Bounds on the Rate of Convergence of the Kacmarz Method in Computed Tomography

by Matthew Boylan
University of Wisconsin
& Peter Caldwell
Western Washington University

Advisor: Professor Don Solmon
Oregon State University
Mathematics REU 1998

Abstract:A conjecture in "The Angles between the Null Spaces of X-rays" by Hamaker and Solmon [HS] states that it is always possible to order N equally spaced projections so that reconstruction using the Kacmarz procedure yields an approximating function that is within 1% relative error within eight iterations. In order to examine this conjecture, we investigate other bounds on the rate of convergence of the procedure. We research bounds given by Deutsch and Hundal [DH], and by Kayalar and Weinert [KW] in addition to developing estimating methods of our own. While this investigation is ongoing, it appears that the bound given by [KW] yields significantly better bounds than those used in [HS].

*__Acknowledgment:__ We would like to thank Dr Solmon for his tremendous support and patience in coaching us through these last eight weeks.

1. Introduction

The Kacmarz method is an iterative scheme for solving systems of linear equations. This method is based on orthogonally projecting a vector onto a series of planes and iterating. As the number of iterations becomes large, the iterated projection of the initial vector converges to the intersection of the planes. (See Halperin [H]). Because the convergence of this process is dependent only on the angles between the planes, we are, in considering the rate of convergence of the method, able to make the simplifying assumption that these planes are subspaces. This allows us to work within the special structure of linear operators.

The focus of this paper is to provide bounds for the rate of convergence of the Kacmarz method in reconstructing a function from a sample set of X-rays. We accomplish this both by developing our own methods of bounding the convergence operator and by applying published bounds on the rate of convergence of the method to the case of tomography. This second technique is made possible by noting the equivalence (in the case of tomography) between the rate of convergence of the Kacmarz method and that of the Method of Alternating Projections (MAP) and the Algebraic Reconstruction Technique (ART).

1.1. Background

1.1.1. The Convergence Operator

Let $L^2(D)$ be the set of all real valued square integrable functions which vanish outside of the unit disk in R^2, and let $L_1, L_2, ... L_N$ be hyperplanes in $L^2(D)$ with intersection I. Denote the projection operator onto L_j by Q_j for $j = 1, 2, ..., N$. If we let the operator of the Kacmarz procedure be denoted $Q = Q_N Q_{N-1} ... Q_1$ and the projection onto the intersection of the planes be Q_I, the norm of the operator expressing the rate of convergence of the technique is given by

$$\|Qg - Q_I g\|,$$

where g is the "initial guess" vector. This operator can be stated more succinctly by orthogonally decomposing g. To that end, let $I = \cap_{j=0}^{N-1} L_j$, and let $g = u + v$, where $u \in I$ and $v \in I^{\perp}$. Then, noting that $Qu = u$, $Q_I u = u$, and that $Q_I v = 0$, the convergence operator can be rewritten as

$$\begin{aligned} \|Q(u+v) - Q_I(u+v)\| &= \|Qu + Qv - Q_I v - u\| \\ &= \|Qv\|. \end{aligned}$$

Thus the norm of the convergence operator for the Kacmarz method is

$$\|Q - Q|_I\| = \|Q|_{I^{\perp}}\|.$$

1.1.2. Projection Operators

In providing bounds on the above operator, we frequently make use of the following basic identities about orthogonal projectional operators. If we let P be an arbitrary orthogonal projection operator, we have that:

1. Each operator P_i is idempotent.
2. Each projection operator is self-adjoint. In other words,

$$\langle P_i u, v\rangle = \langle u, P_i v\rangle .$$

3. The adjoint of P, denoted P^*, is $P^* = P_0 P_1 \dots P_{N-1}$.
4. P^*P is a self-adjoint operator which we denote $Q' = P^*P$.
5. The norm of an operator is given by

$$\|P\| = \sup_{\|x\|=\|y\|=1} \langle Px, y\rangle .$$

   Additionally, it suffices to take y in the range of P, for if y is orthogonal to P, then

$$\langle Px, y\rangle = 0.$$

   Thus

$$\|Q\| = \sup_{\|x\|=1} \|Qx\| .$$

6. The 2-norm of an operator is given by

$$\|Q\|^2 = \sup_{\|x\|=1} \|Qx\|^2 .$$

7. Additionally

$$\|Q^*Q\| = \|Q\|^2$$

   and

$$\|Q\| = \|Q^*\| .$$

1.1.3. Description of $I^\perp$

In order to calculate the norm of the convergence operator, we must first have a clear understanding of the space in which we are working. As mentioned in the introduction, we make the simplifying assumption that all planes are subspaces. Put another way, we assume all X-rays are zero in order to take advantage of the special structure this simplification affords. To that end we define $I^\perp = \left(\cap_{j=0}^{k} N_j\right)^\perp$, where N_j is the set of all functions in $L^2(D)$ with zero X-ray in the direction making angle θ_j with the positive x axis (in contrast to the more general definition in section 1.1.1). By making this simplification, we can make use of the following lemma.

Lemma 1.1. *Let $\theta_1, \theta_2, ..., \theta_N$ be distinct angles mod π; then $I^{\perp} = \sum_{j=0}^{k} N_j^{\perp}$ [HS].*

Thus the problem reduces to classifying $N_j^{\perp}$ for arbitrary j. This is also accomplished through a lemma.

Lemma 1.2. *The null space N_j is a closed subspace of $L^2(D)$ whose orthogonal complement $N_\theta^{\perp}$ consists of the set of all functions constant in $L^2(D)$ on lines making angle $\theta + \frac{\pi}{2}$ (measured counterclockwise) with the x-axis [HS].*

In order to choose a suitable basis for this space, the following definition is invaluable.

Definition 1.3. *The inner product between two functions u and v in $L^2(D)$ is given by*

$$\langle u, v \rangle = \int_{-1}^{1} \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} u(x, y)\, v(x, y)\, dx\ dy.$$

Lemma 1.4. *$N_j^{\perp}$ can be represented by an infinite series of Chebychev polynomials of the second kind subject to the restriction that these functions vanish outside the unit disk,*

$$N_j^{\perp} = \oplus_{m=0}^{\infty} \left[U_m (x \cos\theta + y \sin\theta) \right],$$

where

$$U_m(\cos\theta) = \frac{\sin\left[(m+1)\theta\right]}{\sqrt{\pi}\,(m+1)\sin\theta}.$$

In addition, this basis is orthonormal on $L^2(D)$.

Proof: Let $U_a(x)$ and $U_b(x)$ be two Chebychev polynomials of the second kind defined along the same line. By a rotation of coordinates, we take $\theta_j = 0$. The inner product between the two is defined below:

$$\langle U_a(x), U_b(x) \rangle = \int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} U_a(x)\, U_b(x)\, dy\ dx.$$

However, because $U_a(x)$ and $U_b(x)$ are functions of x alone, we have

$$\langle U_a(x), U_b(x)\rangle = \int_{-1}^{1}\left(2\sqrt{1-x^2}\right) U_a(x)\, U_b(x)\, dx.$$

Making the change of variables $x = \cos\theta$ and rewriting the Chebychevs in terms of their trigonometric identities,

$$\begin{aligned}
\langle U_a(x), U_b(x)\rangle &= 2\int_{\arccos -1}^{\arccos 1} \sqrt{1-\cos^2\theta}\,\frac{\sin[(a+1)\theta]}{\sqrt{\pi}\sin\theta}\,\frac{\sin[(b+1)\theta]}{\sqrt{\pi}\sin\theta}(-\sin\theta d\theta) \\
&= -\frac{2}{\pi}\int_{\pi}^{0}\frac{\sin[(a+1)\theta]\sin[(b+1)\theta]\sin^2\theta}{\sin^2\theta}d\theta \\
&= \frac{2}{\pi}\int_{0}^{\pi}\sin[(a+1)\theta]\sin[(b+1)\theta]\,d\theta.
\end{aligned}$$

This is the coefficient equation for the Fourier sine series over the interval $[0,\pi]$. It is well-known that this function forms an orthonormal basis for the set of functions in [0,1]. Therefore the set of Chebychev polynomials form such a basis for the unit disk.

Henceforth, we will denote $U_m(x\cos\theta_i + y\sin\theta_i)$, where θ_i is the angle in which a given Radon transform is taken, as $U_{m,i}$.

This last result allows us to classify $I^{\perp}$. If $y \in I^{\perp}$, then

$$y = \sum_{j=0}^{N-1}\sum_{m=0}^{\infty} b_{m,i} U_{m,j}$$

for some set of weights $b_{m,i}$. We can further orthogonally decompose $I^{\perp}$, however, by examining the effect of an arbitrary projection operator on $U_{m,i}$.

Lemma 1.5. *For all* $m > 0$,

$$Q_j U_{m,i} = U_{m,i} - \langle U_{m,i}, U_{m,j}\rangle\, U_{m,j}.$$

Proof: Orthogonally decompose $U_{m,i}$ into

$$U_{m,i} = \left(U_{m,i} - \langle U_{m,i}, U_{m,j}\rangle\, U_{m,j}\right) + \left(\langle U_{m,i}, U_{m,j}\rangle\, U_{m,j}\right).$$

Because $\langle U_{m,i}, U_{m,j} \rangle U_{m,j}$ is in $N_j^{\perp}$, it is clear that the first summand, being orthogonal to the second, is in N_j.

