Convex Combinations of Projections

Man-Duen Choi*
Department of Mathematics
University of Toronto
Toronto, Ontario, Canada M5S 1A1

and

Pei Yuan Wu†
Department of Applied Mathematics
National Chiao Tung University
Hsinchu, Taiwan, Republic of China

Submitted by Chandler Davis

ABSTRACT

On an n-dimensional inner-product space, every operator T that satisfies $O \leqslant T \leqslant I$ is a convex combination of as few as $[\log_2 n]+2$ projections, and this number is sharp. If $O \leqslant T \leqslant I$ and trace T is a rational number, then T is an average of projections. Further results are also obtained for the cases when the projections are required to have the same rank and/or to be commuting. In each case, the optimal number of projections is determined.

0. INTRODUCTION

Which linear operator on a complex n-dimensional inner-product space can be expressed as a *convex combination*

$$\lambda_1 P_1 + \cdots + \lambda_m P_m \qquad \text{with} \quad \lambda_j \geqslant 0 \text{ and } \sum \lambda_j = 1$$

*Supported in part by NSERC of Canada.

†The research of this author was done while he was visiting University of Toronto in 1987–88, supported in part by NSC of the Republic of China. He would like to take this opportunity to thank the operator theorists at University of Toronto for their hospitality during this period.

655 Avenue of the Americas, New York, NY 10010 0024-3795/90/$3.50

or an *average* (= arithmetic mean)

$$\frac{1}{m}(P_1 + \cdots + P_m)$$

of finitely many (orthogonal) projections P_j? What is the minimal value of m, the number of projections, required in such an expression? These are among the questions to be addressed in this paper. Note that, here, projections P_j need not be commuting; so the underlying structure theory is inevitably complicated, yet highly intriguing. Readers are also referred to [1–3, 6, 7] for some related research work with judicious manipulations of noncommuting projections.

Actually, we are concerned with the affine structure of the convex compact set

$$\mathscr{C} = \mathscr{C}_n = \{n \times n \text{ positive semidefinite matrices } T \text{ satisfying } T \leqslant I\}.$$

In view of the well-known fact

$$\operatorname{Ext} \mathscr{C} = \{n \times n \text{ projections}\},$$

we proceed to seek a quantitative description for the statement

$$\mathscr{C} = \operatorname{co} \operatorname{Ext} \mathscr{C}.$$

(Here, Ext stands for the extremal set and co stands for the convex hull.) Since $\mathscr{C}$ is a subset of $\{n \times n \text{ hermitian matrices}\}$—a real linear space of real dimension n^2—it follows that, from an elementary classical theorem of Carathéodory, each operator $T \in \mathscr{C}$ is a convex combination of $n^2 + 1$ projections. Nevertheless, a simple diagonalization argument yields a familiar fact: each operator $T \in \mathscr{C}$ is a convex combination of $n + 1$ commuting projections (see Proposition 1.4). To get the ultimate result, we need an optimal manipulation of noncommuting projections; it turns out that the "most economical" way to form convex combination requires as few as $[\log_2 n] + 2$ projections (Theorem 2.4). Along these lines, we also get a description of the averages of projections (Theorem 3.6).

Moreover, Ext $\mathscr{C}$ consists of exactly $n + 1$ components:

$$\operatorname{Ext} \mathscr{C} = \bigcup_{k=0}^{n} \mathscr{R}_k,$$

where $\mathscr{R}_k = \{n \times n \text{ rank-}k \text{ projections}\}$. It is not surprising to see that the affine structure of

$$\operatorname{co} \mathscr{R}_k = \{T \in \mathscr{C} : \operatorname{trace} T = k\}$$

is much more tractable than that of $\mathscr{C}$. Indeed, the class of convex combinations of rank-k projections is exactly the same as the class of averages of rank-k projections. The complete description for this class of operators is given in Theorem 3.5.

Notably, $\mathscr{C} = \{\text{positive semidefinite contractions}\}$ is isomorphic with $\mathscr{C}' = \{\text{hermitian contractions}\}$ under the affine map $T \leftrightarrow 2T - I$. Thus all results in this paper about $\operatorname{Ext} \mathscr{C} = \{\text{projections}\}$ can be appropriately translated to results about $\operatorname{Ext} \mathscr{C}' = \{\text{symmetries}\}$.

For the sake of completeness, we list some preliminary results in Section 1. This section also includes simple structure theorems about commuting projections. Section 2 is devoted to the investigation of convex combinations of noncommuting projections, and Section 3 to averages.

