mathbb{N}$, $n \geq 3$.

Various estimates on the values of δ_n were obtained in the early 1980's. One such estimate established by Salinas [9] states that

$$\frac{1}{2} \leq \delta_n \leq \frac{1}{2} + \frac{1 + \sqrt{n-1}}{2n} \text{ for all } n \in \mathbb{N}.$$

Received by the editors March 30, 2020.
Published online on Cambridge Core April 2, 2020.
Research supported in part by NSERC (Canada).
AMS subject classification: 47A58, 15A99.
Keywords: projection, nilpotent, matrix, operator.



One can note that this upper bound approaches $1/2$ as n tends to infinity, and hence Salinas' inequality leads to an alternative proof that $\delta_{\infty_0} = 1/2$. Herrero [5] subsequently improved upon this upper bound for large values of n by showing that

$$\frac{1}{2} \leq \delta_n \leq \frac{1}{2} + \sin\left(\frac{\pi}{\lfloor \frac{n+1}{2} \rfloor}\right) \text{ for } n \geq 2,$$

where $\lfloor \cdot \rfloor$ denotes the greatest integer function.

For many years, the bounds obtained by Salinas and Herrero remained the best known. In 1995, however, MacDonald [6] established a new upper bound that would improve upon these estimates for all values of n . In order to describe MacDonald's approach, we first make the following remarks.

- (a) Any two projections in $\mathbb{M}_n(\mathbb{C})$ of equal rank are unitarily equivalent, and thus of equal distance to $\mathcal{N}(\mathbb{C}^n)$. As a result, $\delta_n = \min_{1 \leq r \leq n} \nu_{r,n}$, where

$$\nu_{r,n} := \inf \{ \|P - N\| : P \in \mathcal{P}(\mathbb{C}^n), \text{rank}(P) = r, N \in \mathcal{N}(\mathbb{C}^n) \}.$$

- (b) Straightforward estimates show that when computing $\nu_{r,n}$, one need only consider nilpotents of norm at most 2. From here, one may use the compactness of the set of projections in $\mathbb{M}_n(\mathbb{C})$ of rank r and the set of nilpotents in $\mathbb{M}_n(\mathbb{C})$ of norm at most 2 to show that $\nu_{r,n}$ is achieved by some projection-nilpotent pair, and hence so too is δ_n .
- (c) If $\{e_i\}_{i=1}^n$ denotes the standard basis for \mathbb{C}^n , then

$$\nu_{r,n} = \min \{ \|P - N\| : P \in \mathcal{P}(\mathbb{C}^n), \text{rank}(P) = r, N \in \mathcal{T}_n \}$$

where \mathcal{T}_n is the algebra of operators that are strictly upper triangular as matrices written with respect to $\{e_i\}_{i=1}^n$.

The reduction from $\mathcal{N}(\mathbb{C}^n)$ to \mathcal{T}_n described in (c) may seem innocuous at first glance. This alternate formulation, however, allows one to make use of a theorem of Arveson [1] that describes the distance from an operator in $\mathcal{B}(\mathcal{H})$ to a nest algebra. The version of this result that we require was established by Power [8] and is presented below for the algebra \mathcal{T}_n . Note that for vectors $x, y \in \mathbb{C}^n$, the notation $x \otimes y^*$ is used to denote the rank-one operator $z \mapsto \langle z, y \rangle x$ acting on \mathbb{C}^n .

Theorem 1.1 (Arveson Distance Formula). *Let $\{e_i\}_{i=1}^n$ denote the standard basis for \mathbb{C}^n . Define $E_0 = 0$ and $E_k = \sum_{i=1}^k e_i \otimes e_i^*$ for each $k \in \{1, 2, \dots, n\}$. For any $A \in \mathbb{M}_n(\mathbb{C})$,*

$$\text{dist}(A, \mathcal{T}_n) = \max_{1 \leq i \leq n} \|E_{i-1}^\perp A E_i\|.$$

Using Arveson's formula, MacDonald successfully determined the exact value of $\nu_{1,n}$, the distance from a rank-one projection in $\mathbb{M}_n(\mathbb{C})$ to $\mathcal{N}(\mathbb{C}^n)$.

Theorem 1.2 [6, Theorem 1] *For every positive integer n , the distance from the set of rank-one projections in $\mathbb{M}_n(\mathbb{C})$ to $\mathcal{N}(\mathbb{C}^n)$ is*

$$\nu_{1,n} = \frac{1}{2} \sec\left(\frac{\pi}{n+2}\right).$$

The expression for $v_{1,n}$ described above provides an upper bound on δ_n that is sharper than those previously obtained by Herrero and Salinas for all $n \in \mathbb{N}$. In addition, MacDonald proved that this bound is in fact optimal when $n = 3$ [6, Corollary 4]. These results led to the formulation of the following conjecture.

Conjecture 1.3 (MacDonald [6]). *The closest non-zero projections to $\mathcal{N}(\mathbb{C}^n)$ are of rank 1; that is,*

$$\delta_n = v_{1,n} = \frac{1}{2} \sec\left(\frac{\pi}{n+2}\right) \quad \text{for all } n \in \mathbb{N}.$$

Conjecture 1.3 has since been verified for $n = 4$ [7, Theorem 3.4], but remains open for all $n \geq 5$.

MacDonald's success in computing $v_{1,n}$ relied heavily on the simple structure of rank-one projections in $\mathbb{M}_n(\mathbb{C})$. Specifically, the fact that every such projection can be written as a simple tensor $P = e \otimes e^*$ for some unit vector $e \in \mathbb{C}^n$ made it feasible to obtain closed-form expressions for the norms $\|E_{i-1}^\perp P E_i\|$ in terms of the entries of P . MacDonald was then able to prove, using the Arveson Distance Formula, that the rank-one projections of minimal distance to \mathcal{T}_n are such that $\|E_{i-1}^\perp P E_i\| = v_{1,n}$ for all $i \in \{1, 2, \dots, n\}$. An exact expression for $v_{1,n}$ was then derived through algebraic and combinatorial arguments.

Extending this approach to accommodate projections of intermediate ranks appears to be a formidable task; when P is not expressible as a simple tensor $e \otimes e^*$, it becomes significantly more challenging to obtain explicit formulas for $\|E_{i-1}^\perp P E_i\|$. One can note, however, that the simple structure that led to success in the rank-one case can similarly be observed in projections $I - e \otimes e^*$ of rank $n - 1$. It is therefore the goal of this paper to extend MacDonald's approach to determine the exact value of $v_{n-1,n}$.

We accomplish this goal in three stages. Motivated by the analogous result for projections of rank 1, we show in Section 2 that any projection Q of rank $n - 1$ that is of minimal distance to \mathcal{T}_n must be such that $\|E_{i-1}^\perp Q E_i\| = v_{n-1,n}$ for all i . In Section 3, we then apply these equations to determine a list of candidates for $v_{n-1,n}$ via arguments adapted from [6]. Finally, we prove that exactly one such candidate satisfies a certain necessary norm inequality from [7], and therefore deduce that this value must be $v_{n-1,n}$.

In Section 4, we outline a construction of the pairs (Q, T) where $Q \in \mathbb{M}_n(\mathbb{C})$ is a projection of rank $n - 1$, T is an element of \mathcal{T}_n , and $\|Q - T\| = v_{n-1,n}$. We prove that for each $n \in \mathbb{N}$, any two such pairs are, in fact, unitary equivalent. Lastly, in Section 5, we propose a possible formula for $v_{r,n}$ in the case of projections of arbitrary rank, which can be seen to closely resemble numerical estimates for $v_{r,n}$ when n is small. We briefly explain how this formula could be used to answer MacDonald's conjecture in the affirmative.

2 Equality in Arveson's Distance Formula

Fix an integer $n \geq 3$, and let $\{e_1, e_2, \dots, e_n\}$ denote the standard basis for \mathbb{C}^n . Define $E_0 = 0$ and $E_k = \sum_{i=1}^k e_i \otimes e_i^*$ for each $k \in \{1, 2, \dots, n\}$. Throughout, $Q = (q_{ij})$ will

denote a projection in $\mathbb{M}_n(\mathbb{C})$ of rank $n - 1$ that is of minimal distance to \mathcal{T}_n . Additionally, let $e \in \mathbb{C}^n$ be a unit vector such that $Q = I - e \otimes e^*$, and let $P = (p_{ij})$ denote the rank-one projection $e \otimes e^*$.

The goal of this section is to derive a sequence of equations relating the entries of Q to the distance $v_{n-1,n}$. Our strategy will be to use the algebraic relations satisfied by the entries of Q to derive closed-form expressions for the norms $\|E_{i-1}^\perp Q E_i\|$. Next, we will relate these expression to $v_{n-1,n}$ through the Arveson Distance Formula.