This reveals that the set $[U_{m,0}, U_{m,1}, ..., U_{m,N-1}]$ is invariant under operation by any of the projection operators for any m. Additionally, it can be shown that each set of Chebychev polynomials of fixed degree in each of the directions X-rays were taken is orthogonal to other similar sets but of differing degree. This leads us to rewrite $I^{\perp}$ as

$$I^{\perp} = \oplus_{m=0}^{\infty} I_{m,n},$$

where

$$I_{m,n} = [U_{m,0}, U_{m,1}, ..., U_{m,N-1}] \, .$$

Thus

$$\|Q|_{I^{\perp}}\| = \sup_{m \geq 0} \left\|Q|_{I_{m,N}}\right\| .$$

To conclude the classification of $I^{\perp}$, we reduce our spanning set for each component in the decomposition to a basis. In order to accomplish this, we notice that the dimension of the set $[U_{m,0}, U_{m,1}, ..., U_{m,N-1}]$ is the minimum of $m+1$ and N [HS]. This breaks the set $I_{m,N}$ into two parts, each of which we consider seperately. The first is the *lower dimensional subspaces*, where $m+1 < N$, and the second, the *upper dimensional subspaces*, consists of the set of $I_{m,N}$ which have dimension N. This classification allows us to compute the angles between these subspaces.

1.1.4. Angles between Subspaces

Now that we have described $I^{\perp}$ in terms of invariant finite dimensional subspaces with known bases, we may make use of known results to calculate the angle between one subspace of $I_{m,N}$ with basis $[U_{m,0}]$ and the intersection of other subspaces of $I_{m,N}$, represented by the basis $[U_{m,1}, ..., U_{m,k}]$ for $k \in [1, 2, ..., N-1]$.

Since $U_{m,0}$ is one dimensional and of unit length, we can use the properties of right triangles to write

$$\sin^2 \left([U_{m,0}], [U_{m,1}, \cdots, U_{m,k}] \right) = \delta^2 \left(U_{m,0}, [U_{m,1}, \cdots, U_{m,k}] \right),$$

where $\delta\left(a, b\right)$ denotes the distance between a and b. Because the subspace spanned by $[U_{m,1}, \cdots, U_{m,k}]$ is finite dimensional, we can use the following definition to calculate the angle between our subspaces.

Definition 1.6. *The squared distance between a point p and a subspace with basis $[P_1, \cdots, P_N]$ in Hilbert space is given by*

$$\delta^2(p, P) = \frac{G(p, P_{1,} \cdots, P_N)}{G(P_1, \cdots, P_N)},$$

where G denotes the determinant of the Gram matrix with entries such that

$$G(P_1, \cdots, P_N) = \det \begin{bmatrix} \langle P_1, P_1\rangle & \langle P_1, P_1\rangle & \cdots & \langle P_1, P_1\rangle \\ \langle P_2, P_1\rangle & \langle P_2, P_1\rangle & \cdots & \langle P_N, P_1\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle P_N, P_1\rangle & \langle P_N, P_1\rangle & \cdots & \langle P_N, P_1\rangle \end{bmatrix}.$$

The following formula for computing the inner product of $U_{m,i}$ and $U_{m,j}$ allows us to compute this determinant [HS]:

$$\langle U_{m,i}, U_{m,j}\rangle = \frac{\sin[(m+1)(\theta_i - \theta_j)]}{(m+1)\sin(\theta_i, \theta_j)}.$$

With this formula we can finally compute the angles between subspaces generated by X-ray data in given application. This technique is summarized in the following theorem.

Theorem 1.7. *Let $L_0, L_1, ..., L_k$ be subspaces of $I_{m,N}$ for some m and some set of distinct angles $\theta_{1,}, \theta_2, ..., \theta_k$. Then the angle α between L_0 and $\cap_{j=1}^{k} L_j$ for $1 \le k \le N-1$ is given by*

$$\sin^2\alpha = \left(\frac{G(U_{m,0}, U_{m,1}, \ldots, U_{m,k})}{G(U_{m,1}, \ldots, U_{m,k})}\right).$$

where the entries in the Gram matrix are as above.

When the X-rays are equally spaced, the Gram determinant formula simplifies.

Definition 1.8. *A set of angles are said to be equally spaced provided that if $j = 0, 1, ..., N-1$, then $\theta_j = \dfrac{j\pi}{N}$.*

In this case, we have

$$\frac{G(U_{m,0}, U_{m,1}, ..., U_{m,N-1})}{G(U_{m,1}, U_{m,2}, ..., U_{m,N-1})} = \frac{Nu}{m+1} \frac{Uu + N}{Nu + N - v}$$

where $u = \left\lfloor \dfrac{m+1}{N} \right\rfloor \geq 1$ and $v = (m+1) \bmod N$ [HS].

This final formula completes the background needed to calculate angles between subspaces for our specific case.

The results are summarized in the following theorem.

1.1.5. The Angles Theorem – Our Building Block

The following theorem provided us with a basis from which to look for improved bounds.

Theorem 1.9. *Let $Q_1, Q_2, \ldots, Q_{N-1}$ be orthogonal projections onto the closed planes $L_1, L_2, \ldots, L_{N-1}$ of a Hilbert space H with nonempty intersection I. Let Q^n be Q iterated n times. Additionally, define Q_I to be the projection onto the intersection I and $g \in L^2(D)$. Then*

$$\|Q^n g - Q_L g\|^2 \leq c^2 \|g - Q_I g\|^2$$

where

$$c^2 \leq 1 - \prod_{j=1}^{N-1} \sin^2\left(L_j, \cap_{i=j+1}^{N} L_i\right).$$

Note that $\|Q^n g - Q_L g\|^2$ was previously shown to reduce to $\sup_{m\geq 0} \left\| Q^n|_{I_{m,N}} \right\|^2$. Using the above theorem to rewrite (2) and performing the obvious cancellations, the angles bound becomes

$$\begin{aligned} c^2 &\leq 1 - \prod_{j=1}^{N-1} \sin^2\left(L_j, \cap_{i=j+1}^{N} L_i\right) \\ &= 1 - G\left(U_{m,0}, U_{m,1}, \ldots, U_{m,k}\right) \end{aligned}$$

where $k = \min(m+1, N-1)$ as previously mentioned. This bound is clearly easy to describe and to compute. Unfortunately, this estimate is not accurate enough to provide much real information about $\|Q|_{I^\perp}\|$. In particular, when $m \geq N-1$, this estimate is completely independent of the order in which the subspaces are considered. This makes it impossible to use this estimate to evaluate the efficiency of different ordering schemes for X-ray data. In the hopes of finding more useful bounds, we examined the following techniques.

2. Bounding Methods

2.1. Lower Bounds

As a means of deriving a lower bound for $\|Q\|$ on $I_{m,N}$, we looked at the restriction of Q to $\beta = \cap_{j=0}^{N-2} L_j \cap I_{m,n}$. Taking the norm of Q subject to additional restrictions yields lower bounds because the operator may be maximized on a vector in the larger space but not in the restriction. This bound is derived in the following manner.

Let $g \in \beta$. Then

$$\|Q_{N-1}Q_{N-2}...Q_0 g\| = \|Q_{N-1}g\|$$

From identity #5 in section 1.2 we have

$$\|Q_{N-1}g\| = \sup_{\|u\|=\|v\|=1} \langle Q_N u, v \rangle .$$

with v in the range of Q_{N-1}. Now, because Q_{N-1} is self-adjoint, our equation becomes

$$\begin{aligned} \|Q_{N-1}g\| &= \sup_{\|u\|=\|v\|=1} \langle u, Q_{N-1}v \rangle \\ &= \sup_{\|u\|=\|v\|=1} \langle u, v \rangle . \end{aligned}$$

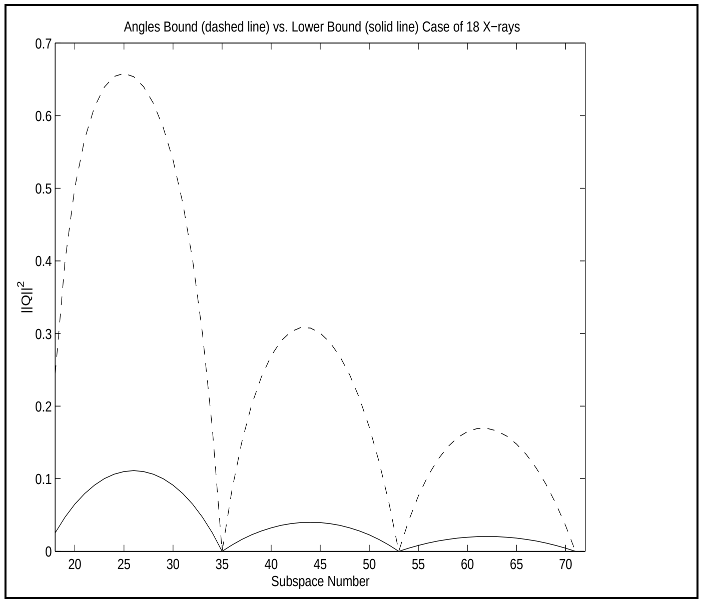
But $u \in \cap_{j=0}^{N-2} L_j \cap I_{m,n}$, $v \in L_{N-1} \cap I_{m,n}$ so the intersection of the spaces containing each vector is $\cap_{j=0}^{N-1} L_j \cap I_{m,n} = 0$. Thus all vectors are perpendicular to the intersection of the spaces containing u and v, so we can describe $\|Q|_\beta\|$ in terms of a cosine,

$$\|Q|_\beta\| = \cos\left(L_{N-1}, {}_{j=0}^{N-2} L_j\right) .$$

Using the machinery for computing this angle developed above,

$$\left\|Q|_{I_{m,N}}\right\| \leq \|Q|_\beta\| = 1 - \frac{G\left(U_{m,0}, ..., U_{m,N-1}\right)}{G\left(U_{m,1}, ..., U_{m,N-1}\right)}.$$

Computation of this bound, however, reveals that $\|Q|_\beta\|$ quickly approaches 0 as m increases. Therefore this technique is not particularly useful in investigating the norms of the invariant subspaces. The following graph is a comparison between the angles bound and this lower bound for 18 evenly spaced X-rays.



