J. OPERATOR THEORY
46(2001), 435–447

CONSTRAINED UNITARY DILATIONS AND NUMERICAL RANGES

MAN-DUEN CHOI and CHI-KWONG LI

Communicated by William B. Arveson

Abstract. It is shown that each contraction A on a Hilbert space $\mathcal{H}$, with $A + A^* \leqslant \mu I$ for some $\mu \in \mathbb{R}$, has a unitary dilation U on $\mathcal{H} \oplus \mathcal{H}$ satisfying $U + U^* \leqslant \mu I$. This is used to settle a conjecture of Halmos in the affirmative: The closure of the numerical range of each contraction A is the intersection of the closures of the numerical ranges of all unitary dilations of A. By means of the duality theory of completely positive linear maps, some further results concerning numerical ranges inclusions and dilations are deduced.

Keywords: *Unitary dilation, numerical range, Hilbert space, completely positive linear map.*

MSC (2000): 15A60, 47A12, 47A30.

1. INTRODUCTION

Let $\mathcal{B}(\mathcal{H})$ be the algebra of all bounded linear operators on a Hilbert space $\mathcal{H}$. When $\mathcal{H}$ is of dimension $n < \infty$, $\mathcal{B}(\mathcal{H})$ can be identified with the algebra of $n \times n$ complex matrices, denoted by M_n. We say that $A \in \mathcal{B}(\mathcal{H}_1)$ has a *dilation* $B \in \mathcal{B}(\mathcal{H}_2)$ if $A = V^*BV$ for some isometry $V : \mathcal{H}_1 \to \mathcal{H}_2$; equivalently, B is unitarily similar to a 2×2 operator-matrix of the form $\begin{pmatrix} A & * \\ * & * \end{pmatrix}$. In [10], Halmos showed explicitly that each contraction $A \in \mathcal{B}(\mathcal{H})$ has a unitary dilation $U \in \mathcal{B}(\mathcal{H} \oplus \mathcal{H})$ of the form

$$U = \begin{pmatrix} A & \sqrt{1 - AA^*} \\ \sqrt{1 - A^*A} & -A^* \end{pmatrix}.$$

This result has generated a lot of research ([1], [11], [12], [18]), including the far reaching Sz.-Nagy dilation theorem ([18]): Each contraction $A \in \mathcal{B}(\mathcal{H})$ has a power unitary dilation; i.e., there is a unitary U satisfying

$$U^k = \begin{pmatrix} A^k & * \\ * & * \end{pmatrix}, \qquad k = 1, 2, \ldots .$$

In this paper, we are concerned about the structure of a contraction $A \in \mathcal{B}(\mathcal{H})$ subject to a constraint $A + A^* \leqslant \mu I$ for some $\mu \in \mathbb{R}$. We prove that such a contraction A always admits a unitary dilation $U \in \mathcal{B}(\mathcal{H} \oplus \mathcal{H})$ subject to the same constraint $U + U^* \leqslant \mu I$ (Theorem 2.1). Obviously, the case $A + A^* \leqslant \mu I$ with $\mu \geqslant 2$ is automatic while the case $A + A^* \leqslant \mu I$ with $\mu < -1$ is vacuous.

It turns out that the constrained unitary dilation is particularly useful in the study of numerical ranges of operators. Recall that the *numerical range* of an operator $A \in \mathcal{B}(\mathcal{H})$, defined by

$$W(A) = \{(Ax, x) : x \in \mathcal{H},\ (x, x) = 1\},$$

is a bounded convex set in $\mathbb{C}$. The numerical range has been studied extensively because of its connections and applications to many different areas (see e.g. [13], [14]).

With the aid of constrained unitary dilations, we prove Theorem 2.4 to affirm the conjecture of Halmos ([11]): For each contraction A on $\mathcal{H}$, the closure of the numerical range of A, denoted by $\overline{W(A)}$, satisfies

$$\overline{W(A)} = \bigcap\{\overline{W(U)} : U \in \mathcal{B}(\mathcal{H} \oplus \mathcal{H}) \text{ is a unitary dilation of } A\}.$$

Of course, the closure signs can be omitted in the finite-dimensional case. But in the infinite-dimensional case, if the closure signs are removed, then the equality may fail even when A is a normal contraction (see [9]).

Furthermore, we use the facility of constrained unitary dilations (the finite-dimensional case only) and the duality theory of completely positive linear maps to deduce Theorem 4.3: If $B \in M_3$ has a non-trivial reducing subspace, then each operator $A \in \mathcal{B}(\mathcal{H})$ with $W(A) \subseteq W(B)$ has a dilation that is unitarily similar to $B \otimes I = B \oplus B \oplus \cdots$. In particular, when $B \in M_3$ is normal or $B = B_1 \oplus [\gamma] \in M_3$ with $\gamma \in W(B_1)$, the statement is reduced to the following known results:

(I) ([15], see also [8], [16]) Let B be the diagonal matrix $\mathrm{Diag}\,(b_1, b_2, b_3) \in M_3$. Then $A \in \mathcal{B}(\mathcal{H})$ satisfies $W(A) \subseteq W(B)$ if and only if A has a dilation of the form $B \otimes I$ (equivalently, there exist positive operators P_1, P_2, P_3 with $P_1 + P_2 + P_3 = I_{\mathcal{H}}$ such that $A = b_1P_1 + b_2P_2 + b_3P_3$).

(II) ([8], see also [2] and Theorem 3.1.1 of [5]) Let $B_1 \in M_2$. Then $A \in \mathcal{B}(\mathcal{H})$ satisfies $W(A) \subseteq W(B_1)$ if and only if A has a dilation of the form $B_1 \otimes I$.

By the examples in [8], we see that Theorem 4.3 need not hold for general $B \in M_3$ or normal $B \in M_4$. Thus, the result is best possible in a certain sense.