In the case of rank-one projections, MacDonald obtained closed-form expressions for the norms in the Arveson Distance Formula by analysing the sequence of partial traces associated with such a projection. Specifically, this sequence $\{a_i\}_{i=0}^n$ is defined by setting $a_0 = 0$ and

$$(2.1) \quad a_k = \sum_{i=1}^k p_{ii} = k - \sum_{i=1}^k q_{ii}, \quad k \in \{1, 2, \dots, n\}.$$

We can then express the entries of e in terms of $\{a_i\}_{i=0}^n$ as

$$e = \left[z_1 \sqrt{a_1 - a_0} \quad z_2 \sqrt{a_2 - a_1} \quad \dots \quad z_n \sqrt{a_n - a_{n-1}} \right]^T,$$

where z_1, z_2, \dots, z_n are complex numbers of modulus 1.

Remark 2.1 By defining $U \in \mathbb{M}_n(\mathbb{C})$ to be the diagonal unitary $U = \text{diag}(z_1, z_2, \dots, z_n)$ and replacing Q with the unitarily equivalent projection $U^* Q U$, we can assume that each of the complex numbers z_i is equal to 1. That is, we can assume that $q_{ij} \leq 0$ for all distinct indices i and j . Since U commutes with each of the projections E_i , the norms $\|E_{i-1}^\perp Q E_i\|$ —and hence $\text{dist}(Q, \mathcal{T}_n)$ —are preserved by this transformation.

Under the assumption of Remark 2.1, one readily obtains useful identity among the entries of P and Q . Notably, the entries on the off-diagonals of these projections can be described entirely by those on the diagonals:

$$(2.2) \quad p_{ij} = \sqrt{p_{ii} p_{jj}} \quad \text{and} \quad q_{ij} = -\sqrt{(1 - q_{ii})(1 - q_{jj})} \quad \text{for all } i \neq j.$$

One can then verify that

$$(2.3) \quad p_{ij} p_{ik} = p_{ii} p_{jk} \quad \text{and} \quad q_{ij} q_{ik} = -q_{jk} (1 - q_{ii}) \quad \text{for all } i, j, k \text{ distinct.}$$

These identities will be used heavily throughout the proof of the following lemma, which serves as the first step toward obtaining closed-form descriptions of the norms $\|E_{i-1}^\perp Q E_i\|$ in terms of the sequence $\{a_i\}_{i=0}^n$.

Lemma 2.2 Let $Q = (q_{ij})$ be a projection in $\mathbb{M}_n(\mathbb{C})$ of rank $n - 1$, and let $\{a_i\}_{i=0}^n$ denote the non-decreasing sequence from equation (2.1). For $k \in \{1, 2, \dots, n\}$, define $Q_k := E_{k-1}^\perp Q E_k$, and let B_k denote the restriction of $Q_k^* Q_k$ to the range of E_k .

(i) If $q_{ij} \leq 0$ for all $i \neq j$, then the entries of $B_k = (b_{ij})$ are given by

$$b_{ij} = \begin{cases} q_{kk} - a_{k-1}(1 - q_{kk}) & \text{if } i = j = k, \\ (1 - a_{k-1})(1 - q_{ii}) & \text{if } i = j \neq k, \\ -(1 - a_{k-1})q_{ij} & \text{if } i, j, k \text{ are distinct,} \\ a_{k-1}q_{ij} & \text{otherwise.} \end{cases}$$

(ii) We have

$$\|Q_k\|^2 = \frac{\text{Tr}(B_k) + \sqrt{2\text{Tr}(B_k^2) - \text{Tr}(B_k)^2}}{2}.$$

Proof First, suppose that $q_{ij} \leq 0$ for all $i \neq j$. Since Q is idempotent, its entries q_{ij} satisfy the equation $q_{ij} = \sum_{\ell=1}^n q_{i\ell}q_{\ell j}$. This equation, together with the identities from (2.3), allows one to compute the entries of B_k directly. Indeed,

$$\begin{aligned} b_{kk} &= q_{kk}^2 + q_{k+1,k}^2 + \dots + q_{nk}^2 \\ &= q_{kk} - q_{1k}^2 - q_{2k}^2 - \dots - q_{k-1,k}^2 \\ &= q_{kk} - \sum_{\ell=1}^{k-1} (1 - q_{\ell\ell})(1 - q_{kk}) = q_{kk} - a_{k-1}(1 - q_{kk}), \end{aligned}$$

and if $i \neq k$, then

$$\begin{aligned} b_{ii} &= q_{ki}^2 + q_{k+1,i}^2 + \dots + q_{ni}^2 \\ &= q_{ii} - q_{1i}^2 - q_{2i}^2 - \dots - q_{k-1,i}^2 \\ &= q_{ii} - q_{ii}^2 - \sum_{\ell=1, \ell \neq i}^{k-1} (1 - q_{\ell\ell})(1 - q_{ii}) \\ &= (1 - q_{ii}) \left((k - 2) - \sum_{\ell=1}^{k-1} q_{\ell\ell} \right) = (1 - a_{k-1})(1 - q_{ii}). \end{aligned}$$

If i, j , and k are all distinct, then

$$\begin{aligned} b_{ij} &= q_{ki}q_{kj} + q_{k+1,i}q_{k+1,j} + \dots + q_{ni}q_{nj} \\ &= q_{ij} - q_{1i}q_{1j} - q_{2i}q_{2j} - \dots - q_{k-1,i}q_{k-1,j} \\ &= q_{ij} - q_{ii}q_{ij} - q_{ji}q_{jj} + \sum_{\ell=1, \ell \neq i, j}^{k-1} q_{ij}(1 - q_{\ell\ell}) \\ &= q_{ij} \left((k - 2) - \sum_{\ell=1}^{k-1} q_{\ell\ell} \right) = -(1 - a_{k-1})q_{ij}. \end{aligned}$$

Lastly, we consider entries b_{ij} for which $i < j = k$ or $j < i = k$. Since $B_k = B_k^*$, it suffices to establish the formula for b_{ij} in the case that $i < j = k$. We have

$$\begin{aligned} b_{ik} &= q_{ki}q_{kk} + q_{k+1,i}q_{k+1,k} + \dots + q_{ni}q_{nk} \\ &= q_{ik} - q_{1i}q_{1k} - q_{2i}q_{2k} - \dots - q_{k-1,i}q_{k-1,k} \end{aligned}$$

$$\begin{aligned}
 &= q_{ik} - q_{ii}q_{ik} + \sum_{\ell=1, \ell \neq i}^{k-1} q_{ik}(1 - q_{\ell\ell}) \\
 &= q_{ik} \left((k-1) - \sum_{\ell=1}^{k-1} q_{\ell\ell} \right) = a_{k-1}q_{ik}.
 \end{aligned}$$

We now turn our attention to the proof of (ii). Note that if P denotes the rank-one projection $I - Q$, then

$$Q_k = E_{k-1}^\perp(I - P)E_k = e_k \otimes e_k^* - E_{k-1}^\perp P E_k.$$

Thus, Q_k —and hence B_k —has rank at most 2. It follows that B_k has at most two non-zero eigenvalues λ_0 and λ_1 , which can be obtained by solving the system of equations

$$\begin{cases} \lambda_0 + \lambda_1 = \text{Tr}(B_k), \\ \lambda_0^2 + \lambda_1^2 = \text{Tr}(B_k^2). \end{cases}$$

The solutions to this system are $\lambda = \frac{1}{2}(\text{Tr}(B_k) \pm \sqrt{2\text{Tr}(B_k^2) - \text{Tr}(B_k)^2})$, and therefore, the result follows. ■

Theorem 2.3 *Let $Q = (q_{ij})$ be a projection in $\mathbb{M}_n(\mathbb{C})$ of rank $n - 1$, and let $\{a_i\}_{i=0}^n$ denote the non-decreasing sequence from equation (2.1). If $f: [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ denotes the function*

$$f(x, y) = \frac{\sqrt{x^2y^2 - 4x^2y + 2xy^2 + 4x^2 - 2xy + y^2 - 2y + 1 - xy - y + 2x + 1}}{2},$$

then for each $k \in \{1, 2, \dots, n\}$, $\|E_{k-1}^\perp Q E_k\|^2 = f(a_{k-1}, a_k)$.