As evident, the area between the angles bound and the lower bound is quite large- the lower bound is not helpful in trapping the actual norm.

2.2. Upper Bounds

Our main goal is to obtain sharper upper bounds for $\|Q^n\|$ by studying $\|Q^n\|_{m,N}$. Since $\dim I_{m,N} = \min(m+1, N)$, the angles bound has the following form [HS]:

$$For\ m+1 \geq N,\ \|Q^n\|_{m,N}^2 = \|(Q^*)^n Q\|_{m,N} \leq (1 - G(U_{m,0}, ..., U_{m,N-1}))^n,$$
$$and\ for\ m+1 < N,\ \|Q^n\|_{m,N}^2 = \|(Q^*)^n Q^n\|_{m,N} \leq (1 - G(U_{m,0}, ..., U_{m,m+1}))^n$$

So we consider the cases $m+1 \geq N$ and $m+1 < N$ separately. The angles bound is derived by dividing Q^n into n equal pieces, $Q_{N-1}...Q_0$, and bounding each piece individually:

$$\|Q^n\| = \|(Q_{N-1}...Q_0)_n \cdots (Q_{N-1}...Q_0)_1\| \leq \|Q_{N-1}...Q_0\|^n$$

Since the angles bound uses only $\|Q_{N-1}...Q_0\|$ to bound $\|Q^n\|$, it does not account for the relationship between operators in different iterations. For example the angles bound does not account for Q_0 following $Q_{N-1}...Q_0$ in each iteration after the first, nor does it account for the Q_1 in the second iteration following $Q_0 Q_{N-1}...Q_0$,i.e, every operator preceding the second Q_1. It only accounts for the second Q_1 following the second $Q_{0.}$ Therefore, although the angles bound has a simple formula, it is not very sharp Before developing our own techniques for finding improved upper bounds, we investigated two published upper bounds on the rate of convergence of the Kacmarz Method in a more abstract setting in which the method is applied to orthogonal projections in subspaces in an arbitrary Hilbert Space. These particular upper bounds interest us because they include varying amounts of information concerning the relationships between operators in distinct iterations, by employing a powerful lemma of Kayalar and Weinert:

Lemma 2.1. *(Kayalar-Weinert [KW]): If $P_1$and P_2 are orthogonal projections onto $L_1$and L_2, subspaces of a Hilbert Space, H, if $P_{L_1 \cap L_2}$ is the orthogonal projection onto $L_1 \bigcap L_2$, and if T is a bounded linear operator on H, then*

$$\|P_2 P_1 T\|^2 \leq \|P_1 T\|^2 \cos^2(L_1, L_2) + \|P_{L_1 \cap L_2} T\|^2 \sin^2(L_1, L_2)$$

The Kayalar-Weinert lemma (KW Lemma) is very useful for bounding norms of operators because it contains information concerning the effects on the norm produced by following $P_1 T$ with P_2.

Deutsch and Hundal used this lemma to include the effects of Q_0 following $Q_{N-1}...Q_0$ at the beginning of each iteration after the first in constructing their bound for $\|Q^n\|$, so the Deutsch-Hundal bound (DH bound) is sharper than the angles bound (see Deutsch and Hundal [DH]). However, like the angles bound, the DH bound does not account for the second Q_1 following $Q_0 Q_{N-1}...Q_0$, for example. It only accounts for the second Q_1 following the second Q_0. When applied to the subspaces $I_{m,N}$, the DH bound yields no improvement over the angles bound for $n = 1$, and minimal improvement for $n \geq 2$. A second upper bound we investigated is a bound of Kayalar and Weinert (KW bound). It uses the KW Lemma much more extensively in order to bound all of the nN operators in $\|Q^n\|$ at once. Therefore, no information is lost in the sense that information is lost by the angles bound and the DH bound. The KW bound acccounts for the relationships between every one of the nN operators and all the operators which precede it [KW]. For example, the KW bound accounts for the Q_5 in the $(n-1)\,st$ iteration following $Q_4...Q_0 Q^{n-2}$, every operator preceding it. The KW bound is the sharpest bound, but as a consequence, the most complicated and the most difficult to apply. Our investigations thus far indicate that the KW bound produces little improvement over the angles bound for $n = 1$, but may produce significant improvement for $n \geq 2$. Hence, calculating the KW bound for$\|Q^n\|_{m,N}$, where $n \geq 2$ is an area of ongoing investigation which may provide insight into methods for proving the conjecture in $[HS]$.

We first describe our investigation of the DH and KW bounds for $m+1 \geq N$.

2.2.1. Subspaces of dimension $= N$

The Deutsch-Hundal Bound

Description of the Bound The first of two upper bounds we considered for $\|Q^n\|_{m,N}$ when $I_{m,N}$ has dimension N is the DH bound. It is the minimum over four constants, two of which equal one when the bound is applied to$\|Q^n\|_{m,N}$, a third, which reduces to the the angles bound. Hence, the DH bound is at least as sharp as the angles bound. We include the case which generates an improved upper bound in our application of interest.

Theorem 2.2. *(Deutsch-Hundal [DH]):*

$$\|(Q_{N-1}...Q_0)^n\|^2 \leq \alpha^2(\beta^2)^{n-1},$$

where

$$\alpha^2 \equiv min\left\{1-\prod_{i=0}^{N-2}\sin^2(L_i,\bigcap_{j=i+1}^{N-1}L_j), 1-\prod_{i=0}^{N-2}\sin^2(L_{N-i},\bigcap_{j=0}^{N-i-1}L_j)\right\},$$

$$\beta^2 \equiv \alpha^2\cos^2(L_0,L_{N-1})+\mu^2\sin^2(L_0,L_{N-1}),$$

$$\mu^2 \equiv \min\left\{\begin{array}{c} 1-\prod_{i=0}^{N-2}\sin^2\left(L_i,\bigcap_{j=i+1}^{N-1}L_j\bigcap L_0\right), \\ 1-\sin^2\left(L_0\bigcap L_{N-1},\bigcap_{j=0}^{N-2}L_j\right)\prod_{i=1}^{N-2}\sin^2(L_{N-i},\bigcap_{j=0}^{N-i-1}L_j) \end{array}\right\}$$

Proof:

$$\begin{aligned}\|(Q_{N-1}...Q_0)^n\|^2 &= \left\|(Q_{N-1}...Q_0)(Q_0Q_{N-1}...Q_0)^{n-1}\right\|^2 \text{ (by the idempotency of } Q_0) \\ &\leq \|(Q_{N-1}...Q_0)\|^2\left(\|(Q_0Q_{N-1}...Q_0)\|^2\right)^{n-1}\end{aligned}$$

α^2 is an upper bound for $\|(Q_{N-1}...Q_0)\|^2$ obtained by using the angles theorem, and β^2 is an upper bound for $\|(Q_0Q_{N-1}...Q_0)\|^2$. Indeed, an application of the KW Lemma yields

$$\begin{aligned}&\|(Q_0Q_{N-1}...Q_0)\|^2 \\ &\leq \|(Q_{N-1}...Q_0)\|^2\cos^2(L_0,L_{N-1})+\left\|(Q_{L_0\cap L_{N-1}}Q_{N-2}...Q_0)\right\|^2\sin^2(L_0,L_{N-1}),\end{aligned}$$

which followed by an application of the angles theorem to the two pieces on the left hand side of the inequality produces β, where $Q_{L_0 \cap L_{N-1}}$ is the projection on $L_0 \bigcap L_{N-1}$.

This derivation shows how the KW Lemma may be used, and it shows that a source of improvement in the upper bound lies in the term $\|(Q_0 Q_{N-1}...Q_0)\|^2$. It contains information concerning how the rate of convergence of Q^n is affected by Q_0 following $Q_{N-1}...Q_0$ in each iteration after the first, information lacking in the angles bound. It is evident that for n=1, the DH bound reduces to the angles bound, so any improvement must occur for n$\geq$ 2.

A calculation of $\alpha^2 \left(\beta^2\right)^{n-1}$ shows that the DH bound yields only minimal improvement in the subspaces $I_{m,N}$. The Gram formula (Theorem 1.7), is very useful in carrying out this calculation.

Calculation of α^2 The first term to calculate is

$$\alpha^2 \equiv min\left\{1 - \prod_{i=0}^{N-2} \sin^2(L_i, \bigcap_{j=i+1}^{N-1} L_j), 1 - \prod_{i=0}^{N-2} \sin^2(L_{N-i}, \bigcap_{j=0}^{N-i-1} L_j)\right\}$$

Using the Gram formula,

$$\begin{aligned} &1 - \prod_{i=0}^{N-2} \sin^2(L_i, \bigcap_{j=i+1}^{N-1} L_j) \\ &= 1 - \frac{G\left(U_{m,0},...,U_{m,N-1}\right)}{G(U_{m,1},...,U_{m,N-1})} \frac{G\left(U_{m,1},...,U_{m,N-1}\right)}{G(U_{m,2},...,U_{m,N-1})} ... \frac{G\left(U_{m,N-2},U_{m,N-1}\right)}{G(U_{m,N-1})} \\ &= 1 - G\left(U_{m,0},...,U_{m,N-1}\right), \end{aligned}$$

observing cancellation [HS]. The other term is computed in exactly the same manner, and the two terms are found to be equal in this case: $1 - \prod_{i=0}^{N-2} \sin^2(L_{N-i}, \bigcap_{j=0}^{N-i-1} L_j) = 1 - G\left(U_{m,0},...,U_{m,N-1}\right)$. Hence:

$$\alpha^2 = 1 - G\left(U_{m,0},...,U_{m,N-1}\right),$$

which is the familiar angles bound.