1. NOTATION AND PRELIMINARIES

In this paper, we deal with matrices of complex entries. A matrix P is called a *projection* if P is self-adjoint and idempotent (i.e., $P = P^* = P^2$). A matrix J is called a *symmetry* if $J = J^* = J^{-1}$. We write O for the zero matrix and I for the identity matrix. We write $S \leqslant T$ or $O \leqslant T - S$ when $T - S$ is a positive semidefinite matrix. We write $\operatorname{Diag}(t_j)_{j=1}^n$ for the $n \times n$ diagonal matrix with diagonal entries $(t_j)_{j=1}^n$. Each hermitian matrix T has a polar decomposition $T = |T|J = J|T|$, where $|T|$ is the positive semidefinite square root of T^2 and $J = f(T)$ is a symmetry defined by the real-valued function f with $f(t) = 1$ if $t \geqslant 0$ and $f(t) = -1$ if $t < 0$.

Now, we collect three trivial lemmas.

LEMMA 1.1. *Suppose T is a convex combination (respectively, an average) of m projections. If l is an integer $\geqslant m$ (l is a positive integer multiple of m), then T also admits an expression as a convex combination (an average) of l projections.*

(Note that herein projections need not be distinct.)

LEMMA 1.2. *A square matrix T is an average of projections iff T is a convex combination of projections with rational coefficients.*

LEMMA 1.3. *A square matrix T is a convex combination (respectively, an average) of m projections iff $I - T$ is so.*

The structure theory for convex combinations of commuting projections is rather simple. The following proposition is probably known to many readers.

PROPOSITION 1.4. *Let T be an $n \times n$ matrix satisfying $O \leqslant T \leqslant I$.*

(1) *Then T admits an expression as a convex combination of "$n+1$" commuting projections.*

(2) *If $n > k$ are positive integers and $\operatorname{trace} T = k$, then T admits an expression as a convex combination of "n" commuting rank-k projections.*

The "quoted" number of commuting projections in each expression is sharp in the sense that it is the smallest integer for the statement to be valid.

Proof. Write $T = \sum_{j=1}^{n} t_j E_j$, where $1 \geqslant t_1 \geqslant \cdots \geqslant t_n \geqslant 0$ and $\{E_j\}_{j=1}^{n}$ are mutually orthogonal rank-one projections.

(1): Let $F_j = E_1 + \cdots + E_j$ $(j = 1, \ldots n)$; then

$$T = (1 - t_1)O + (t_1 - t_2)F_1 + \cdots + (t_{n-1} - t_n)F_{n-1} + t_n F_n$$

is a convex combination of $n+1$ commuting projections.

(2): Now we have the extra assumption $\sum_j t_j = k$. We may assume further that $t_1 + t_n \leqslant 1$ (otherwise, consider $I - T$ and $n - k$ instead of T and k). Let $P = \sum_{j=n-k+1}^{n} E_j$ and $S = T - t_n P$. Then P is a rank-k projection, and rank $S \leqslant n - 1$, trace $(1/(1 - t_n)S) = k$, $0 \leqslant 1/(1 - t_n)S \leqslant I$. By the induction hypothesis, which is obviously valid if $n = 2$, we can write

$$\frac{1}{1 - t_n} S = \sum_{i=1}^{n-1} \lambda_i P_i,$$

where each P_i, as a nonnegative sum of E_j's, is a projection of rank k, and $\sum \lambda_i = 1$, $\lambda_i \geqslant 0$. Hence

$$T = t_n P + \sum_{i=1}^{n-1} (1 - t_n)\lambda_i P_i$$

is a convex combination of n commuting projections of rank k, as desired.

In order to show the sharpness of $n+1$ as the optimal number of projections in (1), let $\{t_1, \ldots, t_n, 1\} \subset [0, 1]$ be linearly independent over the rational field $\mathbf{Q}$ (e.g., $t_i = 2^{1/(i+1)}$), and let T be the $n \times n$ diagonal matrix

$\mathrm{Diag}(t_i)_{i=1}^n$. Suppose $T = \sum_{j=1}^m \lambda_j P_j$, where $\lambda_j \geqslant 0$, $\sum \lambda_j = 1$, and the P_j's are commuting projections; we wish to prove that $m \geqslant n+1$. Since $P_j T = T P_j$ for all j, it follows that

$$P_j = \mathrm{Diag}(\delta_{ij})_{i=1}^n \qquad \text{with} \quad \delta_{ij} \in \{0,1\}.$$

Thus

$$t_i = \sum_j \delta_{ij} \lambda_j \qquad \text{for all } i.$$

This, together with $\sum \lambda_j = 1$, is equivalent to

$$\begin{bmatrix} t_1 \\ \vdots \\ t_n \\ 1 \end{bmatrix} = \begin{bmatrix} \delta_{11} & \cdots & \delta_{1m} \\ \vdots & & \vdots \\ \delta_{n1} & \cdots & \delta_{nm} \\ 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix}.$$

As the linear span of $\{t_1, \ldots, t_n, 1\}$ over $\mathbf{Q}$ is of dimension $n+1$, the linear span of $\{\lambda_1, \ldots, \lambda_m\}$ over $\mathbf{Q}$ is of dimension at least $n+1$. Therefore $m \geqslant n+1$, as desired.