This paper is organized as follows. In Section 2, we present the statement of the main theorem and its consequences, including an affirmative solution to the conjecture of Halmos. Section 3 is devoted to the proof of the main theorem. Section 4 concerns completely positive linear maps in connection with numerical range inclusions and dilation properties. In Section 5, we discuss some related problems and show that the main theorem does not admit further generalizations.

2. MAIN THEOREM AND CONSEQUENCES

The following is the main theorem.

THEOREM 2.1. *Let $A \in \mathcal{B}(\mathcal{H})$ be a contraction such that $A + A^* \leqslant \mu I$ for some $\mu \in \mathbb{R}$. Then A has a unitary dilation $U \in \mathcal{B}(\mathcal{H} \oplus \mathcal{H})$ satisfying $U + U^* \leqslant \mu I$. In the case when $\mathcal{H}$ is of dimension n, the matrix $U \in M_{2n}$ can be chosen such that its $2n$ eigenvalues are $\mathrm{e}^{\pm \mathrm{i}\theta_1}, \dots, \mathrm{e}^{\pm \mathrm{i}\theta_n}$ with $2\cos\theta_j \leqslant \mu$ for all j (i.e., non-real eigenvalues occur in conjugate pairs, real eigenvalues have even multiplicities).*

The proof of Theorem 2.1 will be given in the next section. Meanwhile, we mention some related facts and immediate consequences of the constrained unitary dilations.

First, in the finite-dimensional case, each unitary dilation $U \in M_{2n}$ of a contraction $A \in M_n$ must be of the form $U = \begin{pmatrix} A & C \\ B & D \end{pmatrix}$ with $B = U_1\sqrt{I - A^*A}$, $C = -\sqrt{I - AA^*}U_2$ and $D = U_1 A^* U_2$ for some unitary $U_1, U_2 \in M_n$. Thus U is unitarily similar to

$$(I \oplus U_1^*)U(I \oplus U_1) = \begin{pmatrix} A & -\sqrt{I - AA^*}U_0 \\ \sqrt{I - A^*A} & A^*U_0 \end{pmatrix}$$

with $U_0 = U_2U_1$. In order to get a "constrained" unitary dilation, we need a judicious choice of $U_0 \in M_n$. If A is normal, then we just choose $U_0 = I$. If A is non-normal, there is an algorithm (an account of non-commutative matrix manipulation) to construct U_0 as shown in next section.

Secondly, if $A \in M_n$ is a real matrix, our construction will also yield a real constrained orthogonal dilation $U \in M_{2n}$.

Moreover, we can use Theorem 2.1 to get a spectral decomposition for non-normal constrained contractions in terms of "a non-commutative resolution of the identity". Here we state only the finite-dimensional case:

COROLLARY 2.2. *Suppose $A \in M_n$ is a contraction satisfying $A + A^* \leqslant \mu I_n$ for some $\mu \in \mathbb{R}$. Then there are n real numbers $\theta_1, \dots, \theta_n \in [0, \pi]$ with $2\cos\theta_j \leqslant \mu$ for all j, and positive semidefinite matrices $Q_1, \dots, Q_{2n} \in M_n$ with* $\operatorname{rank} Q_j \leqslant 1$ *for all j, such that*

$$I_n = \sum_{j=1}^{n} (Q_j + Q_{n+j}) \qquad \text{and} \qquad A = \sum_{j=1}^{n} \left(\mathrm{e}^{\mathrm{i}\theta_j} Q_j + \mathrm{e}^{-\mathrm{i}\theta_j} Q_{n+j}\right).$$

Proof. Suppose $A \in M_n$ is a contraction. By Theorem 2.1, there exists a unitary dilation $U = \begin{pmatrix} A & * \\ * & * \end{pmatrix} \in M_{2n}$ with eigenvalues in conjugate pairs. By the spectral decomposition, there exist mutually orthogonal rank 1 projections $P_1, \dots, P_{2n} \in M_{2n}$ such that $\sum\limits_{j=1}^{2n} P_j = I_{2n}$ and $U = \sum\limits_{j=1}^{2n} \alpha_j P_j$, where $\alpha_1, \dots, \alpha_{2n}$ are eigenvalues of U. Write $P_j = \begin{pmatrix} Q_j & * \\ * & * \end{pmatrix}$ with $Q_j \in M_n$. Then $I_n = \sum\limits_{j=1}^{2n} Q_j$ and $A = \sum\limits_{j=1}^{2n} \alpha_j Q_j$ as desired. ∎

The following lemma is a simple fact about numerical ranges.

LEMMA 2.3. *Let* $A \in \mathcal{B}(\mathcal{H})$. *Then* $A + A^* \leqslant \mu I$ *if and only if* $\overline{W(A)}$ *is included in the closed half plane* $\{z \in \mathbb{C} : (z + \overline{z}) \leqslant \mu\}$.

Now, we are ready to settle the conjecture of Halmos.

THEOREM 2.4. *Let* $A \in \mathcal{B}(\mathcal{H})$ *be a contraction. Then*

$$\overline{W(A)} = \bigcap\{\overline{W(U)} : U \in \mathcal{B}(\mathcal{H} \oplus \mathcal{H}) \textit{ is a unitary dilation of } A\}.$$

(*In the finite-dimensional case, the closure signs on the numerical ranges can be omitted.*)

Proof. Let $A \in \mathcal{B}(\mathcal{H})$ be a contraction. It is obvious that $W(A) \subseteq W(B)$ if B is a dilation of A. Thus, we have

$$\overline{W(A)} \subseteq \bigcap\{\overline{W(U)} : U \text{ is a unitary dilation of } A\}.$$

