

ON THE DISTANCE BETWEEN UNITARY ORBITS OF WEIGHTED SHIFTS

LAURENT MARCOUX

ABSTRACT. In this paper, we consider invertible bilateral weighted shift operators acting on a complex separable Hilbert space \mathcal{H} . They have the property that there exist a constant $\tau > 0$ and an orthonormal basis $\{e_i\}_{i \in \mathbb{Z}}$ for \mathcal{H} with respect to which a shift V acts by $We_i = w_i e_{i+1}$, $i \in \mathbb{Z}$ and $|w_i| \geq \tau$. The equivalence class $\mathcal{U}(W) = \{U^* W U : U \in \mathcal{B}(\mathcal{H}), U \text{ unitary}\}$ of weighted shifts with weight sequence (with respect to the basis $\{U^* e_i\}_{i \in \mathbb{Z}}$ for \mathcal{H}) identical to that of W forms the *unitary orbit* of W .

Given two shifts W and V , one can define a distance $d(\mathcal{U}(V), \mathcal{U}(W)) = \inf\{\|X - Y\| : X \in \mathcal{U}(V), Y \in \mathcal{U}(W)\}$ between the unitary orbits of W and V . We establish numerical estimates for upper and lower bounds on this distance which depend upon information drawn from finite dimensional restrictions of these operators.

1. INTRODUCTION

1.0. The notion of Hilbert space isomorphism coincides with that of unitary equivalence. Therefore, two bounded operators A and B can be thought of as representing the same operator if they are unitarily equivalent; that is, if $A = U^* B U$ for some unitary operator U . The equivalence classes $\mathcal{U}(A)$ determined by this relation are called unitary orbits, and have been studied by a number of people.

A characterization of the norm closure $\overline{\mathcal{U}(A)}$ of the unitary orbit of a general operator A in terms of approximate unitary equivalence has been given by Hadwin [Had], using Voiculescu's Theorem [Vcu], (cf. also [Arv]). Complete approximate unitary invariants for particular classes of operators also exist. For example, the work of Gellar and Page [GP] classifies $\mathcal{U}(\overline{N})$ for N normal, while the celebrated Theorem of Brown, Douglas, and Fillmore [BDF] does the same for essentially normal operators (i.e., operators whose image in the Calkin algebra is normal). In [O'Dn], Donal O'Donovan describes the closure of $\mathcal{U}(W)$ for W a weighted shift, and moreover classifies those shifts V which lie in $\mathcal{U}(W)$.

As a natural extension of this work, one might ask the question: if two operators A and B are not unitarily equivalent, then how different are they? More precisely, how far apart are their unitary orbits?

Received by the editors April 1, 1988 and, in revised form, July 17, 1989.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 47AXX.

©1991 American Mathematical Society
0002-9947/91 \$1.00 + \$.25 per page

Results for the general case are due to Davidson [Dav1], who reduces the problem to that of approximating finite pieces of A by finite pieces of B . More precise estimates for the class of normal operators have also been obtained by Davidson [Dav2] using, in particular, the results of Bhatia, Davis, and McIntosh [BDM]. The paper of Azoff and Davis [AD] provides estimates in the case where A and B are selfadjoint.

It is the purpose of this paper to obtain good estimates for upper and lower bounds on the distance between $\mathcal{U}(V)$ and $\mathcal{U}(W)$ in the case where W and V are weighted shifts whose weights are bounded below in modulus by some positive number and bounded above in modulus by 1. Our main result for upper bounds uses a technique of I. D. Berg's ([Brg2], cf. also [Her1, Her2]) with which we can glue together finite dimensional pieces of one shift to approximate another shift. In fact, we obtain estimates of how well we can glue together these finite dimensional restrictions as a function of their dimension. We can use these to obtain estimates on how far W is from being a direct summand of V and vice versa [§2], and in turn this information leads to an estimate of an upper bound on $d(\mathcal{U}(V), \mathcal{U}(W))$ (see also §5). §3 is devoted to obtaining lower bounds for $d(\mathcal{U}(V), \mathcal{U}(W))$ which also depend on finite dimensional restrictions. §4 provides three examples which illustrate various possible phenomena which can occur in the approximation of the finite dimensional pieces, while §5 considers the case of “nonessential” spectrum and mentions results for unilateral shifts.

1.1. Notation. Unless specified otherwise, \mathcal{H} is assumed to be a complex infinite dimensional separable Hilbert space with orthonormal basis $\{e_i\}_{i \in \mathbb{Z}}$. All operators on \mathcal{H} are bounded. W (resp. V) is a weighted shift operator with weights $\{w_i\}_{i \in \mathbb{Z}}$ (resp. $\{v_i\}_{i \in \mathbb{Z}}$) such that $We_i = w_i e_{i+1}$ (resp. $Ve_i = v_i e_{i+1}$), $i \in \mathbb{Z}$. Since we are only concerned with unitary orbits, we may assume all weights are positive (cf. [Shd], for instance). The case of interest is when both A and B have norm less than or equal to 1, since otherwise, the differences in the norm lead to potentially trivial estimates. Henceforth, we assume this. The group of unitary operators in $\mathcal{B}(\mathcal{H})$ is denoted by $\mathcal{U}(\mathcal{H})$, and the unitary orbit of an operator A in $\mathcal{B}(\mathcal{H})$ is $\mathcal{U}(A) = \{U^*AU : U \in \mathcal{U}(\mathcal{H})\}$. The norm closure of $\mathcal{U}(A)$ is denoted $\overline{\mathcal{U}(A)}$. The distance $d(\mathcal{U}(A), \mathcal{U}(B))$ between the unitary orbits of A and B is given by

$$\begin{aligned} d(\mathcal{U}(A), \mathcal{U}(B)) &= \inf_{X \in \mathcal{U}(A), Y \in \mathcal{U}(B)} \|X - Y\| \\ &= \inf_{Y \in \mathcal{U}(B)} \|A - Y\| = d(A, \mathcal{U}(B)), \end{aligned}$$

and is clearly the same as $d(\overline{\mathcal{U}(A)}, \overline{\mathcal{U}(B)})$. A and B are approximately unitarily equivalent (an equivalence relation) if there exist U_n in $\mathcal{U}(\mathcal{H})$, $n \geq 1$, such that $U_n^*AU_n - B = K_n$ is compact and tends to zero in norm as n approaches infinity. This definition is equivalent to saying that $\overline{\mathcal{U}(A)} = \overline{\mathcal{U}(B)}$ (cf. [Vcu]). The notation \cong denotes unitary equivalence, \cong_a denotes approximate

unitary equivalence, and for $\varepsilon > 0$, $A \cong_\varepsilon B$ means that $d(\mathcal{U}(A), \mathcal{U}(B)) \leq \varepsilon$.

When dealing with \mathbb{C}^n or \mathbb{R}^n as a metric space, it will be understood that the underlying metric is the one induced by the supremum norm (l_∞ -norm).

2. BERG'S TECHNIQUE. AN UPPER BOUND

2.1.1. The method of obtaining upper bounds for our estimate of

$$d(\mathcal{U}(V), \mathcal{U}(W))$$

depends heavily upon a very useful technique due to I. D. Berg [Brg2] and modified slightly by Domingo Herrero [Her1]. The version we give below is a minor modification of Davidson, [Dav4, Theorem 1].

2.1.2. Lemma (Berg's Technique). *Let T be an operator on a Hilbert space \mathcal{H} , and let m be a positive integer. Suppose that $\{e_j, f_j\}_{j=0}^m$ is an orthonormal family of vectors in \mathcal{H} and that $\{\alpha_j\}_{j=0}^{m-1}$ is a finite set of complex numbers such that $Te_j = \alpha_j e_{j+1}$ and $Tf_j = \alpha_j f_{j+1}$ for $0 \leq j \leq m-1$. Then there exists an operator S on \mathcal{H} and an orthonormal family $\{g_j, h_j\}_{j=0}^m$ of \mathcal{H} satisfying the following properties:*

- (1) $g_0 = e_0$, $g_m = f_m$, $h_0 = f_0$ and $h_m = -e_m$;
- (2) $Sg_j = \alpha_j g_{j+1}$ and $Sh_j = \alpha_j h_{j+1}$ for $0 \leq j \leq m-1$;
- (3) $Sx = Tx$ if x is orthogonal to $\{e_j, f_j\}_{j=0}^{m-1}$; and
- (4) $\|S - T\| \leq (\pi/2m)(\max_{1 \leq j \leq m-1} |\alpha_j|)$.

Proof. Cf. [Dav4, Theorem 1]. \square

2.1.3. Definition. We say that X is a *weighted cycle* with weights $\{x_1, \dots, x_m\}$ ($m > 1$) and write $X \cong \text{cycle}\{x_1, \dots, x_m\}$ if there exists an m -dimensional complex Hilbert space \mathcal{H}_m with orthonormal basis $\{f_1, \dots, f_m\}$ for which $Xf_i = x_i f_{i+1}$, $1 \leq i \leq m-1$ and $Xf_m = x_m f_1$.

2.1.4. If W is a weighted shift with weights $\{w_i\}_{i \in \mathbb{Z}}$ with respect to an orthonormal basis $\{e_i\}_{i \in \mathbb{Z}}$ of \mathcal{H} , then from the above Lemma we obtain the following:

Proposition. *Assume m is a positive integer, k_0 and k_1 are integers, $|k_0 - k_1| > m$, and let $N = \{k_0 \leq i \leq k_0 + m\}$ and $M = \{k_1 \leq i \leq k_1 + m\}$ be disjoint subsets of \mathbb{Z} . Suppose furthermore that $w_{k_0+j} = w_{k_1+j}$ for $0 \leq j \leq m-1$. Then there exists a weighted shift V with weights $\{w_i\}_{i < k_0} \cup \{w_i\}_{i \geq k_1}$ and a weighted cycle X with weights $\{w_{k_0}, w_{k_0+1}, \dots, w_{k_1-1}\}$ such that*

- (1) $\|W - (V \oplus X)\| < (\max_{0 \leq j \leq m-1} |w_{k_0+j}|)\pi/m$; and
- (2) $W - (V \oplus X) \big| (\text{span}\{e_{k_0+j}, e_{k_1+j}\}_{j=0}^{m-1})^\perp = 0$.

Proof. Cf. [Her2, Chapter 2]. \square

2.1.5. The perturbation of W in the Proposition above can also be viewed as the “glueing” together of the weight sequence $\{w_i\}_{i < k_0}$ to the weight sequence

$\{w_i\}_{i \geq k_1}$ to obtain a new shift V with precisely that weight sequence. The cost of the perturbation is on the order of $1/m$. As a by-product of this process, we also obtain the cycle X .

It is the observation that

$$W - (V \oplus X) \big| (\text{span}\{e_{k_0+j}, e_{k_1+j}\}_{j=0}^{m-1})^\perp = 0$$

that will allow us to repeat Berg's Technique on $(\text{span}\{e_{k_0+j}, e_{k_1+j}\}_{j=0}^{m-1})^\perp$ without the perturbations interfering with each other. In other words, the norm of the two perturbations will be the *maximum* of the norms of the individual perturbations, *not the sum* of their norms.

2.2.1. Consider two operators A and B in $\mathcal{B}(\mathcal{H})$. For T in $\mathcal{B}(\mathcal{H})$, let $T^{(\infty)}$ be the operator $\bigoplus_{i=1}^{\infty} T$ acting on the Hilbert space $\mathcal{H}^{(\infty)} = \bigoplus_{i=1}^{\infty} \mathcal{H}$. We will define a quantity $\delta_A^e(B)$ which measures how far A is from containing $B^{(\infty)}$ as a direct summand, i.e., how far $\mathcal{U}(A)$ is from an operator of the form $J \oplus B^{(\infty)}$. Then one can obtain an estimate for $d(\mathcal{U}(A), \mathcal{U}(B))$ which depends only upon $\delta_A^e(B)$ and $\delta_B^e(A)$. (This will be made more explicit below.) For the remainder of this section, we shall be concerned with generating an upper bound for $\delta_V^e(W)$ which depends only upon finite dimensional information extracted from W and V . Theorem 2.2.11 shows how to use Berg's Technique to "glue together" the finite dimensional approximations, thereby obtaining the desired estimates.