To show the sharpness of n as the optimal number of rank-k projections in (2), we choose $\{t_1, \ldots, t_n\} \subset [0,1]$ to be a linearly independent set over the rational field $\mathbf{Q}$ and $t_1 + \cdots + t_n = k$. [For example, let $t_j = k a_j / (a_1 + \cdots + a_n)$ with $a_j = 1 - (n\pi)^{-j}$.] Following the same argument in the last paragraph, we can prove that $\mathrm{Diag}(t_j)_{j=1}^n$ cannot be written as a convex combination of fewer than n commuting projections of rank k. ■

The structure theory for averages of commuting projections is also very simple. Part (1) of the following proposition has already appeared in [3, Theorem 3].

PROPOSITION 1.5. *Let $n \geqslant k \geqslant 0$ be nonnegative integers, and let T be an $n \times n$ matrix satisfying $O \leqslant T \leqslant I$. Then*

(1) *T admits an expression as an average of commuting projections iff all eigenvalues of T are rational numbers in $[0,1]$;*

(2) *T admits an expression as an average of commuting rank-k projections iff all eigenvalues of T are rational numbers in $[0,1]$ and $\mathrm{trace}\, T = k$.*

The minimal number of commuting projections required in (1) (or in (2) if $n > k > 0$) can be arbitrarily large.

Proof. The "only if" parts are trivial because commuting projections are simultaneously diagonalizable. Conversely, if $T = \sum t_j E_j$ with rational $t_j \in [0,1]$ and mutually orthogonal rank-1 projections E_j (respectively, with the extra condition $\sum t_j = k$), then the proof of Proposition 1.4 shows that T is a convex combination of commuting projections (of commuting projections of rank k) with rational coefficients. By Lemma 1.2, we are done.

To count the optimal number of commuting projections in the average expression, we consider a diagonal $n \times n$ matrix $T = \mathrm{Diag}(t_j)_{j=1}^n$. Suppose $T = (1/m)\sum_{j=1}^m P_j$ is an average of m commuting projections. Then we may assume that the P_j's are also diagonal matrices; thus each entry of T is an integer multiple of $1/m$. In particular, if $t_n = 1/l$, then m must be a positive integer multiple of l, which is large when l is a large integer. ∎

2. CONVEX COMBINATIONS

In this section, we consider the minimal number of projections in convex combinations where noncommuting projections are allowed. We start with manipulations on 2×2 matrices.

LEMMA 2.1. *Suppose a, b, and c are real numbers satisfying $1 \geqslant a \geqslant c \geqslant b \geqslant 0$ and $\frac{1}{2} \geqslant c \geqslant 0$. Then there exist 2×2 matrices P, Q, C such that P and Q are projections, $O \leqslant C \leqslant I$, $\mathrm{rank}\, C \leqslant 1$, $QC = CQ = O$, and*

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} = cP + (1-c)(Q + C).$$

Proof. We may assume $a > b$. (Otherwise $a = b = c$; we can take $P = I$, $Q = C = O$.) We have to consider two possible cases:

(a) $a + b \leqslant 1$. We set

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} = R_1 + R_2$$

with

$$R_1 = \frac{1}{a-b}\begin{bmatrix} a(c-b) & \alpha \\ \alpha & b(a-c) \end{bmatrix}, \quad R_2 = \frac{1}{a-b}\begin{bmatrix} a(a-c) & -\alpha \\ -\alpha & b(c-b) \end{bmatrix},$$

$\alpha = \{ab(a-c)(c-b)\}^{1/2}$. Since R_1 and R_2 are positive semidefinite matrices of rank $\leqslant 1$, and trace $R_1 = c$ and trace $R_2 = a + b - c \leqslant 1 - c$, it follows

that $R_1 = cP$ and $R_2 = (1-c)(Q+C)$, where P is a projection, $Q = O$, and C is of rank 1, and $O \leqslant C \leqslant I$, as desired.

(b) $a+b \geqslant 1$. Then $a+c \geqslant 1 \geqslant b+c$, and $2 \geqslant a+2c \geqslant a+b+c$. We get

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} = R_1 + R_2 + (a+b-1)I$$

with

$$R_1 = \frac{1}{a-b}\begin{bmatrix} (1-b)(a+c-1) & \beta \\ \beta & (1-a)(1-b-c) \end{bmatrix},$$

$$R_2 = \frac{1}{a-b}\begin{bmatrix} (1-b)(1-b-c) & -\beta \\ -\beta & (1-a)(a+c-1) \end{bmatrix},$$

$\beta = \{(1-a)(1-b)(a+c-1)(1-b-c)\}^{1/2}$. Since R_1 and R_2 are positive semidefinite matrices of rank $\leqslant 1$, and trace $R_1 = c$, trace $R_2 = 2-a-b-c$, it follows that $R_1 = cP$, $R_2 = (2-a-b-c)Q$, where P and Q are rank-1 projections. Thus

$$\begin{aligned} R_1 + R_2 + (a+b-1)I &= cP + (1-c)Q + (a+b-1)(I-Q) \\ &= cP + (1-c)Q + (1-c)C \end{aligned}$$

with

$$C = \frac{a+b-1}{1-c}(I-Q) \leqslant \frac{a+b-1}{1-c}I \leqslant I,$$

as desired. ∎

For each real number $x \geqslant 0$, $[x]$ denotes its integral part.

