Designing Structured Tight Frames via an Alternating Projection Method

Joel A. Tropp, *Student Member, IEEE*, Inderjit S. Dhillon, *Member, IEEE*, Robert W. Heath Jr., *Member, IEEE* and Thomas Strohmer

***Abstract*— Tight frames, also known as general Welch-Bound-Equality sequences, generalize orthonormal systems. Numerous applications—including communications, coding and sparse approximation—require finite-dimensional tight frames that possess additional structural properties. This paper proposes an alternating projection method that is versatile enough to solve a huge class of inverse eigenvalue problems, which includes the frame design problem. To apply this method, one only needs to solve a matrix nearness problem that arises naturally from the design specifications. Therefore, it is fast and easy to develop versions of the algorithm that target new design problems. Alternating projection will often succeed even if algebraic constructions are unavailable.**

To demonstrate that alternating projection is an effective tool for frame design, the article studies some important structural properties in detail. First, it addresses the most basic design problem—constructing tight frames with prescribed vector norms. Then, it discusses equiangular tight frames, which are natural dictionaries for sparse approximation. Last, it examines tight frames whose individual vectors have low peak-to-average-power ratio (PAR), which is a valuable property for CDMA applications. Numerical experiments show that the proposed algorithm succeeds in each of these three cases. The appendices investigate the convergence properties of the algorithm.

***Index Terms*— Frames, sequences, algorithms, eigenvalues and eigenfunctions, inverse problems, geometry, extremal problems, code division multiaccess**

I. Introduction

TIGHT FRAMES provide a natural generalization of orthonormal systems, and they arise in numerous practical and theoretical contexts [1]. There is no shortage of tight frames, and applications will generally require that the vectors comprising the frame have some kind of additional structure. For example, it might be necessary for the vectors to have specific Euclidean norms, or perhaps they should have small mutual inner products. Thus arises a design problem: How do you build a structured tight frame?

A. Contributions

To address the design question, this article proposes a numerical method based on alternating projection that builds on our work in [2], [3]. The algorithm alternately finds the nearest tight frame to a given ensemble of structured vectors; then it finds the ensemble of structured vectors nearest to the tight frame; and it repeats the process *ad infinitum*. This technique is analogous to the method of projection on convex sets (POCS) [4], [5], except that the class of tight frames is non-convex, which complicates the analysis significantly. Nevertheless, our alternating projection algorithm affords simple implementations, and it provides a quick route to solve difficult frame design problems. We argue that similar techniques apply to a huge class of inverse eigenvalue problems.

This article demonstrates the elegance and effectiveness of the alternating projection approach with several examples that are motivated by applications. First, we address the most basic frame design problem—building tight frames with prescribed vector norms. This problem arises when constructing signature signatures for direct-spread, synchronous code division multiaccess (DS-CDMA) systems [6]–[8]. Second, we discuss equiangular tight frames, which have the property that each pair of distinct vectors meets at the same (acute) angle. These frames have many applications in coding and communications [9]–[12], and they also form natural dictionaries for sparse approximation [13]–[15]. Third, we examine tight frames whose individual vectors have low peak-to-average-power ratio (PAR), which is another valuable property for DS-CDMA signatures [3]. Our experiments show that alternating projection outperforms some algorithms that were specifically designed to solve these problems.

The appendices investigate the convergence properties of the algorithm. Although alternating projection between subspaces and convex sets has been studied in detail, very few results are available for an alternating projection between two non-convex sets. This paper provides a rigorous treatment of the algorithm's behavior in this general setting by means of the theory of point-to-set maps. In particular, we establish a weak global convergence result, and we show that, under additional hypotheses, the algorithm exhibits stronger local convergence properties.

Note that there is a major conceptual difference between the use of finite models in the numerical calculation of infinite-dimensional frames and the design of finite-dimensional frames. In the former case, the finite model plays only an auxiliary role in the approximate computation of an infinite-dimensional tight frame [1]. In the latter case, the problem under consideration is already finite-dimensional, thus it does not involve discretization issues. In this paper, we consider only finite-dimensional tight frames.

J. A. Tropp is with the Inst. for Computational Engineering and Sciences (ICES), The University of Texas at Austin, Austin, TX 78712 USA, <jtropp@ices.utexas.edu>.

I. S. Dhillon is with the Dept. of Computer Sciences, The University of Texas at Austin, Austin, TX 78712 USA, <inderjit@cs.utexas.edu>.

R. W. Heath is with the Dept. of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712 USA, <rheath@ece.utexas.edu>.

T. Strohmer is with the Dept. of Mathematics, The University of California at Davis, Davis, CA 95616 USA, <strohmer@math.ucdavis.edu>.

B. Previous Work

At root, finite-dimensional frame design is an algebraic problem. It boils down to producing a structured matrix with certain *spectral* properties, which may require elaborate discrete and combinatorial mathematics. In the past, most design methods have employed these techniques. To appreciate the breadth of this literature, one might peruse Sarwate's recent survey paper about tight frames with unit-norm vectors [16]. The last few years have also seen some essentially algebraic algorithms that can construct tight frames with non-constant vector norms [7], [17], [18].

When algebraic methods work, they work brilliantly. A numerical approach like alternating projection can hardly hope to compete with the most profound insights of engineers and mathematicians. On the other hand, algebraic and combinatoric tools are not always effective. For example, we might require a structured tight frame for a vector space whose dimension is not a prime-power. Even in these situations, alternating projection will usually succeed. Moreover, it can help researchers develop the insight necessary for completing an algebraic construction. The power of alternating projection comes from replacing the difficult algebra with a simple analytic question: How does one find an ensemble of structured vectors nearest to a given tight frame? This minimization problem can usually be dispatched with standard tools, such as differential calculus or Karush–Kuhn–Tucker theory.

The literature does not offer many other numerical approaches to frame design. It appears that most current algorithms can be traced to the discovery by Rupf–Massey [6] and Viswanath–Anantharam [7] that tight frames with prescribed column norms are the optimal sequences for DS-CDMA systems. This application prompted a long series of papers, including [19]–[23], that describe iterative methods for constructing tight frames with prescribed column norms. These techniques are founded on an oblique characterization of tight frames as the minimizers of a quantity called *total squared correlation* (TSC). It is not clear how one could generalize these methods to solve different types of frame design problems. Moreover, the alternating projection approach that we propose significantly outperforms at least one of the TSC-minimization algorithms. Two of the algebraic methods that we mentioned above, [7] and [18], were also designed with the DS-CDMA application in mind, while the third algebraic method [17] comes from the *soi-disant* frame community. We are not aware of any other numerical methods for frame design.

C. Outline

Section II continues with a short introduction to tight frames. Then, in Section III, we state two formal frame design problems. Connections among frame design problems, inverse eigenvalue problems and matrix nearness problems are established. This provides a natural segue to the alternating projection algorithm. Afterward, we apply the basic algorithm to design three different types of structured frames, in order of increasing implementation difficulty. Section IV discusses tight frames with prescribed column norms; Section V covers equiangular tight frames; and Section VI constructs tight frames whose individual vectors have low peak-to-average-power ratio. Each of these sections contains numerical experiments. The body of the paper concludes with Section VII, which discusses the method, its limitations and its extensions.

The back matter contains the bulk of the analysis. Appendix I offers a tutorial on point-to-set maps, and Appendix II applies this theory to obtain a rigorous characterization of the algorithm's convergence behavior. The first appendix also contains a brief survey of the alternating projection literature.

II. Tight Frames

This section offers a terse introduction to the properties of tight frames that are essential for our method. For more details, see [1], for example.

A. Frames

Let α and β be positive constants. A *finite frame* for the complex[1] Hilbert space $\mathbb{C}^d$ is a sequence of N vectors $\{\boldsymbol{x}_n\}_{n=1}^N$ drawn from $\mathbb{C}^d$ that satisfies a generalized Parseval condition:

$$\alpha \, \|\boldsymbol{v}\|_2^2 \le \sum_{n=1}^{N} |\langle \boldsymbol{v}, \boldsymbol{x}_n \rangle|^2 \le \beta \, \|\boldsymbol{v}\|_2^2 \quad \text{for all } \boldsymbol{v} \in \mathbb{C}^d. \tag{1}$$

We denote the Euclidean inner product with $\langle \cdot, \cdot \rangle$, and we write $\|\cdot\|_2$ for the associated norm. The numbers α and β are called the *lower* and *upper frame bounds*. The number of vectors in the frame may be no smaller than the dimension of the space (i.e. $N \ge d$).

If it is possible to take $\alpha = \beta$, then we have a *tight frame* or an α*-tight frame*. When the frame vectors all have unit norm, i.e. $\|\boldsymbol{x}_n\|_2 \equiv 1$, the system is called a *unit-norm frame*. Unit-norm tight frames are also known as *Welch-Bound-Equality sequences* [12], [24]. Tight frames with non-constant vector norms have also been called *general Welch-Bound-Equality sequences* [7].