We begin by defining $\delta_A(B)$. Although this definition is different from that of [Dav1], they are shown to be equivalent by the results of that paper.

2.2.2. Definition. Let A and $B \in \mathcal{B}(\mathcal{H})$. We define

$$\delta_A(B) = \inf\{\|U^*AU - J \oplus B\| : U \text{ unitary}, J \in \mathcal{B}(\mathcal{H})\}.$$

The "essential" version of this is

$$\delta_A^e(B) = \inf\{\|U^*AU - J \oplus B^{(\infty)}\| : U \text{ unitary}, J \in \mathcal{B}(\mathcal{H})\}.$$

2.2.3. In general, if an operator A embeds in B as a direct summand, and B embeds in A as a direct summand, little can be said about the distance between their unitary orbits (cf. [Dav1, §5]). However, the quantities $\delta_A^e(B)$ and $\delta_B^e(A)$ are precisely what is needed to produce the following "Schroeder-Bernstein" type of result.

2.2.4. Proposition. Let A and B be elements of $\mathcal{B}(\mathcal{H})$. Then

$$d(\mathcal{U}(A), \mathcal{U}(B)) \leq 2(\delta_A^e(B) + \delta_B^e(A)).$$

Proof. Let $\varepsilon > 0$. Then there exists $J \in \mathcal{B}(\mathcal{H})$ such that

$$A \cong_{(\delta_A^e(B)+\varepsilon)} J \oplus B^{(\infty)} \cong J \oplus B^{(\infty)} \oplus B \cong_{(\delta_B^e(A)+\varepsilon)} A \oplus B.$$

Similarly, $B \cong_{2(\delta_B^e(A)+\varepsilon)} A \oplus B$. Thus $d(\mathcal{U}(A), \mathcal{U}(B)) \leq 2(\delta_A^e(B) + \delta_B^e(A)) + 4\varepsilon$. Since this is true for all $\varepsilon > 0$,

$$d(\mathcal{U}(A), \mathcal{U}(B)) \leq 2(\delta_A^e(B) + \delta_B^e(A)). \quad \square$$

2.2.5. Thus in general for weighted shifts it suffices to estimate $\delta_V^e(W)$ and $\delta_W^e(V)$. In [O'Dn], O'Donovan shows that complete approximate unitary invariants for invertible bilateral weighted shifts are given by what we shall call the n -spectra $\sum_n W = \{(w_{i+1}, w_{i+2}, \dots, w_{i+n}) : i \in \mathbb{Z}\}^{-\|\cdot\|}$, $n \geq 1$ of W . In particular, he shows that two weighted shifts W and V are approximately unitarily equivalent if and only if $\sum_n W = \sum_n V$ for each $n \geq 1$. In our case, we shall also require the notion of an *essential n -spectrum* $\sum_n^e W$ for W . We shall say that a vector $(x_1, \dots, x_n) \in \mathbb{R}^n$ is in $\sum_n^e W$ if for all $\varepsilon > 0$ there exist infinitely many $k \in \mathbb{Z}$ for which $\|(w_{k+1}, \dots, w_{k+n}) - (x_1, \dots, x_n)\| < \varepsilon$. Based on the analogy with normal operators, one would expect $\sum_n V$ and $\sum_n W$ to play a major role in determining the distance $d(\mathcal{U}(V), \mathcal{U}(W))$ between the unitary orbits of W and V .

We shall see that this is in fact the case. But first we must recast the elements $(w_{i+1}, w_{i+2}, \dots, w_{i+n})$ of $\sum_n W$ in a form which will be more useful to us. We begin with a few definitions.

2.2.6. **Definition.** For $t, r \in \mathbb{Z}$, let $E_{(t,r)} \subseteq \mathcal{H}$ be the subspace $E_{(t,r)} = \text{span}\{e_i\}_{t \leq i \leq r}$, $E_{(-\infty, r)} = \overline{\text{span}\{e_i\}_{i \leq r}}$ and $E_{(t, \infty)} = \overline{\text{span}\{e_i\}_{i \geq t}}$. By $E_{(t)}$ we mean $E_{(t,t)} = \text{span}\{e_t\}$. We also define $P_{(t,r)}$, $-\infty \leq t \leq r \leq \infty$, to be the orthogonal projection onto $E_{(t,r)}$.

In order to effect many simultaneous perturbations to an operator while being sure that these perturbations occur on orthogonal spaces and hence do not interfere with each other (a norm controlling device), we shall also require the following technical notion.

2.2.7. **Definition.** Given a Hilbert space \mathcal{H} with orthonormal basis $\{e_i\}_{i \in \mathbb{Z}}$ and a finite dimensional subspace M of \mathcal{H} , we shall say that an operator $T \in \mathcal{B}(\mathcal{H})$ has *finite $\{e_i\}$ -range ($f.e_{i,r}$) with respect to M* if there exists a positive integer q such that $TM \subseteq E_{(-q,q)}$.

Analogously to [Dav1], we shall produce a quantity $\delta_V(W_{(t,r)})$ which measures our ability to estimate finite rank restrictions of W by V . We have

2.2.8. **Definition.** For $t < r \in \mathbb{Z}$, let

$$\delta_V(W_{(t,r)}) = \inf_{\substack{U \in \mathcal{U}(\mathcal{H}) \\ U \text{ has } f.e_{i,r} \text{ with} \\ \text{respect to } E_{(t-1, r+1)}}} \max\{\|(VU - UW)P_{(t,r)}\|, \|(V^*U - UW^*)P_{(t,r)}\|\}.$$

The quantity $\delta_V(W_{(t,r)})$ tells us that for all $\varepsilon > 0$ there exists a unitary operator U and a positive integer q such that

- (1) $UE_{(t-1, r+1)} \subseteq E_{(-q+1, q-1)}$;
- (2) $UE_{(t, r)} \subseteq E_{(-q+2, q-2)}$; and
- (3) there exists a finite dimensional operator C , $\|C\| \leq \delta_V(W_{(t,r)}) + \varepsilon$ such that $\{\text{range } C, \text{range } C^*\} \subseteq E_{(-q, q)}$ and $V + C$ can be decomposed

into the following operator matrix with respect to the decomposition

$$\begin{aligned}\mathcal{H} = & E_{(-\infty, -q-1)} \oplus N_1 \oplus UE_{(t-1)} \oplus UE_{(t, r)} \\ & \oplus UE_{(r+1)} \oplus N_2 \oplus E_{(q+1, \infty)},\end{aligned}$$

where

- (i) $N_1 = E_{(-q, 0)} \ominus UE_{(t-1, r+1)}$;
- (ii) $N_2 = E_{(1, q)} \ominus UE_{(t-1, r+1)}$; and
- (iii) $UE_{(t, r)}$ has orthonormal basis $\{Ue_t, Ue_{t+1}, \dots, Ue_r\}$:

$$V + C \cong \begin{bmatrix} * & * & * & 0 & * & * & 0 \\ * & * & * & 0 & * & * & * \\ * & * & * & 0 & * & * & * \\ 0 & 0 & A_1 & A_2 & 0 & 0 & 0 \\ * & * & * & A_3 & * & * & * \\ * & * & * & 0 & * & * & * \\ 0 & * & * & 0 & * & * & * \end{bmatrix}$$

where

$$A_1 = \begin{bmatrix} w_{t-1} \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \quad A_3 = [0 \cdots 0 \ w_r]$$

and

$$A_2 = \begin{bmatrix} 0 & & & & & & \\ w_t & 0 & & & & & \\ & w_{t+1} & 0 & & & & \\ & & & \ddots & & & \\ & & & & & & \\ & & & & & & w_{r-1} & 0 \end{bmatrix}.$$

With respect to the same decomposition of \mathcal{H} , C has the form

$$C \cong \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 & 0 \\ 0 & * & * & * & * & * & 0 \\ 0 & 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The importance of having C be finite rank lies in the observation that at this point, we have perturbed neither $V|_{E_{(q+1, \infty)}}$ nor $V|_{E_{(-\infty, -q-1)}}$. Thus in the same way that we are approximating $W|_{E_{(t, r)}}$ and $W^*|_{E_{(t, r)}}$ using

$V|_{E_{(-q,q)}}$, we can now consider approximating $W|_{E_{(t',r')}}$ and $W^*|_{E_{(t',r')}}$ using $V|_{\overline{\text{span}\{E_{(q+1,\infty)}, E_{(-\infty,-q-1)}\}}}$ to produce a new finite rank perturbation C' of $V + C$, the action of C' taking place on a subspace orthogonal to $E_{(-q,q)}$. But then we shall have that $\|(V + C) + C' - V\| = \|C + C'\| = \max\{\|C\|, \|C'\|\}$. We can continue this process ad infinitum to obtain a new operator V' whose distance from (a unitarily equivalent copy of) V is the supremum of the norms of the individual perturbations. In other words, *the norms of these perturbations do not add up*.

2.2.9. The “essential” version of $\delta_V(W_{(t,r)})$ asks that $UE_{(t-1,r+1)}$ be orthogonal to an arbitrary given $E_{(-q,q)}$ and can be defined as follows:

Definition. For $t, r \in \mathbb{Z}$, let

$$\begin{aligned} \delta_V^e(W_{(t,r)}) &= \sup_{q \geq 1} \inf_{\substack{U \in \mathcal{U}(\mathcal{H}); \\ UE_{(t-1,r+1)} \perp E_{(-q,q)}; \\ U \text{ has f.e. } r. \\ \text{with respect to } E_{(t-1,r+1)}}} \max\{\|(UV - UW)P_{(t,r)}\|, \|(V^*U - UW^*)P_{(t,r)}\|\}. \end{aligned}$$

2.2.10. At this point we can make a pair of observations.

First, we can replace the notion of orthogonality to $E_{(-q,q)}$ by orthogonality to an arbitrary finite dimensional subspace M . This definition would more closely resemble that of [Dav1]. However, as any finite dimensional subspace M is “almost” contained in $E_{(-q,q)}$ for q sufficiently large, the value we would then obtain for $\delta_V^e(W_{(t,r)})$ would be identical to our present one. The reason for our seemingly more complicated definition is to avoid technical problems later on.

Secondly, as in [Dav1], we also obtain the following piece of information:

$$\delta_V^e(W_{(t,r)}) = \delta_V(W_{(t,r)}^{(\infty)}) = \delta_V^e(W_{(t,r)}^{(\infty)}).$$

We introduce a last bit of terminology before starting. Associated with $\delta_V(W_{(t,r)})$ is the operator

$$\begin{bmatrix} 0 & & & & & \\ w_t & 0 & & & & \\ & w_{t+1} & 0 & & & \\ & & \ddots & \ddots & & \\ & & & & w_{r-1} & 0 \end{bmatrix}$$

and hence we may associate to $\delta_V(W_{(t,r)})$ the weight sequence $\mathbf{w}_{(t,r)} = (w_t, \dots, w_{r-1})$. If $r = t + 2n$ (with n taken to be a fixed constant), then $\mathbf{w}_{(t,t+2n)}$ will also be denoted \mathbf{w}_t . Moreover, $\mathbf{w}_{(t,r)}$ (resp. \mathbf{w}_t) will be thought of as “acting upon” $E_{(t,r)}$ (resp. $E_{(t,t+2n)}$).

The following theorem is the major technical result of this section. It shows how to “glue” together approximations of finite pieces of W in order to obtain integral copies of W in V .

2.2.11. Theorem. *Let $n \geq 2$ and assume $\delta_V^e(W_{(t, t+2n)}) \leq \delta$ for all $t \in \mathbb{Z}$. Then*

$$\delta_V^e(W) \leq \delta + \pi/2(n-1).$$

Proof. Since we are trying to estimate $\delta_V^e(W) = \delta_V(W^{(\infty)})$, we begin by assuming that we are given $W^{(\infty)} = \bigoplus_{i=1}^{\infty} W$ acting on $\mathcal{H}^{(\infty)} = \bigoplus_{i=1}^{\infty} \mathcal{H}$ in the usual way. We shall denote the j th copy of W in $W^{(\infty)}$ by $W[j]$, and it will act upon the j th copy of \mathcal{H} in $\mathcal{H}^{(\infty)}$, $\mathcal{H}_j = \overline{\text{span}\{e_i^j\}_{i \in \mathbb{Z}}}$, $\{e_i^j\}_{i \in \mathbb{Z}}$ an orthonormal basis, by

$$W[j]e_i^j = w_i e_{i+1}^j, \quad i \in \mathbb{Z}, \quad j \geq 1.$$

Then $w_i[j]$ will refer to the weight sequence w_i of $W[j]$. Moreover, $E_{(r,t)}^{[j]}$ will refer to the subspace $\overline{\text{span}\{e_i^j\}_{r \leq i \leq t}}$ of \mathcal{H}_j , and $P_{(r,t)}^{[j]}$ will be the corresponding orthogonal projection.