PROPOSITION 2.2. *Each $n \times n$ matrix T that satisfies $O \leqslant T \leqslant I$ admits an expression as a convex combination of $[\log_2 n]+2$ projections.*

Proof. We prove the proposition by induction. When $n = 1$, $[\log_2 n]+2 = 2$ and the statement is obviously valid. By unitary equivalence, each $n \times n$

matrix T satisfying $O \leqslant T \leqslant I$ can be written as $T = \mathrm{Diag}(t_j)_{j=1}^n$ with $1 \geqslant t_1 \geqslant \cdots \geqslant t_n \geqslant 0$. Let t be the "median" diagonal entry

$$t = \begin{cases} t_{(n+1)/2} & \text{if } n \text{ is odd,} \\ t_{n/2} + t_{(n/2)+1} & \text{if } n \text{ is even.} \end{cases}$$

Without loss of generality, we may assume $0 \leqslant t \leqslant \frac{1}{2}$ (otherwise, consider $I - T$ instead of T by Lemma 1.3). By Lemma 2.1,

$$\begin{bmatrix} t_j & 0 \\ 0 & t_{n-j} \end{bmatrix} = tP_j + (1-t)(Q_j + C_j),$$

where P_j, Q_j, C_j are 2×2 matrices, P_j and Q_j are projections, $\operatorname{rank} C_j \leqslant 1$, and $O \leqslant C_j \leqslant I$ and $Q_j C_j = C_j Q_j = O$. Since

$$T = \bigoplus_{j=1}^{[n/2]} \begin{bmatrix} t_j & 0 \\ 0 & t_{n-j} \end{bmatrix} \qquad \text{if } n \text{ is even}$$

(T has an extra 1×1 matrix direct summand with entry t if n is odd), it follows that

$$T = tP + (1-t)(Q + C),$$

where P and Q are projections, $\operatorname{rank} C \leqslant [n/2]$, $O \leqslant C \leqslant I$, and $QC = CQ = O$. By the induction hypothesis, we can write

$$C = \sum_{j=1}^{m} \lambda_j R_j,$$

where $m = [\log_2 [n/2]] + 2 = [\log_2 n] + 1$, $\lambda_j \geqslant 0$, $\Sigma \lambda_j = 1$, and R_j are projections with Range $R_j \subseteq$ Range C. Therefore

$$\begin{aligned} T &= tP + (1-t)(Q + C) \\ &= tP + \sum_{j=1}^{m} (1-t)\lambda_j (Q + R_j) \end{aligned}$$

is a convex combination of $m + 1 = [\log_2 n] + 2$ projections, as desired. ■

In order to show that the number $[\log_2 n]+2$ is sharp in the proposition above, we need the following known result (see, e.g., [4, p. 182, Corollary 4.3.3] for the proof).

LEMMA 2.3. *Let S and T be $n \times n$ hermitian matrices with eigenvalues $s_1 \geqslant \cdots \geqslant s_n$ and $t_1 \geqslant \cdots \geqslant t_n$ respectively. If $S \leqslant T$, then $s_j \leqslant t_j$ for all j.*

In the next theorem, we write $p(n)$ for the smallest integer m such that *every* $n \times n$ matrix T that satisfies $O \leqslant T \leqslant I$ admits an expression as a convex combination of m projections.

THEOREM 2.4. $p(n) = [\log_2 n]+2$.

Proof. Proposition 2.2 says $p(n) < [\log_2 n]+2$. In order to get the reverse inequality, we first construct, for each $n = 2^N$ (N is a positive integer), an $n \times n$ matrix T such that $O \leqslant T \leqslant I$, but T is not a convex combination of fewer than $N+2$ projections. Specifically, consider the $2^N \times 2^N$ diagonal matrix

$$T = \mathrm{Diag}(t, t^2, t^3, \ldots, t^{2^N}),$$

where t is a small positive real number [e.g., $0 < t < (N2^N)^{-1}$]. Suppose $T = \sum_{j=1}^{m} \lambda_j P_j$, where $m \leqslant N+1$, $\lambda_j \geqslant 0$, $\Sigma \lambda_j = 1$, each P_j is a projection of rank r_j, and $0 \leqslant r_1 \leqslant \cdots \leqslant r_m \leqslant 2^N$. Thus if $t \leqslant \frac{1}{2}$, we get

$$1 > t + t^2 + \cdots + t^{2^N} = \mathrm{trace}\, T = \sum \lambda_j r_j \geqslant \sum \lambda_j r_1 = r_1;$$