Calculation of β^2 Next, we calculate $\beta^2 \equiv \alpha^2 \cos^2\left(L_0, L_{N-1}\right) + \mu^2 \sin^2\left(L_0, L_{N-1}\right) = \alpha^2 + \left(\mu^2 - \alpha^2\right) \sin^2\left(L_0, L_{N-1}\right)$. We have:

$$\sin^2\left(L_0, L_{N-1}\right) = \frac{G\left(U_{m,0}, U_{m,N-1}\right)}{G(U_{m,N-1})} = G\left(U_{m,0}, U_{m,N-1}\right) = 1 - \left\langle U_{m,0}, U_{m,N-1}\right\rangle^2,$$

so it remains to compute

$$\mu^2 \equiv \min \left\{ \begin{array}{c} 1-\prod_{i=0}^{N-2} \sin^2\left(L_i, \bigcap_{j=i+1}^{N-1} L_j \bigcap L_0\right), \\ 1-\sin^2\left(L_0 \bigcap L_{N-1}, \bigcap_{j=0}^{N-2} L_j\right) \prod_{i=1}^{N-2} \sin^2(L_{N-i}, \bigcap_{j=0}^{N-i-1} L_j) \end{array} \right\}$$

Applying the Gram formula to the first term in the same manner as above, yields:

$$1-\prod_{i=0}^{N-2} \sin^2\left(L_i, \bigcap_{j=i+1}^{N-1} L_j \bigcap L_0\right) = 1-\frac{G\left(U_{m,0}, ..., U_{m,N-1}\right)}{G\left(U_{m,0}, U_{m,N-1}\right)}$$

In the second term, a similar computation shows that

$$\prod_{i=1}^{N-2} \sin^2(L_{N-i}, \bigcap_{j=0}^{N-i-1} L_j) = G\left(U_{m,0}, ..., U_{m,N-2}\right).$$

In order to apply the Gram formula to $\sin^2\left(L_0 \bigcap L_{N-1}, \bigcap_{j=0}^{N-2} L_j\right)$,it is necessary to observe that

$$\begin{aligned} &\sin^2\left(L_0 \bigcap L_{N-1}, \bigcap_{j=0}^{N-2} L_j\right) \\ &= \sin^2\left(L_0^{\perp}+L_{N-1}^{\perp}, \sum_{J=0}^{N-2} L_j^{\perp}\right) \\ &= \sin^2\left(L_0^{\perp}+L_{N-1}^{\perp} \ominus \left(\left(L_0^{\perp}+L_{N-1}^{\perp}\right) \bigcap \sum_{J=0}^{N-2} L_j^{\perp}\right), \sum_{J=0}^{N-2} L_j^{\perp}\right) \\ &= \sin^2\left(\left[U_{m,0}, U_{m,N-1}\right] \ominus \left(\left[U_{m,0}, U_{m,N-1}\right] \bigcap \left[U_{m,0}, ..., U_{m,N-2}\right]\right), \left[U_{m,0}, ..., U_{m,N-2}\right]\right) \\ &= \sin^2\left(\left[U_{m,0}, U_{m,N-1}\right] \ominus \left[U_{m,0}\right], \left[U_{m,0}, ..., U_{m,N-2}\right]\right) \end{aligned}$$

Now, since

$$\left[U_{m,0}, U_{m,N-1}\right] \ominus \left[U_{m,0}\right] = \left[Q_0 U_{m,N-1}\right] = \left[U_{m,N-1} - \left\langle U_{m,0}, U_{m,N-1}\right\rangle U_{m,0}\right],$$

we have $\sin^2\left(L_0 \bigcap L_{N-1}, \bigcap_{j=0}^{N-2} L_j\right)$

$$\begin{aligned} &= \sin^2\left(\left[U_{m,N-1} - \left\langle U_{m,0}, U_{m,N-1}\right\rangle U_{m,0}\right], \left[U_{m,0}, ..., U_{m,N-2}\right]\right) \\ &= \sin^2\left(\left[\frac{U_{m,N-1} - \left\langle U_{m,0}, U_{m,N-1}\right\rangle U_{m,0}}{\left\|U_{m,N-1} - \left\langle U_{m,0}, U_{m,N-1}\right\rangle U_{m,0}\right\|}\right], \left[U_{m,0}, ..., U_{m,N-2}\right]\right). \end{aligned}$$

We may now apply the Gram formula since one of the subspaces bounding the angle is one-dimensional and generated by a unit vector. We obtain:

$$\sin^2\left(L_0 \bigcap L_{N-1}, \bigcap_{j=0}^{N-2} L_j\right)$$
$$= \frac{G\left(\frac{U_{m,N-1} - \langle U_{m,0}, U_{m,N-1}\rangle U_{m,0}}{\|U_{m,N-1} - \langle U_{m,0}, U_{m,N-1}\rangle U_{m,0}\|}, U_{m,0}, ..., U_{m,N-2}\right)}{G(U_{m,0}, ..., U_{m,N-2})}$$
$$= \delta^2\left(\frac{U_{m,N-1} - \langle U_{m,0}, U_{m,N-1}\rangle U_{m,0}}{\|U_{m,N-1} - \langle U_{m,0}, U_{m,N-1}\rangle U_{m,0}\|}, [U_{m,0}, ..., U_{m,N-2}]\right),$$

(viewing the ratio of Gram determinants as a distance)

$$= \frac{\delta^2\left(U_{m,N-1} - \langle U_{m,0}, U_{m,N-1}\rangle U_{m,0}, [U_{m,0}, ..., U_{m,N-2}]\right)}{\|U_{m,N-1} - \langle U_{m,0}, U_{m,N-1}\rangle U_{m,0}\|},$$

(using a property of distances)

$$= \frac{\delta^2\left(U_{m,N-1} - \langle U_{m,0}, U_{m,N-1}\rangle U_{m,0}, [U_{m,0}, ..., U_{m,N-2}]\right)}{\sqrt{1 - \langle U_{m,0}, U_{m,N-1}\rangle^2}},$$
$$= \frac{G\left(U_{m,N-1} - \langle U_{m,0}, U_{m,N-1}\rangle U_{m,0}, U_{m,0}, ..., U_{m,N-2}\right)}{G(U_{m,0}, ..., U_{m,N-2})\sqrt{1 - \langle U_{m,0}, U_{m,N-1}\rangle^2}}$$
$$= \frac{G\left(U_{m,N-1} - \langle U_{m,0}, U_{m,N-1}\rangle U_{m,0}, U_{m,0}, ..., U_{m,N-2}\right)}{G(U_{m,0}, ..., U_{m,N-2})\sqrt{G\left(U_{m,0}, U_{m,N-1}\right)}}.$$

Thus, the second term in the calculation of μ^2 is

$$1 - \frac{G\left(U_{m,N-1} - \langle U_{m,0}, U_{m,N-1}\rangle U_{m,0}, U_{m,0}, ..., U_{m,N-2}\right)}{G(U_{m,0}, ..., U_{m,N-2})\sqrt{G\left(U_{m,0}, U_{m,N-1}\right)}}.G\left(U_{m,0}, ..., U_{m,N-2}\right)$$
$$= 1 - \frac{G\left(U_{m,N-1} - \langle U_{m,0}, U_{m,N-1}\rangle U_{m,0}, U_{m,0}, ..., U_{m,N-2}\right)}{\sqrt{G\left(U_{m,0}, U_{m,N-1}\right)}}.$$

Hence, we have that

$$\mu^2 = \min\left\{\begin{array}{c} 1 - \frac{G\left(U_{m,0}, ..., U_{m,N-1}\right)}{G\left(U_{m,0}, U_{m,N-1}\right)}, \\ 1 - \frac{G\left(U_{m,N-1} - \langle U_{m,0}, U_{m,N-1}\rangle U_{m,0}, U_{m,0}, ..., U_{m,N-2}\right)}{\sqrt{G\left(U_{m,0}, U_{m,N-1}\right)}} \end{array}\right\}$$

After showing that $G\left(U_{m,0},...,U_{m,N-1}\right)=G\left(U_{m,N-1}-\left\langle U_{m,0},U_{m,N-1}\right\rangle U_{m,0},U_{m,0},...,U_{m,N-2}\right)$, these terms become easy to compare.