Proof By Remark 2.1, we can conjugate Q by a diagonal unitary if necessary and assume that $q_{ij} \leq 0$ for all $i \neq j$. Fix an integer $k \in \{1, 2, \dots, n\}$, define $Q_k := E_{k-1}^\perp Q E_k$, and let $B_k = (b_{ij})$ denote the restriction of $Q_k^* Q_k$ to the range of E_k . By Lemma 2.2(ii),

$$\|Q_k\|^2 = \frac{\text{Tr}(B_k) + \sqrt{2\text{Tr}(B_k^2) - \text{Tr}(B_k)^2}}{2}.$$

Using the expressions for the entries of B_k derived in Lemma 2.2(i), we find that

$$\begin{aligned}
 \text{Tr}(B_k) &= \sum_{i=1}^{k-1} b_{ii} + b_{kk} \\
 &= \sum_{i=1}^{k-1} (1 - a_{k-1})(1 - q_{ii}) + q_{kk} - a_{k-1}(1 - q_{kk}) \\
 &= a_{k-1}(1 - a_{k-1}) + q_{kk} - a_{k-1}(1 - q_{kk}) \\
 &= q_{kk} + a_{k-1}(q_{kk} - a_{k-1}) \\
 &= q_{kk} + a_{k-1}(1 - a_k).
 \end{aligned}$$

Moreover, if $B_k^2 = (c_{ij})$, then

$$\begin{aligned} c_{kk} &= b_{kk}^2 + \sum_{\ell=1}^{k-1} b_{k\ell}^2 \\ &= (q_{kk} - a_{k-1}(1 - q_{kk}))^2 + \sum_{\ell=1}^{k-1} a_{k-1}^2 q_{k\ell}^2 \\ &= (q_{kk} - a_{k-1}(1 - q_{kk}))^2 + \sum_{\ell=1}^{k-1} a_{k-1}^2 (1 - q_{kk})(1 - q_{\ell\ell}) \\ &= (q_{kk} - a_{k-1}(1 - q_{kk}))^2 + a_{k-1}^3 (1 - q_{kk}), \end{aligned}$$

and for $i \leq k - 1$,

$$\begin{aligned} c_{ii} &= b_{ii}^2 + b_{ik}^2 + \sum_{\ell=1, \ell \neq i}^{k-1} b_{i\ell}^2 \\ &= (1 - a_{k-1})^2 (1 - q_{ii})^2 + a_{k-1}^2 q_{ik}^2 + \sum_{\ell=1, \ell \neq i}^{k-1} (1 - a_{k-1})^2 q_{i\ell}^2 \\ &= (1 - a_{k-1})^2 (1 - q_{ii})^2 + a_{k-1}^2 (1 - q_{ii})(1 - q_{kk}) \\ &\quad + \sum_{\ell=1, \ell \neq i}^{k-1} (1 - a_{k-1})^2 (1 - q_{ii})(1 - q_{\ell\ell}) \\ &= a_{k-1}(1 - q_{ii})((1 - a_{k-1})^2 + a_{k-1}(1 - q_{kk})). \end{aligned}$$

Thus,

$$\begin{aligned} \text{Tr}(B_k^2) &= c_{kk} + \sum_{\ell=1}^{k-1} a_{k-1}(1 - q_{\ell\ell})((1 - a_{k-1})^2 + a_{k-1}(1 - q_{kk})) \\ &= (q_{kk} - a_{k-1}(1 - q_{kk}))^2 + a_{k-1}^3 (1 - q_{kk}) \\ &\quad + a_{k-1}^2 ((1 - a_{k-1})^2 + a_{k-1}(1 - q_{kk})). \end{aligned}$$

These descriptions of $\text{Tr}(B_k)$ and $\text{Tr}(B_k^2)$ allow one to express $\|Q_k\|^2$ as a function of a_{k-1} , a_k , and q_{kk} . The desired formula for $\|Q_k\|^2$ can now be obtained by writing $q_{kk} = 1 - (a_k - a_{k-1})$. ■

Our first goal of this section is now complete: we have derived a closed-form expression for each norm $\|E_{i-1}^\perp Q E_i\|$ in terms of the sequence $\{a_i\}_{i=0}^n$. In order to show that every such norm is equal to $v_{n-1, n}$, we must first investigate the properties of the function f from Theorem 2.3.

Lemma 2.4 *If $f : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ denotes the function*

$$f(x, y) = \frac{\sqrt{x^2 y^2 - 4x^2 y + 2xy^2 + 4x^2 - 2xy + y^2 - 2y + 1 - xy - y + 2x + 1}}{2},$$

then f is increasing in x and decreasing in y . Moreover, $0 \leq f(x, y) \leq 1$ whenever $0 \leq x \leq y \leq 1$.

Proof Define $g : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ by

$$g(x, y) = x^2y^2 - 4x^2y + 2xy^2 + 4x^2 - 2xy + y^2 - 2y + 1,$$

so $f(x, y) = \frac{1}{2}(\sqrt{g(x, y)} - xy - y + 2x + 1)$. We begin by proving that $g(x, y)$ is non-negative on its domain and zero only at $(0, 1)$. We will therefore verify that f is well defined on $[0, 1] \times [0, 1]$, and that the partial derivatives of f exist at all points $(x, y) \neq (0, 1)$.

Observe that for each fixed $y \in [0, 1]$, the map

$$x \mapsto g(x, y) = (2 - y)^2x^2 - 2y(1 - y)x + (1 - y)^2$$

defines a convex quadratic on $[0, 1]$ with vertex at $x_0 = y(1 - y)/(2 - y)^2$. If $y \in [0, 1]$, then

$$g(x_0, y) = g\left(\frac{y(1 - y)}{(2 - y)^2}, y\right) = \frac{4(1 - y)^3}{(2 - y)^2} > 0.$$

Consequently, $g(x, y) > 0$ for all $(x, y) \in [0, 1] \times [0, 1]$. Note as well that at $y = 1$ we have $g(x, 1) = x^2$. It follows that $g(0, 1) = 0$ and $g(x, y) > 0$ for all other values of $(x, y) \in [0, 1] \times [0, 1]$. Thus, f is well defined, and the partial derivatives

$$\begin{aligned} f_x(x, y) &= \frac{g_x(x, y)}{4\sqrt{g(x, y)}} + \frac{2 - y}{2}, & f_{xx}(x, y) &= \frac{2(1 - y)^3}{(g(x, y))^{3/2}}, \\ f_y(x, y) &= \frac{g_y(x, y)}{4\sqrt{g(x, y)}} - \frac{x + 1}{2}, & f_{yy}(x, y) &= \frac{2x^3}{(g(x, y))^{3/2}} \end{aligned}$$

exist for all $(x, y) \neq (0, 1)$.

Our next task is to prove that $f(x, y)$ is increasing in x . First observe that $f(x, 1) = x$ is clearly increasing. Furthermore, for every fixed $y \in [0, 1]$, $f_{xx}(x, y)$ is well defined and strictly positive for all x . Hence,

$$f_x(x, y) = \frac{xy^2 - 4xy + y^2 + 4x - y}{2\sqrt{g(x, y)}} + \frac{2 - y}{2}$$

is an increasing function of x . We conclude that $f_x(x, y) \geq f_x(0, y) = 1 - y > 0$ for every $x \in [0, 1]$. Thus, f is an increasing function of x on $[0, 1]$.

We now use a similar argument to show that f is a decreasing function of y . For $x = 0$, we have that $f(0, y) = 1 - y$ is clearly decreasing. Now given a fixed $x \in (0, 1]$, it is clear from above that $f_{yy}(x, y)$ is well defined and strictly positive for all y . It follows that the partial derivative

$$f_y(x, y) = \frac{x^2y - 2x^2 + 2xy - x + y - 1}{2\sqrt{g(x, y)}} - \frac{x + 1}{2}$$

is an increasing function of y on $[0, 1]$. Hence, $f_y(x, y) \leq f_y(x, 1) = -x < 0$ for every $y \in [0, 1]$. This proves that f is a decreasing function of y on $[0, 1]$, as desired.

For the final claim, suppose that $0 \leq x \leq y \leq 1$. Consider the sequence $\{a_i\}_{i=0}^3$ defined by $a_0 = 0$, $a_1 = x$, $a_2 = y$, and $a_3 = 1$, as well as the vector

$e = [\sqrt{x} \ \sqrt{y-x} \ \sqrt{y}]^T$. Note that $Q := I - e \otimes e^*$ is a rank-two projection in $\mathbb{M}_3(\mathbb{C})$ with partial trace sequence given by $\{a_i\}_{i=0}^3$. It then follows from Theorem 2.3 that

$$f(x, y) = f(a_1, a_2) = \|E_1^\perp Q E_2\|^2,$$

and hence $0 \leq f(x, y) \leq 1$. ■

Theorem 2.5 *If $Q \in \mathbb{M}_n(\mathbb{C})$ is a projection of rank $n - 1$ that is of minimal distance to \mathcal{T}_n , then $\|E_{i-1}^\perp Q E_i\| = \|E_{j-1}^\perp Q E_j\|$ for all i and j .*

Proof By Remark 2.1, we can assume without loss of generality that $Q = (q_{ij})$ where $q_{ij} \leq 0$ for all $i \neq j$. Let $\{a_i\}_{i=0}^n$ denote the non-decreasing sequence from equation (2.1), and for each $i \in \{1, 2, \dots, n\}$, define $Q_i := E_{i-1}^\perp Q E_i$. Suppose to the contrary that not all values of $\|Q_i\|$ are equal. Define

$$\mu := \max_{1 \leq i \leq n} \|Q_i\|,$$

and let j denote the largest index in $\{1, 2, \dots, n\}$ such that $\|Q_j\| = \mu$.