B. Associated Matrices

Form a $d \times N$ matrix with the frame vectors as its columns:

$$\mathsf{X} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 & \dots & \boldsymbol{x}_N \end{bmatrix}.$$

This matrix is referred to as the *frame synthesis operator*, but we shall usually identify the synthesis operator with the frame itself.

Two other matrices arise naturally in connection with the frame. We define the *Gram matrix* as $\mathsf{G} \stackrel{\text{def}}{=} \mathsf{X}^* \mathsf{X}$. (The symbol * indicates conjugate transposition of matrices and vectors.) The diagonal entries of the Gram matrix equal the squared norms of the frame vectors, and the off-diagonal entries of the Gram matrix equal the inner products between distinct frame vectors. The Gram matrix is always Hermitian and positive semi-definite, and it has rank d.

[1] We work with complex vectors for complete generality. The adaptations for real vectors are transparent.

The positive-definite matrix $\mathsf{X}\mathsf{X}^*$ is usually called the *frame operator*. Since

$$\boldsymbol{v}^*(\mathsf{X}\mathsf{X}^*)\,\boldsymbol{v} = \sum_{n=1}^{N} |\langle \boldsymbol{v}, \boldsymbol{x}_n \rangle|^2$$

we can rewrite (1) as

$$\alpha \le \frac{\boldsymbol{v}^*(\mathsf{X}\mathsf{X}^*)\,\boldsymbol{v}}{\boldsymbol{v}^*\boldsymbol{v}} \le \beta. \tag{2}$$

That is, any Rayleigh quotient of $\mathsf{X}\mathsf{X}^*$ must lie between α and β. It follows from the Courant–Fischer Theorem [25] that each eigenvalue of the frame operator lies in the interval $[\alpha, \beta]$.

When the frame is α-tight, the condition (2) is equivalent to the statement that $\mathsf{X}\mathsf{X}^* = \alpha\,\mathsf{I}_d$. Three other characterizations of an α-tight frame follow immediately.

Proposition 1: A $d \times N$ frame X is α-tight if and only if it satisfies one (hence all) of the following conditions.

1) All d non-zero singular values of X equal $\sqrt{\alpha}$.
2) All d non-zero eigenvalues of the Gram matrix $\mathsf{X}^*\mathsf{X}$ equal α.
3) The rows of $\alpha^{-1/2}\,\mathsf{X}$ form an orthonormal family.

These properties undergird our method for constructing tight frames. It is now clear that the being a tight frame is a *spectral* requirement on the matrix X.

C. Norms of Frame Vectors

Throughout this article, we shall denote the squared norms of the frame vectors as

$$c_n \stackrel{\text{def}}{=} \|\boldsymbol{x}_n\|_2^2\,.$$

There is an intimate relationship between the tightness parameter of an α-tight frame and the norms of its columns. The computation is straightforward:

$$\alpha\, d = \operatorname{Tr} \mathsf{X}\mathsf{X}^* = \operatorname{Tr} \mathsf{X}^*\mathsf{X} = \sum_{n=1}^{N} \|\boldsymbol{x}_n\|_2^2 = \sum_{n=1}^{N} c_n. \tag{3}$$

The notation $\operatorname{Tr}(\cdot)$ represents the matrix trace operator, which sums the diagonal entries of its argument.

A related point is that one cannot construct a tight frame with an arbitrary set of column norms. According to the Schur–Horn Theorem, a Hermitian matrix can exist if and only if its diagonal entries majorize[2] its eigenvalues [25], [26]. If X is a $d \times N$ tight frame, the d non-zero eigenvalues of its Gram matrix all equal $\sum_n c_n / d$. Meanwhile, the diagonal entries of the Gram matrix are $c_1, \dots, c_N$. In this case, the majorization condition is equivalent to the system of inequalities

$$0 \le c_k \le \frac{1}{d} \sum_{n=1}^{N} c_n \qquad \text{for each } k = 1, \dots, N. \tag{4}$$

It follows that a tight frame with squared column norms $c_1, \dots, c_N$ exists if and only if (4) holds. For an arbitrary set of column norms, the frames that are "closest" to being tight have been characterized in [7], [27].

[2]The literature equivocates about the direction of the majorization relation. We adopt the sense used by Horn and Johnson [25].

III. Design via Alternating Projections

This section begins with formal statements of two frame design problems. Next we establish a connection with inverse eigenvalue problems. It becomes clear that an alternating projection algorithm offers a simple and natural approach to general inverse eigenvalue problems, including both frame design problems. We then solve the matrix nearness problems that arise when implementing the proposed algorithm. The section concludes with a discussion of the algorithm's convergence properties.

A. Structured Tight Frames

Define the collection of $d \times N$ α-tight frames:

$$\mathscr{X}_\alpha \stackrel{\text{def}}{=} \{\mathsf{X} \in \mathbb{C}^{d \times N} : \mathsf{X}\mathsf{X}^* = \alpha\,\mathsf{I}_d\}. \tag{5}$$

We fix the tightness parameter α for simplicity. It is easy to extend our methods to situations where the tightness is not predetermined, and one can apply similar ideas to construct frames with prescribed upper and lower frame bounds, viz. the parameters α and β in (1). It is worth noting that $\mathscr{X}_\alpha$ is essentially the Stiefel manifold, which consists of all sets of d orthonormal vectors in $\mathbb{C}^N$ [28].

Let $\mathscr{S}$ denote a closed[3] collection of $d \times N$ matrices that possess some structural property. In the sequel, we shall explore several different structural constraints that have arisen in electrical engineering contexts. Section IV considers tight frames with specified column norms, and Section VI shows how to construct tight frames whose individual vectors have a low peak-to-average-power ratio. Many other structural properties are possible.

Each constraint set $\mathscr{S}$ raises a basic question.

Problem 1: Find a matrix in $\mathscr{S}$ that is minimally distant from $\mathscr{X}_\alpha$ with respect to a given norm.

If the two sets intersect, any solution to this problem is a structured tight frame. Otherwise, the problem requests a structured matrix that is "most nearly tight." A symmetric problem is to find a tight frame that is "most nearly structured."

B. Structured Gram Matrices

If the structural constraints restrict the inner products between frame vectors, it may be more natural to work with Gram matrices. Define a collection that contains the Gram matrices of all $d \times N$ α-tight frames:

$$\begin{aligned}\mathscr{G}_\alpha \stackrel{\text{def}}{=} \{\mathsf{G} \in \mathbb{C}^{N \times N} &: \mathsf{G} = \mathsf{G}^* \text{ and} \\ &\mathsf{G} \text{ has eigenvalues } (\underbrace{\alpha, \dots, \alpha}_{d}, 0, \dots, 0)\}.\end{aligned} \tag{6}$$

The set $\mathscr{G}_\alpha$ is essentially the Grassmannian manifold that consists of d-dimensional subspaces of $\mathbb{C}^N$ [29]. One may also identify the matrices in $\mathscr{G}_\alpha$ as rank-d orthogonal projectors, scaled by α. (An orthogonal projector can be defined as an

[3]We equip $\mathbb{C}^{d \times N}$ and $\mathbb{C}^{N \times N}$ with the topology induced by the Frobenius norm, which is identical with every other norm topology [25].

idempotent, Hermitian matrix. The rank of a projector equals the dimension of its range.)

Let $\mathscr{H}$ be a closed collection of $N \times N$ Hermitian matrices that possess some structural property. In Section V, for example, we shall consider equiangular tight frames. The Gram matrices of these frames have off-diagonal entries with identical moduli, and it is an important challenge to construct them.

Once again, a fundamental question arises.

Problem 2: Find a matrix in $\mathscr{G}_\alpha$ that is minimally distant from $\mathscr{H}$ with respect to a given norm.

If the two sets intersect, any solution to this problem will lie in the intersection. Otherwise, the problem requests a tight frame whose Gram matrix is "most nearly structured." We do not mention the problem of producing a matrix in $\mathscr{H}$ that is nearest to $\mathscr{G}_\alpha$ because it is not generally possible to factor a matrix in $\mathscr{H}$ to obtain a frame with dimensions $d \times N$.

C. Inverse Eigenvalue Problems

We view Problems 1 and 2 as *inverse eigenvalue problems* (IEPs). As Chu explains in [30], an IEP is an inquiry about structured matrices with prescribed spectral properties. These spectral properties may include restrictions on eigenvalues, eigenvectors, singular values or singular vectors. According to Proposition 1, the defining characteristic of a tight frame is its spectrum, so frame design is an IEP.

In the study of IEPs, the two fundamental questions are *solvability* and *computability*. The former problem is to find necessary or sufficient conditions under which a given IEP has a solution. The latter problem is how to produce a matrix that has given spectral properties and simultaneously satisfies a structural constraint. The solvability and computability of some classical IEPs have been studied extensively by the matrix analysis community, although many open problems still remain. The articles [30], [31] survey this literature.