Lemma 2.3. *:*

$$G\left(U_{m,0},...,U_{m,N-1}\right)=G\left(U_{m,N-1}-\left\langle U_{m,0},U_{m,N-1}\right\rangle U_{m,0},U_{m,0},...,U_{m,N-2}\right)$$

Proof: It suffices to compare

$$\frac{G\left(U_{m,0},...,U_{m,N-1}\right)}{G(U_{m,0},...,U_{m,N-2})}=\delta^2\left(U_{m,N-1},\left[U_{m,0},...,U_{m,N-2}\right]\right)$$

and

$$\frac{G\left(U_{m,N-1}-\left\langle U_{m,0},U_{m,N-1}\right\rangle U_{m,0},U_{m,0},...,U_{m,N-2}\right)}{G(U_{m,0},...,U_{m,N-2})}$$
$$=\delta^2\left(U_{m,N-1}-\left\langle U_{m,0},U_{m,N-1}\right\rangle U_{m,0},\left[U_{m,0},...,U_{m,N-2}\right]\right)$$

We are comparing the distance of two points to the same subspace, $[U_{m,0},...,U_{m,N-2}]$. This distance is invariant under a change of basis for the subspace, so we apply the Gram-Schmidt process to the basis vectors $U_{m,0},...,U_{m,N-2}$ to obtain an orthonormal basis $V_0,V_1,...,V_{N-2}$,where $V_0=U_{m,0}$,and $V_k=\dfrac{U_{m,k}-\sum_{j=0}^{k-1}\left\langle V_j,U_{m,k}\right\rangle V_j}{\left\|U_{m,k}-\sum_{j=0}^{k-1}\left\langle V_j,U_{m,k}\right\rangle V_j\right\|}$, $k=1,2,...N-2$. Since a Gram determinant with orthonormal generators is equal to one, we have

$$\begin{aligned}
&\delta^2\left(U_{m,N-1},\left[U_{m,0},...,U_{m,N-2}\right]\right)\\
&=\delta^2\left(U_{m,N-1},\left[V_0,...,V_{N-2}\right]\right)\\
&=\frac{G\left(V_0,...,V_{N-2},U_{m,N-1}\right)}{G(V_0,...,V_{N-2})}\\
&=G\left(V_0,...,V_{N-2},U_{m,N-1}\right)\\
&=\begin{vmatrix}
1 & \langle V_0,U_{m,N-1}\rangle & \langle V_1,U_{m,N-1}\rangle & \cdots & \langle V_{N-2},U_{m,N-1}\rangle\\
\langle V_0,U_{m,N-1}\rangle & 1 & 0 & \cdots & 0\\
\langle V_1,U_{m,N-1}\rangle & 0 & 1 & \cdots & 0\\
\vdots & \vdots & \vdots & \ddots & \vdots\\
\langle V_{N-2},U_{m,N-1}\rangle & 0 & 0 & \cdots & 1
\end{vmatrix}\\
&=1-\sum_{j=0}^{N-2}(-1)^j\left\langle V_j,U_{m,N-1}\right\rangle^2.
\end{aligned}$$

Then also, we have

$$\begin{aligned}
&\delta^2\left(U_{m,N-1}-\left\langle U_{m,0},U_{m,N-1}\right\rangle U_{m,0},\left[U_{m,0},...,U_{m,N-2}\right]\right)\\
&=\delta^2\left(U_{m,N-1}-\left\langle V_0,U_{m,N-1}\right\rangle V_0,\left[V_0,...,V_{N-2}\right]\right)\\
&=\frac{G\left(U_{m,N-1}-\left\langle V_0,U_{m,N-1}\right\rangle V_0,V_0,...,V_{N-2}\right)}{G(V_0,...,V_{N-2})}\\
&=G\left(U_{m,N-1}-\left\langle V_0,U_{m,N-1}\right\rangle V_0,V_0,...,V_{N-2}\right)\\
&=\begin{vmatrix}
1-\left\langle V_0,U_{m,N-1}\right\rangle & 0 & \left\langle V_1,U_{m,N-1}\right\rangle & \cdots & \left\langle V_{N-2},U_{m,N-1}\right\rangle\\
0 & 1 & 0 & \cdots & 0\\
\left\langle V_1,U_{m,N-1}\right\rangle & 0 & 1 & \cdots & 0\\
\vdots & \vdots & \vdots & \ddots & \vdots\\
\left\langle V_{N-2},U_{m,N-1}\right\rangle & 0 & 0 & \cdots & 1
\end{vmatrix}\\
&=1-\sum_{j=0}^{N-2}(-1)^j\left\langle V_j,U_{m,N-1}\right\rangle^2.
\end{aligned}$$

Hence,

$$\frac{G\left(U_{m,0},...,U_{m,N-1}\right)}{G(U_{m,0},...,U_{m,N-2})}=\frac{G\left(U_{m,N-1}-\left\langle U_{m,0},U_{m,N-1}\right\rangle U_{m,0},U_{m,0},...,U_{m,N-2}\right)}{G(U_{m,0},...,U_{m,N-2})},$$

and the lemma is proved.

Now it follows that since

$$G\left(U_{m,0},U_{m,N-1}\right)\leq\left\|U_{m,0}\right\|\left\|U_{m,N-1}\right\|=1,$$

then

$$G\left(U_{m,0},U_{m,N-1}\right)\leq\sqrt{G\left(U_{m,0},U_{m,N-1}\right)},$$

and then also

$$\frac{G\left(U_{m,0},...,U_{m,N-1}\right)}{G\left(U_{m,0},U_{m,N-1}\right)}\geq\frac{G\left(U_{m,N-1}-\left\langle U_{m,0},U_{m,N-1}\right\rangle U_{m,0},U_{m,0},...,U_{m,N-2}\right)}{\sqrt{G\left(U_{m,0},U_{m,N-1}\right)}}.$$

Therefore,

$$1-\frac{G\left(U_{m,0},...,U_{m,N-1}\right)}{G\left(U_{m,0},U_{m,N-1}\right)}\leq1-\frac{G\left(U_{m,N-1}-\left\langle U_{m,0},U_{m,N-1}\right\rangle U_{m,0},U_{m,0},...,U_{m,N-2}\right)}{\sqrt{G\left(U_{m,0},U_{m,N-1}\right)}},$$

so we find that

$$\mu^2=1-\frac{G\left(U_{m,0},...,U_{m,N-1}\right)}{G\left(U_{m,0},U_{m,N-1}\right)}.$$

Comparison: DH Bound Versus Angles Bound Recall that

$$G\left(U_{m,0},U_{m,N-1}\right)=1-\left\langle U_{m,0},U_{m,N-1}\right\rangle^{2}=1-\frac{\sin^{2}\left(kv\pi/N\right)}{\left(m+1\right)^{2}\sin^{2}\left(k\pi/N\right)}$$

and

$$G\left(U_{m,0},...,U_{m,N-1}\right)=\left(\frac{uN}{m+1}\right)^{N-v}\left(\frac{uN+N}{m+1}\right)^{v},$$

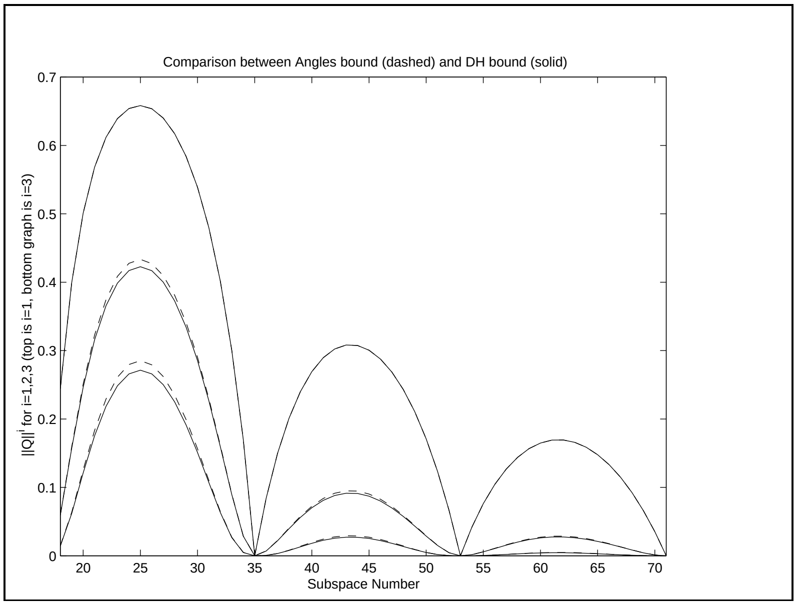
where $m+1=uN+v$, with $0\leq v\leq N-1$, and $k\pi/N=\theta_{N-1}$. The result is:

$$\begin{aligned}
&\left\|\left(Q_{N-1}...Q_{0}\right)^{n}\right\|^{2}\\
&\leq\alpha^{2}(\beta^{2})^{n-1}\\
&=\left[1-\left(\frac{uN}{m+1}\right)^{N}\left(1+\frac{1}{u}\right)^{v}\right]\left[1-\left(\frac{uN}{m+1}\right)^{N}\left(1+\frac{1}{u}\right)^{v}\left(1+\frac{\sin^{2}\left(kv\pi/N\right)}{\left(m+1\right)^{2}\sin^{2}\left(k\pi/N\right)}\right)\right]
\end{aligned}$$

The angles bound is

$$\left[1-\left(\frac{uN}{m+1}\right)^{N}\left(1+\frac{1}{u}\right)^{v}\right]^{n},$$

so the DH bound is an improvement since for every m$\geq$ 0, and every $k=1,2,...,N-1$, $\dfrac{\sin^{2}\left(kv\pi/N\right)}{\left(m+1\right)^{2}\sin^{2}\left(k\pi/N\right)}>0$. The following is a graphical comparison of the two bounds which shows that the improvement over the angles bound is slight.