To prove the reverse inclusion, we consider any particular $\zeta \notin \overline{W(A)}$. Since $\overline{W(A)}$ is a compact convex set, there exists $\theta \in [0, 2\pi)$ and $\mu \in \mathbb{R}$ such that $\mathrm{e}^{\mathrm{i}\theta}\zeta + \mathrm{e}^{-\mathrm{i}\theta}\overline{\zeta} > \mu$, while $\mathrm{e}^{\mathrm{i}\theta}\overline{W(A)} = \overline{W(\mathrm{e}^{\mathrm{i}\theta}A)}$ is included in the closed half plane $\{z \in \mathbb{C} : z + \overline{z} \leqslant \mu\}$. By Theorem 2.1, there is a unitary dilation U of A such that $\mathrm{e}^{\mathrm{i}\theta}U + \mathrm{e}^{-\mathrm{i}\theta}U^* \leqslant \mu I_{2n}$. By Lemma 2.3 again, $\overline{W(\mathrm{e}^{\mathrm{i}\theta}U)} \subseteq \{z \in \mathbb{C} : z + \overline{z} \leqslant \mu\}$. Hence $\mathrm{e}^{\mathrm{i}\theta}\zeta \notin \overline{W(\mathrm{e}^{\mathrm{i}\theta}U)}$ and $\zeta \notin \overline{W(U)}$. ∎

3. PROOF OF THE MAIN THEOREM

We begin with several lemmas.

LEMMA 3.1. *Let* $A \in \mathcal{B}(\mathcal{H})$ *be a contraction. Suppose* v *is a unit vector satisfying* $\|Av\| = 1$ *and* $(A + A^*)v = (2\cos\theta)v$ *for some* $\theta \in [0, \pi]$, *but* v *is not an eigenvector for* A. *Then* $\mathcal{V} = \operatorname{span}\{v, Av\}$ *is a* 2*-dimensional reducing subspace of* A, *and the restriction of* A *on* $\mathcal{V}$ *is a unitary operator with* $\mathrm{e}^{\pm\mathrm{i}\theta}$ *as the two eigenvalues.*

Proof. Write $Av = \lambda v + \sqrt{1 - |\lambda|^2}u$, where u is a unit vector orthogonal to v and $|\lambda| \neq 1$. With respect to the orthonormal basis $\{v, u\}$, the compression of A onto $\mathcal{V}$ is a 2×2 matrix $X = \begin{pmatrix} \lambda & b \\ \sqrt{1 - |\lambda|^2} & c \end{pmatrix}$ for some $b, c \in \mathbb{C}$. Since $(X + X^*)v = (2\cos\theta)v$, we see that $\lambda + \overline{\lambda} = 2\cos\theta$ and $b = -\sqrt{1 - |\lambda|^2}$. From $X^*X \leqslant I$, it follows that $c = \overline{\lambda}$. Hence, X is unitary, and X is a direct summand of the contraction A. Furthermore, from the matrix form of X, we see that the two eigenvalues of X are $\mathrm{e}^{\pm\mathrm{i}\theta}$. ∎

Lemma 3.2. *Let $H \in M_n$ be the leading principal submatrix of a Hermitian matrix $\widetilde{H} \in M_{n+1}$. Suppose there exists a unit vector $u \in \mathbb{C}^{n+1}$ with nonzero $(n+1)$st entry such that $\widetilde{H}u = \xi u$. If $H \leqslant \xi I_n$, then $\widetilde{H} \leqslant \xi I_{n+1}$.*

Proof. Suppose $H, \widetilde{H}$ and u satisfy the hypotheses of the lemma. Let $w \in \mathbb{C}^{n+1}$ be a unit vector orthogonal to u. Then there exist $t_1, t_2 \in \mathbb{C}$ such that $t_1 \neq 0$, $|t_1|^2 + |t_2|^2 = 1$, and $v = t_1 w + t_2 u \in \mathbb{C}^{n+1}$ is a unit vector whose $(n+1)$st entry is zero. Thus, $|t_1|^2(\widetilde{H}w, w) + |t_2|^2\xi = (\widetilde{H}v, v)$ is a quadratic form of H on a unit vector, and hence not larger than ξ. It follows that $(\widetilde{H}w, w) \leqslant \xi$ and $\widetilde{H} \leqslant \xi I_{n+1}$ as asserted. ∎

Lemma 3.3. *Let $A \in M_n$ be a contraction such that $2\cos\theta$ is the largest eigenvalue for $A + A^*$. Then A has a contractive dilation $\widetilde{A} \in M_{n+1}$ such that $\widetilde{A} + \widetilde{A}^* \leqslant (2\cos\theta)I_{n+1}$ and $\mathrm{e}^{\pm\mathrm{i}\theta}$ are two eigenvalues for $\widetilde{A}$.*

Proof. Suppose $\mathrm{e}^{\mathrm{i}\theta}$ or $\mathrm{e}^{-\mathrm{i}\theta}$ is an eigenvalue for A. Then evidently A has a desired dilation.

Next, suppose neither $\mathrm{e}^{\mathrm{i}\theta}$ nor $\mathrm{e}^{-\mathrm{i}\theta}$ is an eigenvalue for A. Let v be a unit vector such that $(A + A^*)v = (2\cos\theta)v$. By Lemma 3.1, we have $\|Av\| < 1$. Since

$$v^*(AA^* - A^*A)v = v^*\{A(A+A^*) - (A+A^*)A\}v = v^*A(2\cos\theta)v - (2\cos\theta)v^*Av = 0,$$

we have $\|A^*v\|^2 = \|Av\|^2$. Let $d = \sqrt{1 - \|Av\|^2} = \sqrt{1 - \|A^*v\|^2}$. Then $x = \sqrt{I_n - A^*A}v/d$ and $y = \sqrt{I_n - AA^*}v/d$ are unit vectors in $\mathbb{C}^n$. Write

$$X = \begin{pmatrix} I_n & 0 \\ \mathbf{0} & x \end{pmatrix}, \quad Y = \begin{pmatrix} I_n & 0 \\ \mathbf{0} & y \end{pmatrix}, \quad Z = \begin{pmatrix} A & -\sqrt{I_n - AA^*} \\ \sqrt{I_n - A^*A} & A^* \end{pmatrix},$$

and

$$\widetilde{A} = X^*ZY = \begin{pmatrix} A & -(I_n - AA^*)v/d \\ v^*(I_n - A^*A)/d & x^*A^*y \end{pmatrix} \in M_{n+1}.$$