this proves that $r_1 = 0$ and $P_1 = O$. From

$$0 + 1 + 2 + \cdots + 2^{N-1} < 2^N = \mathrm{rank}\, T = \mathrm{rank}\left(\sum_{j=1}^{m} \lambda_j P_j \right) \leqslant 0 + r_2 + \cdots + r_m$$

and $m \leqslant N+1$, there exists an integer $k > 1$ such that $r_k \geqslant 2^{k-2} + 1$ but $r_j \leqslant 2^{j-2}$ for all $j < k$. Since $\sum_{j=1}^{k-1} \lambda_j P_j \leqslant T$ and

$$\mathrm{rank}\left(\sum_{j=1}^{k-1} \lambda_j P_j \right) \leqslant \sum_{j=1}^{k-1} r_j \leqslant \sum_{j=2}^{k-1} 2^{j-2} = 2^{k-2} - 1,$$

it follows that, by Lemma 2.3, the sum of the largest $2^{k-2} - 1$ eigenvalues of

$\sum_{j=1}^{k-1}\lambda_j P_j$ is less than or equal to the sum of those of T; i.e.,

$$\sum_{j=1}^{k-1}\lambda_j r_j = \operatorname{trace}\left(\sum_{j=1}^{k-1}\lambda_j P_j\right) \leqslant \sum_{j=1}^{2^{k-2}-1} t^j.$$

Similarly, from the fact $\lambda_j P_j \leqslant T$, we get $\lambda_j \leqslant t^{r_j}$ by Lemma 2.3. Thus

$$\sum_{j=1}^{2^N} t^j = \operatorname{trace} T = \sum_{j=1}^{k-1}\lambda_j r_j + \sum_{j=k}^{m}\lambda_j r_j \leqslant \sum_{j=1}^{2^{k-2}-1} t^j + \sum_{j=k}^{m} t^{r_j} r_j;$$

and hence

$$t^{2^{k-2}} \leqslant \sum_{j=2^{k-2}}^{2^N} t^j \leqslant \sum_{j=k}^{m} t^{r_j} r_j \leqslant t^{r_k}(m-k+1)2^N \leqslant t^{2^{k-2}+1} N\cdot 2^N,$$

which leads to $t \geqslant (N\cdot 2^N)^{-1}$, a contradiction. Therefore $p(2^N) = N+2$.

For a general positive integer $n > 1$, say $2^N \leqslant n < 2^{N+1}$ for some $N \geqslant 1$, we have

$$p(n) \leqslant [\log_2 n] + 2 = N + 2 = p(2^N).$$

It is clear from definition that $p(2^N) \leqslant p(n)$, and so $p(n) = [\log_2 n] + 2$, as desired. ■

A simplified version of the argument above can be used to prove the following result about nonnegative real linear combinations of projections. We leave the details to the reader.

COROLLARY 2.5. *Each $n \times n$ positive semidefinite matrix is a linear combination of $[\log_2 n]+1$ projections with nonnegative real coefficients. For each $n \geqslant 1$, the number $[\log_2 n]+1$ is sharp.*

Note that the first half of Corollary 2.5 has been essentially proved by Nakamura [6, p. 135]. The difference of the numbers of projections in the preceding two theorems reflects the fact that convex combinations require one extra constraint on the coefficients: their sums must be one.

3. AVERAGES

This section is devoted to the study of the averages of finitely many projections. We need simple manipulations on pairs of projections.

LEMMA 3.1. *Suppose two $n \times n$ matrices P and Q are projections of the same rank. Then there exists a symmetry J such that $Q = JPJ$.*

Proof. First assume that $P + Q - I$ is invertible. Then

$$J = |P + Q - I|(P + Q - I)^{-1}$$

is a symmetry. Since

$$(P + Q - I)Q = P(P + Q - I), \qquad (P + Q - I)P = Q(P + Q - I),$$

it follows that $(P + Q - I)^2$ commutes with Q, and thus $|P + Q - I|$ commutes with Q. Therefore

$$JP = |P + Q - I|(P + Q - I)^{-1}P = |P + Q - I|Q(P + Q - I)^{-1}$$

$$= Q|P + Q - I|(P + Q - I)^{-1} = QJ,$$

and $Q = JPJ$, as desired.