The Kayalar-Weinert Bound I

Description of the Bound A second upper bound for $\|Q^n\|_{I_{m,N}}$ when $I_{m,N}$ has dimension N is the KW bound. As a bound on the nth iteration, it is the minimum over four constants, two of which are the product of nN factors, and two of which are the product of 2nN factors. These constants originate from bounding expressions equal to $\|Q^n\|$, i.e., bounding $\|Q^n\|$ itself, $\|(Q^n)^*\|$, $\sqrt{\|(Q^n)^*(Q^n)\|}$,and $\sqrt{\|(Q^n)(Q^n)^*\|}$. As mentioned previously, this bound is very complicated because it accounts for the relationships between each projection and all the preceding projections in Q^n [KW]. Its complexity and the fact that the number of calculations increases with increasing n make it an unwieldy computational tool. However, we have applied this bound to our situation in computed tomography for $n = 1$ and for a small number of x-rays when $n = 2$. Our results indicate that the sharpness of the bound increases with increasing n. This will be shown explicitly in the case of three x-rays. The bound is stated as follows:

Theorem 2.4. *(Kayalar-Weinert [KW]):*

$$\|(Q_{N-1}...Q_0)^n\| \leq \min\left\{\prod_{i=0}^{nN-1} p_i^f, \sqrt{\prod_{i=0}^{2nN-1} p_i^f}, \prod_{i=0}^{nN-1} p_i^b, \sqrt{\prod_{i=0}^{2nN-1} p_i^b}\right\}$$

where f denotes the forward ordering of the projections, $Q_{N-1}...Q_0 = Q$, and b denotes the backward ordering ,$Q_0...Q_{N-1} = Q^*$, and where for $0 \leq i \leq 2nN-1$,

$$\left(p_i^f\right)^2 = \min\left\{1, \left[\begin{array}{l} \cos^2\left(L_{i-1,}L_i\right)^f + \dfrac{\prod_{j=i-1}^{i-1}\sin^2\left(L_{j,}\bigcap_{k=j+1}^{i}L_k\right)^f\cos^2\left(L_{i-2,}L_{i-1}\bigcap L_i\right)^f}{\left(p_{i-1}^f\right)^2} + \\ +\dfrac{\prod_{j=i-2}^{i-1}\sin^2\left(L_{j,}\bigcap_{k=j+1}^{i}L_k\right)^f\cos^2\left(L_{i-3,}\bigcap_{k=i-2}^{i}L_k\right)^f}{\left(p_{i-1}^f p_{i-2}^f\right)^2} + ... \\ +\dfrac{\prod_{j=1}^{i-1}\sin^2\left(L_{j,}\bigcap_{k=j+1}^{i}L_k\right)^f\cos^2\left(L_{0,}\bigcap_{k=1}^{i}L_k\right)^f}{\left(p_{i-1}^f p_{i-2}^f...p_1^f\right)^2} + \\ +\dfrac{\prod_{j=0}^{i-1}\sin^2\left(L_{j,}\bigcap_{k=j+1}^{i}L_k\right)^f}{\left(p_{i-1}^f p_{i-2}^f...p_0^f\right)^2}\left\|Q_{\bigcap_{j=0}^{i}L_j}\right\|^2 \end{array}\right]\right\}$$

The formula for $p_i^b$is exactly the same with the f's and b's interchanged.

Since f denotes a forward ordering and b denotes a backward ordering, b is a reordering of f. It follows that the formula for the $\left(p_i^f\right)^2$'s is identical to the formula for the $\left(p_i^b\right)^2$'s applied to a reverse ordering of the projections, so it suffices to consider only the $\left(p_i^f\right)^2$'s.

How the superscript f is used in the formula is best explained by the following example. If we consider three x-rays $(N = 3)$,and want to use the KW bound for $n = 2$, our projections are Q_0, Q_1, Q_2, which project onto L_0, L_1, L_2. For the forward ordering f,

$$\begin{aligned}\left(Q^*\right)^2 Q^2 &= Q_0Q_1Q_2Q_0Q_1Q_2Q_2Q_1Q_0Q_2Q_1Q_0 \\ &= Q_{11}^fQ_{10}^fQ_9^fQ_8^fQ_7^fQ_6^fQ_5^fQ_4^fQ_3^fQ_2^fQ_1^fQ_0^f,\end{aligned}$$

where the subscripts in the third expression constitute a renumbering of the projections Q_0, Q_1, Q_2 corresponding to their order of occurrence in $\left(Q^*\right)^2 Q^2$, i.e., $Q_4^f = Q_{0,}$ $Q_5^f = Q_1, Q_6^f = Q_2$, etc. The corresponding subspaces $L_0^f, L_1^f,...,L_{11}^f$, which appear in the formulas for $\left(p_0^f\right)^2, \left(p_1^f\right)^2,...,\left(p_{11}^f\right)^2$may be similarly interpreted: $L_4^f = L_0$, $L_5^f = L_1$, $L_6^f = L_2$, etc. If, for example, the expression

$\cos^2\left(L_{11},\bigcap_{i=9}^{10} L_i\right)^f$ occurs in the formula for $\left(p_1^f\right)^2$,then

$$cos^2\left(L_{11},\bigcap_{i=9}^{10} L_i\right)^f = cos^2\left(L_{11}^f, L_9^f\bigcap L_{10}^f\right) = cos^2\left(L_0, L_2\bigcap L_1\right),$$

which is something we understand [KW].

Calculating the Bound When we compute with this bound, certain summands in the expressions for p_i^f are found to be zero. For instance, when $L_i \subseteq L_j$,then $\cos^2(L_i, L_j)$ =$\cos^2(\pi/2) = 0$. Also, when $i \geq N-1$, $\left\|Q_{\bigcap_{j=0}^{i} L_j}\right\|^2 = 0$ since each L_j has dimension $N-1$ [KW]. Deciding which such terms are zero and deciphering the complicated notation makes calculating the p^f's difficult.

However, we have obtained formulas for the bound applied to the subspaces $I_{m,N}$ in terms of Gram determinants for the case of one iteration. Applying the Gram formula,we find that formulas for $\left(p_0^f\right)^2, ..., \left(p_{2N-1}^f\right)^2$ are as follows:

$$For\ 0 \leq i \leq N-2, \left(p_i^f\right)^2 = 1.$$

$$For\ i = N-1, \left(p_{N-1}^f\right)^2 = 1 - G\left(U_{m,0}, ..., U_{m,N-1}\right)$$

If for $x \geq y$, we define $\left(\Gamma^f(x,y)\right)^2 = \left(p_x^f p_{x-1}^f ... p_y^f\right)^2$,then

$$for\ 2 \leq j \leq N-1,\ we\ have\ \left(p_{N+j-1}^f\right)^2$$

$$= min\left(\begin{array}{c} 1, 1 - G\left(U_{m,N-j}, U_{m,N-j+1}\right) + \\ + \sum_{l=1}^{j-2}\left[\frac{G\left(U_{m,N-j}, ..., U_{m,N-j+l}\right) - G\left(U_{m,N-j}, ..., U_{m,N-j+l+1}\right)}{\left(\Gamma^f\left(N+j-2, N+j-l-1\right)\right)^2}\right] + \\ + \sum_{l=j-1}^{N-2}\left[\frac{G\left(U_{m,N-l-1}, ..., U_{m,N-1}\right) - G\left(U_{m,N-l-2}, ..., U_{m,N-1}\right)}{\left(\Gamma^f\left(N+j-2, N-l-1\right)\right)^2}\right] \end{array}\right)$$

$$For\ i = 2N-1, \left(p_{2N-1}^f\right)^2$$

$$= min\left(1, 1 - G\left(U_{m,0}, U_{m,1}\right) + \sum_{l=1}^{N-2}\left[\frac{G\left(U_{m,0}, ..., U_{m,l}\right) - G\left(U_{m,0}, ..., U_{m,l+1}\right)}{\left(\Gamma^f\left(2N-2, 2N-l-1\right)\right)^2}\right]\right)$$

For the case of one iteration, the bound reduces to

$$\|Q\|_{m,N} \leq \min \left\{ 1 - G\left(U_{m,0}, ..., U_{m,N-1}\right), \sqrt{\prod_{i=N-1}^{2N-1} p_i^f} \right\},$$

which will be sharper than the the angles bound if $1 - G\left(U_{m,0}, U_{m,1}, ..., U_{m,N-1}\right) > \prod_{i=N+1}^{2N-1} p_i^f$. This seems to be true in general, and can be easily verified analytically for the case of three x-rays ($N = 3$).

A Comparison of Three Bounds Formulas for the p_i^f's become increasingly more complicated as n increases, so our investigation of the bound for larger numbers of iterations has been very limited. However, the case of three x-rays is relatively easy to analyze since

$$G\left(U_{m,0}, U_{m,1}\right) = G\left(U_{m,0}, U_{m,2}\right) = G\left(U_{m,1}, U_{m,2}\right) = \frac{\sin^2\left(\left(m+1\right)\pi/3\right)}{\left(m+1\right)^2 \sin^2\left(\pi/3\right)} = \frac{1}{\left(m+1\right)^2}.$$

We now compare the angles bound , the DH bound, and the KW bound as bounds for $\|Q\|_{3,3}$ and $\|Q^2\|_{3,3}$:

$$\begin{aligned} \|Q\|_{3,3} &\leq .39528 \text{ (angles bound)} \\ &\leq .39528 \text{ (DH bound)} \\ &\leq .38121 \text{ (KW bound)} \\ \left\|Q^2\right\|_{3,3} &\leq .15625 \text{ (angles bound)} \\ &\leq .14937 \text{ (DH bound)} \\ &\leq .10035 \text{ (KW bound)} \end{aligned}$$

The KW bound is significantly better than the others in the case of two iterations. In fact, the KW bound in this case, $.10035 \leq \|Q\|_{3,3}^2 = .13104$. Such significant improvement suggests that the sharpness of the KW bound increases as n increases. It may be possible to show that the Kacmarz Method converges at a faster than linear rate in some cases of three or more projections. For the case of two x-rays, the best possible bound is known and the convergence is linear.