Then $X^*X = I_{n+1} = Y^*Y$, $Z^*Z = I_{2n}$ and hence $\widetilde{A}$ is a contractive dilation of A. Write $\widetilde{v} = \begin{pmatrix} v \\ 0 \end{pmatrix} \in \mathbb{C}^{n+1}$. Then

$$\widetilde{A}\widetilde{v} = \begin{pmatrix} Av \\ v^*(I_n - A^*A)v/d \end{pmatrix} = \begin{pmatrix} Av \\ d \end{pmatrix}$$

is a unit vector because $d = \sqrt{1 - \|Av\|^2}$, and

$$(\widetilde{A} + \widetilde{A}^*)\widetilde{v} = \begin{pmatrix} (A + A^*)v \\ v^*(AA^* - A^*A)v/d \end{pmatrix} = \begin{pmatrix} (2\cos\theta)v \\ 0 \end{pmatrix} = (2\cos\theta)\widetilde{v}$$

because $\|Av\| = \|A^*v\|$. It follows from Lemma 3.1 that $\mathcal{V} = \mathrm{span}\,\{\widetilde{v}, \widetilde{A}\widetilde{v}\}$ is a reducing subspace of $\widetilde{A}$ and the restriction of $\widetilde{A}$ on $\mathcal{V}$ has $\mathrm{e}^{\pm\mathrm{i}\theta}$ as the two eigenvalues. So, $\widetilde{A}\widetilde{v} = \begin{pmatrix} Av \\ d \end{pmatrix}$ is also an eigenvector of $\widetilde{A} + \widetilde{A}^*$ corresponding to the eigenvalue $2\cos\theta$. Note that the last entry of $\widetilde{A}\widetilde{v}$ is $d \neq 0$. Applying Lemma 3.2 with $H = A + A^*$, $\widetilde{H} = \widetilde{A} + \widetilde{A}^*$ and $\xi = 2\cos\theta$, we conclude that $\widetilde{A} + \widetilde{A}^* \leqslant (2\cos\theta)I_{n+1}$. Thus, $\widetilde{A}$ is a desired dilation. ∎

Proof of Theorem 2.1 *for the finite-dimensional case.* Let $A \in M_n$ be a contraction with $A + A^* \leqslant \mu I_n$. We prove by induction that A has a unitary dilation $U \in M_{2n}$ such that $U + U^* \leqslant \mu I_{2n}$ and the eigenvalues of U occur in conjugate pairs.

If $n = 1$ and $A = [\lambda]$, then the matrix $\begin{pmatrix} \lambda & -\sqrt{1-|\lambda|^2} \\ \sqrt{1-|\lambda|^2} & \overline{\lambda} \end{pmatrix}$ is a desired dilation.

Now suppose $n > 1$, and the result is true for the cases up to $n-1$. Assume the largest eigenvalue of $A + A^*$ is $2\cos\theta \leqslant \mu$. By Lemma 3.3, A has a contractive dilation that is unitarily similar to $[\mathrm{e}^{\mathrm{i}\theta}] \oplus [\mathrm{e}^{-\mathrm{i}\theta}] \oplus T \in M_{n+1}$, where $T \in M_{n-1}$ is a contraction satisfying $T + T^* \leqslant (2\cos\theta)I_{n-1} \leqslant \mu I_{n-1}$. By the induction assumption, T has a unitary dilation $U_1 \in M_{2n-2}$ with all required properties. Therefore, A has a dilation that is unitarily similar to $[\mathrm{e}^{\mathrm{i}\theta}] \oplus [\mathrm{e}^{-\mathrm{i}\theta}] \oplus U_1 \in M_{2n}$ as desired. ∎

To prove Theorem 2.1 for the infinite-dimensional case, we need one proposition.

PROPOSITION 3.4. *Suppose A and B in $\mathcal{B}(\mathcal{H})$ satisfy $A^*A + B^*B \leqslant I_{\mathcal{H}}$ and $A + A^* \leqslant \mu I_{\mathcal{H}}$ for some $\mu \in \mathbb{R}$. Then there exist C and D in $\mathcal{B}(\mathcal{H})$ so that*

$$Z = \begin{pmatrix} A & C \\ B & D \end{pmatrix} \in \mathcal{B}(\mathcal{H} \oplus \mathcal{H})$$

is a contraction with $Z + Z^* \leqslant \mu I_{\mathcal{H}\oplus\mathcal{H}}$.

Proof. First, we consider the case when $\dim \mathcal{H} = n$. By the finite-dimensional case of Theorem 2.1, there exists a unitary dilation $U \in M_{2n}$ of A in the form

$$U = \begin{pmatrix} A & * \\ \sqrt{I - A^*A} & * \end{pmatrix}$$

with $U + U^* \leqslant \mu I_{2n}$. Since $A^*A + B^*B \leqslant I_n$, we see that $B = J\sqrt{I - A^*A}$ for some contraction $J \in M_n$. Let

$$V = \begin{pmatrix} I_n & \mathbf{0}_n \\ \mathbf{0}_n & J^* \\ \mathbf{0}_n & -\sqrt{I_n - JJ^*} \end{pmatrix} \quad \text{and} \quad \widetilde{U} = U \oplus (-I_n).$$

Then $V^*V = I_{2n}$ and $Z = V^*\widetilde{U}V$ is a desired matrix.