Although specific IEPs may require carefully tailored numerical methods, there are a few general tools available. One scheme is the coordinate-free Newton method, which has been explored in [32]–[34]. Newton-type algorithms do not apply to most problems, and they only converge locally. Another general method is the projected gradient approach developed by Chu and Driessel in [35]. This technique involves numerical integration of a matrix differential equation, which relies on advanced software that may not be readily available. Another problem with Newton methods and projected gradient methods is that they may not handle repeated singular values well. This shortcoming makes them a poor candidate for constructing tight frames, which have only two distinct singular values.

This article concentrates on alternating projection, which has occasionally been used to solve inverse eigenvalue problems (in [36] and [37], for example). But alternating projection has not been recognized as a potential method for solving *any* type of inverse eigenvalue problem. The most general treatment of alternating projection in the IEP literature is probably [38], but the authors do not offer a serious analysis of their algorithm's behavior.

Here is the basic idea behind alternating projection. We seek a point of intersection between the set of matrices that satisfy a structural constraint and the set of matrices that satisfy a spectral constraint. An alternating projection begins at a matrix in the first set, from which it computes a matrix of minimal distance in the second set. Then the algorithm reverses the roles of the two sets and repeats the process *ad infinitum*. Alternating projection is easy to apply, and it is usually globally convergent in a weak sense, as we show later.

D. Alternating Projections

Let us continue with a formal presentation of the generic alternating projection method for solving inverse eigenvalue problems. Suppose that we have two collections, $\mathscr{Y}$ and $\mathscr{Z}$, of matrices with identical dimensions. Of course, we are imagining that one collection of matrices incorporates a spectral constraint while the other collection incorporates a structural constraint. To ensure that the algorithm is well-posed, assume that one collection is closed and the other is compact.

Algorithm 1 (Alternating Projection):

INPUT:

- An (arbitrary) initial matrix Y_0 with appropriate dimensions
- The number of iterations, J

OUTPUT:

- A matrix $\overline{Y}$ in $\mathscr{Y}$ and a matrix $\overline{Z}$ in $\mathscr{Z}$

PROCEDURE:

1) Initialize $j = 0$.
2) Find a matrix Z_j in $\mathscr{Z}$ such that
$$Z_j \in \arg\min_{Z \in \mathscr{Z}} \|Z - Y_j\|_F .$$
We use $\|\cdot\|_F$ to indicate the Frobenius norm.
3) Find a matrix Y_{j+1} in $\mathscr{Y}$ such that
$$Y_{j+1} \in \arg\min_{Y \in \mathscr{Y}} \|Y - Z_j\|_F .$$
4) Increment j by one.
5) Repeat Steps 2–4 until $j = J$.
6) Let $\overline{Y} = Y_J$ and $\overline{Z} = Z_{J-1}$.

A solution to the optimization problem in Step 2 is called a *projection* of Y_j onto $\mathscr{Z}$ in analogy with the case where $\mathscr{Z}$ is a linear subspace. Step 3 computes the projection of Z_j onto $\mathscr{Y}$. In a Hilbert space, it can be shown geometrically that a given point has a unique projection onto each closed, convex set. Projections onto general sets may not be uniquely determined, which fiercely complicates the analysis of Algorithm 1.

von Neumann, in 1933, was the first to consider alternating projection methods. He showed that if $\mathscr{Y}$ and $\mathscr{Z}$ are closed, linear subspaces of a Hilbert space, then alternating projection converges to the point in $\mathscr{Y} \cap \mathscr{Z}$ nearest to Y_0 [39]. In 1959, Cheney and Goldstein demonstrated that alternating projection between two compact, convex subsets of a Hilbert space always yields a point in one set at minimal distance from the opposite set [4]. These two results inspired the application of Algorithm 1 to the inverse eigenvalue problems, Problems 1 and 2. Of course, the constraint sets that we consider are generally *not* convex. For a more extensive discussion of the literature on alternating projection, turn to Appendix I-G.

To implement the alternating projection algorithm, one must first solve the minimization problems in Steps 2 and 3. For obvious reasons, these optimizations are called the *matrix nearness problems* associated with $\mathscr{Y}$ and $\mathscr{Z}$. Already there is an extensive literature on matrix nearness problems. See, for example, the articles [40]–[42], the survey [43] and many sections of the book [25]. Section III-F exhibits solutions to the nearness problems associated with the spectral constraints $\mathscr{X}_\alpha$ and $\mathscr{G}_\alpha$. Even when it is necessary to solve a new nearness problem, this task often reduces to an exercise in differential calculus. This is one of the great advantages of Algorithm 1. In this article, we shall always measure the distance between matrices using the Frobenius norm $\|\cdot\|_{\mathrm{F}}$ because it facilitates the solution of matrix nearness problems. Of course, one could develop a formally identical algorithm using other norms, metrics or divergences.

Since the constraint sets are generally non-convex, alternating projection may not converge as well as one might wish. This explains why we have chosen to halt the algorithm after a fixed number of steps instead of waiting for $\|Y_j - Y_{j+1}\|_{\mathrm{F}}$ to decline past a certain threshold. Indeed, it is theoretically possible that the sequence of iterates will not converge in norm. In practice, it appears that norm convergence always occurs. Section III-G provides a skeletal discussion of the theoretical convergence of alternating projection. We do not flesh out the analysis until Appendices I and II because a proper treatment requires some uncommon mathematics.

E. Application to Problems 1 and 2

To solve Problem 1, we propose an alternating projection between the structural constraint set $\mathscr{S}$ and the spectral constraint set $\mathscr{X}_\alpha$. Two matrix nearness problems arise. In the next subsection, we demonstrate how to find a tight frame in $\mathscr{X}_\alpha$ nearest to an arbitrary matrix. Sections IV and VI contain detailed treatments of two different structural constraints.

To solve Problem 2, we alternate between the spectral constraint $\mathscr{G}_\alpha$ and the structural constraint $\mathscr{H}$. In the next subsection, we show how to produce a matrix in $\mathscr{G}_\alpha$ that is nearest to an arbitrary matrix. In Section V, we analyze a specific structural constraint $\mathscr{H}$. After performing the alternating projection, it may still be necessary to extract a frame X from the output Gram matrix. This is easily accomplished with a rank-revealing QR factorization or with an eigenvalue decomposition. Refer to [44] for details.

F. Nearest Tight Frames

Standard tools of numerical linear algebra can be used to calculate an α-tight frame that is closest to an arbitrary matrix in Frobenius norm.

Theorem 2: Let $N \geq d$, and suppose that the $d \times N$ matrix Z has singular value decomposition $U \Sigma V^*$. With respect to the Frobenius norm, a nearest α-tight frame to Z is given by $\alpha\, UV^*$. Note that UV^* is the unitary part of a polar factorization of Z.

Assume in addition that Z has full row-rank. Then $\alpha\, UV^*$ is the *unique* α-tight frame closest to Z. Moreover, one may compute UV^* using the formula $(ZZ^*)^{-1/2}\, Z$.

Proof: The proof of this well-known result is similar to that of Theorem 3, which appears below. See also pp. 431–432 of [25]. Classical references on related problems include [45], [46]. The formula for the polar factor may be verified with a direct calculation. ∎

It is also straightforward to compute a matrix in $\mathscr{G}_\alpha$ nearest to an arbitrary Hermitian matrix. This theorem appears to be novel, so we provide a short demonstration.

Theorem 3: Suppose that Z is an $N \times N$ Hermitian matrix with a unitary factorization $U \Lambda U^*$, where the entries of Λ are arranged in algebraically non-increasing order. Let U_d be the $N \times d$ matrix formed from the first d columns of U. Then $\alpha\, U_d {U_d}^*$ is a matrix in $\mathscr{G}_\alpha$ that is closest to Z with respect to the Frobenius norm. This closest matrix is unique if and only if λ_d strictly exceeds λ_{d+1}.

Proof: We must minimize $\|Z - \alpha\, G\|_{\mathrm{F}}$ over all rank-d orthogonal projectors G. Square and expand this objective function:

$$\|Z - \alpha\, G\|_{\mathrm{F}}^2 = \|Z\|_{\mathrm{F}}^2 + \alpha^2\, \|G\|_{\mathrm{F}}^2 - 2\alpha\, \mathrm{Re}\,\mathrm{Tr}\, G^* Z.$$

The squared Frobenius norm of an orthogonal projector equals its rank, so we only need to maximize the (negation of) the last term.