We shall find it sufficient to restrict our attention to those $w_i[j]$ for which $t \in n\mathbb{Z} + 1$. The observation that

$$\delta_V^e(W_{(kn+1, (k+2)n+1)}) = \delta_V(W_{(kn+1, (k+2)n+1)}^{(\infty)})$$

leads us to consider a set L which contains infinitely many copies of each w_{kn+1} , $k \in \mathbb{Z}$, each copy acting on a space orthogonal to that of any other copy. We take

$$L = \{w_1[1], w_1[2], w_{-n+1}[3], w_{n+1}[4], w_1[5], w_{-n+1}[6], \\ w_{n+1}[7], w_{-2n+1}[8], w_{2n+1}[9], w_1[10], \dots\}$$

which we relabel as $L = \{z_1, z_2, z_3, \dots\}$ as an ordered list.

That this choice for L possesses the property above is clear. We now proceed by induction to successively approximate each z_j , $j \geq 1$, using mutually orthogonal finite dimensional subspaces of V .

To that end, let $0 < \varepsilon < 1$. Consider $z_1 = w_1[1]$. Since $\delta_V^e(W_{(1, 2n+1)}) \leq \delta$, there exists a unitary $U_1: \mathcal{H}_1 \rightarrow \mathcal{H}$ and $q(1) > 0$ such that,

- (1) $\max\{\|(VU_1 - U_1 W[1])P_{(1, 2n+1)}^{[1]}\|, \|(V^*U_1 - U_1 W[1]^*)P_{(1, 2n+1)}^{[1]}\|\} < \delta + \varepsilon$;
- (2) $U_1 E_{(0, 2n+2)}^{[1]} \subseteq E_{(-q(1)+1, q(1)-1)}$;
- (3) $U_1 E_{(1, 2n+1)}^{[1]} \subseteq E_{(-q(1)+2, q(1)-2)}$; and
- (4) there exists a finite dimensional operator C_1 , $\|C_1\| < \delta + \varepsilon$, such that $\{\text{range } C_1, \text{range } C_1^*\} \subseteq E_{(-q(1), q(1))}$ and $V + C_1$ can be decomposed into an operator matrix of the form:

$$V + C_1 \cong \begin{bmatrix} * & * & * & 0 & * & * & 0 \\ * & * & * & 0 & * & * & * \\ * & * & * & 0 & * & * & * \\ 0 & 0 & A_1 & B_1 & 0 & 0 & 0 \\ * & * & * & D_1 & * & * & * \\ * & * & * & 0 & * & * & * \\ 0 & * & * & 0 & * & * & * \end{bmatrix}$$

with respect to the decomposition of \mathcal{H} as in §2.2.8. Here,

$$A_1 = \begin{bmatrix} w_0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \quad D_1 = [0 \cdots 0 \ w_{2n+1}]$$

and

$$B_1 = \begin{bmatrix} 0 & & & & & & \\ w_1 & 0 & & & & & \\ & w_2 & 0 & & & & \\ & & w_3 & & & & \\ & & & \ddots & & & \\ & & & & & & \\ & & & & & w_{2n} & 0 \end{bmatrix}$$

In general, for $j > 1$, \mathbf{z}_j corresponds to $\mathbf{w}_{s(j)}[j]$ for some $s(j) \in n\mathbb{Z} + 1$. Since $\delta_V^e(W_{(s(j), s(j)+2n)}) \leq \delta$, there exists a unitary $U_j: \mathcal{H}_j \rightarrow \mathcal{H}$ and $q(j) > 0$ such that

- (1) $\max\{\|(VU_j - U_jW[j])P_{(s(j), s(j)+2n)}^{[j]}\|, \|(V^*U_j - U_jW[j]^*)P_{(s(j), s(j)+2n)}^{[j]}\|\} < \delta + \varepsilon$;
- (2) $U_jE_{(s(j)-1, s(j)+2n+1)}^{[j]} \subseteq E_{(-q(j)+1, q(j)-1)}$;
- (3) $U_jE_{(s(j), s(j)+2n)}^{[j]} \subseteq E_{(-q(j)+2, q(j)-2)}$;
- (4) $U_jE_{(s(j)-1, s(j)+2n+1)}^{[j]}$ is orthogonal to $E_{(-2q(j-1), 2q(j-1))}$.

It is condition (4) which allows us to assert that we can perturb V by a finite rank operator C_j , $\|C_j\| < \delta + \varepsilon$ and $\{\text{range } C_j, \text{range } C_j^*\} \subseteq E_{(-q(j), q(j))} \ominus E_{(-2q(j-1), 2q(j-1))}$ (thus making $\text{range } C_j$ and $\text{range } C_j^*$ orthogonal to $\text{span}_{1 \leq i \leq j-1} \{\text{range } C_i, \text{range } C_i^*\}$), so that $V + C_j$ can be decomposed, as in

§2.2.8, into an operator matrix of the form

$$V + C_j \cong \begin{bmatrix} * & * & * & 0 & * & * & 0 \\ * & * & * & 0 & * & * & * \\ * & * & * & 0 & * & * & * \\ 0 & 0 & A_j & B_j & 0 & 0 & 0 \\ * & * & * & D_j & * & * & * \\ * & * & * & 0 & * & * & * \\ 0 & * & * & 0 & * & * & * \end{bmatrix}.$$

Again

$$A_j = \begin{bmatrix} w_{s(j)-1} \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}, \quad D_j = [0 \ 0 \ \cdots \ 0 \ w_{s(j)+2n}]$$

and

$$B_j = \begin{bmatrix} 0 & & & & & & \\ w_{s(j)} & 0 & & & & & \\ & w_{s(j)+1} & 0 & & & & \\ & & w_{s(j)+2} & & & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & & w_{s(j)+2n-1} & 0 \end{bmatrix}.$$

We also consider the partial isometry

$$U_0: \mathcal{H}^{(\infty)} \rightarrow \mathcal{H}, \quad U_0 = \text{strong-limit}_{m \rightarrow \infty} \sum_{j=1}^m U_j P_{(s(j), s(j)+2n)}^{[j]}$$

and let U be any extension of U_0 to a unitary in $\mathcal{B}(\mathcal{H}^{(\infty)}, \mathcal{H})$. The construction of U was made possible by condition (4) above, since $U_j E_{(s(j)-1, s(j)+2n+1)}^{[j]}$ orthogonal to $E_{(-2q(j-1), 2q(j-1))}$ ensured that U_0 has infinite cokernel.

Now $VU|E_{(s(j), s(j)+2n)}^{[j]} = VU_j|E_{(s(j), s(j)+2n)}^{[j]}$ (as all U_j 's have orthogonal domains and ranges when restricted to $E_{(s(j), s(j)+2n)}^{[j]}$). Moreover, since each C_j acts upon a space orthogonal to that acted upon by any other $C_{j'}$, $j \neq j'$, we may consider the operator

$$C = \text{strong-limit}_{m \rightarrow \infty} \left(\sum_{j=1}^m C_j \right),$$

and note that $\|C\| = \sup\{\|C_j\| : j \geq 1\} \leq \delta + \varepsilon$.

We then observe that

$$\begin{aligned}
 & \|((V + C)U - UW^{(\infty)})P_{(s(j), s(j)+2n)}^{[j]}\| \\
 &= \|((V + C)U_j - U_j W^{(\infty)})P_{(s(j), s(j)+2n)}^{[j]}\| \\
 &= \|((V + C)^* U_j - U_j W^{(\infty)*})P_{(s(j), s(j)+2n)}^{[j]}\| \\
 &= \|((V + C)^* U - UW^{(\infty)*})P_{(s(j), s(j)+2n)}^{[j]}\| \\
 &= 0, \quad j \geq 1.
 \end{aligned}$$

This implies that U has simultaneously embedded infinitely many copies of each w_{kn+1} , $k \in \mathbb{Z}$ into V . For given $k \in \mathbb{Z}$, it follows from our construction of L that $s^{-1}(kn+1)$ is an infinite set, say $s^{-1}(kn+1) = \{j_t : t \geq 1\}$. Furthermore, the spaces $F_{(k,t)} = UE_{(kn+1, (k+2)n+1)}^{[j_t]}$, $t \geq 1$, are mutually orthogonal with bases $\{f_1^{(k,t)}, \dots, f_{2n+1}^{(k,t)}\}$ where $f_r^{(k,t)} = Ue_{kn+r}^{j_t}$ on which $(V + C)$ acts by

$$(V + C)f_r^{(k,t)} = w_{kn+r}f_{r+1}^{(k,t)}, \quad 1 \leq r \leq 2n, \quad k \in \mathbb{Z}, \quad t \geq 1.$$

Now the action of $(V + C)$ restricted to $\{f_{n+1}^{(k,t)}, \dots, f_{2n}^{(k,t)}\}$ and $\{f_1^{(k+1,t)}, \dots, f_n^{(k+1,t)}\}$ satisfies the hypotheses of Berg's Theorem as described in §2.1.4, and hence we may apply his process of gradual change simultaneously on each of those mutually orthogonal pieces to obtain a new shift operator V' which acts upon an orthonormal set $\{g_{(k+1)n+1}^{(t)}, g_{(k+1)n+2}^{(t)}, \dots, g_{(k+1)n+n}^{(t)}\}$ with weights $\{w_{(k+1)n+1}, w_{(k+1)n+2}, \dots, w_{(k+1)n+n}\}$ where

- (1) $g_{(k+1)n+1}^{(t)} = f_{n+1}^{(k,t)}$ and $g_{(k+1)n+n}^{(t)} = f_n^{(k+1,t)}$; and
- (2) $g_{(k+1)n+r}^{(t)} \in \text{span}\{f_{n+r}^{(k,t)}, f_r^{(k+1,t)}\}$.

Of course, this technique also produces a second operator J which acts upon $\mathcal{H}^{(\infty)} \ominus \text{span}\{g_i^{(t)}\}_{i \in \mathbb{Z}, t \geq 1}$, J a direct sum of cycles and other shifts. However, the precise structure of J does not come into play here.

We also know from Berg's Technique that $\delta_{(V+C)}(V') < \pi/2(n-1)$. In fact, closer examination of V' reveals that for fixed t , V' restricted to the span of $\{g_i^{(t)}\}_{i \in \mathbb{Z}}$ is a weighted shift with weight sequence $\{w_i\}_{i \in \mathbb{Z}}$ as well as being a direct summand of V' , and as such $V'|_{\text{span}\{g_i^{(t)}\}_{i \in \mathbb{Z}}} \cong W$. Since this is true for each $t \geq 1$, $V' \cong W^{(\infty)}$.

Combining these results, we get

$$V =_{(\delta+\varepsilon)} V + C \cong_{\pi/2(n-1)} V' \oplus J \cong W^{(\infty)} \oplus J.$$

Thus $\delta_V^e(W) \leq (\delta + \pi/2(n-1)) + \varepsilon$. Since $\varepsilon > 0$ was arbitrary, $\delta_V^e(W) \leq \delta + \pi/2(n-1)$. \square

As to calculating an upper bound for $\delta_V^e(W_{(t, t+2n)})$, we note that an elementary one is obtained from considering $d((w_{t-1}, w_t, w_{t+1}, \dots, w_{t+2n}), \sum_{2n+2}^e V) = \sup_{\mathbf{v}_{(i, i+2n+1)} \in \sum_{2n+2}^e V} \|w_{(t-1, t+2n)} - \mathbf{v}_{(i, i+2n+1)}\|$. It is strong enough to produce

the following results, of which Corollary 2.2.13 and Proposition 2.2.14 were already known.