In general, $P + Q - I$ need not be invertible. Let $\mathscr{H}$ be the underlying Hilbert space, and let

$$\mathscr{H}_0 = (P + Q - I)\mathscr{H},$$

$$\mathscr{H}_1 = \{x \in \mathscr{H} : Px = x \text{ and } Qx = 0\},$$

$$\mathscr{H}_2 = \{x \in \mathscr{H} : Px = 0 \text{ and } Qx = x\}.$$

Then

$$y \in \mathscr{H}_0^{\perp} \quad \Leftrightarrow \quad (P + Q - I)y = 0$$

$$\Leftrightarrow \quad y = Py + Qy \quad \text{with } QPy = PQy = 0$$

$$\Leftrightarrow \quad y \in \mathscr{H}_1 + \mathscr{H}_2;$$

thus $\mathscr{H}_0$, $\mathscr{H}_1$, and $\mathscr{H}_2$ are mutually orthogonal subspaces of $\mathscr{H}$, and

$\mathscr{H} = \mathscr{H}_0 + \mathscr{H}_1 + \mathscr{H}_2$. With respect to this orthogonal decomposition of $\mathscr{H}$, we can write

$$P = \begin{bmatrix} P_0 & O & O \\ O & I & O \\ O & O & O \end{bmatrix}, \qquad Q = \begin{bmatrix} Q_0 & O & O \\ O & O & O \\ O & O & I \end{bmatrix}.$$

By the argument before, there exists a symmetry $J_0 \in \mathscr{L}(\mathscr{H}_0)$ such that $Q_0 = J_0 P_0 J_0$. Since P and Q are of the same rank, it follows that $\dim \mathscr{H}_1 = \dim \mathscr{H}_2$. Therefore

$$J = \begin{bmatrix} J_0 & O & O \\ O & O & I \\ O & I & O \end{bmatrix}$$

will satisfy $Q = JPJ$. ∎

LEMMA 3.2. *Suppose two $n \times n$ matrices P and Q are projections of rank k. If $a \geqslant b \geqslant c \geqslant d \geqslant 0$ are real numbers satisfying $a + d = b + c$, then there exist rank-k projections R_1 and R_2 such that $aP + dQ = bR_1 + cR_2$.*

Proof. By Lemma 3.1, there exists a symmetry J such that $Q = JPJ$. With respect to the decomposition of

$$J = \begin{bmatrix} I & O \\ O & -I \end{bmatrix}$$

(here I and $-I$ need not have the same dimension), write

$$P = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix}, \qquad Q = \begin{bmatrix} A & -B \\ -B^* & C \end{bmatrix}.$$

Let $R_j = U_j^* P U_j$ with

$$U_j = \begin{bmatrix} I & O \\ O & e^{i\theta_j} I \end{bmatrix}$$

and $\theta_j \in [0, 2\pi)$ to be determined. Then each R_j is a projection of rank k.

The condition $aP + dQ = bR_1 + cR_2$ is equivalent to

$$a - d = be^{i\theta_1} + ce^{i\theta_2}.$$

The real number θ_2 is determined if we can make

$$c = |a - d - be^{i\theta_1}|,$$

which, by direct computation, is the same as

$$2(a-d)b\cos\theta_1 = (a-d)^2 + b^2 - c^2;$$

θ_1 is realizable because the hypotheses $a + d = b + c$ and $a \geqslant b \geqslant c \geqslant d \geqslant 0$ together imply

$$0 \leqslant (a-d)^2 + b^2 - c^2 = 2(a-d)b - 4(a-b)d \leqslant 2(a-d)b.$$

Therefore, the required projections R_j $(j = 1,2)$ can be constructed. ∎

The following corollary is analogous to a result of Kadison and Pedersen [5, Corollary 15] about averages of unitary operators in any C^*-algebra.

COROLLARY 3.3. *If an $n \times n$ matrix T is a convex combination of m rank-k projections, then T also admits an expression as an average of m rank-k projections.*

Proof. Let $T = \sum_{j=1}^{m} \lambda_j P_j$, where the P_j's are projections of rank k, $1 \geqslant \lambda_1 \geqslant \cdots \geqslant \lambda_m \geqslant 0$, and $\sum \lambda_j = 1$. Then $\lambda_1 \geqslant 1/m \geqslant \lambda_m$, and by Lemma 3.2, we can replace P_1 and P_m by two other projections of rank k, and replace the pair of coefficients (λ_1, λ_m) by $(1/m, \lambda_1 + \lambda_m - 1/m)$. Continuing this process in finitely many steps, we can change all coefficients to $1/m$. ∎

COROLLARY 3.4. *If an $n \times n$ matrix T is an average of m rank-k projections and l is an integer larger than m, then T also admits an expression as an average of l rank-k projections.*

Proof. If $T = (1/m)\sum_{j=1}^{m} P_j$ is the average of m projections of equal rank, then, obviously, $T = \sum_{j=1}^{l} \lambda_j P_j$ $[\lambda_1 = \cdots = \lambda_{m-1} = 1/m,\ \lambda_m = \cdots =$

$\lambda_l = 1/m(l-m+1)$, $P_m = \cdots = P_l]$ is a convex combination of l projections of equal rank. An application of Corollary 3.3 proves the assertion. ∎

THEOREM 3.5. *Let $n \geqslant k \geqslant 0$ be integers, and let T be an $n \times n$ matrix. The following are equivalent*:

(1) *T admits an expression as an average of rank-k projections.*
(2) *T admits an expression as a convex combination of rank-k projections.*
(3) *$O \leqslant T \leqslant I$ and* trace $T = k$.