2.2.2. Subspaces of Dimension $< N$

A second, distinct aspect of our problem is to find sharper upper bounds on subspaces $I_{m,N}$ of dimension $<N$. To this end, we modified the angles bound, and we again employed the KW bound.

Description of the Subspaces To see why the subspaces of dimension $<N$ must be treated separately, we investigate their structure. Again, we consider projections $Q_0, ..., Q_{N-1}$,and the subspaces $L_0, ..., L_{N-1}$, onto which they project. Since $I_{m,N}$ has dimension m+1, any subset of m+1 $U_{m,i}$'s taken from $\{U_{m,0}, ..., U_{m,N-1}\}$ is a basis for the subspace $I_{m,N}$. Restricted to $I_{m,N}$, Q_j projects onto $L_j = I_{m,N} \ominus [U_{m,j}]$, a subspace of codimension 1. Therefore, $\bigcap_{i_0,...i_m} L_{i_j} = \{0\}$ for any subset of m+1 subspaces, and $\bigcap_{i_0,...i_k} L_{i_j} \neq \{0\}$ for any subset of $k \leq m$ subspaces, and in this case, $\|Q_{i_k}...Q_{i_0}\| = 1$, showing that the angles theorem may only be applied to a subset of at least m+1 projections. But since any subset of $U_{m,i}$'s consisting of more than m+1 functions is linearly dependent, the Gram determinant generated by these functions is zero. Hence, a simple bound may be obtained from the angles theorem as follows:

$$\begin{aligned}\|Q_{N-1}...Q_{m+1}Q_m...Q_0\|^2_{m,N} &\leq \|Q_{N-1}...Q_{m+1}\|^2_{m,N} \|Q_m...Q_0\|^2_{m,N} \\ &\leq \|Q_{N-1}...Q_{m+1}\|^2_{m,N} \left(1 - G\left(U_{m,0}, ...U_{m,m+1}\right)\right) \\ &\leq 1 - G\left(U_{m,0}, ...U_{m,m+1}\right)\end{aligned}$$

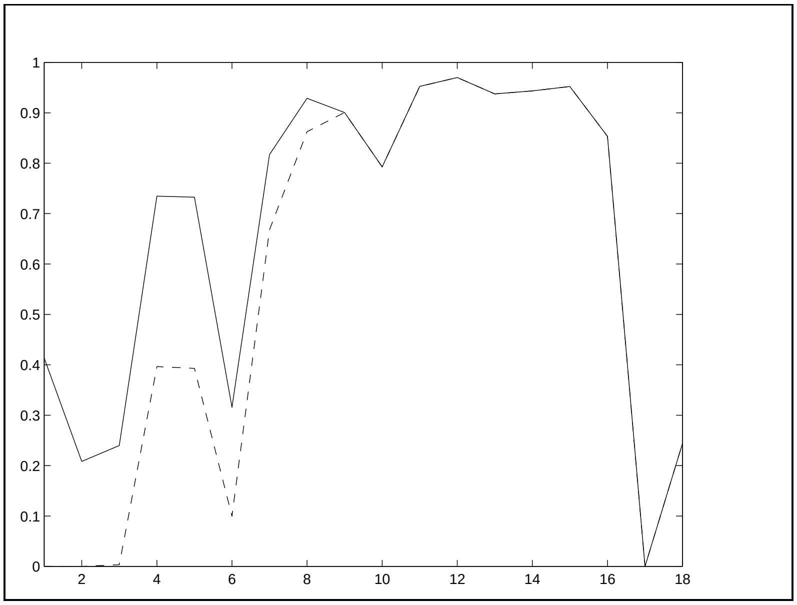
since $\|Q_{N-1}...Q_{m+1}\|^2_{m,N} \leq 1$. There is more than one way to choose m+1 projections from a set of N when m+1<N, so the bound is order dependent [HS]. However, it yields a particularly poor estimate since it includes no information about the final $N - m - 1$ projections.

Improved Angles Bounds

The Block Method The following examples show how to modify this approach to obtain an improved estimate. In order to bound $\|Q_{17}...Q_0\|^2_{4,18}$, for example, we partition the operators into as many groups of m+1 =5 as possible:

$$\begin{aligned}\|Q_{17}...Q_0\|^2_{4,18} &\leq \|Q_{17}Q_{16}Q_{15}\|^2_{4,18} \|Q_{14}...Q_{10}\|^2_{4,18} \|Q_9...Q_5\|^2_{4,18} \|Q_4...Q_0\|^2_{4,18} \\ &\leq \left(1 - G\left(U_{m,14}, ...U_{m,10}\right)\right)\left(1 - G\left(U_{m,9}, ...U_{m,5}\right)\right)\left(1 - G\left(U_{m,4}, ...U_{m,0}\right)\right)\end{aligned}$$

which is better than the angles bound, $1 - G\left(U_{m,4}, ...U_{m,0}\right)$. We find that the improvement gained from this block method is minor in most cases because of our inability to bound the short leftover block consisting of $k < m + 1$ projections. Furthermore, it is only useful when $m + 1 \leq \lfloor N/2 \rfloor$. For example, in the graph below, the block method cannot be used in subspaces $I_{10,18}, ..., I_{17,18}$ since $10 > 18/2 = 9, ..., 17 > 9$.



The Angles bound (solid) vs. the Block method (dashed)

.

The Block Method with Idempotency We can improve the block method by using the idempotency of the Q_j's. In a few special cases this allows us to eliminate the short leftover block. As an example, using the block method and the idempotency of the projections to bound $\|Q_{17}...Q_0\|^2_{4,18}$, we find:

$$\|Q_{17}...Q_0\|^2_{4,18} \leq \|Q_{17}...Q_{13}\|^2_{4,18} \|Q_{12}...Q_8\|^2_{4,18} \|Q_8...Q_4\|^2_{4,18} \|Q_4...Q_0\|^2_{4,18}$$
$$\leq (1 - G(U_{m,17}, ...U_{m,13}))(1 - G(U_{m,12}, ...U_{m,8}))$$
$$(1 - G(U_{m,8}, ...U_{m,4}))(1 - G(U_{m,4}, ...U_{m,0}))$$

Hence, if it is possible to use idempotency to eliminate the short block, we gain an extra factor of $1 - G(U_{m,N-m-2}, ...U_{m,N-1})$, which improves the estimate.

The improved angles technique using the block method and the idempotency of the projections produces only small improvements over the angles bound, and then, only when $m + 1 \leq \lfloor N/2 \rfloor$.

A Variation of the Angles Bound Another possible strategy for bounding $\|Q\|_{m,N}$ when $\dim I_{m,N} < N$, and the number of projections is $m + k + 1,\ k < m$, is to use $Q_{\bigcap_{j=m+1}^{m+k} L_j}$, the projection on $\bigcap_{j=m+1}^{m+k} L_j$ to obtain:

$$\begin{aligned}\|Q_{m+k}...Q_0 u\|^2 &= \left\|Q_{m+k}...Q_0 u - Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_0 u + Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_0 u\right\|^2 \\ &= \left\|Q_{m+k}...Q_0 u - Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_0 u\right\|^2 + \left\|Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_0 u\right\|^2\end{aligned}$$

The term $2\left\langle Q_{m+k}...Q_0 u - Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_0 u, Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_0 u\right\rangle = 0$ by the invariance of $\bigcap_{j=m+1}^{m+k} L_j$ under $Q_{m+k}...Q_{m+1}$ Now, letting $Q_m...Q_0 u = v$,and $Q_{m+k}...Q_{m+1} = Q'$, estimate the first summand in the above expression by using the angles theorem:

$$\begin{aligned}&\left\|Q_{m+k}...Q_0 u - Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_0 u\right\|^2 \\ &= \left\|(Q_{m+k}...Q_{m+1})(Q_m...Q_0 u) - Q_{\bigcap_{j=m+1}^{m+k} L_j}(Q_m...Q_0 u)\right\|^2 \\ &= \left\|Q' v - Q_{\bigcap_{j=m+1}^{m+k} L_j} v\right\|^2 \\ &\leq \left(1 - \prod_{i=m+1}^{m+k-1} \sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right)\left\|v - Q_{\bigcap_{j=m+1}^{m+k} L_j} v\right\|^2 \\ &= \left(1 - \prod_{i=m+1}^{m+k-1} \sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right)\left(\|v\|^2 - \left\|Q_{\bigcap_{j=m+1}^{m+k} L_j} v\right\|^2\right) \\ &= \left(1 - \prod_{i=m+1}^{m+k-1} \sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right)\left(\|Q_m...Q_0 u\|^2 - \left\|Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_0 u\right\|^2\right)\end{aligned}$$

Therefore, we have

$$\|Q_{m+k}...Q_0u\|^2 \leq \left(1 - \prod_{i=m+1}^{m+k-1} \sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right).$$

$$\cdot\left(\|Q_m...Q_0u\|^2 - \left\|Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_0u\right\|^2\right) +$$

$$+ \left\|Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_0u\right\|^2$$

$$= \left(1 - \prod_{i=m+1}^{m+k-1} \sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right) \|Q_m...Q_0u\|^2 +$$

$$+ \prod_{i=m+1}^{m+k-1} \sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right) \left\|Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_0u\right\|^2$$

$$= \left(1 - \prod_{i=m+1}^{m+k-1} \sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right) \left(1 - \prod_{i=0}^{m-1} \sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right) \|u\|^2 +$$

$$+ \prod_{i=m+1}^{m+k-1} \sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right) \left\|Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_0u\right\|^2$$