First consider the case in which $j = n$. Let k denote the largest index in $\{1, 2, \dots, n - 1\}$ such that $\|Q_k\| < \mu$. With f as in Theorem 2.3, we have that

$$f(a_{k-1}, a_k) = \|Q_k\|^2 < \|Q_{k+1}\|^2 = f(a_k, a_{k+1}).$$

Thus, if $g: [a_{k-1}, a_k] \rightarrow \mathbb{R}$ is given by

$$g(x) = f(a_{k-1}, x) - f(x, a_{k+1}),$$

then $g(a_k) = f(a_{k-1}, a_k) - f(a_k, a_{k+1}) < 0$, while $g(a_{k-1}) = 1 - f(a_{k-1}, a_{k+1}) \geq 0$ by Lemma 2.4. Since g is continuous on its domain, the Intermediate Value Theorem gives rise to some $a'_k \in [a_{k-1}, a_k]$ such that $g(a'_k) = 0$. By replacing a_k with a'_k in the sequence $\{a_i\}_{i=0}^n$, one may equate $\|Q_k\|$ and $\|Q_{k+1}\|$ while leaving the remaining norms $\|Q_i\|$ unchanged. Most importantly, since $a'_k \leq a_k$, Lemma 2.4 implies that the new common value of $\|Q_k\|$ and $\|Q_{k+1}\|$ is strictly less than μ .

This argument can now be repeated to successively reduce the norms $\|Q_i\|$ for $i > k$ to values strictly less than μ . At the end of this process, either the new largest index j at which the maximum norm occurs is strictly less than n , or the maximum μ decreases. Of course, the latter cannot happen as Q was assumed to be of minimal distance to \mathcal{T}_n .

Thus, we can assume that the largest index j at which μ occurs is strictly less than n . In this case, we have that

$$f(a_j, a_{j+1}) = \|Q_{j+1}\|^2 < \|Q_j\|^2 = f(a_{j-1}, a_j).$$

As above, we may invoke the Intermediate Value Theorem to obtain a root a'_j of the continuous function

$$h(x) := f(a_{j-1}, x) - f(x, a_{j+1})$$

on the interval $[a_j, a_{j+1}]$. By replacing a_j with a'_j in the sequence $\{a_i\}_{i=0}^n$, one can equate $\|Q_j\|$ and $\|Q_{j+1}\|$ while preserving all other norms $\|Q_i\|$. Since $a'_j \geq a_j$, Lemma 2.4 demonstrates that the new common value of $\|Q_j\|$ and $\|Q_{j+1}\|$ is strictly less

than μ . Thus, this process either decreases the largest index j at which the maximum norm occurs, or reduces the value of μ . Since this argument can be repeated for smaller and smaller values of j , eventually μ must decrease—a contradiction. ■

3 Computing the Distance

We will now use the results of Section 2 to determine the precise value of $v_{n-1,n}$. The first step in this direction is the following proposition, which applies Theorem 2.5 to obtain a recursive description of the sequence $\{a_i\}_{i=0}^n$.

Proposition 3.1 *Let $Q \in \mathbb{M}_n(\mathbb{C})$ be a projection of rank $n - 1$ that is of minimal distance to \mathcal{T}_n . If $\{a_i\}_{i=0}^n$ denotes the non-decreasing sequence from equation (2.1), then*

$$a_k = \frac{-v_{n-1,n}^4 + 2v_{n-1,n}^2 a_{k-1} + v_{n-1,n}^2 - a_{k-1}}{v_{n-1,n}^2 a_{k-1} + v_{n-1,n}^2 - a_{k-1}}$$

for each $k \in \{1, 2, \dots, n\}$.

Proof Since the distance from Q to \mathcal{T}_n is minimal, Theorems 1.1 and 2.5 imply that $\|E_{k-1}^\perp Q E_k\| = v_{n-1,n}$ for all $k \in \{1, 2, \dots, n\}$. Thus, with f as in Theorem 2.3, we have that

$$f(a_{k-1}, a_k) = \|E_{k-1}^\perp Q E_k\|^2 = v_{n-1,n}^2.$$

The desired formula can now be obtained by solving this equation for a_k . ■

The recursive formula for a_k described in Proposition 3.1 will be the key to computing $v_{n-1,n}$. Our goal will be to use this formula and some basic properties of the sequence $\{a_i\}_{i=0}^n$ to determine a list of candidates for $v_{n-1,n}^2$. A careful analysis of these candidates will reveal that exactly one of them satisfies a certain necessary norm inequality from [7]. This value must therefore be $v_{n-1,n}^2$.

To simplify notation, let $t = v_{n-1,n}^2$ and define the function $h_t : [0, 1] \rightarrow \mathbb{R}$ by

$$(3.1) \quad h_t(x) = \frac{-t^2 + 2tx + t - x}{tx + t - x}.$$

Proposition 3.1 states that for each $k \in \{1, 2, \dots, n\}$,

$$a_k = \frac{-t^2 + 2ta_{k-1} + t - a_{k-1}}{ta_{k-1} + t - a_{k-1}} = h_t(a_{k-1}).$$

Since $h_t(0) = (t - t^2)/t = 1 - t = a_1$, this formula can be expressed as $a_k = h_t^{(k)}(0)$ for all $k \in \{1, 2, \dots, n\}$. Upon taking into account the condition $a_n = \text{Tr}(P) = 1$, we are interested in identifying the values of $t \in [\frac{1}{4}, 1]$ that satisfy the equation $h_t^{(n)}(0) = 1$.

Notice that each expression $h_t^{(k)}(0)$ is a rational function of t . For each $k \geq 1$, let $p_{k-1}(t)$ and $q_{k-1}(t)$ denote polynomials in t such that

$$h_t^{(k)}(0) = \frac{p_{k-1}(t)}{q_{k-1}(t)}.$$

It then follows that

$$\frac{p_k(t)}{q_k(t)} = h_t(h_t^{(k)}(0)) = h_t\left(\frac{p_{k-1}(t)}{q_{k-1}(t)}\right) = \frac{-t^2q_{k-1}(t) + 2tp_{k-1}(t) + tq_{k-1}(t) - p_{k-1}(t)}{tp_{k-1}(t) + tq_{k-1}(t) - p_{k-1}(t)},$$

and hence, we obtain the relations

$$(3.2) \quad p_k(t) = t(1-t)q_{k-1}(t) + (2t-1)p_{k-1}(t),$$

$$(3.3) \quad q_k(t) = tq_{k-1}(t) - (1-t)p_{k-1}(t).$$

We can replace $p_{k-1}(t)$ in (3.3) using equation (3.2), thereby leading to a recurrence expressed only in the $q_k(t)$'s. Specifically, we have that

$$\begin{aligned} q_k(t) &= tq_{k-1}(t) - (1-t)p_{k-1}(t) \\ &= tq_{k-1}(t) - (1-t)[t(1-t)q_{k-2}(t) + (2t-1)p_{k-2}(t)] \\ &= tq_{k-1}(t) - t(1-t)^2q_{k-2}(t) - (2t-1)[tq_{k-2}(t) - q_{k-1}(t)] \\ &= (3t-1)q_{k-1}(t) - t^3q_{k-2}(t) \end{aligned}$$

for all integers $k \geq 2$. Notice as well that since

$$\frac{p_0(t)}{q_0(t)} = h_t(0) = 1-t \quad \text{and} \quad \frac{p_1(t)}{q_1(t)} = h_t(h_t(0)) = \frac{-3t^2 + 4t - 1}{-t^2 + 3t - 1},$$

we have initial terms $q_0(t) = 1$ and $q_1(t) = -t^2 + 3t - 1$.