Next, suppose $A, B \in \mathcal{B}(\mathcal{H})$, where $\mathcal{H}$ is of infinite dimension, such that $A^*A + B^*B \leqslant I_{\mathcal{H}}$ and $A + A^* \leqslant \mu I_{\mathcal{H}}$. Without loss of generality, we may assume that $A = (a_{pq})$ and $B = (b_{pq})$ are infinite matrices with respect to a countable orthonormal basis so that indices p and q run through all positive integers. For each positive integer n, let $A_n = (a_{pq})_{1\leqslant p,q\leqslant n}$ and $B_n = (b_{pq})_{1\leqslant p,q\leqslant n}$ be the finite sections of A and B. Then $A_n^*A_n + B_n^*B_n \leqslant I_n$ and $A_n + A_n^* \leqslant \mu I_n$. By the finite-dimensional result, there exists a contraction

$$T_n = \begin{pmatrix} A_n & C_n \\ B_n & D_n \end{pmatrix} \in M_{2n}$$

with $T_n + T_n^* \leqslant \mu I_{2n}$. Let

$$\widetilde{T}_n = \begin{pmatrix} A_n & \mathbf{0} & \vdots & C_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \vdots & \mathbf{0} & \mathbf{0} \\ \cdots & \cdots & & \cdots & \cdots \\ B_n & \mathbf{0} & \vdots & D_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \vdots & \mathbf{0} & \mathbf{0} \end{pmatrix} \in \mathcal{B}(\mathcal{H} \oplus \mathcal{H}).$$

Then $\widetilde{T}_n$ is a contraction with $\widetilde{T}_n + \widetilde{T}_n^* \leqslant \mu I_{\mathcal{H}\oplus\mathcal{H}}$. Taking the weak-operator limit of a convergent subsequence of the bounded sequence $\{\widetilde{T}_n\}$, we get a contraction

$$Z = \begin{pmatrix} A & C \\ B & D \end{pmatrix} \in \mathcal{B}(\mathcal{H} \oplus \mathcal{H})$$

satisfying $Z + Z^* \leqslant \mu I_{\mathcal{H}\oplus\mathcal{H}}$. ∎

Proof of Theorem 2.1 *for the infinite-dimensional case.* Suppose A is a contraction acting on an infinite-dimensional Hilbert space $\mathcal{H}$ such that $A + A^* \leqslant \mu I$. Applying Proposition 3.4 with $B = \sqrt{I - A^*A}$, we see that there is a contraction

$$Z_1 = \begin{pmatrix} A & C \\ \sqrt{I - A^*A} & D \end{pmatrix} \in \mathcal{B}(\mathcal{H} \oplus \mathcal{H})$$

satisfying $Z_1 + Z_1^* \leqslant \mu I$ and $\|Z_1 v\| = \|v\|$ for all vectors $v \in \mathcal{H} \oplus \mathcal{O}$, where $\mathcal{O}$ is the zero subspace in $\mathcal{H}$.

Repeating the argument on Z_1, we get a contractive dilation of Z_1 of the form

$$Z_2 = \begin{pmatrix} Z_1 & \widetilde{C} \\ \sqrt{I - Z_1^* Z_1} & \widetilde{D} \end{pmatrix} \in \mathcal{B}(\mathcal{H} \oplus \mathcal{H} \oplus \mathcal{H} \oplus \mathcal{H})$$

such that $Z_2 + Z_2^* \leqslant \mu I$ and $\|Z_2 v\| = \|v\|$ for all vectors $v \in \mathcal{H} \oplus \mathcal{H} \oplus \mathcal{O} \oplus \mathcal{O}$. Continuing this process for further dilations, we obtain eventually a contraction Z_∞, denoted by U, acting on $\mathcal{H} \oplus \mathcal{H} \oplus \cdots$ so that $U + U^* \leqslant \mu I$ and $\|Uv\| = 1$ for all unit vectors v. Identifying $\mathcal{O} \oplus \mathcal{H}$ with $\mathcal{O} \oplus \mathcal{H} \oplus \mathcal{H} \oplus \cdots$, we may regard U as an isometry acting on $\mathcal{H} \oplus \mathcal{H}$ while A still acts on $\mathcal{H} \oplus \mathcal{O}$.

Next, we show that U is invertible. To this end, we presume $\mu < 2$ (otherwise, the constraint is automatic); hence $U + U^* \leqslant \mu I < 2I$ implies that $1 \notin \overline{W(U)}$. Since $\sigma(U) \subseteq \overline{W(U)}$, we deduce that $I - U$ is invertible. By a similar argument, or by the fact that $I - U^* = (I - U)^*$, we see that $I - U^*$ is also invertible. Thus, $U - I = U - U^*U = (I - U^*)U$ implies that $U = (I - U^*)^{-1}(U - I)$ is invertible. Therefore, U is unitary and U is a desired dilation. ∎

4. NUMERICAL RANGE INCLUSION AND DILATION

In this section, we establish relations between the conditions (i) $W(A) \subseteq W(B)$, and (ii) A has a dilation of the form $B \otimes I = B \oplus B \oplus \cdots$. The tool is the facility of constrained unitary dilations in conjunction with the theory of completely positive linear maps (see e.g. [4], [5], [7], [17] for general background). Let $\mathcal{S}$ be a subspace of M_n satisfying $I_n \in \mathcal{S}$, and $A^* \in \mathcal{S}$ whenever $A \in \mathcal{S}$. A linear map $\varphi : \mathcal{S} \to \mathcal{B}(\mathcal{H})$ is said to be *positive* if it maps positive elements to positive elements; φ is said to be *completely positive* if $\varphi_k : M_k(\mathcal{S}) \to M_k(\mathcal{B}(\mathcal{H}))$, defined by

$$(A_{ij})_{1 \leqslant i,j \leqslant k} \mapsto (\varphi(A_{ij}))_{1 \leqslant i,j \leqslant k},$$

is positive for $k = 1, 2, \ldots$. We begin with a lemma of some standard results (cf. [3], Section 1.4, [4] and [5]).