(by applying the angles theorem to $\|Q_m...Q_0u\|^2$)

Since:

$$\left\|Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_0u\right\|^2 = \left\|Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_kQ_{k-1}...Q_0u\right\|^2$$

$$\leq \left\|Q_{\bigcap_{j=m+1}^{m+k} L_j} Q_m...Q_k\right\|^2 \|Q_{k-1}...Q_0u\|^2$$

$$\leq \left(1 - \prod_{i=m+1}^{m+k-1} \sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right) \|Q_{k-1}...Q_0u\|^2$$

We now have:

$$\|Q_{m+k}...Q_0u\|^2 \leq \left(1-\prod_{i=m+1}^{m+k-1}\sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right)\left(1-\prod_{i=0}^{m-1}\sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right)\|u\|^2 +$$
$$+\prod_{i=m+1}^{m+k-1}\sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right).$$
$$.\left(1-\prod_{i=m+1}^{m+k-1}\sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right)\|Q_{k-1}...Q_0u\|^2$$
$$\leq \left(1-\prod_{i=m+1}^{m+k-1}\sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right)\left(1-\prod_{i=0}^{m-1}\sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right)+$$
$$+\prod_{i=m+1}^{m+k-1}\sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\left(1-\prod_{i=m+1}^{m+k-1}\sin^2\left(L_i, \bigcap_{j=i+1}^{m+k} L_j\right)\right)$$
$$(replacing\ \|Q_{k-1}...Q_0u\|^2\ by\ 1)$$
$$= \left(1-G\left(U_{m,m+1},...U_{m,m+k}\right)\right)\left(1-G\left(U_{m,0},...U_{m,m}\right)\right)+$$
$$+G\left(U_{m,m+1},...U_{m,m+k}\right)\left(1-G\left(U_{m,k},...U_{m,m+k}\right)\right)$$

which will be a sharper upper bound than $1-G\left(U_{m,0},...U_{m,m}\right)$,the angles bound, iff $G\left(U_{m,k},...U_{m,m+k}\right) > G\left(U_{m,0},...U_{m,m}\right)$. This approach does not seem worth pursuing unless we can find a better upper bound for $\|Q_{k-1}...Q_0u\|^2$. Orthogonal decompositions of $I_{m,N}$ may lead to improvements. If $I_{m,N} = \left(I_{m,N}\ominus\left[U_{m,0},...,U_{m,k-1}\right]\right)\oplus \left[U_{m,0},...,U_{m,k-1}\right]$,then

$$\|Q_{k-1}...Q_0u\|^2 = \|Q_{k-1}...Q_0v\|^2 + \|Q_{k-1}...Q_0w\|^2$$
$$\left(\text{ where } v\in\left(I_{m,N}\ominus\left[U_{m,0},...,U_{m,k-1}\right]\right) \text{ and } w\in\left[U_{m,0},...,U_{m,k-1}\right]\ \right)$$

since the components of the orthogonal decomposition of $I_{m,N}$ are invariant under $Q_{k-1}...Q_0$. We can estimate $\|Q_{k-1}...Q_0w\|^2$ using the angles theorem:

$$\|Q_{k-1}...Q_0w\|^2 \leq 1-\prod_{i=0}^{k-2}\sin^2\left(L_i, \bigcap_{j=i+1}^{k-1} L_j\right) = 1-G(U_{m,0},...,U_{m,k-1}),$$

Then it remains to bound $\|Q_{k-1}...Q_0v\|^2$, that is, $\|Q_{k-1}...Q_0\|^2$ restricted to $I_{m,N}\ominus\left[U_{m,0},...,U_{m,k-1}\right]$. The particular decomposition above may not be the best decomposition. It is an example showing how to estimate using orthogonal decompositions of the subspace $I_{m,N}$. Other, more clever decompositions may exist.

The Kayalar-Weinert Bound II We also applied the KW bound to subspaces $I_{m,N}$ of dimension $< N$ in the case of one iteration. We find that, (using the same notation as in the previous discussion of this bound):

$$For\ 0 \le i \le N-2, \left(p_i^f\right)^2 = 1.$$

$$For\ 0 \le i \le N-m-1,\ \left(p_{m+i}^f\right)^2 =$$

$$min\left(\begin{array}{c} 1, 1 - G\left(U_{m,m+i-1}, U_{m,m+i}\right) + \\ + \sum_{l=1}^{m-2}\left[\frac{G\left(U_{m,m+i-l}, ..., U_{m,m+i}\right) - G\left(U_{m,m+i-l-1}, ..., U_{m,m+i}\right)}{\left(\Gamma^f\left(m+i-1, m+i-l\right)\right)^2}\right]\end{array}\right)$$

We have yet to calculate formulas for the p_j^f's when $N \le j \le 2N-1$. Nevertheless, $\prod_{i=m}^{N-1} p_i^f$ is an upper bound. It is interesting to note that p_m^f is the angles bound. Therefore, if $p_m^f > \prod_{i=m}^{N-1} p_i^f$, i.e., if at least one of $p_{m+1}^f, ..., p_{N-1}^f$ is less than one, then the KW bound is an improvement. For example, a necessary and sufficient condition that $p_{m+1}^f < 1$ is that

$$G\left(U_{m,2}, ...U_{m,m+1}\right) > G\left(U_{m,m}, U_{m,m+1}\right) G\left(U_{m,1}, ...U_{m,m}\right)$$

However, the actual sense of this inequality is dependent on the ordering of the projections, so it is not clear whether the KW bound gives an improved estimate for $n = 1$ on the subspaces of dimension $< N$. We conjecture, that for $n = 1$, the improvement is small, but increases rapidly with increasing n . These lower dimensional subspaces are of particular interest to us, since our data for the case of 18 x-rays indicates that for inefficient orderings of the projections, i.e., orderings for which convergence of the Kacmarz Method is slow, $\sup_{m \ge 0} \|Q\|_{m,N}$ occurs in some subspace $I_{m,N}$ of dimension $< N$.

3. The Eigenvalue Problem

An additional goal of ours is to find an algebraic formula for $\|Q\|_{N,N}^2$. To do this, we compute the maximum eigenvalue of Q^*Q restricted to $I_{N,N}$. The computation is facilitated by using the basis $V_0, ..., V_{N-1}$, where $V_i = (-1)^{n_i} U_{m,i}$,

$\theta_j = n_j\pi/N$, and $Q_jV_i = \left(V_i - \left(\frac{1}{m+1}\right)V_j\right)$ and $Q_jV_j = 0$ [HS]. For example, when $N = 3$,

$$Q_0 = \begin{bmatrix} 0 & -1/4 & -1/4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, Q_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1/4 & 0 & -1/4 \\ 0 & 0 & 1 \end{bmatrix}, Q_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1/4 & -1/4 & 0 \end{bmatrix},$$

$$\textit{and } Q^*Q = \begin{bmatrix} 0 & -25/1024 & -37/1024 \\ 0 & 13/256 & 9/256 \\ 0 & 3/64 & 7/64 \end{bmatrix}$$

The maximum eigenvalue of Q^*Q is found to be $\frac{41+3\sqrt{73}}{512} = .13104$. The matrix representation of Q^*Q becomes more complicated as N increases. For $N = 4$,

$$Q^*Q = \begin{bmatrix} 0 & -1281/78125 & -2101/78125 & -2501/78125 \\ 0 & 461/15625 & 256/15625 & 156/15625 \\ 0 & 84/3125 & 189/3125 & 164/3125 \\ 0 & 16/625 & 36/625 & 61/625 \end{bmatrix}$$

The maximum eigenvalue in this case is .1428 Its algebraic form is too long to be listed here. This shows that although the Q_j's are represented by sparse matrices, an exact formula for the largest eigenvalue of Q^*Q quickly becomes very complicated. Numerical computations show that an interesting feature of the eigenvalues of Q^*Q in $I_{N,N}$ is that the maximum eigenvalue is significantly larger than the other $N-1$ eigenvalues. For example, in $I_{18,18}$,the maximum eigenvalue is .1816, and the remaining 17 eigenvalues are all less than .037.

4. Conclusion

The KW bound is the best known upper bound on the rate of convergence of the Kacmarz Method. It seems to sharpen as the number of iterations increases. Therefore, we would like to further investigate the KW bound in the subspaces $I_{m,N}$. An analysis of the KW bound would be particularly valuable in the subspaces of dimension <N, where the angles bound and the DH bound are poorest. If the KW bound is not significantly better than the angles bound and the DH bound in the subspaces of dimension <N, it may be improved by using a clever orthogonal decomposition of $I_{m,N}$, similar to the example given in the section on the variant angles bound. Perhaps the KW bound can be used to prove the conjecture in [HS]. Since an impediment to the study of the KW bound is its

complexity, perhaps there is a simpler bound which retains a large degree of the accuracy of the KW bound.

5. References

[DH] F.Deutsch and H. Hundal, "The Rate of Convergence for the Method of Alternating Projections, II," J. Math. Anal. Appl. 205 (1997) 381-405.

[H] I. Halperin, "The Product of Projection Operators," *Acta. Sci. Math. (Szeged)* **23** (1962), 96-99.

[HS] C.Hamaker and D.Solmon, "The Angles Between the Null Spaces of X-Rays," J. Math. Anal. Appl., 62 (1978) 1-23.

[GG] S. Gohberg and I. Goldberg, *Basic Operator Theory,* Birkhauser, Basel, 1981.

[KW] S.Kayalar and H.Weinert, "Error Bounds for the Method of Alternating Projections," Math. Control Signals Systems 1 (1988) 43-59.