Every rank-d orthogonal projector G can be written as $G = VV^*$, where the $N \times d$ matrix V satisfies $V^* V = \mathrm{I}_d$. Meanwhile, we may factor Z into its eigenvalue decomposition $U \Lambda U^*$, where U is unitary and Λ is a diagonal matrix with non-increasing, real entries. Using the properties of the trace operator, we calculate that

$$\begin{aligned} \mathrm{Re}\,\mathrm{Tr}\, G^* Z &= \mathrm{Re}\,\mathrm{Tr}\,(VV^*)(U \Lambda U^*) \\ &= \mathrm{Re}\,\mathrm{Tr}\,(U^* VV^* U)\, \Lambda \\ &= \mathrm{Re} \sum_{n=1}^{N} (U^* VV^* U)_{nn}\, \lambda_n. \end{aligned}$$

Observe that $U^* VV^* U$ is a positive semi-definite matrix whose eigenvalues do not exceed one. Therefore, the diagonal entries of $U^* VV^* U$ are real numbers that lie between zero and one inclusive. Moreover, these diagonal entries must sum to d because

$$\mathrm{Tr}\, U^* VV^* U = \|U^* V\|_{\mathrm{F}}^2 = \|V\|_{\mathrm{F}}^2 = d.$$

It follows that

$$\max_{G}\, \mathrm{Re}\,\mathrm{Tr}\, G^* Z \quad \leq \quad \sum_{n=1}^{d} \lambda_n.$$

The bound is met whenever the diagonal of $U^* VV^* U$ contains d ones followed by $(N - d)$ zeroes. This sufficient condition for attainment can be written

$$U^* VV^* U \;=\; \mathrm{I}_d \,\oplus\, 0_{N-d}.$$

Furthermore, if $\lambda_d > \lambda_{d+1}$, this condition is also necessary.

Form a matrix U_d by extracting the first d columns of U. Then the sufficient condition holds whenever $G = VV^* =$

$U_d U_d^*$. That is, G is the orthogonal projector onto any d-dimensional subspace spanned by eigenvectors corresponding to the d algebraically largest eigenvalues of Z. If $\lambda_d > \lambda_{d+1}$, this subspace is uniquely determined. The orthogonal projector onto a fixed subspace is always unique, and the uniqueness claim follows. ∎

It may be valuable to know that there are specialized algorithms for performing the calculations required by Theorems 2 and 3. For example, Higham has developed stable numerical methods for computing the polar factor of a matrix [47], [48] that are more efficient than computing a singular value decomposition or applying the formula $(ZZ^*)^{-1/2} Z$.

G. Basic Convergence Results

It should be clear that alternating projection never increases the distance between successive iterates. This does not mean that it will locate a point of minimal distance between the constraint sets. It can be shown, however, that Algorithm 1 is globally convergent in a weak sense.

Define the distance between a point M and a set $\mathscr{Y}$ via

$$\operatorname{dist}(M, \mathscr{Y}) = \inf_{Y \in \mathscr{Y}} \|Y - M\|_{\mathrm{F}}.$$

Theorem 4 (Global Convergence of Algorithm 1): Let $\mathscr{Y}$ and $\mathscr{Z}$ be closed sets, one of which is bounded. Suppose that alternating projection generates a sequence of iterates $\{(Y_j, Z_j)\}$. This sequence has at least one accumulation point.

- Every accumulation point lies in $\mathscr{Y} \times \mathscr{Z}$.
- Every accumulation point $(\overline{Y}, \overline{Z})$ satisfies
$$\left\|\overline{Y} - \overline{Z}\right\|_{\mathrm{F}} = \lim_{j \to \infty} \|Y_j - Z_j\|_{\mathrm{F}}.$$
- Every accumulation point $(\overline{Y}, \overline{Z})$ satisfies
$$\left\|\overline{Y} - \overline{Z}\right\|_{\mathrm{F}} = \operatorname{dist}(\overline{Y}, \mathscr{Z}) = \operatorname{dist}(\overline{Z}, \mathscr{Y}).$$

For a proof of Theorem 4, turn to Appendix II-A. In some special cases, it is possible to develop stronger convergence results and characterizations of the fixed points. We shall mention these results where they are relevant. The convergence of Algorithm 1 is geometric at best [49]–[52]. This is the major shortfall of alternating projection methods.

Note that the sequence of iterates may have many accumulation points. Moreover, the last condition does not imply that the accumulation point $(\overline{Y}, \overline{Z})$ is a fixed point of the alternating projection. So what are the potential accumulation points of a sequence of iterates? Since the algorithm never increases the distance between successive iterates, the set of accumulation points includes every pair of matrices in $\mathscr{Y} \times \mathscr{Z}$ that lie at minimal distance from each other. Therefore, we say that the algorithm tries to solve Problems 1 and 2.

IV. Prescribed Column Norms

As a first illustration of alternating projection, let us consider the most basic frame design problem: How does one build a tight frame with prescribed column norms?

This question has arisen in the context of constructing optimal signature sequences for direct-spread synchronous code-division multiple-access (DS-CDMA) channels. There are some finite algorithms available that yield a small number of solutions to the problem [7], [18]. These methods exploit the connection between frames and the Schur–Horn Theorem. They work by applying plane rotations to an initial tight frame to adjust its column norms while maintaining its tightness. Casazza and Leon have also developed a finite method that seems different in spirit [17].

To construct larger collections of frames, some authors have proposed iterative algorithms [19]–[23]. These techniques attempt to minimize the *total squared correlation* (TSC) of an initial matrix subject to constraints on its column norms. The TSC of a matrix is defined as

$$\operatorname{TSC}(S) \stackrel{\text{def}}{=} \|S^* S\|_{\mathrm{F}}^2 = \sum_{m,n} |\langle \boldsymbol{s}_m, \boldsymbol{s}_n \rangle|^2.$$

If we fix the squared column norms of S to be $c_1, \dots, c_N$, a short algebraic manipulation shows that minimizing the TSC is equivalent to solving

$$\min_S \|SS^* - \alpha \, \mathrm{I}_d\|_{\mathrm{F}}$$

where $\alpha = \sum_n c_n / d$. In words, minimizing the TSC is equivalent to finding a frame with prescribed column norms that is closest in Frobenius norm to a tight frame [53].

In comparison, alternating projection affords an elegant way to produce many tight frames with specified column norms. It focuses on the essential property of a tight frame—its singular values—to solve the problem. In this special case, we provide a complete accounting of the behavior of the alternating projection. Moreover, experiments show that it outperforms some of the other iterative algorithms that were developed specifically for this problem.

A. Constraint Sets and Nearness Problems

The algorithm will alternate between the set of matrices with fixed column norms and the set of tight frames with an appropriate tightness parameter.

Let the positive numbers $c_1, \dots, c_N$ denote the squared column norms that are desired. We do not require that these numbers satisfy the majorization inequalities given in (4), although one cannot hope to find a tight frame if these inequalities fail. In that case, we would seek a matrix with the prescribed column norms that is closest to being a tight frame. In the DS-CDMA application, the column norms depend on the users' power constraints [6], [7].

In light of (3), the relation between the tightness parameter and the column norms, it is clear that α must equal $\sum_n c_n / d$. The spectral constraint set becomes

$$\mathscr{X}_\alpha \stackrel{\text{def}}{=} \{X \in \mathbb{C}^{d \times N} : XX^* = (\textstyle\sum_n c_n / d) \, \mathrm{I}_d\}.$$

Given an arbitary $d \times N$ matrix, one may compute the closest tight frame in $\mathscr{X}_\alpha$ using Theorem 2.

The structural constraint set contains matrices with the correct column norms.

$$\mathscr{S} \stackrel{\text{def}}{=} \{S \in \mathbb{C}^{d \times N} : \|\boldsymbol{s}_n\|_2^2 = c_n\}.$$

It is straightforward to solve the matrix nearness problem associated with this collection.

Proposition 5: Let Z be an arbitrary matrix with columns $\{\boldsymbol{z}_n\}$. A matrix in $\mathscr{S}$ is closest to Z in Frobenius norm if and only if it has the columns

$$\boldsymbol{s}_n = \begin{cases} c_n\,\boldsymbol{z}_n / \left\|\boldsymbol{z}_n\right\|_2, & \boldsymbol{z}_n \neq \boldsymbol{0} \quad \text{and} \\ c_n\,\boldsymbol{u}_n, & \boldsymbol{z}_n = \boldsymbol{0}, \end{cases}$$

where $\boldsymbol{u}_n$ represents an arbitrary unit vector. If the columns of Z are all non-zero, then the solution to the nearness problem is unique.

Proof: We must minimize $\|S - Z\|_{\mathrm{F}}$ over all matrices S from $\mathscr{S}$. Square and rewrite this objective function:

$$\|S - Z\|_{\mathrm{F}}^2 = \sum_{n=1}^{N} \|\boldsymbol{s}_n - \boldsymbol{z}_n\|_2^2 .$$

We can minimize each summand separately. Fix an index n, and expand the n-th term using $\|\boldsymbol{s}_n\|_2^2 = c_n$.

$$\|\boldsymbol{s}_n - \boldsymbol{z}_n\|_2^2 = c_n + \|\boldsymbol{z}_n\|_2^2 - 2\sqrt{c_n}\ \mathrm{Re}\left\langle \frac{\boldsymbol{s}_n}{\|\boldsymbol{s}_n\|_2}, \boldsymbol{z}_n \right\rangle.$$

If $\boldsymbol{z}_n \neq \boldsymbol{0}$, the unique maximizer of $\mathrm{Re}\langle \boldsymbol{u}, \boldsymbol{z}_n\rangle$ over all unit vectors is $\boldsymbol{u} = \boldsymbol{z}_n / \|\boldsymbol{z}_n\|_2$. If $\boldsymbol{z}_n = \boldsymbol{0}$, then every unit vector $\boldsymbol{u}$ maximizes the inner product. ∎

B. Convergence Results

In this setting, alternating projection converges in a fairly strong sense.