The requirement that $h_t^{(n)}(0) = 1$ is equivalent to asking that $p_{n-1}(t) = q_{n-1}(t)$. Using the relations above, this equation can be restated as $tq_{n-2}(t) = p_{n-2}(t)$, or equivalently, $q_{n-1}(t) = t^2q_{n-2}(t)$ by (3.3). Thus, we wish to determine the values of $t \in [\frac{1}{4}, 1]$ that satisfy

$$q_{n-1}(t) = t^2q_{n-2}(t),$$

where

$$q_0(t) = 1, \quad q_1 = -t^2 + 3t - 1, \quad \text{and} \quad q_k(t) = (3t-1)q_{k-1}(t) - t^3q_{k-2}(t) \quad \text{for } k \geq 2.$$

A solution to this problem will require closed-form expressions for the polynomials $q_{n-1}(t)$ and $q_{n-2}(t)$, which can be obtained via diagonalization arguments akin to those in [6]. Our analysis reveals that with

$$y := \sqrt{4t-1}, \quad \lambda_1 := \frac{3t-1+(1-t)iy}{2}, \quad \text{and} \quad \lambda_2 := \frac{3t-1-(1-t)iy}{2},$$

we have

$$\begin{aligned} q_{n-1}(t) &= \frac{t(\lambda_1^n - \lambda_2^n) - \lambda_2\lambda_1^n + \lambda_1\lambda_2^n}{t(1-t)iy}, \\ q_{n-2}(t) &= \frac{t(\lambda_1^{n-1} - \lambda_2^{n-1}) - \lambda_2\lambda_1^{n-1} + \lambda_1\lambda_2^{n-1}}{t(1-t)iy}. \end{aligned}$$

These expressions for $q_{n-1}(t)$ and $q_{n-2}(t)$ can now be used to identify the desired values of t . Indeed, when $q_{n-1}(t) = t^2 q_{n-2}(t)$, we have that

$$\begin{aligned} t(\lambda_1^n - \lambda_2^n) - \lambda_2 \lambda_1^n + \lambda_1 \lambda_2^n &= t^2(t(\lambda_1^{n-1} - \lambda_2^{n-1}) - \lambda_2 \lambda_1^{n-1} + \lambda_1 \lambda_2^{n-1}) \\ \implies \lambda_1^n(t - \lambda_2) - \lambda_2^n(t - \lambda_1) &= t^2(\lambda_1^{n-1}(t - \lambda_2) - \lambda_2^{n-1}(t - \lambda_1)) \\ \implies \lambda_2^{n-1}(t^2 - \lambda_2)(t - \lambda_1) &= \lambda_1^{n-1}(t^2 - \lambda_1)(t - \lambda_2), \end{aligned}$$

and therefore,

$$(3.4) \quad \left(\frac{\lambda_2}{\lambda_1}\right)^{n-1} \left(\frac{t^2 - \lambda_2}{t^2 - \lambda_1}\right) \left(\frac{t - \lambda_1}{t - \lambda_2}\right) = 1.$$

This equation can be simplified using the following identities that relate the values of t , λ_1 , and λ_2 . Verification of these identities is straightforward, and thus their proofs are left to the reader.

Lemma 3.2 *If $y = \sqrt{4t - 1}$, $\lambda_1 = (3t - 1 + (1 - t)iy)/2$, and $\lambda_2 = (3t - 1 - (1 - t)iy)/2$, then*

- (i) $t - \lambda_1 = (1 - t)\left(\frac{1 - iy}{2}\right)$ and $t - \lambda_2 = (1 - t)\left(\frac{1 + iy}{2}\right)$;
- (ii) $t^2 - \lambda_1 = (1 - t)\left(\frac{1 - 2t - iy}{2}\right)$ and $t^2 - \lambda_2 = (1 - t)\left(\frac{1 - 2t + iy}{2}\right)$;
- (iii) $\frac{1 + iy}{1 - iy} = \frac{1 - 2t + iy}{2t}$ and $\frac{1 - iy}{1 + iy} = \frac{1 - 2t - iy}{2t}$;
- (iv) $\frac{\lambda_2}{\lambda_1} = \left(\frac{1 + iy}{1 - iy}\right)^3$.

One can apply the identities above to simplify equation (3.4) as follows:

$$\begin{aligned} 1 &= \left(\frac{\lambda_2}{\lambda_1}\right)^{n-1} \left(\frac{t^2 - \lambda_2}{t^2 - \lambda_1}\right) \left(\frac{t - \lambda_1}{t - \lambda_2}\right) \\ &= \left(\frac{1 + iy}{1 - iy}\right)^{3(n-1)} \left(\frac{1 - 2t + iy}{1 - 2t - iy}\right) \left(\frac{1 - iy}{1 + iy}\right) \\ &= \left(\frac{1 + iy}{1 - iy}\right)^{3n-3} \left(\frac{1 + iy}{1 - iy}\right)^2 \left(\frac{1 - iy}{1 + iy}\right) \\ &= \left(\frac{1 + iy}{1 - iy}\right)^{3n-2}. \end{aligned}$$

We therefore conclude that $\frac{1+iy}{1-iy} = \rho_m^k$, where $m := 3n - 2$, $\rho_m := e^{2\pi i/m}$, and k is an integer.

We are now in a position to determine the possible values of t . By solving for y in the equation above, we obtain

$$y = \frac{1 \rho_m^k - 1}{i \rho_m^k + 1} = \frac{1 \rho_m^{k/2} (\rho_m^{k/2} - \rho_m^{-k/2})}{i \rho_m^{k/2} (\rho_m^{k/2} + \rho_m^{-k/2})}$$

$$\begin{aligned}
 &= \frac{\rho_m^{k/2} - \rho_m^{-k/2}}{2i} \frac{2}{\rho_m^{k/2} + \rho_m^{-k/2}} \\
 &= \frac{\sin(k\pi/m)}{\cos(k\pi/m)} = \tan\left(\frac{k\pi}{m}\right).
 \end{aligned}$$

Since $y = \sqrt{4t - 1}$, we have

$$t = \frac{1}{4} \left(\tan^2\left(\frac{k\pi}{m}\right) + 1 \right) = \frac{1}{4} \sec^2\left(\frac{k\pi}{3n-2}\right) \text{ for some } k \in \mathbb{Z}.$$

That is, the distance $v_{n-1,n}$ from Q to \mathcal{T}_n must belong to the set $\left\{ \frac{1}{2} \sec\left(\frac{k\pi}{3n-2}\right) : k \in \mathbb{Z} \right\}$.

It remains to determine which element of this set represents $v_{n-1,n}$. We will accomplish this task by appealing to the following result of MacDonald concerning a lower bound on the distance from a projection to a nilpotent.

Proposition 3.3 [7, Lemma 3.3] *If $P \in \mathbb{M}_n(\mathbb{C})$ is a projection of rank r and $N \in \mathbb{M}_n(\mathbb{C})$ is nilpotent, then*

$$\|P - N\| \geq \sqrt{\frac{r}{2n} \left(1 + \frac{r}{n}\right)}.$$

In the analysis that follows, we will demonstrate that the only value in $\left\{ \frac{1}{2} \sec\left(\frac{k\pi}{3n-2}\right) : k \in \mathbb{Z} \right\}$ that respects the lower bound of Proposition 3.3 for projections of rank $r = n - 1$ occurs when $k = n - 1$. We begin with the following lemma, which proves that MacDonald’s lower bound is indeed satisfied for this choice of k .

Lemma 3.4 *For every integer $n \geq 3$,*

$$\frac{n-1}{2n} \left(1 + \frac{n-1}{n}\right) \leq \frac{1}{4} \sec^2\left(\frac{(n-1)\pi}{3n-2}\right) \leq 1.$$

Proof Define $\alpha_n := (3n - 2)/(n - 1)$. By considering reciprocals, this problem is equivalent to that of establishing the inequalities

$$\frac{1}{4} \leq \cos^2\left(\frac{\pi}{\alpha_n}\right) \leq \frac{n^2}{2(n-1)(2n-1)} \text{ for all } n \in \mathbb{Z}, n \geq 3.$$

In the computations that follow, it will be helpful to view n as a continuous variable on $[3, \infty)$.

To establish the inequality $1/4 \leq \cos^2(\pi/\alpha_n)$, simply note that π/α_n is an increasing function of n tending to $\pi/3$, $\cos(x)$ is decreasing on $[0, \pi/3]$, and $\cos(\pi/3) = 1/2$. The second inequality will require a bit more work. Since $(2n - \frac{3}{2})^2 \geq 2(n - 1)(2n - 1)$ for all n , it suffices to prove that

$$\cos^2\left(\frac{\pi}{\alpha_n}\right) \leq \frac{n^2}{(2n - \frac{3}{2})^2} \text{ for } n \in [3, \infty).$$

Note that this inequality holds if and only if the function

$$f(n) := \frac{2n}{4n - 3} - \cos\left(\frac{\pi}{\alpha_n}\right)$$

is non-negative on $[3, \infty)$.