LEMMA 4.1. *Let $A \in \mathcal{B}(\mathcal{H})$ and $B \in M_n$. Define a linear map*

$$\varphi : \operatorname{span}\{I, B, B^*\} \to \operatorname{span}\{I, A, A^*\}$$

by

$$\varphi(aI_3 + bB + cB^*) = aI_{\mathcal{H}} + bA + cA^*, \qquad \textit{for any } a, b, c \in \mathbb{C}.$$

We have (i) $\Leftrightarrow$ (ii) $\Leftarrow$ (iii) $\Leftrightarrow$ (iv) *for the following conditions.*

(i) *φ is a positive linear map.*

(ii) *The inclusion relation $W(A) \subseteq W(B)$ holds.*

(iii) *The operator A has a dilation that is unitarily similar to $B \otimes I$.*

(A finite number of copies of B is enough for the dilation if $\mathcal{H}$ is of finite dimension.)

(iv) *φ is a completely positive linear map.*

Proof. The implication (iv) $\Rightarrow$ (i) and the implication (iii) $\Rightarrow$ (ii) are clear.

(i) $\Leftrightarrow$ (ii): Note that $W(A) \subseteq W(B)$ if and only if every closed half plane including $W(B)$ also includes $W(A)$. Since $W(T)$ is included in the closed half plane

$$\{z \in \mathbb{C} : a + bz + c\overline{z} \geqslant 0\}$$

if and only if T satisfies $aI + bT + cT^* \geqslant \mathbf{0}$, the result follows.

(iv) $\Leftrightarrow$ (iii): Suppose (iv) holds. By Arveson's extension theorem (Theorem 1.2.3 of [4]), φ can be extended to a completely positive linear map Φ from M_n to $\mathcal{B}(\mathcal{H})$. Then there is an isometry V such that $\Phi(X) = V^*(X \otimes I)V$ for all $X \in M_n$; thus $A = \varphi(B) = \Phi(B) = V^*(B \otimes I)V$ and Condition (iii) holds. As shown in [7], if $\mathcal{H}$ is of finite dimension, the number of copies of B needed is finite.

Conversely, suppose (iii) holds, i.e., $A = V^*(B \otimes I)V$ for some isometry V. Then φ is just the completely positive linear map $X \mapsto V^*(X \otimes I)V$. ∎

Usually, a general positive linear map is "far away" from being completely positive (see [6]). Nevertheless, we show below that there is a special instance for a whole class of positive maps to be completely positive. Conceivably, the hypothesis of Proposition 4.2 may be the only situation in the non-commutative case that the two classes of linear maps coincide.

PROPOSITION 4.2. *Suppose $B \in M_3$ has a reducing subspace. Then each unital positive linear map from* span$\{I, B, B^*\}$ *to* $\mathcal{B}(\mathcal{H})$ *is completely positive.*

Proof. Let φ be a unital positive linear map from $\mathcal{S} = \operatorname{span}\{I, B, B^*\}$ to $\mathcal{B}(\mathcal{H})$. Suppose B is normal; we may write B in a diagonal form. If $\mathcal{S}$ is of dimension one, i.e., B is scalar, then the result is trivial. If $\mathcal{S}$ is of dimension two, then $\mathcal{S}$ is spanned by I and a hermitian matrix $H = \operatorname{Diag}(0, r, 1)$ with $0 \leqslant r \leqslant 1$. Let $\{e_1, e_2, e_3\}$ be the standard basis for $\mathbb{C}^3$. Write

$$V = e_1 \otimes \sqrt{I_{\mathcal{H}} - \varphi(H)} + e_3 \otimes \sqrt{\varphi(H)}.$$

Then V is an isometry and $\varphi(X) = V^*(X \otimes I_{\mathcal{H}})V$ is complete positive. If $\mathcal{S}$ is of dimension three, then $\mathcal{S}$ is just the set of all diagonal matrices in M_3. Write

$$V = e_1 \otimes \sqrt{\varphi(\operatorname{Diag}(1,0,0))} + e_2 \otimes \sqrt{\varphi(\operatorname{Diag}(0,1,0))} + e_3 \otimes \sqrt{\varphi(\operatorname{Diag}(0,0,1))}.$$

Then V is an isometry and $\varphi(X) = V^*(X \otimes I_{\mathcal{H}})V$ is completely positive.

Suppose B is not normal; we may write $B = \begin{pmatrix} a & b \\ c & a \end{pmatrix} \oplus [d]$ with $|b| \neq |c|$. Then

$$\{\bar{b}(B - aI) - c(B - aI)^*\}/(|b|^2 - |c|^2) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \oplus [r\mathrm{e}^{\mathrm{i}\theta}]$$

for some $r \geqslant 0$. Multiplying the last matrix by $\mathrm{e}^{-\mathrm{i}\theta}$ and using the fact that $\begin{pmatrix} 0 & \mathrm{e}^{-\mathrm{i}\theta} \\ 0 & 0 \end{pmatrix}$ is unitarily similar to $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, we may assume

$$B_0 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \oplus [r] \in \mathcal{S}.$$

Hence $\mathcal{S} = \operatorname{span}\{I, B_0, B_0^*\}$ because the latter is evidently of dimension 3. For simplicity, we assume that $B_0 = B$ from now on. To prove that φ is completely positive, we need to show that for any $R, S, T \in M_n$ satisfying

$$R \otimes I_3 + (S \otimes B + T \otimes B^*) \geqslant \mathbf{0}_{3n}, \tag{4.1}$$

we have

$$R \otimes I_{\mathcal{H}} + (S \otimes \varphi(B) + T \otimes \varphi(B)^*) \geqslant \mathbf{0}_n \otimes I_{\mathcal{H}} \tag{4.2}$$

(cf. Section 1.4 of [3]). It is clear that (4.1) is just the same as

$$\begin{pmatrix} R & S \\ T & R \end{pmatrix} \geqslant \mathbf{0}_{2n} \quad \text{and} \quad R + r(S + T) \geqslant \mathbf{0}_n. \tag{4.3}$$