2.2.12. Corollary. *Let $n \geq 2$ and let V be an invertible bilateral weighted shift satisfying $\sum_{2n+2} V = \sum_{2n+2}^e V$. Then*

$$d(\mathcal{U}(V), \mathcal{U}(V^{(\infty)})) \leq \pi/(n-1).$$

Proof. We note that $\delta_V^e(V_{(t, t+2n)}) = 0$ for all $t \in \mathbb{Z}$ as is made clear by the above remarks. By our Theorem, $\delta_V^e(V) = \delta_V(V^{(\infty)}) \leq \pi/2(n-1)$. Thus

$$\begin{aligned} V &\cong_{\pi/2(n-1)} V^{(\infty)} \oplus J \cong V^{(\infty)} \oplus V^{(\infty)} \oplus J \\ &\cong_{\pi/2(n-1)} V^{(\infty)} \oplus V \cong V^{(\infty)}. \end{aligned}$$

We conclude that $d(\mathcal{U}(V), \mathcal{U}(V^{(\infty)})) \leq \pi/(n-1)$. \square

2.2.13. Corollary. *Let V be an invertible bilateral weighted shift satisfying $\sum_n V = \sum_n^e V$ for all $n \geq 1$. Then $V \cong_a V^{(\infty)}$.*

Proof. By the above corollary, $d(\mathcal{U}(V), \mathcal{U}(V^{(\infty)})) \leq \pi/(n-1)$ for all $n \geq 2$. Thus $d(\mathcal{U}(V), \mathcal{U}(V^{(\infty)})) = 0$. By Voiculescu's Theorem [Vcu], given $\varepsilon > 0$ there exists U unitary and K compact, $\|K\| < \varepsilon$ such that $V = U^* V^{(\infty)} U + K$. \square

2.2.14. Proposition (O'Donovan). *Let W and V be invertible bilateral weighted shifts satisfying $\sum_n W = \sum_n V$ for all $n \geq 1$. Then $W \cong_a V$.*

Proof. We first consider the case where $\sum_n V \neq \sum_n^e V$ for some $n > 1$. Assume the vector (x_1, \dots, x_n) occurs only m times as an element of $\sum_n^\circ V = \sum_n V \setminus \sum_n^e V$. That is, there exist exactly m distinct integers t_1, t_2, \dots, t_m for which $(v_{t_i}, v_{t_i+1}, \dots, v_{t_i+n-1}) = (x_1, \dots, x_n)$, $1 \leq i \leq m$ and there also exists $\tau > 0$ such that for $k \notin \{t_1, t_2, \dots, t_m\}$, $\|(v_k, v_{k+1}, \dots, v_{k+n}) - (x_1, \dots, x_n)\| > \tau$. Then by choosing N large enough, ($N > \max_{1 \leq j \leq m} |t_j| + n$ will do), (v_{-N}, \dots, v_N) contains all occurrences of (x_1, \dots, x_n) . Moreover, this means that (v_{-N}, \dots, v_N) occurs only once in $\sum_{2N+1} V$ and if $k \neq -N$ then $\|(v_k, v_{k+1}, \dots, v_{k+2N}) - (v_{-N}, \dots, v_N)\| > \tau$. In fact, any string of weights (v_r, \dots, v_s) where $r \leq -N$ and $s \geq N$ is also unique in $\sum_{s-r+1} V$, since it contains (v_{-N}, \dots, v_N) . Now (v_{-N}, \dots, v_N) matches up with (w_t, \dots, w_{t+2N}) for some $t \in \mathbb{Z}$ by assumption. If the choice of t is not unique, let $t' \in \mathbb{Z}$ such that $(v_{-N}, \dots, v_N) = (w_{t'}, \dots, w_{t'+2N})$ and $t' \neq t$. Choose $M > (|t'| + |t| + 2N)$. Then (w_{-M}, \dots, w_M) is a string which contains two distinct substrings arbitrarily close to (v_{-N}, \dots, v_N) . But from above, any string of length $2M+1$ in $\sum_{2M+1} V$ contains at most one copy of (v_{-N}, \dots, v_N) , a contradiction. Thus the choice of t is unique. But it is now not too difficult to see that this pairing forces the pairing of (v_r, \dots, v_s) , $r \leq -N$, $s \geq N$ with its unique counterpart in $\sum_{s-r+1} W$. Clearly then, all

the weights of W and V must match up (i.e., let $s \rightarrow \infty$, $r \rightarrow -\infty$) and so $W \cong V$.

If $\sum_n V = \sum_n^e V = \sum_n^e W = \sum_n W$ for all $n \geq 1$, then $\delta_V^e(W_{(t, t+2n)}) = 0$ and $\delta_W^e(V_{(t, t+2n)}) = 0$ for all $n \geq 1$. By Theorem 2.2.11, we may conclude that

$$\delta_V^e(W) = 0 = \delta_W^e(V).$$

Applying Proposition 2.2.4,

$$d(\mathcal{U}(V), \mathcal{U}(W)) \leq 2(\delta_V^e(W) + \delta_W^e(V)) = 0.$$

But then $V \cong_a W$ by Voiculescu's Theorem [Vcu]. \square

That the converse of this Proposition is true follows from the work of the next section on lower bounds for $\delta_V^e(W)$.¹ The above results combine to give us our main Theorem for upper bounds. For $X, Y \subseteq \mathbb{C}^n$, let $d^E(X, Y) = \sup_{x \in X} \inf_{y \in Y} \|x - y\|_\infty$.

2.2.15. Theorem (Main result for upper bounds). *Let W and V be invertible bilateral weighted shifts and let*

$$\delta_n = \max\{d^E(\sum_{2n+2} W, \sum_{2n+2}^e V), d^E(\sum_{2n+2} V, \sum_{2n+2}^e W)\}.$$

Then

$$d(\mathcal{U}(V), \mathcal{U}(W)) \leq \inf_{n \geq 2} (4\delta_n + 2\pi/(n-1)).$$

Proof. Fix n temporarily. Now $d^E(\sum_{2n+2} W, \sum_{2n+2}^e V) \leq \delta_n$ implies $\delta_V^e(W_{(t, t+2n)}) \leq \delta_n$ and $d^E(\sum_{2n+2} V, \sum_{2n+2}^e W) \leq \delta_n$ implies $\delta_W^e(V_{(t, t+2n)}) \leq \delta_n$. By Theorem 2.2.11, $\delta_V^e(W) \leq \delta_n + \pi/2(n-1)$ and $\delta_W^e(V) \leq \delta_n + \pi/2(n-1)$. From Proposition 2.2.4,

$$d(\mathcal{U}(V), \mathcal{U}(W)) \leq 2(2\delta_n + \pi/(n-1)).$$

Since this is true for all $n \geq 2$, the result follows. \square

2.2.16. Corollary. *Let W and V be invertible bilateral weighted shifts satisfying $\sum_n W = \sum_n^e W$ and $\sum_n V = \sum_n^e V$ for all $n \geq 1$. If d_H represents Hausdorff distance in the l^∞ -norm on \mathbb{C}^n , then*

$$d(\mathcal{U}(V), \mathcal{U}(W)) \leq 4 \inf_{n \geq 2} (d_H(\sum_{2n+2} V, \sum_{2n+2} W) + \pi/2(n-1)).$$

Proof. Immediate. \square

3. A LOWER BOUND

3.1.1. In general, it is clear that for $t, r \in \mathbb{Z}$,

$$\delta_V(W) \geq \delta_V(W_{(t, r)}),$$

¹ See the remark at the end of §3.1.7.

since the same unitary that serves to embed W in V will embed the restriction of W to a finite dimensional subspace. Moreover, if $W \cong_a W^{(\infty)}$, then it is immediate that

$$\delta_V(W) = \delta_V^e(W) = \delta_V(W^{(\infty)}),$$

and

$$\delta_V(W_{(t,r)}) = \delta_V^e(W_{(t,r)}) = \delta_V(W_{(t,r)}^{(\infty)}).$$

(In fact, the latter equations hold true if we only require $(w_{t-1}, w_t, \dots, w_r) \in \sum_{r-t+2}^e W$.) We note that for any unitary operator $U_1: \mathcal{H} \rightarrow \mathcal{H}$,

$$\delta_V(W) = \inf_{\substack{U \text{ unitary} \\ J \in \mathcal{B}(\mathcal{H})}} \|V - U^*(W \oplus J)U\| \leq \|V - U_1^* W U_1\|.$$

(Consider J acting on a zero-dimensional space, i.e., J absent from the left-hand side of this equation.) Since this relation holds true for $\delta_W(V)$ as well, it is clear that

$$d(\mathcal{U}(V), \mathcal{U}(W)) \geq \max(\delta_V(W), \delta_W(V)).$$

Thus any lower bound for $\delta_V(W_{(t,r)})$ is automatically a lower bound for $d(\mathcal{U}(V), \mathcal{U}(W))$.

The purpose of this section is to find a lower bound for $\delta_V(W_{(t,r)})$ based once again on the distance $d(\sum_{r-t+2} W, \sum_{r-t+2} V)$ between the unitarily invariant subsets $\sum_{r-t+2} W$ and $\sum_{r-t+2} V$ of \mathbb{R}^{r-t+2} .

3.1.2. The unitary U in the norm estimate $\|U^* V U - W\| = \|V U - U W\|$ intertwines V and W and sends the original basis $\{e_i\}_{i \in \mathbb{Z}}$ to a new basis $\{f_i = U e_i\}_{i \in \mathbb{Z}}$. Each f_i can be thought of as an averaging of the e_i 's, and consequently the action of V on f_i "averages" the weights of V , although it does more.

One might therefore expect that if, for example, $w_1 = 3/4$ and $v_1 = 1/2$, $v_2 = 1$, then one could use a unitary U which takes e_1 to, say, $f_1 = \sqrt{7/12}e_1 + \sqrt{5/12}e_2$ so that $\|V U e_1\| = \|V f_1\| = 3/4 = \|U W e_1\| = \|W e_1\|$. That is, U averages the values $1/2$ and 1 by the proper l_2 -combination to produce a weight of $3/4$. The next Lemma shows us that averaging weights v_j of V which are far away from a given w_i in order to approximate w_i produces a large value for $\|V^* U - U W^*\|$.

3.1.3. Lemma. *Let $f = \sum_{j \in \mathbb{Z}} \lambda_j e_j$, $\|f\| = 1$. Assume that there exist $\varepsilon > 0$ and $\delta > 0$ for which $(\sum_{|v_j - w_i| \geq \delta} |\lambda_j|^2)^{1/2} \geq \varepsilon$. Then $\|V^* V f - w_i^2 f\| \geq (\|V^{-1}\|^{-1} + \|W^{-1}\|^{-1})\delta\varepsilon$.*

Proof. First note that $V^* V f = \sum_{j \in \mathbb{Z}} \lambda_j v_j^2 e_j$. Thus

$$\|V^* V f - w_i^2 f\| \geq \left(\sum_{|v_j - w_i| \geq \delta} |\lambda_j|^2 (v_j^2 - w_i^2)^2 \right)^{1/2}.$$

But $|v_j - w_i| \geq \delta$ implies that

$$|v_j^2 - w_i^2| \geq (v_j + w_i)|v_j - w_i| \geq (\|V^{-1}\|^{-1} + \|W^{-1}\|^{-1})\delta.$$

Thus

$$\begin{aligned} \|V^*Vf - w_i^2f\| &\geq (\|V^{-1}\|^{-1} + \|W^{-1}\|^{-1})\delta \left(\sum_{|v_j - w_i| \geq \delta} |\lambda_j| \right)^{1/2} \\ &\geq (\|V^{-1}\|^{-1} + \|W^{-1}\|^{-1})\delta\varepsilon. \quad \square \end{aligned}$$

3.1.4. Notation. Because we shall have frequent occasion to use it in the sequel, we write $\beta = (1/2)(\|V^{-1}\|^{-1} + \|W^{-1}\|^{-1})$.

3.1.5. Lemma. Let U be a unitary operator taking e_i to $f_i = Ue_i$, $t \leq i \leq r$, and assume

- (1) there exists i_0 , $t \leq i_0 < r$ such that $f_{i_0} = \sum_{j \in \mathbb{Z}} \lambda(i_0, j)e_j$; and
- (2) there exist $\varepsilon > 0$ and $\delta > 0$ such that $(\sum_{|v_j - w_{i_0}| \geq \delta} |\lambda(i_0, j)|^2)^{1/2} \geq \varepsilon$.