Theorem 6: Let S_0 have full rank and non-zero columns, and suppose that the alternating projection generates a sequence of iterates $\{(S_j, X_j)\}$. This sequence possesses at least one accumulation point, say $(\overline{S}, \overline{X})$.

- Both $\overline{S}$ and $\overline{X}$ have full rank and non-zero columns.
- The pair $(\overline{S}, \overline{X})$ is a fixed point of the alternating projection. In other words, if we applied the algorithm to $\overline{S}$ or to $\overline{X}$ every pair of iterates would equal $(\overline{S}, \overline{X})$.
- The accumulation point satisfies
$$\left\|\overline{S} - \overline{X}\right\|_{\mathrm{F}} = \lim_{j\to\infty} \|S_j - X_j\|_{\mathrm{F}} .$$
- The component sequences are asymptotically regular, i.e.
$$\|S_{j+1} - S_j\|_{\mathrm{F}} \to 0 \quad \text{and} \quad \|X_{j+1} - X_j\|_{\mathrm{F}} \to 0.$$
- *Either* the component sequences both converge in norm,
$$\left\|S_j - \overline{S}\right\|_{\mathrm{F}} \to 0 \quad \text{and} \quad \left\|X_j - \overline{X}\right\|_{\mathrm{F}} \to 0,$$
or the set of accumulation points forms a continuum.

Proof: See Appendix II-C. ∎

In the present case, it is also possible to characterize completely the fixed points of the algorithm that lie in $\mathscr{S}$.

Proposition 7: A full-rank matrix S from $\mathscr{S}$ is a fixed point of the alternating projection between $\mathscr{S}$ and $\mathscr{X}_\alpha$ if and only if its columns are all eigenvectors of SS^*. That is, $SS^*S = S\Lambda$, where $\Lambda \in \mathbb{C}^{N\times N}$ is diagonal and positive with no more than d distinct entries.

Proof: Refer to Appendix II-D. ∎

Many of the fixed points in $\mathscr{S}$ do not lie at minimal distance from $\mathscr{X}_\alpha$, so they are not solutions to Problem 1. Nevertheless, the fixed points still have a tremendous amount of structure. Each fixed point can be written as a union of tight frames for mutually orthogonal subspaces of $\mathbb{C}^d$, and the set of fixed points is identical with the set of critical points of the TSC functional subject to the column norm constraint [23], [53]. The Ulukus-Yates algorithm, another iterative method for designing tight frames with specified column norms, has identical fixed points [20].

C. Numerical Examples

We offer a few simple examples to illustrate that the algorithm succeeds, and we provide some comparisons with the Ulukus-Yates algorithm.

Suppose first that we wish to construct a unit-norm tight frame for $\mathbb{R}^3$ consisting of five vectors. Initialized with a 3×5 matrix whose columns are chosen uniformly at random from the surface of the unit sphere, the algorithm returns

$$\overline{S} = \begin{bmatrix} 0.1519 & 0.4258 & -0.7778 & 0.0160 & -0.9258 \\ 0.9840 & -0.6775 & 0.1882 & 0.3355 & -0.3024 \\ -0.0926 & 0.5998 & 0.5997 & 0.9419 & -0.2269 \end{bmatrix}.$$

Each column norm of the displayed matrix equals one to machine precision, and the singular values are identical in their first eight digits. In all the numerical examples, the algorithm was terminated on the condition that $\|S_{j+1} - S_j\|_{\mathrm{F}} < 10^{-8}$. Implemented in Matlab, the computation took 65 iterations, which lasted 0.0293 seconds on a 1.6 GHz machine.

Now let us construct a tight frame for $\mathbb{R}^3$ whose five vectors have norms 0.75, 0.75, 1, 1.25 and 1.25. With random initialization, we obtain

$$\overline{S} = \begin{bmatrix} -0.1223 & 0.1753 & -0.7261 & 0.0128 & -1.0848 \\ 0.7045 & -0.6786 & 0.6373 & 0.0972 & -0.6145 \\ -0.2263 & 0.2670 & 0.2581 & 1.2461 & -0.0894 \end{bmatrix}.$$

The column norms are correct to machine precision, and the singular values are identical to seven digits. The computation took 100 iterations, which lasted 0.0487 seconds.

Next we examine a case where the column norms do not satisfy the majorization condition. Suppose that we seek a "nearly tight" frame with column norms 0.5, 0.5, 1, 1 and 2. Random initialization yields

$$\overline{S} = \begin{bmatrix} -0.1430 & 0.1353 & -0.4351 & -0.0941 & -1.8005 \\ 0.4293 & -0.4213 & 0.7970 & -0.2453 & -0.7857 \\ -0.2127 & 0.2329 & 0.4189 & 0.9649 & -0.3754 \end{bmatrix}.$$

The column norms are all correct, but, as predicted, the frame is not tight. Nevertheless, the last vector is orthogonal to the first four vectors, which form a tight frame for their span. This is an optimal solution to the frame design problem. The calculation required 34 iterations over 0.0162 seconds.

Of course, alternating projection can produce complex-valued tight frames, as well as larger frames in higher-dimensional spaces. Such ensembles are too large to display in these columns. To give a taste of the algorithm's general performance, we have compared it with our implementation of the Ulukus-Yates algorithm [20]. To construct unit-norm

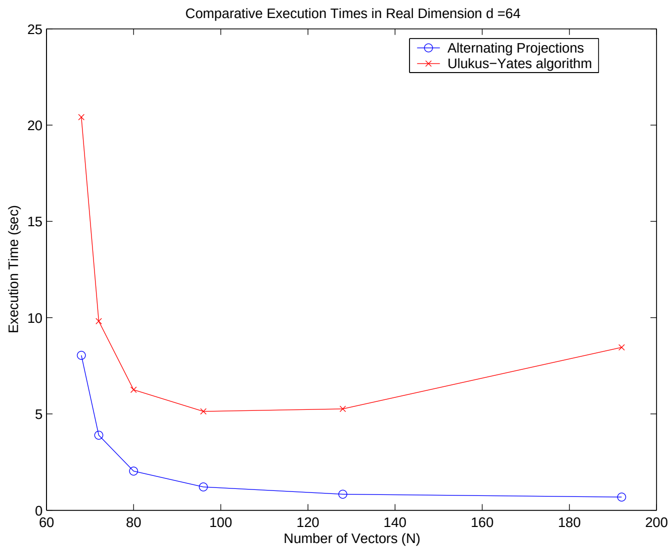

Fig. 1. Comparison of alternating projection with the Ulukus-Yates algorithm in 64 real dimensions.

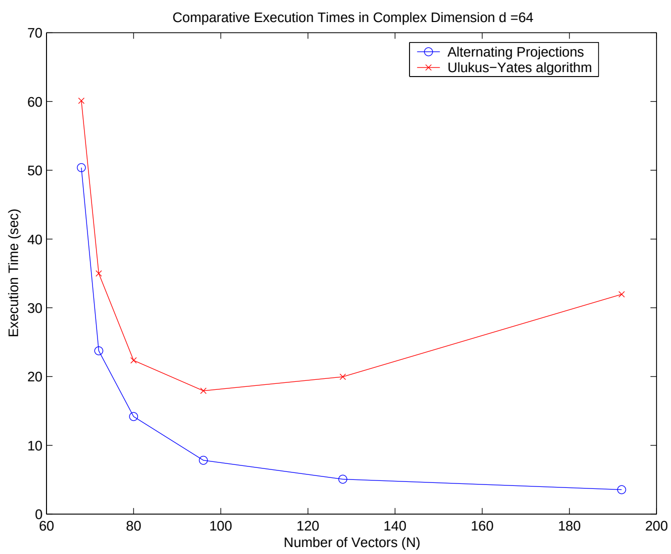

Fig. 2. Comparison of alternating projection with the Ulukus-Yates algorithm in 64 complex dimensions.

tight frames of various lengths, we initialized each algorithm with the same random matrix. Then we plotted the comparative execution times. Figure 1 shows the results for 64 real dimensions, and Figure 2 shows the results for 64 complex dimensions. Note the different scales on the time axes.

Both algorithms perform slowly when N is small because tight frames are relatively scarce, which makes them difficult to find. Indeed, it is known that (modulo rotations) there exists a unique tight frame of $(d+1)$ vectors in d-dimensional space [54]. Another reason that the alternating projection algorithm performs better as the problem grows is that a collection of N uniformly random unit-vectors converges almost surely to a tight frame as N tends to infinity [55]. It is therefore perplexing that the Ulukus-Yates algorithm performs more and more slowly. One might attribute this behavior to the fact that the algorithm does not act to equalize the singular values of the frame.