Thus $T = S^*$ and $R \geqslant \mathbf{0}$; we may further assume that $R > \mathbf{0}$. (Otherwise, replace R by $R + \delta I_n$ for some $\delta > 0$ to get the result, and then take the limit for $\delta \to 0$.) Write $J = R^{-1/2} S R^{-1/2}$. Then (4.3) is equivalent to

$$\begin{pmatrix} I_n & J \\ J^* & I_n \end{pmatrix} \geqslant \mathbf{0}_{2n} \quad \text{and} \quad I_n + r(J + J^*) \geqslant \mathbf{0}_n;$$

i.e., J is a contraction with $-(J + J^*) \leqslant I_n/r$. (If $r = 0$, then all constraints concerning r are automatically valid.) By Theorem 2.1, J has a unitary dilation

$U \in M_{2n}$ with $-(U+U^*) \leqslant I_{2n}/r$. Hence for each eigenvalue of $\mathrm{e}^{\mathrm{i}\theta}$ of U, we have $-(\mathrm{e}^{\mathrm{i}\theta}+\mathrm{e}^{-\mathrm{i}\theta}) \leqslant 1/r$; i.e., $1+2r\cos\theta \geqslant 0$. Thus

$$I_3 + \left(\mathrm{e}^{\mathrm{i}\theta}B + \mathrm{e}^{-\mathrm{i}\theta}B^*\right) = \begin{pmatrix} 1 & \mathrm{e}^{\mathrm{i}\theta} \\ \mathrm{e}^{-\mathrm{i}\theta} & 1 \end{pmatrix} \oplus [1+2r\cos\theta] \geqslant \mathbf{0}_3.$$

Since φ is positive, it follows that

$$I_{\mathcal{H}} + (\mathrm{e}^{\mathrm{i}\theta}\varphi(B) + \mathrm{e}^{-\mathrm{i}\theta}\varphi(B)^*) \geqslant \mathbf{0}_{\mathcal{H}};$$

so

$$I_{2n} \otimes I_{\mathcal{H}} + (U \otimes \varphi(B) + U^* \otimes \varphi(B)^*) \geqslant \mathbf{0}_{2n} \otimes \mathbf{0}_{\mathcal{H}};$$

thus, for the compression on $\mathbb{C}^n \otimes \mathcal{H}$,

$$I_n \otimes I_{\mathcal{H}} + (J \otimes \varphi(B) + J^* \otimes \varphi(B)^*) \geqslant \mathbf{0}_n \otimes I_{\mathcal{H}}.$$

Finally, (4.2) follows form the fact $S = R^{1/2}JR^{1/2}$ and $T = S^*$. ∎

We are now ready to state the major result of this section.

THEOREM 4.3. *Suppose $B \in M_3$ has a non-trivial reducing subspace, and let $A \in \mathcal{B}(\mathcal{H})$. Then $W(A) \subseteq W(B)$ if and only if A has a dilation that is unitarily similar to $B \otimes I$.*

Proof. By Lemma 4.1, we need only to show that Condition (i) implies Condition (iv). This follows immediately from Proposition 4.2. ∎

Note that if $B \in M_3$ is normal or if $B = B_1 \oplus [\gamma] \in M_3$ with $\gamma \in W(B_1)$, then Theorem 4.3 reduces to known results mentioned in the Introduction. Here we are dealing with the more difficult case where $B = B_1 \oplus [\gamma] \in M_3$ such that B_1 need not be normal; so $W(B)$ is the convex hull of an elliptical region and one point. Consequently, Theorem 4.3 can be restated as follows.

Suppose $K \subseteq \mathbb{C}$ is the the convex hull of an ellipse and a point. Then there exists $B = B_1 \oplus [\gamma] \in M_3$ such that $W(B) = K$. Moreover, we have that $A \in \mathcal{B}(\mathcal{H})$ satisfies $W(A) \subseteq K$ if and only if A has a dilation of the form $B \otimes I$.

As mentioned in [5], Section 2.5, and [8], it is pertinent to study other geometrically significant regions K in $\mathbb{C}$ (or appropriate model matrices B) so that each operator $A \in \mathcal{B}(\mathcal{H})$ with $W(A) \subseteq K$ (or $W(A) \subseteq W(B)$) has a dilation of the form $B \otimes I$.

5. REMARKS AND EXAMPLES

In this section, we discuss some related problems and show that Theorem 2.1 does not admit further generalizations.

Notably, the fact "each contraction $A \in M_n$ has a unitary dilation $U \in M_{2n}$" can be deduced directly from the following simple and yet stronger result.

Proposition 5.1. *Let $A \in M_n$ be a contraction. Then $A = (U_1 + U_2)/2$ for some unitary matrices $U_1, U_2 \in M_n$; consequently, $A = V^*(U_1 \oplus U_2)V$ where* $V = \frac{1}{\sqrt{2}}\begin{pmatrix} I_n \\ I_n \end{pmatrix}$.

Proof. By the polar decomposition, $A = UP$ where U is unitary and P is positive semidefinite. Since A is a contraction, so is P. Let $U_1 = U(P + \mathrm{i}\sqrt{I - P^2})$ and $U_2 = U(P - \mathrm{i}\sqrt{I - P^2})$. Then $A = (U_1 + U_2)/2$, and the proposition follows. ∎

It may be plausible to conjecture that each contraction $A \in M_n$, with $A + A^* \leqslant \mu I_n$ for some $\mu \in \mathbb{R}$, can be written as a convex combination of unitary $U \in M_n$ satisfying $U + U^* \leqslant \mu I$. This is not true as shown in the following.