In this case, the number of rank-k projections required in the expressions can be as few as

$$\begin{cases} 1 & \text{if } k = 0 \text{ or } n, \\ k+1 & \text{if } n/2 \leqslant k < n, \\ n-k+1 & \text{if } 0 < k \leqslant n/2, \end{cases}$$

and the number is sharp for each given pair (n, k).

Proof. (1) $\Rightarrow$ (2) $\Rightarrow$ (3) is obvious, and (2) $\Rightarrow$ (1) follows from Corollary 3.3. It remains to prove (3) $\Rightarrow$ (2) and the assertion on the number of projections, by induction.

Obviously, the statement is valid for $n = 1$, or $k = 0$, or $n = k$. Now let T be an $n \times n$ matrix such that $O \leqslant T \leqslant I$ and trace $T = k$. Write $T = \sum_{j=1}^{n} t_j E_j$, where $1 \geqslant t_1 \geqslant \cdots \geqslant t_n \geqslant 0$, $\sum t_j = k$, and the E_j's are mutually orthogonal rank-1 projections.

We first consider the special case $n = 2k > 0$. We may assume that $t_1 + t_n \leqslant 1$ (otherwise, consider $I - T$ instead of T). Since

$$t_1 + t_{k+1} \geqslant \frac{1}{k}(t_1 + \cdots + t_k) + \frac{1}{k}(t_{k+1} + \cdots + t_n) = 1,$$

we have $t_n \leqslant 1 - t_1 \leqslant t_{k+1}$. Applying Lemma 3.2, we obtain

$$t_{k+1}E_{k+1} + t_n E_n = (1-t_1)R_1 + (t_1 + t_{k+1} + t_n - 1)R_2,$$

where R_1 and R_2 are rank-1 projections with Range $R_j \subseteq$ Range $E_{k+1} +$ Range E_n. Let $P = E_2 + \cdots + E_k + R_1$ and

$$\begin{aligned} S &= T - t_1 E_1 - (1-t_1)P \\ &= \sum_{j=2}^{k} (t_1 + t_j - 1)E_j + \sum_{j=k+2}^{n-1} t_j E_j + (t_1 + t_{k+1} + t_n - 1)R_2. \end{aligned}$$

Then P is a rank-k projection, and

$$O \leqslant \frac{1}{t_1} S \leqslant I, \qquad \operatorname{trace}\left(\frac{1}{t_1} S\right) = k - 1, \qquad \operatorname{rank}\left(\frac{1}{t_1} S\right) \leqslant n - 2.$$

By the induction hypothesis, we can write $(1/t_1)S = \sum_{j=1}^{k} \lambda_j Q_j$, where $\lambda_j \geqslant 0$, $\sum \lambda_j = 1$, and the Q_j's are rank-$(k-1)$ projections with Range $Q_j \subseteq$ Range S. Therefore

$$T = t_1 E_1 + (1 - t_1)P + t_1 \sum_{j=1}^{k} \lambda_j Q_j = (1 - t_1)P + \sum_{j=1}^{k} t_1 \lambda_j (E_1 + Q_j)$$

is a convex combination of $k + 1$ projections of rank k.

To complete the induction, we turn to the case $n > 2k > 0$. We have

$$k(t_k + t_n) \leqslant t_1 + \cdots + t_k + \cdots + t_n = k; \qquad \text{i.e.,} \quad t_k + t_n \leqslant 1.$$

Let $P = \sum_{j=1}^{k-1} E_j + E_n$ and $S = T - t_n P$. Then P is a rank-k projection and

$$O \leqslant \frac{1}{1 - t_n} S \leqslant I, \qquad \operatorname{trace}\left(\frac{1}{1 - t_n} S\right) = k, \qquad \operatorname{rank}\left(\frac{1}{1 - t_n} S\right) \leqslant n - 1.$$

By the induction hypothesis, we can write

$$\frac{1}{1 - t_n} S = \sum_{j=1}^{n-k} \lambda_j Q_j$$

as a convex combination of $n - k$ rank-k projections. Therefore,

$$T = t_n P + \sum_{j=1}^{n-k} (1 - t_n) \lambda_j Q_j$$

is a convex combination of $n - k + 1$ rank-k projections. The case $n < 2k < 2n$ may be proved by considering $I - T$ and Lemma 1.3. This completes the induction.