V. Equiangular Tight Frames

In this section, we shall consider a frame design problem that leads to a simple structural constraint on the Gram matrix. The goal of the alternating projection will be to design a suitable Gram matrix, from which the frame may be extracted afterward.

A tight frame is a generalization of an orthonormal basis because they share the Parseval property. But orthonormal bases have other characteristics that one may wish to extend. In particular, every orthonormal basis is *equiangular*. That is, each pair of distinct vectors has the same inner product, namely zero. This observation suggests that one seek out equiangular tight frames. The underlying intuition is that these frames will contain vectors maximally separated in space.

Define an *equiangular tight frame* to be a unit-norm tight frame in which each pair of columns has the same *absolute* inner product. Since we are considering unit-norm tight frames, the absolute inner product between two frame vectors equals the cosine of the acute angle between the one-dimensional subspaces spanned by the two vectors. For this reason are the frames called equiangular. One can show that each inner product in an equiangular tight frame has modulus

$$\mu \stackrel{\text{def}}{=} \sqrt{\frac{N-d}{d\,(N-1)}}. \tag{7}$$

It is a remarkable fact that *every* ensemble of N unit vectors in d dimensions contains a pair whose absolute inner product strictly exceeds μ, *unless* the vectors form an equiangular tight frame. Unfortunately, equiangular tight frames only exist for rare combinations of d and N. In particular, a real, equiangular tight frame can exist only if $N \le \frac{1}{2}\, d\,(d+1)$, while a complex, equiangular tight frame requires that $N \le d^2$ [12]. The paper [56] contains detailed necessary conditions on real, equiangular tight frames and on equiangular tight frames over finite alphabets.

One can view equiangular tight frames as a special type of *Grassmannian frame*. In finite dimensions, Grassmannian frames are unit-norm frames whose largest absolute inner product is minimal for a given d and N [12]. Their name is motivated by the fact that they correspond with sphere packings in the Grassmannian manifold of all one-dimensional subspaces of a vector space [29]. Grassmannian frames have applications in coding theory and communications engineering [9]–[12]. They also provide a natural set of vectors to use for sparse approximation [13]–[15].

In general, it is torturous to design Grassmannian frames. Not only is the optimization difficult, but there is no general procedure for deciding when a frame solves the optimization problem unless it meets a known lower bound. Most of the current research has approached the design problem with algebraic tools. A notable triumph of this type is the construction of Kerdock codes over $\mathbb{Z}_2$ and $\mathbb{Z}_4$ due to Calderbank *et al.* [57]. Other explicit constructions are discussed in the articles [10], [12]. In the numerical realm, Sloane has used his Gosset software to produce and study sphere packings in real Grassmannian spaces [58]. Sloane's algorithms have been extended to complex Grassmannian spaces in [59]. We are not aware of any other numerical methods.

In this article, we shall construct equiangular tight frames for real and complex vector spaces using alternating projection. The method can easily be extended to compute other finite Grassmannian frames and packings in higher Grassmannian manifolds, but that is another paper for another day [60].

A. Constraint Sets and Nearness Problems

The signal of an equiangular tight frame is that each inner product between distinct vectors has the same modulus. Since the Gram matrix of a tight frame displays all of the inner products, it is more natural to construct the Gram matrix of an equiangular tight frame than to construct the frame synthesis matrix directly. Therefore, the algorithm will alternate between the collection of Hermitian matrices that have the correct spectrum and the collection of Hermitian matrices that have sufficiently small off-diagonal entries.

Since we are working with unit-norm tight frames, the tightness parameter α must equal N/d. This leads to the spectral constraint set

$$\mathscr{G}_\alpha \stackrel{\text{def}}{=} \{ G \in \mathbb{C}^{N \times N} : G = G^* \text{ and } G \text{ has eigenvalues } (\underbrace{N/d, \dots, N/d}_{d}, 0, \dots, 0) \}.$$

Theorem 3 shows how to find a matrix in $\mathscr{G}_\alpha$ nearest to an arbitrary Hermitian matrix.

In an equiangular tight frame, each vector has unit norm but no two vectors have inner product larger than μ. Therefore, we define the structural constraint set

$$\mathscr{H}_\mu \stackrel{\text{def}}{=} \{ H \in \mathbb{C}^{N \times N} : H = H^*, \operatorname{diag} H = \mathbf{1} \text{ and } \max_{m \neq n} |h_{mn}| \le \mu \}.$$

It may seem more natural to require that the off-diagonal entries have modulus exactly equal to μ, but our experience indicates that the present formulation works better, perhaps because $\mathscr{H}_\mu$ is convex. The following proposition shows how to produce the nearest matrix in $\mathscr{H}_\mu$.

Proposition 8: Let Z be an arbitrary matrix. With respect to Frobenius norm, the unique matrix in $\mathscr{H}_\mu$ closest to Z has a unit diagonal and off-diagonal entries that satisfy

$$h_{mn} = \begin{cases} z_{mn} & \text{if } |z_{mn}| \le \mu \text{ and} \\ \mu \, \mathrm{e}^{\mathrm{i} \arg z_{mn}} & \text{otherwise.} \end{cases}$$

We use i to denote the imaginary unit.

Proof: The argument is straightforward. ■

B. Convergence Results

The general convergence result, Theorem 4, applies to the alternating projection between $\mathscr{G}_\alpha$ and $\mathscr{H}_\mu$. We also obtain a local convergence result.

Theorem 9: Assume that the alternating projection between $\mathscr{G}_\alpha$ and $\mathscr{H}_\mu$ generates a sequence of iterates $\{(G_j, H_j)\}$, and suppose that there is an iteration J during which $\|G_J - H_J\|_{\mathrm{F}} < N/(d\sqrt{2})$. The sequence of iterates possesses at least one accumulation point, say $(\overline{G}, \overline{H})$.

- The accumulation point lies in $\mathscr{G}_\alpha \times \mathscr{H}_\mu$.
- The pair $(\overline{G}, \overline{H})$ is a fixed point of the alternating projection. In other words, if we applied the algorithm to $\overline{G}$ or to $\overline{H}$, every iterate would equal $(\overline{G}, \overline{H})$.
- The accumulation point satisfies

$$\|\overline{G} - \overline{H}\|_{\mathrm{F}} = \lim_{j \to \infty} \|G_j - H_j\|_{\mathrm{F}}.$$

- The component sequences are asymptotically regular, i.e.

$$\|G_{j+1} - G_j\|_{\mathrm{F}} \to 0 \quad \text{and} \quad \|H_{j+1} - H_j\|_{\mathrm{F}} \to 0.$$

- *Either* the component sequences both converge in norm,

$$\|G_j - \overline{G}\|_{\mathrm{F}} \to 0 \quad \text{and} \quad \|H_j - \overline{H}\|_{\mathrm{F}} \to 0,$$

or the set of accumulation points forms a continuum.

Proof: See Appendix II-B. ■

C. Numerical Examples

First, let us illustrate just how significant a difference there is between vanilla tight frames and equiangular tight frames. Here is the Gram matrix of a six-vector, unit-norm tight frame for $\mathbb{R}^3$:

$$\begin{bmatrix} 1.0000 & 0.2414 & -0.6303 & 0.5402 & -0.3564 & -0.3543 \\ 0.2414 & 1.0000 & -0.5575 & -0.4578 & 0.5807 & -0.2902 \\ -0.6303 & -0.5575 & 1.0000 & 0.2947 & 0.3521 & -0.2847 \\ 0.5402 & -0.4578 & 0.2947 & 1.0000 & -0.2392 & -0.5954 \\ -0.3564 & 0.5807 & 0.3521 & -0.2392 & 1.0000 & -0.5955 \\ -0.3543 & -0.2902 & -0.2847 & -0.5954 & -0.5955 & 1.0000 \end{bmatrix}.$$

Notice that the inner-products between vectors are quite disparate, ranging in magnitude between 0.2392 and 0.6303. These inner products correspond to acute angles of $76.2°$ and $50.9°$. In fact, this tight frame is pretty tame; the largest inner products in a unit-norm tight frame can be arbitrarily close to one[4]. The Gram matrix of a six-vector, equiangular tight frame for $\mathbb{R}^3$ looks quite different:

$$\begin{bmatrix} 1.0000 & 0.4472 & -0.4472 & 0.4472 & -0.4472 & -0.4472 \\ 0.4472 & 1.0000 & -0.4472 & -0.4472 & 0.4472 & -0.4472 \\ -0.4472 & -0.4472 & 1.0000 & 0.4472 & 0.4472 & -0.4472 \\ 0.4472 & -0.4472 & 0.4472 & 1.0000 & -0.4472 & -0.4472 \\ -0.4472 & 0.4472 & 0.4472 & -0.4472 & 1.0000 & -0.4472 \\ -0.4472 & -0.4472 & -0.4472 & -0.4472 & -0.4472 & 1.0000 \end{bmatrix}.$$

Every pair of vectors meets at an acute angle of $63.4°$. The vectors in this frame can be interpreted as the diagonals of an icosahedron [29].