Then

$$\max\{\|(VU - UW)P_{(t,r)}\|, \|(V^*U - UW^*)P_{(t,r)}\|\} \geq \beta\delta\varepsilon.$$

Proof. From the previous Lemma, we know that

$$\|V^*Vf_{i_0} - w_{i_0}^2f_{i_0}\| \geq 2\beta\delta\varepsilon.$$

Thus $\|(V^*VU - UW^*W)e_{i_0}\| \geq 2\beta\delta\varepsilon$, or equivalently

$$\|(U^*V^*VU - W^*W)e_{i_0}\| \geq 2\beta\delta\varepsilon.$$

It follows that

$$\|(U^*V^*U)(U^*VU - W)e_{i_0} + (U^*V^*U - W^*)We_{i_0}\| \geq 2\beta\delta\varepsilon$$

so that

$$\max\{\|U^*V^*U(U^*VU - W)e_{i_0}\|, \|(U^*V^*U - W^*)We_{i_0}\|\} \geq \beta\delta\varepsilon.$$

But $e_{i_0+1} \in E_{(t,r)}$ and $w_{i_0} \leq 1$ so that

$$\max\{\|(U^*VU - W)P_{(t,r)}\|, \|(U^*V^*U - W^*)P_{(t,r)}\|\} \geq \beta\delta\varepsilon,$$

which is equivalent to our claim. \square

3.1.6. Thus what goes wrong is not our inability to approximate the weight w_{i_0} , but rather to simultaneously approximate both the weight and the shifting from f_{i_0} to f_{i_0+1} . These observations lead to the major result of this section.

Theorem. Let W and V be invertible bilateral weighted shifts and let $n > 0$. Assume $d((w_1, \dots, w_n), \sum_n V) = \delta > 0$. Then

$$\delta_V(W_{(1, n+1)}) \geq C_V^n(W) \min\{1/n, \delta/\sqrt{n}\},$$

where

$$C_V^n(W) = \frac{\beta \|V^{-(n-1)}\|^{-1}}{1 + \beta \|W^{-(n-2)}\| \|W^{-1}\|}.$$

Proof. First we recall that $d((w_1, \dots, w_n), \sum_n V)$ is taken in the l^∞ -norm of \mathbb{R}^n . We now suppose that we are given $\tau > 0$ and a unitary U such that

$$\begin{aligned} \max\{\|(VU - UW)P_{(1, n+1)}\|, \|(V^*U - UW^*)P_{(1, n+1)}\|\} \\ = \varepsilon < (1 + \tau)\delta_V(W_{(1, n+1)}). \end{aligned}$$

Let $f_i = Ue_i$, $i \in \mathbb{Z}$. In particular, we shall be concerned with

$$f_i = \sum_{j \in \mathbb{Z}} \lambda(i, j) e_{j+(i-1)}, \quad 1 \leq i \leq n+1.$$

We associate with each index j the sequence $(v_j, v_{j+1}, \dots, v_{j+(n-1)})$ of n -consecutive weights from V . Clearly, since $d((w_1, \dots, w_n), \sum_n V) = \delta$, for each $j \in \mathbb{Z}$ we can find a number $m(j)$, $0 \leq m(j) \leq n-1$ for which $|v_{j+m(j)} - w_{1+m(j)}| \geq \delta$. This leads us to consider the quantity

$$\|((v_j \cdots v_{j+m(j)-1}/w_1 \cdots w_{m(j)})\lambda(1, j))_{j \in \mathbb{Z}} - (\lambda(1+m(j), j))_{j \in \mathbb{Z}}\|.$$

Now

$$\begin{aligned} & \left| (v_j \cdots v_{j+m(j)-1}/w_1 \cdots w_{m(j)})\lambda(1, j) - \lambda(1+m(j), j) \right|^2 \\ & \leq \left[\sum_{i=1}^{m(j)-1} |v_{j+i} \cdots v_{j+m(j)-1}/w_{i+1} \cdots w_{m(j)}| \left| (v_{j+(i-1)}/w_i)\lambda(i, j) - \lambda(i+1, j) \right| \right. \\ & \quad \left. + |v_{j+m(j)-1}/w_{m(j)}|\lambda(m(j), j) - \lambda(1+m(j), j)| \right]^2 \\ & \leq m(j) \|W^{-m(j)+1}\|^2 \left[\sum_{i=1}^{m(j)} \left| (v_{j+(i-1)}/w_i)\lambda(i, j) - \lambda(i+1, j) \right|^2 \right] \\ & \quad \text{(by the Cauchy-Schwarz inequality)} \\ & \leq (n-1) \|W^{-(n-2)}\|^2 \left[\sum_{i=1}^{n-1} \left| (v_{j+(i-1)}/w_i)\lambda(i, j) - \lambda(i+1, j) \right|^2 \right]. \end{aligned}$$

At this point we observe that $\|VUe_i - UW e_i\| \leq \varepsilon$, that is, $\|Vf_i - w_i f_{i+1}\| \leq \varepsilon$ for $1 \leq i \leq n$. Thus for $1 \leq i \leq n$,

$$\left[\sum_{j \in \mathbb{Z}} |v_{j+(i-1)}\lambda(i, j) - w_i\lambda(i+1, j)|^2 \right]^{1/2} \leq \varepsilon$$

or equivalently,

$$\sum_{j \in \mathbb{Z}} |(v_{j+(i-1)}/w_i)\lambda(i, j) - \lambda(i+1, j)|^2 \leq w_i^{-2} \varepsilon^2 \leq \|W^{-1}\|^2 \varepsilon^2.$$

Continuing from above,

$$\begin{aligned} & \|((v_j \cdots v_{j+m(j)-1}/w_1 \cdots w_{m(j)})\lambda(1, j))_{j \in \mathbb{Z}} - (\lambda(1+m(j), j))_{j \in \mathbb{Z}}\|^2 \\ &= \sum_{j \in \mathbb{Z}} |(v_j \cdots v_{j+m(j)-1}/w_1 \cdots w_{m(j)})\lambda(1, j) - \lambda(1+m(j), j)|^2 \\ &\leq \sum_{j \in \mathbb{Z}} (n-1) \|W^{-(n-2)}\|^2 \left[\sum_{i=1}^{n-1} |(v_{j+(i-1)}/w_i)\lambda(i, j) - \lambda(i+1, j)|^2 \right] \\ &\leq (n-1) \|W^{-(n-2)}\|^2 \left[\sum_{i=1}^{n-1} \sum_{j \in \mathbb{Z}} |(v_{j+(i-1)}/w_i)\lambda(i, j) - \lambda(i+1, j)|^2 \right] \\ &\leq (n-1) \|W^{-(n-2)}\|^2 \sum_{i=1}^{n-1} \|W^{-1}\|^2 \varepsilon^2 \\ &\leq (n-1)^2 \|W^{-(n-2)}\|^2 \|W^{-1}\|^2 \varepsilon^2; \end{aligned}$$

so that

$$\begin{aligned} & \|((v_j \cdots v_{j+m(j)-1}/w_1 \cdots w_{m(j)})\lambda(1, j))_{j \in \mathbb{Z}} - (\lambda(1+m(j), j))_{j \in \mathbb{Z}}\| \\ &\leq (n-1) \|W^{-(n-2)}\| \|W^{-1}\| \varepsilon. \end{aligned}$$

Furthermore,

$$\|((v_j \cdots v_{j+m(j)-1}/w_1 \cdots w_{m(j)})\lambda(1, j))_{j \in \mathbb{Z}}\| \geq \|V^{-(n-1)}\|^{-1}.$$

Combining these two inequalities and applying the triangle inequality,

$$\|(\lambda(1+m(j), j))_{j \in \mathbb{Z}}\| \geq \|V^{-(n-1)}\|^{-1} - (n-1) \|W^{-(n-2)}\| \|W^{-1}\| \varepsilon.$$

Now $0 \leq m(j) \leq n-1$. For $1 \leq i \leq n$, let $F(i) = \{j : m(j) = i-1\}$. This allows us to rewrite $(\lambda(1+m(j), j))_{j \in \mathbb{Z}}$ as $(\lambda(i, F(i)))_{i=1}^n$. Since

$$\|(\lambda(i, F(i)))_{i=1}^n\| \geq \|V^{-(n-1)}\|^{-1} - (n-1) \|W^{-(n-2)}\| \|W^{-1}\| \varepsilon,$$

there must exist i_0 , $1 \leq i_0 \leq n$ for which

$$\left[\sum_{j \in F(i_0)} |\lambda(i_0, j)|^2 \right]^{1/2} \geq (1/\sqrt{n}) (\|V^{-(n-1)}\|^{-1} - (n-1) \|W^{-(n-2)}\| \|W^{-1}\| \varepsilon).$$

We now apply Lemma 3.1.5 to f_{i_0} to conclude that

$$\begin{aligned} \varepsilon &= \max\{\|(U^* V U - W)P_{(1, n+1)}\|, \|(U^* V^* U - W^*)P_{(1, n+1)}\|\} \\ &\geq \beta(\delta/\sqrt{n}) [\|V^{-(n-1)}\|^{-1} - (n-1) \|W^{-(n-2)}\| \|W^{-1}\| \varepsilon], \end{aligned}$$

or equivalently,

$$\varepsilon \geq \beta \delta \|V^{-(n-1)}\|^{-1} / [\sqrt{n} \cdot (1 + \beta \delta \sqrt{n} \|W^{-(n-2)}\| \|W^{-1}\|)].$$

Then if $\delta \geq 1/\sqrt{n}$,

$$\varepsilon \geq \beta \|V^{-(n-1)}\|^{-1} / [(\sqrt{n}/\delta) + \beta n \|W^{-(n-2)}\| \|W^{-1}\|] \geq C_V^n(W)/n.$$

If $\delta < 1/\sqrt{n}$,

$$\begin{aligned} \varepsilon &\geq \beta \delta \|V^{-(n-1)}\|^{-1} / [\sqrt{n}(1 + \beta \delta \sqrt{n} \|W^{-(n-2)}\| \|W^{-1}\|)] \\ &\geq C_V^n(W) \delta / \sqrt{n}. \end{aligned}$$

Finally, since $\varepsilon < (1 + \tau) \delta_V(W_{(1, n+1)})$, by letting $\tau \rightarrow 0$ we find that

$$\delta_V(W_{(1, n+1)}) \geq C_V^n(W) \min\{1/n, \delta/\sqrt{n}\},$$

as required. \square

3.1.7. Recall that $d^E(X, Y) = \sup_{x \in X} \inf_{y \in Y} \|x - y\|_\infty$, $X, Y \subseteq \mathbb{C}^n$. The analogue of Theorem 2.2.15 for lower bounds then takes on the form:

Theorem. Let W and V be invertible bilateral weighted shifts and let $\delta_n = \max\{d^E(\sum_n W, \sum_n V), d^E(\sum_n V, \sum_n W)\} = d_H(\sum_n W, \sum_n V)$. Then

$$d(\mathcal{U}(V), \mathcal{U}(W)) \geq \sup_{n \geq 2} \{\min\{C_V^n(W), C_W^n(V)\} \cdot \min\{1/n, \delta_n/\sqrt{n}\}\}.$$

Proof. Immediate from the previous theorem and the comments of §3.1.1. \square

The need to consider $\min\{C_V^n(W), C_W^n(V)\}$ arises from the fact that, in general, for $m \neq n$ it is possible to have $d^E(\sum_n W, \sum_n V) > d^E(\sum_n V, \sum_n W)$ and $d^E(\sum_m W, \sum_m V) < d^E(\sum_m V, \sum_m W)$. The value δ_n of Theorem 3.1.6 is then a measure of $\delta_V(W)$, while δ_m is a measure of $\delta_W(V)$. We must therefore apply Theorem 3.1.6 to estimate the correct embedding.

Remark. The converse of Proposition 2.2.14 (O'Donovan's Theorem) now follows immediately from Theorem 3.1.7. For if there exists $n > 0$ for which $\sum_n W \neq \sum_n V$, then $\delta_n > 0$ and so $d(\mathcal{U}(V), \mathcal{U}(W)) > 0$, from which we conclude that W is not approximately unitarily equivalent to V .