We will prove that $f'(n) < 0$ for all $n \in [3, \infty)$, so that f is decreasing on this interval. Since

$$\lim_{n \rightarrow \infty} f(n) = 0 \quad \text{and} \quad f(3) = \frac{2}{3} - \cos\left(\frac{2\pi}{7}\right) \approx 0.043 > 0,$$

this will demonstrate that $f(n) \geq 0$ for all $n \geq 3$. To this end, we compute

$$f'(n) = \frac{16\pi \sin\left(\frac{\pi}{\alpha_n}\right)n^2 - 24\pi \sin\left(\frac{\pi}{\alpha_n}\right)n + 9\pi \sin\left(\frac{\pi}{\alpha_n}\right) - 54n^2 + 72n - 24}{(4n - 3)^2(3n - 2)^2}.$$

Of course $(4n - 3)^2(3n - 2)^2 \geq 0$, so the sign of $f'(n)$ depends only on the sign of

$$g(n) := 16\pi \sin\left(\frac{\pi}{\alpha_n}\right)n^2 - 24\pi \sin\left(\frac{\pi}{\alpha_n}\right)n + 9\pi \sin\left(\frac{\pi}{\alpha_n}\right) - 54n^2 + 72n - 24.$$

But since $\pi/\alpha_n \in [\pi/4, \pi/3]$ for $n \geq 3$, we have that $\sin(\pi/\alpha_n) \in [\sqrt{2}/2, \sqrt{3}/2]$ for all such n , and hence

$$\begin{aligned} g(n) &\leq 16\pi\left(\frac{\sqrt{3}}{2}\right)n^2 - 24\pi\left(\frac{\sqrt{2}}{2}\right)n + 9\pi\left(\frac{\sqrt{3}}{2}\right) - 54n^2 + 72n - 24 \\ &= (8\sqrt{3}\pi - 54)n^2 - (12\sqrt{2} - 72)n + \left(\frac{9\sqrt{3}}{2} - 24\right). \end{aligned}$$

This upper bound for g is a concave quadratic whose larger root occurs at $n \approx 1.8105$. It follows that g is negative on $[3, \infty)$, and therefore so too is f' . ■

Lemma 3.5 For any integer $n \geq 3$, the set

$$\left\{ \frac{1}{4} \sec^2\left(\frac{k\pi}{3n-2}\right) : k \in \mathbb{Z} \right\}$$

contains exactly one value in $\left[\frac{n-1}{2n}\left(1 + \frac{n-1}{n}\right), 1\right]$, and it occurs when $k = n - 1$.

Proof Fix an integer $n \geq 3$. We wish to prove that $\mathcal{A} := \left\{ \cos^2\left(\frac{k\pi}{3n-2}\right) : k \in \mathbb{Z} \right\}$ contains exactly one value in the interval $\left[\frac{1}{4}, \frac{n^2}{2(n-1)(2n-1)}\right]$. Since Lemma 3.4 demonstrates that this is the case when $k = n - 1$, it suffices to show that no other values in \mathcal{A} are within distance

$$\beta(n) := \frac{n^2}{2(n-1)(2n-1)} - \frac{1}{4}$$

of $\cos^2((n - 1)\pi/(3n - 2))$.

Note, however, that not all values of $k \in \mathbb{Z}$ need to be considered. In particular, since the function $k \mapsto \cos^2(k\pi/(3n - 2))$ is periodic, it suffices to check only its values at the integers $k \in \{0, 1, \dots, 3n - 2\}$. Additionally, since

$$\cos^2\left(\frac{((3n-2)-k)\pi}{3n-2}\right) = \cos^2\left(\frac{k\pi}{3n-2}\right) \text{ for all } k,$$

we can restrict our attention to $k \in \{0, 1, 2, \dots, \lfloor (3n - 2)/2 \rfloor\}$.

Although we are solely concerned with the integer values of k described above, it will be useful to view k as a continuous real variable. With this in mind, define the function $f_n : [0, (3n - 2)/2] \rightarrow \mathbb{R}$ by

$$f_n(k) := \sin\left(\frac{(n - k - 1)\pi}{3n - 2}\right) \sin\left(\frac{(n + k - 1)\pi}{3n - 2}\right).$$

It follows from the identity $\cos^2(x) - \cos^2(y) = -\sin(x - y)\sin(x + y)$ that

$$\left| \cos^2\left(\frac{k\pi}{3n - 2}\right) - \cos^2\left(\frac{(n - 1)\pi}{3n - 2}\right) \right| < \beta(n) \iff |f_n(k)| < \beta(n).$$

Notice, however, that

$$f'_n(k) = \left(\frac{-\pi}{3n - 2}\right) \sin\left(\frac{2k\pi}{3n - 2}\right),$$

so $f'_n(k) < 0$ on $[0, (3n - 2)/2]$, and hence f_n is decreasing on its domain. Since $f_n(n - 1) = 0$, it therefore suffices to prove that $f_n(n - 2) > \beta(n)$ and $-f_n(n) > \beta(n)$. We will demonstrate that these inequalities hold via application of Taylor's Theorem.

Consider the approximation of $\sin(x)$ by $x - x^3/6$, its third degree MacLauren polynomial. On $[0, \pi/6]$, the error in this approximation is at most

$$E(x) = \frac{\sin(\pi/6)}{4!} |x|^4 = \frac{x^4}{48}.$$

Thus, since $1/n \leq \pi/(3n - 2) \leq \pi/6$, we have

$$\begin{aligned} \sin\left(\frac{\pi}{3n - 2}\right) &\geq \sin\left(\frac{1}{n}\right) \geq \left(\frac{1}{n} - \frac{1}{6n^3} - E\left(\frac{1}{n}\right)\right) \\ &\geq \left(\frac{1}{n} - \frac{1}{6n} - \frac{1}{48n}\right) = \frac{13}{16n}. \end{aligned}$$

It is routine to verify that $\sin((2x - 1)\pi/(3x - 2))$ is an increasing function of x on $[3, \infty)$. Consequently, this function is bounded below by $\sin(5\pi/7)$, its value at $x = 3$. We deduce that

$$-f_n(n) = \sin\left(\frac{\pi}{3n - 2}\right) \sin\left(\frac{(2n - 1)\pi}{3n - 2}\right) \geq \frac{13}{16n} \sin\left(\frac{5\pi}{7}\right) \geq \frac{13}{16n} \cdot \frac{3}{4} = \frac{39}{64n}.$$

Lastly, one can show directly that

$$\frac{39}{64n} > \beta(n) \quad \text{whenever} \quad n > \frac{101 + \sqrt{5521}}{60} \approx 2.9217,$$

and hence $-f_n(n) > \beta(n)$ for our fixed integer $n \geq 3$.

A similar analysis can now be used to prove that $f_n(n - 2) > \beta(n)$. Indeed, it is straightforward to verify that $\sin((2n - 3)\pi/(3n - 2))$ is bounded below by $\sin(2\pi/3)$, and therefore

$$\begin{aligned} f_n(n - 2) &= \sin\left(\frac{\pi}{3n - 2}\right) \sin\left(\frac{(2n - 3)\pi}{3n - 2}\right) \\ &\geq \frac{13}{16n} \sin\left(\frac{2\pi}{3}\right) = \frac{13}{16n} \cdot \frac{\sqrt{3}}{2} \geq \frac{13}{16n} \cdot \frac{3}{4} = \frac{39}{64n}. \end{aligned}$$

It now follows from the arguments of the previous case that $f_n(n - 2) > \beta(n)$. ■

With the above analysis complete, we can now present the main result of this paper: the distance from a projection in $\mathbb{M}_n(\mathbb{C})$ of rank $n - 1$ to the set $\mathcal{N}(\mathbb{C}^n)$ is

$$v_{n-1,n} = \frac{1}{2} \sec\left(\frac{(n-1)\pi}{3n-2}\right).$$

Interestingly, this expression can be rewritten to bear an even stronger resemblance to MacDonald's formula in the rank-one case.

Theorem 3.6 *For every integer $n \geq 2$, the distance from the set of projections in $\mathbb{M}_n(\mathbb{C})$ of rank $n - 1$ to $\mathcal{N}(\mathbb{C}^n)$ is*

$$v_{n-1,n} = \frac{1}{2} \sec\left(\frac{\pi}{\frac{n}{n-1} + 2}\right).$$

4 Closest Projection-Nilpotent Pairs

Given a projection Q in $\mathbb{M}_n(\mathbb{C})$ of rank $n - 1$ that is of distance $v_{n-1,n}$ to \mathcal{T}_n , the following theorem provides a means for determining an element $T \in \mathcal{T}_n$ that is closest to Q . As we will see in Theorem 4.2, this element of \mathcal{T}_n is unique to Q .

Theorem 4.1 [2, 7] *Fix $\gamma \in [0, \infty)$. An operator $A \in \mathbb{M}_n(\mathbb{C})$ is such that $\|E_{i-1}^\perp A E_i\| = \gamma$ for all $i \in \{1, 2, \dots, n\}$ if and only if there exist $T \in \mathcal{T}_n$ and a unitary $U \in \mathbb{M}_n(\mathbb{C})$ such that $A - T = \gamma U$. Furthermore, if $\|E_{i-1}^\perp A E_i\| = \gamma$ and $\|E_i^\perp A E_i\| < \gamma$ for all $i \in \{1, 2, \dots, n - 1\}$, then the operators T and U are unique.*

With this result in hand, we are now able to describe all closest pairs (Q, N) where Q is a projection of rank $n - 1$ and $N \in \mathcal{N}(\mathbb{C}^n)$.