We have used alternating projection to compute equiangular tight frames, both real and complex, in dimensions two through six. The algorithm performed poorly when initialized with random vectors, which led us to adopt a more sophisticated approach. We begin with many random vectors and winnow this collection down by repeatedly removing whatever vector has the largest inner product against another vector. It is fast and easy to design starting points in this manner, yet

	d						d				
N	2	3	4	5	6	N	2	3	4	5	6
3	$\mathbb{R}$	$\mathbb{R}$	..	..	..	20	..	..	..	.	.
4	$\mathbb{C}$	$\mathbb{R}$	$\mathbb{R}$	..	..	21	..	..	..	$\mathbb{C}$	.
5	..	.	$\mathbb{R}$	$\mathbb{R}$	..	22	..	..	..	.	.
6	..	$\mathbb{R}$	.	$\mathbb{R}$	$\mathbb{R}$	23	..	..	..	.	.
7	..	$\mathbb{C}$	$\mathbb{C}$	.	$\mathbb{R}$	24	..	..	..	.	.
8	..	.	$\mathbb{C}$	.	.	25	..	..	..	$\mathbb{C}$	.
9	..	$\mathbb{C}$	.	.	$\mathbb{C}$	26	..	..	..	..	.
10	..	..	.	$\mathbb{R}$	.	27	..	..	..	..	.
11	..	..	.	$\mathbb{C}$	$\mathbb{C}$	28	..	..	..	..	.
12	..	..	.	.	$\mathbb{C}$	29	..	..	..	..	.
13	..	..	$\mathbb{C}$	.	.	30	..	..	..	..	.
14	..	..	.	.	.	31	..	..	..	..	$\mathbb{C}$
15	..	..	.	.	.	32	..	..	..	..	.
16	..	..	$\mathbb{C}$	.	$\mathbb{R}$	33	..	..	..	..	.
17	..	..	..	.	.	34	..	..	..	..	.
18	..	..	..	.	.	35	..	..	..	..	.
19	..	..	..	.	.	36	..	..	..	..	$\mathbb{C}$

TABLE I

EQUIANGULAR TIGHT FRAMES

The notations $\mathbb{R}$ and $\mathbb{C}$ respectively indicate that alternating projection was able to compute a real, or complex, equiangular tight frame. Note that every real, equiangular tight frame is automatically a complex, equiangular tight frame. One period (.) means that no real, equiangular tight frame exists, and two periods (..) mean that no equiangular tight frame exists at all.

the results are impressive. These calculations are summarized in Table I.

Alternating projection can locate every real, equiangular tight frame in dimensions two through six; algebraic considerations eliminate all the remaining values of N [12], [56]. Moreover, the method computes these ensembles very efficiently. For example, the algorithm produced a six-vector, equiangular tight frame for $\mathbb{R}^3$ after a single trial. In this case, 70 iterations lasting 0.4573 seconds were sufficient to determine the first eight decimal places of the inner products.

In the complex case, the algorithm was able to compute every equiangular tight frame that we know of. Unfortunately, no one has yet developed necessary conditions on the existence of complex, equiangular tight frames aside from the upper bound, $N \leq d^2$, and so we have been unable to rule out the existence of other ensembles. Some of the computations progressed quite smoothly. After 1000 iterations and 18.75 seconds, alternating projection delivered a collection of 25 vectors in five dimensions whose inner products were identical in the first eight decimal places. On the other hand, it took 5000 iterations and 85.75 seconds to produce 21 vectors in five dimensions whose inner products reached the same level of accuracy. Even worse, we were unable to locate the 31-vector equiangular tight frame in $\mathbb{C}^6$ until we had performed two dozen random trials that lasted several minutes each. It is some consolation that the authors of [59] indicate their algorithm could not compute this ensemble at all.

It seems clear that some equiangular tight frames are much easier to find than others. We have encountered less success at constructing equiangular tight frames in higher dimensions. But we have neither performed extensive experiments nor have we attempted to fine-tune the method.

[4]To see this, consider a tight frame that contains two copies of an orthonormal basis, where one copy is rotated away from the other by an arbitrarily small angle.

VI. PEAK-TO-AVERAGE-POWER RATIO

Finally, let us present a situation in which the matrix nearness problem is much more difficult.

As we have mentioned, tight frames with prescribed vector norms coincide with signature sequences that maximize sum capacity in the uplink of direct-spread, synchronous code division multiple access (DS-CDMA) systems [6]–[8]. Unfortunately, general tight frames can have properties that are undesirable in practice. In particular, the individual frame vectors may have large *peak-to-average-power ratio* (PAR).

The PAR of an analog signal measures how the largest value of the signal compares with its average power. Signals with large PAR require higher dynamic range on the analog-to-digital converters and the digital-to-analog converters. They may also require more linear (and thus higher cost) power amplifiers. In DS-CDMA systems, the PAR is normally of concern only in the downlink (see e.g. [61]), where linear combinations of signatures can conspire to have tremendous peak power. On the uplink, the PAR problem is fundamentally different because it only involves individual signatures. Conventionally, the uplink PAR has not received attention because most systems use binary spreading sequences, which always have unit PAR. If general sum-capacity-optimal sequences are to be used in real systems, then PAR side constraints should be included in the design problem. Therefore, we shall consider how to construct tight frames whose columns have prescribed norms and low peak-to-average-power ratios.

As discussed in Section IV, many algorithms have already been developed for constructing tight frames with prescribed vector norms, such as [7], [19], [20], [22]. Unfortunately, these methods cannot accept additional constraints on the vectors, and thus they are not suitable for finding tight frames whose vectors have low PAR. We show that alternating projection provides a way to produce these ensembles. The PAR problem makes an interesting test case because it induces a matrix nearness problem that is considerably more challenging than those we have examined in previous sections.

A. Constraint Sets and Matrix Nearness Problems

The PAR in a digital communication system is fundamentally related to the analog waveforms that are generated. From the perspective of sequence design, it usually suffices to consider the PAR defined directly from the discrete sequence. The discrete PAR of a vector $\boldsymbol{z}$ is the quantity

$$\mathrm{PAR}(\boldsymbol{z}) \stackrel{\text{def}}{=} \frac{\max_m |z_m|^2}{\sum_m |z_m|^2 / d}.$$

Note that $1 \leq \mathrm{PAR}(\boldsymbol{z}) \leq d$. The lower extreme corresponds to a vector whose entries have identical modulus, while the upper bound is attained only by (scaled) canonical basis vectors.

Suppose that we require the columns of the frame to have squared norms $c_1, \dots, c_N$. In the DS-CDMA application, these numbers depend on the users' power constraints [6], [7]. It follows from (3) that $\alpha = \sum_n c_n/d$. The spectral constraint set becomes

$$\mathscr{X}_\alpha \stackrel{\text{def}}{=} \{X \in \mathbb{C}^{d,N} : XX^* = (\textstyle\sum_n c_n/d)\, \mathrm{I}_d\}.$$

Theorem 2 delivers the solution to the associated matrix nearness problem.

Let ρ denote the upper bound on the PAR of the frame elements. Then the structural constraint set will be

$$\mathscr{S} \stackrel{\text{def}}{=} \{S \in \mathbb{C}^{d\times N} : \mathrm{PAR}(\boldsymbol{s}_n) \le \rho \text{ and } \|\boldsymbol{s}_n\|_2^2 = c_n\}.$$

Given an arbitrary matrix Z, we must compute the nearest element of $\mathscr{S}$. Since the structural constraint on each column is independent and the Frobenius norm is separable, each column yields an independent optimization problem. For each column $\boldsymbol{z}_n$ of the input matrix, we claim that the following algorithm returns $\boldsymbol{s}_n$, the corresponding column of a nearest matrix S from $\mathscr{S}$.

Algorithm 2 (Nearest Vector with Low PAR):

INPUT:

- An input vector $\boldsymbol{z}$ from $\mathbb{C}^d$
- A positive number c, the squared norm of the solution vector
- A number ρ from $[1, d]$, which equals the maximum permissible PAR

OUTPUT:

- A vector $\boldsymbol{s}$ from $\mathbb{C}^d$ that solves

$$\min_{\boldsymbol{s}} \|\boldsymbol{s} - \boldsymbol{z}\|_2 \quad \text{subj. to} \quad \mathrm{PAR}(\boldsymbol{s}) \le \rho \text{ and } \|\boldsymbol{s}\|_2^2 = c.$$

PROCEDURE:

1) Scale $\boldsymbol{z}$ to have unit norm; define $\delta = \sqrt{c\,\rho/d}$; and initialize $k = 0$.
2) Let $\mathscr{M}$ index $(d-k)$ components of $\boldsymbol{z}$ with least magnitude. If this set is not uniquely determined, increment k and repeat Step 2.
3) If $z_m = 0$ for each m in $\mathscr{M}$, a solution vector is

$$\boldsymbol{s} = \begin{cases} \sqrt{\frac{c-k\,\delta^2}{d-k}} & \text{for } m \in \mathscr{M}, \text{ and} \\ \delta\, \mathrm{e}^{\mathrm{i}\,\arg z_m} & \text{for } m \notin \mathscr{M}. \end{cases}$$

4) Otherwise, let

$$\gamma = \sqrt{\frac{c - k\,\delta^2}{\sum_{m\in\mathscr{M}} |z_m|^2}}.$$

5) If $\gamma\, z_m > \delta$ for any m in $\mathscr{M}$, increment k and return to Step 2.
6) The unique solution vector is

$$\boldsymbol{s} = \begin{cases} \gamma\, z_m & \text{for } m \in \mathscr{M}, \text{ and} \\ \delta\, \mathrm{e}^{\mathrm{i}\,\arg z_m} & \text{for } m \notin \mathscr{M}. \end{cases}$$

When $\rho = 1$, the output of the algorithm is a unimodular vector whose entries have the same phase as the corresponding entries of $\boldsymbol{z}$. On the other hand, when $\rho = d$, the output vector equals $\boldsymbol{z}$. We now prove that the algorithm is correct.