3.1.8. The lower bound produced by Theorem 3.1.7 is unfortunately not always the greatest lower bound (cf. Example 1, Example 2). The difficulty stems from trying to determine how much one can “average” the weights of V to minimize the estimate of $\|(VU - UW)P_{(t, r)}\|$ without running into too much difficulty with the adjoint estimate of $\|(V^*U - UW^*)P_{(t, r)}\|$ (cf. Lemma 3.1.5). If the n -spectrum of V is a tractable set, then one can achieve a stronger result which subsumes Theorem 3.1.6, namely,

3.1.9. Theorem. Let W and V be invertible bilateral weighted shifts and let $n > 0$. Assume $d((w_1, w_2, \dots, w_n), \sum_n V) \geq \delta_n \geq 0$. For each $j \in \mathbb{Z}$, let

$B_n(j) = \{i : 1 \leq i \leq n, |v_{j+(i-1)} - w_i| \geq \delta_n\}$. Also, let A_n be a subset of $\{1, 2, \dots, n\}$ which satisfies

- (1) $B_n(j) \cap A_n \neq \emptyset$ for each $j \in \mathbb{Z}$; and
- (2) if A'_n is another subset of $\{1, 2, \dots, n\}$ which satisfies condition (1), then $|A'_n| \geq |A_n|$.

We may conclude that

$$\delta_V(W) \geq \delta_V(W_{(1, n+1)}) \geq C_V^n(W) \min\{1/n, \delta_n/\sqrt{|A_n|}\}.$$

Proof. As in Theorem 3.1.6, let $\tau > 0$ and choose a unitary U such that

$$\max\{\|(VU - UW)P_{(1, n+1)}\|, \|(V^*U - UW^*)P_{(1, n+1)}\|\} = \varepsilon < (1 + \tau)\delta_V(W).$$

As before, let $Ue_i = f_i = \sum_{j \in \mathbb{Z}} \lambda(i, j)e_{j+(i-1)}$, $1 \leq i \leq n+1$. For each $j \in \mathbb{Z}$, select the smallest element of $B_n(j) \cap A_n$ and denote that element $m_n(j)$. As in the proof of Theorem 3.1.6, we can obtain the inequality

$$\|(\lambda(1 + m_n(j), j))_{j \in \mathbb{Z}}\| \geq \|V^{-(n-1)}\|^{-1} - (n-1)\|W^{-(n-2)}\| \|W^{-1}\| \varepsilon.$$

Since $|\{m_n(j) : j \in \mathbb{Z}\}| \leq |A_n|$, there must exist $i_0 \in \{m_n(j) : j \in \mathbb{Z}\}$ such that

$$\|(\lambda(i_0, m_n^{-1}(i_0)))\| \geq \frac{\|V^{-(n-1)}\|^{-1} - (n-1)\|W^{-(n-2)}\| \|W^{-1}\| \varepsilon}{\sqrt{|A_n|}}.$$

By Lemma 3.1.5,

$$\varepsilon \geq \left[\frac{\|V^{-(n-1)}\|^{-1} - (n-1)\|W^{-(n-2)}\| \|W^{-1}\| \varepsilon}{\sqrt{|A_n|}} \right] \beta \delta_n.$$

Let $\delta'_n = \delta_n \sqrt{n/|A_n|}$. Then the proof of Theorem 3.1.6 shows that

$$\begin{aligned} \delta_V(W_{(1, n+1)}) &\geq C_V^n(W) \min\{1/n, \delta'_n/\sqrt{n}\} \\ &= C_V^n(W) \min\{1/n, \delta_n/\sqrt{|A_n|}\}. \quad \square \end{aligned}$$

3.1.10. For any $0 < \tau < 1$ we can choose $k = k(\tau) \in \mathbb{Z}$ such that

$$d((w_k, w_{k+1}, \dots, w_{k+(n-1)}), \sum_n V) \geq \tau d^E(\sum_n W, \sum_n V).$$

We then see that we can replace δ_n in Theorem 3.1.9 by $\tau d^E(\sum_n W, \sum_n V)$ to obtain a lower bound for $\delta_V(W)$. By then letting τ tend to 1 and by taking the supremum over n , we obtain the estimate

$$\delta_V(W) \geq \sup_n \{C_V^n(W) \min\{1/n, d^E(\sum_n W, \sum_n V)/\sqrt{|A_n|}\}.$$

Moreover, if we replace $|A_n|$ by n , the maximum value $|A_n|$ can achieve in any case, then we reobtain Theorem 3.1.6. In this sense, Theorem 3.1.9 subsumes Theorem 3.1.6.

The seemingly awkward formulation of Theorem 3.1.9 does not prevent it from being of practical use as we shall see in the examples of the next section, where it yields “optimal” lower bounds—“optimal” up to a factor of $\min\{C_V^n(W), C_W^n(V)\}$.

4. EXAMPLES

4.1.1. As mentioned in §3.1.8, the lower bound of Theorem 3.1.6 is not always optimal. Our first example illustrates a case where our general upper bound is on the correct order of magnitude, but for which we must turn to Theorem 3.1.9 for a good lower bound.

4.1.2. **Example 1.** Let $\delta > 0$, $m > 2$ (m even for the sake of convenience), and consider the shifts W and V whose weight sequences are as below:

$$\begin{aligned} w_i &= (1 - \delta), & i \in (2m - 1)\mathbb{Z}; & & v_i &= (1 - \delta), & i \in 2m\mathbb{Z}; \\ w_i &= 1, & i \notin (2m - 1)\mathbb{Z}; & & v_i &= 1, & i \notin 2m\mathbb{Z}. \end{aligned}$$

As usual, $We_i = w_i e_{i+1}$, $Ve_i = v_i e_{i+1}$ for our orthonormal basis $\{e_i\}_{i \in \mathbb{Z}}$ and we extend W and V by linearity and continuity. Then W and V have periodic weight sequences with respective periods:

$$W : (1 - \delta)1 \ 1 \ \cdots \ 1 \quad (\text{with } 2m - 2 \text{ consecutive } 1\text{'s}),$$

and

$$V : (1 - \delta)1 \ 1 \ \cdots \ 1 \quad (\text{with } 2m - 1 \text{ consecutive } 1\text{'s}).$$

It is then immediate that $\sum_n W = \sum_n^e W$ and $\sum_n V = \sum_n^e V$ for all $n \geq 1$. We first examine the question of obtaining upper bounds.

4.1.3. To apply Theorem 2.2.15, we note that for $n \geq 2$,

$$\begin{aligned} \delta_n &= \max\{d^E(\sum_{2n} W, \sum_{2n}^e V), d^E(\sum_{2n} V, \sum_{2n}^e W)\} \\ &= \begin{cases} 0, & n < m; \\ \delta, & n \geq m. \end{cases} \end{aligned}$$

Thus

$$\begin{aligned} d(\mathcal{U}(V), \mathcal{U}(W)) &\leq \min \left\{ \min_{2 \leq n < m} (0 + 2\pi/(n - 1)), \inf_{n \geq m} (4\delta + 2\pi/(n - 1)) \right\} \\ &\leq \min\{2\pi/(m - 2), 4\delta\}. \end{aligned}$$

Direct observation shows that $\|W - V\| = \delta$, and so in fact we obtain

$$d(\mathcal{U}(V), \mathcal{U}(W)) \leq \min\{2\pi/(m - 2), \delta\}.$$

4.1.4. This example admits an interesting analysis. If, on the one hand, the minimum is δ , then what we are saying is that $1/m$ is relatively large compared to δ , i.e., m is small and so we do not have a long enough string of consecutive 1's to apply Berg's technique effectively. In other words, it “costs less” to simply change the weights. On the other hand, if the minimum is $2\pi/(m - 2)$, then we have a situation where $(1 - \delta)$ is relatively far from 1. This time the length

of the string of consecutive 1's is long enough to apply Berg's Technique as in Theorem 2.2.11 effectively without necessarily changing the weights at all.

4.1.5. As for calculating a lower bound for the distance $d(\mathcal{U}(V), \mathcal{U}(W))$, consider the estimate $d(\mathbf{w}_{(0, 2m-1)} = (1 - \delta, 1, \dots, 1, 1 - \delta), \sum_{2m} V)$. Now $\sum_{2m} V$ contains only blocks identical to

$$\begin{aligned} \mathbf{v}_{(2m, 4m-1)} &= (1 - \delta, 1, 1, \dots, 1) \\ \mathbf{v}_{(2m-1, 4m-2)} &= (1, 1 - \delta, 1, \dots, 1) \\ &\vdots \\ \mathbf{v}_{(1, 2m)} &= (1, 1, \dots, 1, 1 - \delta) \end{aligned}$$

while $\mathbf{w}_{(0, 2m-1)} = (1 - \delta, 1, \dots, 1, 1 - \delta)$. Here then is an example of when the $2m$ -spectrum of W is tractable enough to apply Theorem 3.1.9.

Borrowing the notation from that Theorem, if we let $A_{2m} = \{1, 2m\}$, then it is clear that

- (1) $B_{2m}(j) \cap A_{2m} \neq \emptyset$ for $j \in \mathbb{Z}$; and
- (2) there is no set A'_{2m} of cardinality 1 which satisfies condition (1).

Applying Theorem 3.1.9 to our case,

$$d(\mathcal{U}(V), \mathcal{U}(W)) \geq \delta_V(W) \geq C_V^{2m}(W) \min\{1/2m, \delta/\sqrt{2}\},$$

where

$$\begin{aligned} C_V^{2m}(W) &= \frac{\beta \|V^{-(2m-1)}\|^{-1}}{1 + \beta \|W^{-(2m-2)}\| \|W^{-1}\|} \\ &= \frac{(1 - \delta)(1 - \delta)}{1 + (1 - \delta)(1 - \delta)^{-1}(1 - \delta)^{-1}} = \frac{(1 - \delta)^3}{(2 - \delta)}. \end{aligned}$$

Thus,

$$d(\mathcal{U}(V), \mathcal{U}(W)) \geq \frac{(1 - \delta)^3}{(2 - \delta)} \min\{1/2m, \delta/\sqrt{2}\}.$$

4.1.6. Combining this with the upper bound,

$$\frac{(1 - \delta)^3}{(2 - \delta)} \min\{1/2m, \delta/\sqrt{2}\} \leq d(\mathcal{U}(V), \mathcal{U}(W)) \leq \min\{2\pi/(m - 2), \delta\}.$$

Modulo the (seemingly unavoidable) factor of $(1 - \delta)^3/(2 - \delta)$ arising from $C_V^{2m}(W)$, the upper and lower bounds are on the same order of magnitude. Note that Theorem 3.1.6 would have given us

$$d(\mathcal{U}(V), \mathcal{U}(W)) \geq (1 - \delta)^3/(2 - \delta) \cdot \min\{1/2m, \delta/\sqrt{2m}\},$$

a much less accurate estimate.

4.2.1. We now wish to examine an example where the value of $\delta_V(W_{(t, t+2n)})$ depends upon our ability to "average" the weights of V .

4.2.2. Example 2. We consider the following $(n + m)$ -tuples $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ ($m > n > 0$), each meant to represent an element both of $\sum_{(n+m)} W$ and of $\sum_{(n+m)} V$. The last m weights of \mathbf{b}_j , $1 \leq j \leq n$, are all $1/2$.

$$\begin{aligned} \mathbf{b}_1 &= (1 - \delta) & 1 & 1 & \cdots & 1 & \frac{1}{2} & \frac{1}{2} & \cdots & \frac{1}{2} \\ \mathbf{b}_2 &= & 1 & (1 - \delta) & 1 & \cdots & 1 & \frac{1}{2} & \frac{1}{2} & \cdots & \frac{1}{2} \\ & \vdots & & & & & & & & & \\ \mathbf{b}_n &= & 1 & 1 & \cdots & 1 & (1 - \delta) & \frac{1}{2} & \frac{1}{2} & \cdots & \frac{1}{2} \end{aligned}$$

Also consider the $(n + m)$ -tuple

$$\mathbf{b}_0 = 1 \quad 1 \quad 1 \quad \cdots \quad 1 \quad \frac{1}{2} \quad \frac{1}{2} \quad \cdots \quad \frac{1}{2},$$

which will be an element of $\sum_{(n+m)} W$, but not of $\sum_{(n+m)} V$. Again, the last m weights of \mathbf{b}_0 are $\frac{1}{2}$. As in the previous example, $\delta > 0$ and m is even.