Example 5.2. *Consider $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then $A + A^* \leqslant I$, but A is not a convex combination of 2×2 unitary matrices U_j with $U_j + U_j^* \leqslant I$.*

Proof. Suppose A is a convex combination of unitary matrices $U_1, \ldots, U_k \in M_2$ with $U_j + U_j^* \leqslant I$ for each j. Then the $(1,2)$ entry of U_j must be one, and thus $U_j = \begin{pmatrix} 0 & 1 \\ \mu_j & 0 \end{pmatrix}$ with $|\mu_j| = 1$. The condition $U_j + U_j^* \leqslant I$ forces $\mu_j + \overline{\mu}_j \leqslant -1$, in contradiction to the fact that the $(2,1)$ entry of A is a convex combination of μ_j's. ∎

As mentioned in the Introduction, each contraction $A \in \mathcal{B}(\mathcal{H})$ has a unitary power dilation U. This leads to important consequences such as the von Neumann inequality. However, the analogous statement for the constrained unitary power dilation is not valid as shown in the following simple example.

Example 5.3. *Let A be the zero matrix in M_n. Then obviously $A + A^* \leqslant \mathbf{0}$ and A is contractive. However, there is no unitary dilation U of A such that $U + U^* \leqslant \mathbf{0}$ and*

$$U^k = \begin{pmatrix} A^k & * \\ * & * \end{pmatrix}, \qquad k = 1 \text{ and } 2.$$

Proof. Suppose $A = \mathbf{0}$ has a unitary dilation $U = \begin{pmatrix} \mathbf{0} & C \\ B & D \end{pmatrix}$ with

$$U + U^* = \begin{pmatrix} \mathbf{0} & B^* + C \\ B + C^* & D + D^* \end{pmatrix} \leqslant \mathbf{0}.$$

It follows that $C = -B^* \neq \mathbf{0}$ and thus $U^2 = \begin{pmatrix} -B^*B & * \\ * & * \end{pmatrix}$ cannot have the zero matrix at the upper left corner. ∎

Theorem 2.1 shows that if $A \in M_n$ is a contraction subject to a *single* affine constraint, then A has a unitary dilation subject to the same constraint. It is natural to ask whether more constraints can be added to the statement above. The following example shows that in general it is not possible to find a normal dilation subject to *two* constraints.

EXAMPLE 5.4. *Let* $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. *Then* A *is a contraction satisfying* $-I \leqslant A + A^* \leqslant I$. *However,* A *has no normal dilation* N *satisfying* $-I \leqslant N + N^* \leqslant I$.

Proof. Suppose $N = \begin{pmatrix} A & C \\ B & D \end{pmatrix}$ is a normal dilation of A such that $-I \leqslant N + N^* \leqslant I$. Since the leading 2×2 submatrix of $N + N^*$ is unitary, we deduce that $N + N^* = (A + A^*) \oplus H$. Thus, $B = -C^* \neq \mathbf{0}$. It follows that $NN^* \neq N^*N$ by comparing the leading 2×2 submatrices on both sides. ∎

Acknowledgements. Research of the first author was supported by an NSERC grant.

Research of the second author was supported by an NSF grant. This paper was done while he was visiting the University of Toronto in the academic year 1998–1999 supported by a Faculty Research Grant of the College of William and Mary. He would like to thank the staff of the University of Toronto for their warm hospitality.

REFERENCES

1. T. ANDO, On a pair of commuting contractions, *Acta Sci. Math. (Szeged)* **24**(1963), 88–90.
2. T. ANDO, Structure of operators with numerical radius one, *Acta Sci. Math. (Szeged)* **34**(1973), 11–15.
3. T. ANDO, *Operator-Theoretic Methods for Matrix Inequalities*, Research monograph, Hokusei Gakuen Univ., 1998.
4. W.B. ARVESON, Subalgebras of C^*-algebras, *Acta Math.* **123**(1969), 141–224.
5. W.B. ARVESON, Subalgebras of C^*-algebras. II, *Acta Math.* **128**(1972), 271–308.
6. M.D. CHOI, Positive linear maps on C^*-algebras, *Canad. J. Math.* **24**(1972), 520–529.
7. M.D. CHOI, Completely positive linear maps on complex matrices, *Linear Algebra Appl.* **10**(1975), 285–290.
8. M.D. CHOI, C.K. LI, Numerical ranges and dilations, *Linear and Multilinear Algebra*, to appear.
9. E. DURSZT, On the numerical range of normal operators, *Acta Sci. Math. (Szeged)* **25**(1964), 262–265.
10. P.R. HALMOS, Normal dilations and extension of operators, *Summa Brasil Math.* **2**(1950), 125–134.
11. P.R. HALMOS, Numerical ranges and normal dilations, *Acta Sci. Math. (Szeged)* **25**(1964), 1–5.
12. P.R. HALMOS, Ten problems in Hilbert space, *Bull. Amer. Math. Soc.* **25**(1970), 887–933.
13. P.R. HALMOS, *A Hilbert Space Problem Book,* 2nd ed., Springer-Verlag, New York 1982.

14. R.A. HORN, C.R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, New York 1991.
15. B.A. MIRMAN, Numerical range and norm of a linear operator, *Voronezh Gos. Univ. Trudy Sem. Funktsional Anal.* **10**(1968), 51–55.
16. Y. NAKAMURA, Numerical range and norm, *Math. Japon.* **27**(1982), 149–150.
17. V.I. PAULSEN, *Completely Bounded Maps and Dilations*, Longman, Harlow 1986.
18. B. SZ.-NAGY, Unitary dilations of Hilbert space operators and related topics, *CBMS Regional Conf. Ser. in Math.*, vol. 19, Amer. Math. Soc., Providence 1971.

MAN-DUEN CHOI
Department of Mathematics
University of Toronto
Toronto
M5S 3G3 Ontario
CANADA
E-mail: choi@math.toronto.edu

CHI-KWONG LI
Department of Mathematics
The College of William and Mary
P.O. Box 8795
Williamsburg, Virginia 23187–8795
USA
E-mail: ckli@math.wm.edu

Received April 19, 1999.