Theorem 4.2 *Fix a positive integer $n \geq 2$. Let $\{a_i\}_{i=0}^n$ be the sequence given by $a_0 = 0$ and*

$$a_k = \frac{-v_{n-1,n}^4 + 2v_{n-1,n}^2 a_{k-1} + v_{n-1,n}^2 - a_{k-1}}{v_{n-1,n}^2 a_{k-1} + v_{n-1,n}^2 - a_{k-1}} \text{ for } k \geq 1.$$

Let $\{z_i\}_{i=1}^n$ be a sequence of complex numbers of modulus 1, define

$$e = [z_1 \sqrt{a_1 - a_0} \quad z_2 \sqrt{a_2 - a_1} \quad \dots \quad z_n \sqrt{a_n - a_{n-1}}]^T,$$

and let $Q = I - e \otimes e^*$.

- (i) Q is a projection of rank $n - 1$ such that $\text{dist}(Q, \mathcal{T}_n) = v_{n-1,n}$. Moreover, every projection of rank $n - 1$ that is of minimal distance to \mathcal{T}_n is of this form.
- (ii) There is a unique operator $T \in \mathcal{T}_n$ of minimal distance to Q , and this T is such that $Q - T = v_{n-1,n} U$ for some unitary $U \in \mathbb{M}_n(\mathbb{C})$. Thus, if $q_k = Q e_k$ and $t_k = T e_k$ denote the columns of Q and T , respectively, then one can iteratively determine columns t_k by solving the system of linear equations

$$\begin{cases} \langle q_1 - t_1, q_k - t_k \rangle = 0 \\ \langle q_2 - t_2, q_k - t_k \rangle = 0 \\ \vdots \\ \langle q_{k-1} - t_{k-1}, q_k - t_k \rangle = 0 \end{cases}$$

for $k \in \{2, 3, \dots, n\}$.

Proof Statement (i) follows immediately from the results of §2 and §3. For statement (ii), the existence of T and U is guaranteed by Theorems 2.5 and 4.1. All that remains to show is the uniqueness of these operators.

To accomplish this task, note that it suffices to prove uniqueness in the case that $z_i = 1$ for all i (i.e., when $q_{ij} \leq 0$ for all $i \neq j$). For $k \in \{1, 2, \dots, n\}$, let Q_k denote the restriction of $E_{k-1}^\perp Q E_k$ to the range of E_k , and define $B_k := Q_k^* Q_k$. Let $Q'_k = E_k^\perp Q_k$, so that

$$Q_k = \begin{bmatrix} v_k^* \\ Q'_k \end{bmatrix},$$

where $v_k := [q_{k1} \ q_{k2} \ \dots \ q_{kk}]^T$.

We will demonstrate that $\|Q'_k\| < \|Q_k\|$ for all $k \in \{1, 2, \dots, n-1\}$, and therefore obtain the uniqueness of T and U via Theorem 4.1. Observe that this inequality holds when $k = 1$, as

$$\|Q_1\|^2 - \|Q'_1\|^2 = q_{11}^2 = v_{n-1,n} > 0.$$

Suppose now that $k \in \{2, 3, \dots, n-1\}$ is fixed, and define $B'_k := Q_k'^* Q'_k = B_k - v_k v_k^*$. One can determine the entries of $B'_k = (b'_{ij})$ using the formulas for the entries of $B_k = (b_{ij})$ from Lemma 2.2(i). Indeed,

$$\begin{aligned} b'_{kk} &= b_{kk} - q_{kk}^2 \\ &= q_{kk} - a_{k-1}(1 - q_{kk}) - q_{kk}^2 \\ &= (q_{kk} - a_{k-1})(1 - q_{kk}) = (1 - a_k)(1 - q_{kk}), \end{aligned}$$

and for if $i < k$,

$$\begin{aligned} b'_{ii} &= b_{ii} - q_{ki}^2 \\ &= (1 - a_{k-1})(1 - q_{ii}) - (1 - q_{kk})(1 - q_{ii}) \\ &= (q_{kk} - a_{k-1})(1 - q_{ii}) = (1 - a_k)(1 - q_{ii}). \end{aligned}$$

If i, j , and k are all distinct, then

$$\begin{aligned} b'_{ij} &= b_{ij} - q_{ki} q_{kj} \\ &= -(1 - a_{k-1})q_{ij} + q_{ij}(1 - q_{kk}) \\ &= -(q_{kk} - a_{k-1})q_{ij} = -(1 - a_k)q_{ij}. \end{aligned}$$

Finally, either $i < j = k$ or $j < i = k$. In the case of former, we have

$$\begin{aligned} b'_{ik} &= b_{ik} - q_{ki}q_{kk} \\ &= a_{k-1}q_{ik} - q_{ik}q_{kk} \\ &= -(q_{kk} - a_{k-1})q_{ik} = -(1 - a_k)q_{ik}. \end{aligned}$$

The fact that B'_k is self-adjoint implies that $b'_{kj} = -(1 - a_k)q_{kj}$ for all $j < k$ as well.

The above expressions for the entries b'_{ij} reveal that $B'_k = (1 - a_k)(I - \hat{Q})$, where $\hat{Q} \in \mathbb{M}_k(\mathbb{C})$ denotes the k -th leading principal submatrix of Q . Since Q has rank $n - 1$, it follows that $I - \hat{Q}$ has rank at most 1, and hence B'_k has at most one non-zero eigenvalue. Consequently,

$$\|B'_k\| = \text{Tr}(B'_k) = \sum_{\ell=1}^k (1 - a_k)(1 - q_{\ell\ell}) = a_k(1 - a_k).$$

Now let $f: [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ denote the function from Theorem 2.3, so that $\|Q_k\|^2 = f(a_{k-1}, a_k)$. Suppose for the sake of contradiction that $\|B_k\| = \|B'_k\|$, and hence $f(a_{k-1}, a_k) = a_k(1 - a_k)$. One can verify that for this equation to hold, we necessarily have that $a_k = 1$ or $a_k = a_{k-1}$.

If the former is true, then $a_j = 1$ for all $j \geq k$. In particular, $a_{n-1} = a_n$. From this it follows that $q_{nn} = 1 - (a_n - a_{n-1}) = 1$, and hence $\|Q_n\| \geq 1$. This contradicts the minimality of $\text{dist}(Q, \mathcal{T}_n)$. If instead $a_k = a_{k-1}$, then $q_{kk} = 1$, and thus $\|Q_k\| \geq 1$. Again we reach a contradiction. We therefore conclude that $\|B'_k\| < \|B_k\|$, hence $\|Q'_k\| < \|Q_k\|$. ■

To save the reader from lengthy computations, we have included a few examples of pairs (Q, T) where $Q \in \mathbb{M}_n(\mathbb{C})$ is a projection of rank $n - 1$, T belongs to \mathcal{T}_n , and $\|Q - T\| = \nu_{n-1,n}$. Theorem 4.2 implies that if (Q', T') is any other projection-nilpotent pair such that $\text{rank}(Q') = n - 1$ and $\|Q' - T'\| = \nu_{n-1,n}$, then there is a unitary $V \in \mathbb{M}_n(\mathbb{C})$ such that $Q' = V^*QV$ and $T' = V^*TV$. In each case, the entries of Q and T have been rounded to the fifth decimal place.

$n = 3$

$$Q = \begin{bmatrix} 0.64310 & -0.31960 & -0.35689 \\ -0.31960 & 0.71379 & -0.31960 \\ -0.35689 & -0.31960 & 0.64310 \end{bmatrix},$$

$$T = \begin{bmatrix} 0 & -0.49697 & -0.80194 \\ 0 & 0 & -0.49697 \\ 0 & 0 & 0 \end{bmatrix};$$

$n = 4$

$$Q = \begin{bmatrix} 0.72361 & -0.24860 & -0.24860 & -0.27639 \\ -0.24860 & 0.77639 & -0.22361 & -0.24860 \\ -0.24860 & -0.22361 & 0.77639 & -0.24860 \\ -0.27639 & -0.24860 & -0.24860 & 0.72361 \end{bmatrix},$$

$$T = \begin{bmatrix} 0 & -0.34356 & -0.46094 & -0.65836 \\ 0 & 0 & -0.34164 & -0.46094 \\ 0 & 0 & 0 & -0.34356 \\ 0 & 0 & 0 & 0 \end{bmatrix};$$

$n = 5$

$$Q = \begin{bmatrix} 0.77471 & -0.20512 & -0.19907 & -0.20512 & -0.22528 \\ -0.20512 & 0.81324 & -0.18126 & -0.18676 & -0.20512 \\ -0.19907 & -0.18126 & 0.82409 & -0.18126 & -0.19907 \\ -0.20512 & -0.18676 & -0.18126 & 0.81324 & -0.20512 \\ -0.22528 & -0.20512 & -0.19907 & -0.20512 & 0.77472 \end{bmatrix},$$

$$T = \begin{bmatrix} 0 & -0.26477 & -0.32678 & -0.41846 & -0.55566 \\ 0 & 0 & -0.26373 & -0.32453 & -0.41846 \\ 0 & 0 & 0 & -0.26373 & -0.32678 \\ 0 & 0 & 0 & 0 & -0.26477 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