In fact, we shall construct our shifts W (resp. V) from the blocks $\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_n$ (resp. $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$) in such a way that

- (1) $\mathbf{b}_j \in \sum_{(n+m)} W$ (resp. $\sum_{(n+m)} V$) for $0 \leq j \leq n$ (resp. $1 \leq j \leq n$), and
- (2) $\sum_{(n+m)} W = \sum_{(n+m)}^e W$ (resp. $\sum_{(n+m)} V = \sum_{(n+m)}^e V$).

We construct V , for example, in the following way.

There are n^t distinct strings of length t (i.e., of ordered t -tuples) of the blocks \mathbf{b}_j , $1 \leq j \leq n$, each, of course, corresponding to a string of length $t(n + m)$ of weights for V . Let L_t be a listing of the n^t strings of length t : $L_t = \{l_1^t, l_2^t, \dots, l_{n^t}^t\}$. By concatenating the strings l_r^t , $1 \leq r \leq n^t$, we obtain a new string $S_t = l_1^t \cdots l_{n^t}^t$ of weights for V of length $n^t \cdot t \cdot (n + m)$. We then let V be the weighted shift whose weight sequence is obtained by concatenating the S_t 's as follows:

$$\cdots S_5 \quad S_4 \quad S_3 \quad S_2 \quad S_1 \quad S_2 \quad S_3 \quad S_4 \quad S_5 \cdots$$

We leave it to the reader to verify that this works. A similar construction which this time includes the block \mathbf{b}_0 produces W .

The point of this example is to show that if our sequence of m consecutive weights of $1/2$ is long enough, say $m > 2\delta^{-1}\sqrt{n} + 4$, then although the Hausdorff distance $d_H(\sum_n^e W, \sum_n^e V) = \delta$, (for $(1, 1, \dots, 1) \in \sum_n^e W$ and $d((1, 1, \dots, 1), \sum_n^e V) = \delta$), by averaging the weights of V in an appropriate manner, we can obtain an estimate for $d(\mathcal{U}(V), \mathcal{U}(W))$ which is on the order of δ/\sqrt{n} .

4.2.3. We first consider $\delta_W^e(V_{(t, t+m-2)})$. (We are assuming, for the sake of convenience, that both m and n are even.) Let $\mathbf{x} \in \sum_m^e V = \sum_m V$. Since V is constructed from sequences of the \mathbf{b}_j 's ($1 \leq j \leq n$), and since any such sequence must also appear infinitely often as a weight sequence for W , $\delta_W^e(\mathbf{x}) = 0$. That is, $\delta_W^e(V_{(t, t+m-2)}) = 0$ for all $t \in \mathbb{Z}$. From Theorem 2.2.11,

$$\delta_W^e(V) \leq \pi/2((m-2)/2 - 1) = \pi/(m-4) \leq (\pi/2)(\delta/\sqrt{n}).$$

As for $y \in \sum_m W = \sum_m^e W$, either $y \in \sum_m^e V$ as above, in which case $\delta_V^e(y) = 0$, or y must contain a string of n consecutive 1's. In the latter case, y looks like:

$$y = \frac{1}{2} \quad \frac{1}{2} \quad \cdots \quad \frac{1}{2} \quad 1 \quad 1 \quad \cdots \quad 1 \quad \frac{1}{2} \quad \frac{1}{2} \quad \cdots \quad \frac{1}{2}$$

when the first r weights (for some $0 \leq r \leq m - n$) are $\frac{1}{2}$, the next n weights are 1, and the last $m - (r + n)$ are $\frac{1}{2}$.

We wish to show that $\delta_V^e(y)$ is on the order of δ/\sqrt{n} . To that purpose, choose $q(j)$ ($1 \leq j \leq n$) $\in \mathbb{Z}$ such that $c_i = v_{(q(i)+1, q(i)+m)}$, $1 \leq i \leq n$, yielding

$$\begin{aligned} c_1 &= \frac{1}{2} \quad \frac{1}{2} \quad \cdots \quad \frac{1}{2} \quad (1 - \delta) \quad 1 \quad 1 \quad \cdots \quad 1 \quad \frac{1}{2} \quad \frac{1}{2} \quad \cdots \quad \frac{1}{2} \\ c_2 &= \frac{1}{2} \quad \frac{1}{2} \quad \cdots \quad 1 \quad \frac{1}{2} \quad (1 - \delta) \quad 1 \quad \cdots \quad 1 \quad \frac{1}{2} \quad \frac{1}{2} \quad \cdots \quad \frac{1}{2} \\ &\vdots \\ c_n &= \frac{1}{2} \quad \frac{1}{2} \quad \cdots \quad \frac{1}{2} \quad 1 \quad 1 \quad \cdots \quad 1 \quad (1 - \delta) \quad \frac{1}{2} \quad \frac{1}{2} \quad \cdots \quad \frac{1}{2}. \end{aligned}$$

Again, in each case the first r weights are $\frac{1}{2}$. Our construction of V allows us to assume that the underlying spaces for the c_j 's, $1 \leq j \leq n$ are mutually orthogonal. That is, we assume

$$|q(j) - q(j')| > 2m, \quad 1 \leq j \neq j' \leq n.$$

Now consider the vectors

$$f_s = \sum_{j=1}^n (1/\sqrt{n}) e_{q(j)+s}, \quad 1 \leq s \leq (m+1).$$

Then $\|f_s\| = 1$, $1 \leq s \leq m+1$, and $\langle f_s, f_{s'} \rangle = 0$, $s \neq s'$ since $|q(j) - q(j')| > 2m$, $j \neq j'$. We wish to examine the action of V on the space $F = \text{span}_{1 \leq s \leq (m+1)} \{f_s\}$ relative to this basis. We obtain the following:

$$\begin{aligned} V f_s &= 1/2 f_{s+1}, & 1 \leq s \leq r, \\ V f_s &= 1/2 f_{s+1}, & (r+n+1) \leq s \leq m. \end{aligned}$$

Similarly

$$\begin{aligned} V^* f_{s+1} &= 1/2 f_s, & 1 \leq s \leq r, \\ V^* f_{s+1} &= 1/2 f_s, & (r+n+1) \leq s \leq m. \end{aligned}$$

As for $r+1 \leq s \leq r+n$, we see that if $s = r+p$, $1 \leq p \leq n$, then

$$\begin{aligned} V f_s &= V \left(\sum_{j=1, j \neq p}^n (1/\sqrt{n}) e_{q(j)+(r+p)} \right) + V((1/\sqrt{n}) e_{q(p)+(r+p)}) \\ &= \sum_{j=1, j \neq p}^n (1/\sqrt{n}) e_{q(j)+(r+p+1)} + (1 - \delta) (1/\sqrt{n}) e_{q(p)+(r+p+1)}. \end{aligned}$$

Thus

$$\|V f_s - f_{s+1}\| = \|(\delta/\sqrt{n}) e_{q(p)+(r+p+1)}\| = \delta/\sqrt{n}.$$

Also,

$$\|V^* f_{s+1} - f_s\| = \|(\delta/\sqrt{n}) e_{q(p)+(r+p)}\| = \delta/\sqrt{n}.$$

If we let $Q = \max_{1 \leq j \leq n} \{|q(j)| + 2m\}$, then we can make the further observation that $VF \subseteq \text{span}\{e_i\}_{i=-Q}^Q$, a “finite- e_i -space”. Now $y = (w_t, w_{t+1}, \dots, w_{t+(m-1)})$ for some $t \in \mathbb{Z}$, and if we let U be a unitary which takes $e_i \rightarrow f_{i-t+1}$, $t \leq i \leq t+m$, then we have shown that

$$\|(VU - UW)P_{(t+1, t+m)}\| \leq \delta/\sqrt{n},$$

and

$$\|(V^*U - UW^*)P_{(t+1, t+m)}\| \leq \delta/\sqrt{n}.$$

Since $\sum_m V = \sum_m^e V$, we conclude that $\delta_V^e(W_{(t+1, t+m)}) \leq \delta/\sqrt{n}$ for all $t \in \mathbb{Z}$. Appealing once more to Theorem 2.2.11, we conclude that

$$\delta_V^e(W) \leq \delta/\sqrt{n} + \pi/2((m-2)/2 - 1) \leq (2 + \pi)\delta/2\sqrt{n}.$$

Then by Proposition 2.2.4,

$$\begin{aligned} d(\mathcal{U}(V), \mathcal{U}(W)) &\leq 2(\pi\delta/2\sqrt{n} + (2 + \pi)\delta/2\sqrt{n}) \\ &\leq (2 + 2\pi)\delta/\sqrt{n}. \end{aligned}$$

An easy, but nonetheless significant observation, is that

$$\sum_{n-1} W = \sum_{n-1}^e W = \sum_{n-1}^e V = \sum_{n-1} V,$$

so that

$$\delta_V^e(W_{(t, t+(n-3))}) = 0 = \delta_W^e(V_{(t, t+(n-3))})$$

for all $t \in \mathbb{Z}$. From Theorem 2.2.11,

$$\delta_V^e(W) \leq \pi/2((n-2)/2 - 1) = \pi/(n-4), \quad \delta_W^e(V) \leq \pi/(n-4).$$

Again, from Proposition 2.2.4,

$$d(\mathcal{U}(V), \mathcal{U}(W)) \leq 2(\pi/(n-4) + \pi/(n-4)) = 4\pi/(n-4).$$

In conclusion then,

$$d(\mathcal{U}(V), \mathcal{U}(W)) \leq \min \{(2 + 2\pi)\delta/\sqrt{n}, 4\pi/(n-4)\}.$$

4.2.4. To calculate a lower bound for this example, we simply appeal to Theorem 3.1.6. In our case, $d(\sum_n^e W, \sum_n^e V) = \delta$. Thus

$$d(\mathcal{U}(V), \mathcal{U}(W)) \geq C_V^n(W) \min \{1/n, \delta/\sqrt{n}\}$$

where

$$\begin{aligned} C_V^n(W) &= \frac{\beta \|V^{-(n-1)}\|^{-1}}{1 + \beta \|W^{-(n-2)}\| \|W^{-1}\|} \\ &= \frac{(1/2)(1/2)^{n-1}}{1 + (1/2)(2^{n-2})(2)} = \frac{1}{(2^n)(1 + 2^{n-2})}. \end{aligned}$$

That is,

$$d(\mathcal{U}(V), \mathcal{U}(W)) \geq \frac{1}{(2^n)(1 + 2^{n-2})} \min \{1/n, \delta/\sqrt{n}\}.$$

Again, we see that modulo the factor of $C_V^n(W)$, the lower bound and upper bounds are on the same order.

5. THE NONESSENTIAL n -SPECTRUM

5.1.1. Proposition 2.2.4 provides us with an estimate for $d(\mathcal{U}(A), \mathcal{U}(B))$ which depends solely upon $\delta_A^e(B)$ and $\delta_B^e(A)$. It is possible, of course, to have two weighted shifts W and V for which $d(\mathcal{U}(V), \mathcal{U}(W))$ is small, and yet $\delta_V^e(W)$ is large. An easy example is found by considering $W = V$ where $Ve_i = 1/2e_{i+1}$, $i \neq 0$ and $Ve_0 = e_1$. Clearly $\delta_V^e(V) = 1/2$ whereas $d(\mathcal{U}(V), \mathcal{U}(V)) = 0$.

In [Dav1, §5] Davidson produces an example of operators A and B for which $\delta_A(B)$ and $\delta_B(A)$ are small, and yet $d(\mathcal{U}(A), \mathcal{U}(B))$ is large. His example can easily be converted to an example of invertible weighted shifts satisfying this property.