To see that the number of rank-k projections required in convex combinations is sharp for the case $n \geqslant 2k > 0$, we let $T = \mathrm{Diag}(t_j)_{j=1}^n$ with

$$t_j = \begin{cases} 1 & \text{if } j < k, \\ 1/(n-k+1) & \text{if } j \geqslant k. \end{cases}$$

Then $O \leqslant T \leqslant I$ and $\mathrm{trace}\, T = k$. Suppose $T = \sum_{j=1}^m \lambda_j P_j$, where $\sum \lambda_j = 1$, $\lambda_1 \geqslant \cdots \geqslant \lambda_m \geqslant 0$, and the P_j's are rank-k projections. Then $T \geqslant \lambda_1 P_1 \geqslant (1/m)P_1$; thus, by Lemma 2.3, the kth largest eigenvalue of T is not less than that of $(1/m)P_1$; i.e., $1/(n-k+1) \geqslant 1/m$, so $m \geqslant n-k+1$, as desired. For the case $2k > n$, the assertion follows by symmetry. ∎

Finally, we consider averages of projections where the projections need not be of same rank. (Cf. Fillmore's result about sums of projections [2, Theorem 1].)

THEOREM 3.6. *Let T be an $n \times n$ matrix. Then T is an average of projections iff $O \leqslant T \leqslant I$ and* $\mathrm{trace}\, T$ *is rational. The minimal number of projections required in the average expression can be arbitrarily large for each fixed n.*

Proof. The "only if" part is obvious. Conversely, suppose $O \leqslant T \leqslant I$ and $\mathrm{trace}\, T$ is rational. Write $T = \sum_{j=1}^n t_j E_j$, where the E_j's are mutually orthogonal rank-1 projections with $1 \geqslant t_1 \geqslant \cdots \geqslant t_n \geqslant 0$ and $\sum_{j=1}^n t_j = p/q$ (p and q are positive integers). Write

$$qt_j = k_j + s_j \qquad \text{with} \quad k_j \in \mathbf{Z} \text{ and } s_j \in [0,1).$$

Then $q \geqslant k_1 \geqslant k_2 \geqslant \cdots \geqslant k_n \geqslant 0$ and

$$n > \sum_{j=1}^n s_j = q\sum t_j - \sum k_j = p - \sum k_j = k, \qquad \text{say.}$$

By Theorem 3.5, we can write the operator $\sum_{j=1}^n s_j E_j$ as an average of rank-k projections $(1/m)\sum_{i=1}^m Q_i$. Let $F_j = E_1 + \cdots + E_j$ $(j = 1, \ldots, n)$. If $1 > t_1$,

then $q > k_1$ and

$$T = \frac{1}{q}\sum_{j=1}^{n}(k_j + s_j)E_j$$

$$= \frac{1}{q}\{(q-1-k_1)O + (k_1-k_2)F_1 + \cdots + (k_{n-1}-k_n)F_{n-1} + k_nF_n\}$$

$$+ \frac{1}{q}\sum_{i=1}^{m}\frac{1}{m}Q_i$$

is a convex combination of projections with rational coefficients. If $1 = t_1 = \cdots = t_l > t_{l+1}$, then $q > k_{l+1}$, $s_1 = \cdots = s_l = 0$, and $Q_i \perp F_l$, thus

$$T = F_l + \frac{1}{q}\sum_{j=l+1}^{n}(k_j + s_j)E_j$$

$$= \frac{1}{q}\{(q-1-k_{l+1})F_l + (k_{l+1}-k_{l+2})F_{l+1}$$

$$+ \cdots + (k_{n-1}-k_n)F_{n-1} + k_nF_n\}$$

$$+ \frac{1}{q}\sum_{i=1}^{m}\frac{1}{m}(Q_i + F_l)$$

is a convex combination of projections with rational coefficients. By Lemma 1.2, we conclude that T is an average of projections.

To count the optimal number of projections in the average expressions, we consider any $n \times n$ matrix T satisfying $O \leqslant T \leqslant I$ and trace $T = p/q$, where p and q are positive integers with no common factor. Suppose $T = (1/m)\sum_{j=1}^{m} P_j$ is an average of m projections; then $p/q = (1/m)\sum_{j=1}^{m} \operatorname{rank} P_j$, so m must be a positive integer multiple of q. Since we can choose T so that q is large, it follows that the number of projections in the average expression can be arbitrarily large. ■

Note added in proof. The equivalence of (2) and (3) of Theorem 3.5 has appeared in [P. A. Fillmore and J. P. Williams, Some convexity theorems for matrices, *Glasgow Math. J.* 12:110–116 (1971).]

REFERENCES

1 C. Davis, Separation of two linear subspaces, *Acta Sci. Math.* 19:172–187 (1958).
2 P. A. Fillmore, On sums of projections, *J. Funct. Anal.* 4:146–152 (1969).

3 C. K. Fong and G. J. Murphy, Averages of projections, *J. Operator Theory* 13:219–225 (1985).
4 R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge U.P., Cambridge, 1985.
5 R. V. Kadison and G. K. Pedersen, Means and convex combinations of unitary operators, *Math. Scand.* 57:249–266 (1985).
6 Y. Nakamura, Any Hermitian matrix is a linear combination of four projections, *Linear Algebra Appl.* 61:133–139 (1984).
7 K. Nishio, The structure of a real linear combination of two projections, *Linear Algebra Appl.* 66:169–176 (1985).

Received 1 May 1989; final manuscript accepted 19 May 1989