It is interesting to note that each projection above is symmetric about its *anti-diagonal*, the diagonal from the $(n, 1)$ -entry to the $(1, n)$ -entry. This symmetry is in fact, always present in the optimal projection $Q = (q_{ij})$ from Theorem 4.2 obtained by taking $z_i = 1$ for all i . To see this, first observe that the function h_t from equation (3.1) satisfies the identity

$$h_t(x) + h_t^{-1}(1 - x) = 1, \quad x \in [0, 1].$$

From here we have that $a_1 + a_{n-1} = h_t(0) + h_t^{-1}(1) = 1$, and by induction,

$$a_k + a_{n-k} = h_t(a_{k-1}) + h_t^{-1}(a_{n-k+1}) = h_t(a_{k-1}) + h_t^{-1}(1 - a_{k-1}) = 1$$

for all $k \in \{1, 2, \dots, n\}$. Consequently,

$$\begin{aligned} q_{kk} &= 1 - (a_k - a_{k-1}) \\ &= a_{n-k} + a_{k-1} \\ &= a_{n-k} + (1 - a_{n-k+1}) \\ &= 1 - (a_{n-k+1} - a_{n-k}) = q_{n-k+1, n-k+1} \end{aligned}$$

for all k . We now turn to the identity $q_{ij} = -\sqrt{(1 - q_{ii})(1 - q_{jj})}$ to conclude that that $q_{ij} = q_{n-j+1, n-i+1}$ for all i and j , which is exactly the statement that Q is symmetric about its anti-diagonal. An analogous argument using the formulas from [6] demonstrates a similar phenomenon for optimal projections of rank 1.

5 Conclusion

The distance $v_{r,n}$ from the set of projections in $\mathbb{M}_n(\mathbb{C})$ of rank r to the set of nilpotent operators on \mathbb{C}^n , as well as the corresponding closest projection-nilpotent pairs, are now well understood when $r = 1$ or $r = n - 1$. Of course, it is natural to wonder about the value of $v_{r,n}$ for r strictly between 1 and $n - 1$.

The difficulty in extending the above arguments to projections P of intermediate ranks lies in deriving closed-form expressions for $\|E_{i-1}^{\perp} P E_i\|$. Computing these norms for projections of rank $r = 1$ or $r = n - 1$ was made possible by the simple structure afforded by such projections. In particular, we made frequent use of equations (2.2) and (2.3) throughout the proofs of Lemma 2.2 and Theorem 2.3 to relate the off-

diagonal entries of such projections to those on the diagonal. Analogous equations for projections of intermediate ranks become significantly more complex.

For small values of r and n , the mathematical programming software *Maple* was used to construct examples of rank r projections $P_{r,n}$ in $\mathbb{M}_n(\mathbb{C})$ which we believe are of minimal distance to \mathcal{T}_n . To ease the computations, the program was tasked with minimizing the maximum norm $\|E_{i-1}^\perp P E_i\|$ over all projections P of rank r with real entries and symmetry about the anti-diagonal. While it may not always be possible for such conditions to be met by an optimal projection of rank r , the computations that follow may still shed light on a potential formula for $v_{r,n}$.

The smallest value of n for which $\mathcal{P}(\mathbb{C}^n)$ contains projections of intermediate ranks is $n = 4$. In this case, the intermediate-rank projections are those of rank 2. We found that

$$P_{2,4} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix}$$

is an optimal projection of rank 2 satisfying the conditions above. It is easy to see that

$$\|E_{i-1}^\perp P_{2,4} E_i\| = 1/\sqrt{2} = v_{1,2} \quad \text{for all } i,$$

and hence $P_{2,4}$ is a direct sum of optimal rank-one projections in $\mathbb{M}_2(\mathbb{C})$.

In $\mathbb{M}_5(\mathbb{C})$, the intermediate-rank projections are those of rank $r = 2$ or $r = 3$. For such r , we obtained

$$P_{2,5} = \begin{bmatrix} 0.42602 & -0.07632 & 0.22568 & 0.42334 & -0.09248 \\ -0.07632 & 0.42127 & 0.23481 & -0.06022 & 0.42334 \\ 0.22568 & 0.23481 & 0.30541 & 0.23481 & 0.22568 \\ 0.42334 & -0.06022 & 0.23481 & 0.42127 & -0.07632 \\ -0.09248 & 0.42334 & 0.22568 & -0.07632 & 0.42602 \end{bmatrix}$$

$$P_{3,5} = \begin{bmatrix} 0.58296 & -0.29271 & -0.10684 & 0.12213 & 0.36209 \\ -0.29271 & 0.62479 & -0.33169 & -0.15433 & 0.12213 \\ -0.10684 & -0.33169 & 0.58448 & -0.33169 & -0.10684 \\ 0.12213 & -0.15433 & -0.33169 & 0.62479 & -0.29271 \\ 0.36209 & 0.12213 & -0.10684 & -0.29271 & 0.58296 \end{bmatrix},$$

with entries rounded to the fifth decimal place. Again, the norms $\|E_{i-1}^\perp P_{r,n} E_i\|$ share a common value, with

$$\|E_{i-1}^\perp P_{2,5} E_i\| \approx 0.65270 \approx \frac{1}{2} \sec\left(\frac{\pi}{\frac{5}{2} + 2}\right) \quad \text{for all } i, \text{ and}$$

$$\|E_{i-1}^\perp P_{3,5} E_i\| \approx 0.76352 \approx \frac{1}{2} \sec\left(\frac{\pi}{\frac{5}{3} + 2}\right) \quad \text{for all } i.$$

In light of these findings, as well as the distance formulas that exist for projections of rank 1 or $n - 1$, we propose the following generalized distance formula for projections of arbitrary rank.

Conjecture 5.1 For every $n \in \mathbb{N}$ and each $r \in \{1, 2, \dots, n\}$, the distance from the set of projections in $\mathbb{M}_n(\mathbb{C})$ of rank r to $\mathcal{N}(\mathbb{C}^n)$ is

$$v_{r,n} = \frac{1}{2} \sec\left(\frac{\pi}{\frac{n}{r} + 2}\right).$$

Using a random walk process implemented by the computer algebra system *PARI/GP*, we estimated the values of $v_{r,n}$ for all $r \leq n \leq 10$ without the additional assumptions described above. We observed only minute differences between these estimates and the expression from Conjecture 5.1. In many cases, these quantities differed by no more than 1×10^{-3} .

The proposed formula from Conjecture 5.1 merits several interesting consequences. First, this formula suggests that $v_{r,n} = v_{kr,kn}$ for every positive integer k , meaning that a closest projection of rank kr to \mathcal{T}_{kn} could be obtained as a direct sum of k closest projections of rank r to \mathcal{T}_n . Notice as well that if the equation $v_{r,n} = v_{kr,kn}$ were true, it would follow that $v_{1,n} = v_{r,rn} \leq v_{r,n}$ for each n and r . Thus, a proof of Conjecture 5.1—or of the formula $v_{r,n} = v_{kr,kn}$ —would validate Conjecture 1.3.

Acknowledgment The author would like to thank Paul Skoufranis for many stimulating conversations and Boyu Li for providing the *PARI/GP* script used to estimate $v_{r,n}$.

References

- [1] W. B. Arveson, *Interpolation problems in nest algebras*. J. Functional Analysis 20(1975), 208–233. [https://doi.org/10.1016/0022-1236\(75\)90041-5](https://doi.org/10.1016/0022-1236(75)90041-5)
- [2] D. Bini, Y. Eidelman, L. Gemignani, and I. Gohberg, *The unitary completion and QR iterations for a class of structured matrices*. Math. Comp. 77(2008), 353–378. <https://doi.org/10.1090/S0025-5718-07-02004-2>
- [3] J. H. Hedlund, *Limits of nilpotent and quasinilpotent operators*. Michigan Math. J. 19(1972), 249–255.
- [4] D. A. Herrero, *Normal limits of nilpotent operators*. Indiana Univ. Math. J. 23(1974), 1097–1108.
- [5] D. A. Herrero, *Unitary orbits of power partial isometries and approximation by block-diagonal nilpotents*. In: Topics in modern operator theory. Springer, 1981, pp. 171–210.
- [6] G. W. MacDonald, *Distance from projections to nilpotents*. Canad. J. Math. 47(1995), 841–851. <https://doi.org/10.4153/CJM-1995-043-3>
- [7] G. W. MacDonald, *Distance from idempotents to nilpotents*. Canad. J. Math. 59(2007), 638–657. <https://doi.org/10.4153/CJM-2007-027-x>
- [8] S. C. Power, *The distance to upper triangular operators*. Math. Proc. Cambridge Philos. Soc. 88(1980), 327–329. <https://doi.org/10.1017/S030500410057637>
- [9] N. Salinas, *On the distance to the set of compact perturbations of nilpotent operators*. J. Operator Theory 3(1980), 179–194.

Department of Pure Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1

e-mail: zcramer@uwaterloo.ca