5.1.2. To be precise, we are considering the adjoint of the example given in the aforementioned paper. Similar to the case there, we let $J = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \in M_2(\mathbb{C})$ and $n > 1$ be a positive integer. Let $F_n = \bigoplus_{k=1}^{n-1} (k/(n-1))J$ and $G_n = 0_1 \oplus F_n$ in $M_{2n-2}(\mathbb{C})$ and $M_{2n-1}(\mathbb{C})$ respectively. Let $A_n = J^{(\infty)} \oplus F_n \oplus J^{(\infty)}$ and $B_n = J^{(\infty)} \oplus G_n \oplus J^{(\infty)}$. It is shown there that for all $n > 1$,

- (1) $\delta_{B_n}(A_n) = 0$;
- (2) $\delta_{A_n}(B_n) \leq 1/(n-1)$; and
- (3) $d(\mathcal{U}(A_n), \mathcal{U}(B_n)) \geq 1/\sqrt{5}$.

Choose $n > 101$ so that $0 < 1/(n-1) < 1/100$. Now A_n can be thought of as a noninvertible weighted shift, every second weight of which is zero. Similarly, B_n is a weighted shift with zero weights at every second position, except on the space where $J \oplus G_n$ acts, where two consecutive weights are zero. Consider the weighted shifts W_n and V_n derived from A_n and B_n by perturbing the zero weights to $1/(n-1)$. Then W_n and V_n have the weight sequence—for W_n :

$$\begin{array}{cccccccc} \cdots & w_{-3} & w_{-2} & w_{-1} & w_0 & & & \\ & \frac{1}{(n-1)} & 1 & \frac{1}{(n-1)} & 1 & & & \\ & & w_1 & w_2 & w_3 & w_4 & w_5 & \cdots \\ & & \frac{1}{(n-1)} & \frac{1}{(n-1)} & \frac{1}{(n-1)} & \frac{2}{(n-1)} & \frac{1}{(n-1)} & \\ & \cdots & w_{2n-4} & w_{2n-3} & w_{2n-2} & w_{2n-1} & w_{2n} & \cdots \\ & & \frac{n-2}{(n-1)} & \frac{1}{(n-1)} & 1 & \frac{1}{(n-1)} & 1 & \end{array}$$

and for V_n :

$$\begin{array}{cccccccc}
 \cdots & v_{-3} & v_{-2} & v_{-1} & v_0 & & & \\
 & \frac{1}{(n-1)} & 1 & \frac{1}{(n-1)} & 1 & & & \\
 & & v_1 & v_2 & v_3 & v_4 & v_5 & v_6 \quad \cdots \\
 & & \frac{1}{(n-1)} & \frac{1}{(n-1)} & \frac{1}{(n-1)} & \frac{1}{(n-1)} & \frac{2}{(n-1)} & \frac{1}{(n-1)} \quad \cdots \\
 & \cdots & v_{2n-3} & v_{2n-2} & v_{2n-1} & v_{2n} & v_{2n+1} & \cdots \\
 & & \frac{n-2}{(n-1)} & \frac{1}{(n-1)} & 1 & \frac{1}{(n-1)} & 1 & \cdots
 \end{array}$$

If we subtract from W_n a shift X whose weight sequence $\{x_i\}_{i \in \mathbb{Z}}$ is

$$\begin{aligned}
 x_i &= 1/(n-1), & i \text{ odd}; \\
 x_i &= 0, & i \text{ even};
 \end{aligned}$$

then we find that $W_n - X \cong A_n$. Clearly $\|X\| = 1/(n-1)$. Similarly, $d(V_n, \mathcal{U}(B_n)) \leq 1/(n-1)$. Thus

$$\begin{aligned}
 d(\mathcal{U}(W_n), \mathcal{U}(V_n)) \\
 &\geq d(\mathcal{U}(A_n), \mathcal{U}(B_n)) - d(\mathcal{U}(W_n), \mathcal{U}(A_n)) - d(\mathcal{U}(V_n), \mathcal{U}(B_n)) \\
 &\geq 1/\sqrt{5} - 2/(n-1).
 \end{aligned}$$

On the other hand,

$$\begin{aligned}
 W_n &\cong_{1/(n-1)} A_n \\
 &\cong_{1/(n-1)} B_n \oplus J \quad \text{for some } J \in \mathcal{B}(\mathcal{H}) \\
 &\cong_{1/(n-1)} V_n \oplus J
 \end{aligned}$$

and so $\delta_{W_n}(V_n) \leq 3/(n-1)$. Similarly, $\delta_{V_n}(W_n) \leq 2/(n-1)$.

5.1.3. This demonstrates that an operator-theoretic analogue of the Schroeder-Bernstein Theorem for embeddings of operators into each other fails even for weighted shifts. The question of determining optimal estimates for $d(\mathcal{U}(V), \mathcal{U}(W))$ in the case $\sum_n W \neq \sum_n^e W$ and $\sum_n V \neq \sum_n^e V$ is therefore seen to be quite subtle.

5.1.4. We conclude by mentioning that results for unilateral shifts have also been obtained. In the case where the weights $\{w_i\}_{i=1}^\infty$ and $\{v_i\}_{i=1}^\infty$ are all bounded below, one must also take care to approximately match the kernels of W and V , and hence any intertwining unitary must send the basis vector e_1 (acted upon by w_1) approximately to f_1 (acted upon by v_1).

We state these results without proof. The interested reader is referenced to [Mar]. But first we define some notation.

5.1.5. Notation. Given a unilateral weighted shift W which we may assume possesses nonnegative weights $\{w_i\}_{i \geq 1}$, we let $\mathbf{w}_{(1,n)} = (w_1, \dots, w_n)$.

We also let $\sum_n W = \{\mathbf{w}_{(i, i+(n-1))} : i \geq 1\}^{-\|\cdot\|}$ and $\sum_n^e W$ consists of the set of accumulation points of $\sum_n W$.

Finally, we set $W_{\inf}^n = \inf_{i \geq 1} \{w_i \cdots w_{i+(n-1)}\}$, in analogy to $\|V^{-n}\|^{-1}$ for invertible bilateral weighted shifts V .

Armed with these definitions and notations, we obtain the following results:

5.1.6. Proposition. Let W and V be unilateral weighted shifts and set $\beta = (V_{\inf}^1 + W_{\inf}^1)/2$. For each $n \geq 2$, let $\delta_n = d_H(\sum_n V, \sum_n W)$ and let

$$D_V^n(W) = \frac{\beta(V_{\inf}^{n-1})}{2\sqrt{2}(1 + \beta(W_{\inf}^{n-2})^{-1}(W_{\inf})^{-1})}.$$

Then

$$\begin{aligned} d(\mathcal{U}(V), \mathcal{U}(W)) &\geq \max \left\{ \sup_{n \geq 2} (\max\{D_V^n(W), D_W^n(V)\}) (\min\{1/(n-1), \|\mathbf{w}_{(1,n)} - \mathbf{v}_{(1,n)}\|\}) \right. \\ &\quad \left. \sup_{n \geq 2} \left(\sqrt{2} \cdot (\min\{D_V^n(W), D_W^n(V)\}) (\min\{1/n, \delta_n/\sqrt{n}\}) \right) \right\}. \end{aligned}$$

5.1.7. Theorem. Let V and W be unilateral weighted shifts such that for all $n \geq 1$, $\sum_n V = \sum_n^e V$ and $\sum_n W = \sum_n^e W$. Let

$$\delta_n = \max\{d^E(\sum_{2n} V, \sum_{2n} W), d^E(\sum_{2n} W, \sum_{2n} V)\}.$$

Then

$$d(\mathcal{U}(V), \mathcal{U}(W)) \leq \inf_{n \geq 8} \{17\pi/2(n-8) + 3 \max\{\delta_n, \|\mathbf{w}_{(1,n)} - \mathbf{v}_{(1,n)}\|\}\}.$$

ACKNOWLEDGMENT

This paper was part of the author's Ph.D. thesis. I would like to thank Kenneth R. Davidson for his many insightful and helpful comments. His encouragement has been greatly appreciated.

REFERENCES

- [AHHK] W. B. Arveson, D. W. Hadwin, T. B. Hoover, and E. E. Kymala, *Circular operators*, Indiana Univ. Math. J. **33** (1984), 583–595.
- [AD] E. Azoff and C. Davis, *On the distance between unitary orbits of self-adjoint operators*, Acta Sci. Math. (Szeged) **47** (1984), 419–439.
- [Arv] W. B. Arveson, *Notes on extensions of C^* -algebras*, Duke Math. J. **44** (1977), 329–355.
- [Brg1] I. D. Berg, *On approximation of normal operators by weighted shifts*, Michigan Math. J. **21** (1974), 377–383.
- [Brg2] —, *Index theory for perturbations of direct sums of normal operators and weighted shifts*, Canad. J. Math. **30** (1978), 1152–1165.

- [BDM] R. Bhatia, C. Davis, and A. McIntosh, *Perturbation of spectral subspaces and solution of linear operator equations*, Linear Algebra Appl. **52:53** (1983), 45–67.
- [BDF] L. G. Brown, R. G. Douglas, and P. A. Fillmore, *Unitary equivalence modulo the compact operators and extensions of C^* -algebras*, (Proc. Conf. Operator Theory, Halifax, Nova Scotia, 1973), Lecture Notes in Math., vol. 345, Springer-Verlag, Berlin, Heidelberg and New York, 1973, pp. 58–128.
- [Dav1] K. R. Davidson, *Estimating the distance between unitary orbits*, J. Operator Theory (to appear).
- [Dav2] —, *The distance between unitary orbits of normal operators*, Acta Sci. Math. (Szeged) **50** (1986), 213–223.
- [Dav3] —, *The distance between unitary orbits of normal elements in the Calkin algebra*, Proc. Roy. Soc. Edinburgh **99A** (1984), 35–43.
- [Dav4] —, *Berg's technique and irrational rotation algebras*, Proc. Roy. Irish Acad. **84A** (1984), 117–123.
- [Dve] A. M. Davie, *Classification of essentially normal operators*, Lecture Notes in Math., vol. 512, Springer-Verlag, 1975, pp. 29–55.
- [Gel] R. Gellar, *Circular symmetric operators and subnormal operators*, J. Analyse Math. **32** (1977), 93–117.
- [GP] R. Gellar and L. Page, *Limits of unitarily equivalent normal operators*, Duke Math. J. **41** (1974), 319–322.
- [Had] D. W. Hadwin, *An operator-valued spectrum*, Indiana Univ. Math. J. (2) **26** (1977), 329–340.
- [HH] D. W. Hadwin and T. B. Hoover, *Weighted translation and weighted shift operators*, Lecture Notes in Math., vol. 693, Springer-Verlag, 1978, pp. 93–99.
- [Hal1] P. R. Halmos, *Limits of shifts*, Acta Sci. Math. (Szeged) **34** (1973), 131–139.
- [Hal2] —, *A Hilbert space problem book*, Graduate Texts in Math., vol. 19, Springer-Verlag, 1974.
- [Her1] D. A. Herrero, *On quasidiagonal weighted shifts and approximation of operators*, Indiana Univ. Math. J. (4) **33** (1984), 549–571.
- [Her2] —, *Approximation of Hilbert space operators*, Research Notes in Math. 72, Pitman, 1982.
- [Lam] A. Lambert, *Unitary equivalence and reducibility of invertibly weighted shifts*, Bull. Austral. Math. Soc. **5** (1971).
- [Mar] L. Marcoux, Ph. D. Thesis, Univ. of Waterloo, 1988.
- [McP] A. McIntosh and A. Pryde, *The solution of systems of operator equations using Clifford algebras*, Proc. Centre Math. Anal. Austral. Nat. Univ., Canberra, vol. 9, 1985, pp. 212–222.
- [MM] B. B. Morrell and P. S. Muhly, *Centered operators*, Studia Math. **51** (1974), 251–263.
- [O'Dn] D. O'Donovan, *Weighted shifts and covariance algebras*, Trans. Amer. Math. Soc. **208** (1975), 1–25.
- [Shd] A. L. Shields, *Weighted shift operators and analytic function theory*, Math. Surveys, no. 13, Amer. Math. Soc., Providence, R. I., 1974, pp. 49–128.
- [Vcu] D. Voiculescu's, *A non-commutative Weyl-von Neumann theorem*, Rev. Roumaine Math. Pures Appl. **21** (1976), 97–113.
- [Vgt] J. Voigt, *Perturbation theory for commutative m -tuples of self-adjoint operators*, J. Funct. Anal. **25** (1977), 317–334.