Proof: We must solve the optimization problem

$$\min_{\boldsymbol{s}} \|\boldsymbol{s} - \boldsymbol{z}\|_2^2 \quad \text{subject to} \quad \mathrm{PAR}(\boldsymbol{s}) \le \rho \text{ and } \|\boldsymbol{s}\|_2^2 = c.$$

Let us begin with some major simplifications. First, rewrite the PAR constraint by enforcing the norm requirement and rearranging to obtain the equivalent condition

$$\max_m |s_m| \le \sqrt{c\,\rho/d}.$$

In the rest of the argument, the symbol δ will abbreviate the quantity $\sqrt{c\,\rho/d}$. The PAR constraint becomes $|s_m| \le \delta$ for each $m = 1, \dots, d$.

Now expand the objective function and enforce the norm constraint again to obtain

$$\min_{\boldsymbol{s}} \left[c - 2\,\mathrm{Re}\,\langle \boldsymbol{s}, \boldsymbol{z}\rangle + \|\boldsymbol{z}\|_2^2 \right].$$

Observe that it is necessary and sufficient to minimize the second term. It follows that the optimizer does not depend on the scale of the input vector $\boldsymbol{z}$. So take $\|\boldsymbol{z}\|_2 = 1$ without loss of generality.

Next observe that the PAR constraint and the norm constraint do not depend on the phases of the components in $\boldsymbol{s}$. Therefore, the components of an optimal $\boldsymbol{s}$ must have the same phases as the components of the input vector $\boldsymbol{z}$. In consequence, we may assume that both $\boldsymbol{s}$ and $\boldsymbol{z}$ are non-negative real vectors.

We have reached a much more straightforward optimization problem. Given a non-negative vector $\boldsymbol{z}$ with unit norm, we must solve

$$\max_{\boldsymbol{s}} \langle \boldsymbol{s}, \boldsymbol{z}\rangle \quad \text{subject to} \quad \langle \boldsymbol{s}, \boldsymbol{s}\rangle = c \text{ and } 0 \le s_m \le \delta,$$

Observe that every point of the feasible set is a regular point, i.e. the gradients of the constraints are linearly independent. Therefore, Karush-Kuhn-Tucker (KKT) theory will furnish necessary conditions on an optimizer [62].

We form the Lagrangian function

$$\begin{aligned} L(\boldsymbol{s}, \lambda, \boldsymbol{\mu}, \boldsymbol{\nu}) = -\langle \boldsymbol{s}, \boldsymbol{z}\rangle + \tfrac{1}{2}\,\lambda\,(\langle \boldsymbol{s}, \boldsymbol{s}\rangle - c) \\ - \langle \boldsymbol{s}, \boldsymbol{\mu}\rangle + \langle \boldsymbol{s} - \delta\,\mathbf{1}, \boldsymbol{\nu}\rangle\,. \end{aligned}$$

The Lagrange multipliers $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ are non-negative because they correspond to the lower and upper bounds on $\boldsymbol{s}$. Meanwhile, the multiplier λ is unrestricted because it is associated with the equality constraint.

The first-order KKT necessary condition on a regular local maximum $\boldsymbol{s}^\star$ is that

$$\begin{aligned} \mathbf{0} &= (\nabla_{\boldsymbol{s}}\, L)(\boldsymbol{s}^\star, \lambda^\star, \boldsymbol{\mu}^\star, \boldsymbol{\nu}^\star) \\ &= -\boldsymbol{z} + \lambda^\star\, \boldsymbol{s}^\star - \boldsymbol{\mu}^\star + \boldsymbol{\nu}^\star, \end{aligned} \tag{8}$$

where $\mu_m^\star > 0$ only if $s_m^\star = 0$ and $\nu_m^\star > 0$ only if $s_m^\star = \delta$. Notice that one of $\mu_m^\star$ or $\nu_m^\star$ must be zero because they correspond to mutually exclusive constraints. The second-order KKT necessary condition on a regular local maximum is that

$$\begin{aligned} 0 &\le \boldsymbol{y}^T\,(\nabla_{\boldsymbol{s}}^2\, L)(\boldsymbol{s}^\star, \lambda^\star, \boldsymbol{\mu}^\star, \boldsymbol{\nu}^\star)\,\boldsymbol{y} \\ &= \lambda^\star\, \boldsymbol{y}^T\,\boldsymbol{y} \end{aligned}$$

for every vector $\boldsymbol{y}$ in the subspace of first-order feasible variations. This subspace is non-trivial, so $\lambda^\star \ge 0$.

Solve Equation (8) to obtain

$$\lambda^\star \boldsymbol{s}^\star = \boldsymbol{z} + \boldsymbol{\mu}^\star - \boldsymbol{\nu}^\star.$$

Whenever $\mu_m^\star > 0$, both $s_m^\star = 0$ and $\nu_m^\star = 0$. This combination is impossible because $z_m \geq 0$. Therefore, we may eliminate $\boldsymbol{\mu}^\star$ to reach

$$\lambda^\star \boldsymbol{s}^\star = \boldsymbol{z} - \boldsymbol{\nu}^\star.$$

The cases $\lambda^\star = 0$ and $\lambda^\star > 0$ require separate consideration.

If $\lambda^\star = 0$, it is clear that $\boldsymbol{\nu}^\star = \boldsymbol{z}$. Since $\nu_m^\star > 0$ only if $s_m^\star = \delta$, we must have $s_m^\star = \delta$ whenever $z_m > 0$. Suppose that k components of $\boldsymbol{s}^\star$ equal δ. The remaining $(d-k)$ components are not uniquely determined by the optimization problem. From the many solutions, we choose one such that

$$s_m^\star = \sqrt{\frac{c - k\,\delta^2}{d-k}} \qquad \text{for } m \text{ where } z_m = 0.$$

This formula ensures that $\boldsymbol{s}^\star$ has the correct norm and that none of its entries exceeds δ.

When $\lambda^\star > 0$, the solution has the form

$$\boldsymbol{s}^\star = [\gamma\,\boldsymbol{z}]_\delta\,,$$

where γ is positive and the operator $[\cdot]_\delta$ truncates to δ components of its argument that exceed δ. It is clear that the largest components of $\boldsymbol{z}$ are all truncated at the same time. We only need to determine *which* components these are.

To that end, observe that $\gamma \mapsto \left\|[\gamma\,\boldsymbol{z}]_\delta\right\|_2$ is a strictly increasing function on $[0, \delta/z_{\min}]$, where $z_{\min}$ is the least positive component of $\boldsymbol{z}$. For at most one value of γ, therefore, does the vector $[\gamma\,\boldsymbol{z}]_\delta$ have norm $\sqrt{c}$. If this norm value were not attained, then $\lambda^\star$ would equal zero. Let k be the number of entries of $\boldsymbol{s}^\star$ that equal δ, and suppose that $\mathscr{M}$ indexes the remaining $(d-k)$ components. Then

$$c = \|\boldsymbol{s}^\star\|_2^2 = k\,\delta^2 + \gamma^2 \sum_{m\in\mathscr{M}} |z_m|^2.$$

Recall that γ is positive. Therefore, is impossible that $k\,\delta^2 > c$. When $k\,\delta^2 = c$, it follows that $z_m = 0$ for each m in $\mathscr{M}$. Otherwise, z_m must be non-zero for some m in $\mathscr{M}$. Then the value of γ must be

$$\gamma = \sqrt{\frac{c - k\,\delta^2}{\sum_{m\in\mathscr{M}} |z_m|^2}}.$$

■

B. Convergence

For the alternating projection between the PAR constraint set and the set of α-tight frames, we have not proven a more elaborate convergence theorem than the basic result, Theorem 4, because it is not easy to guarantee that the solution to the PAR matrix nearness problem is unique. We have been able to provide a sufficient condition on the fixed points of the iteration that lie in the PAR constraint set $\mathscr{S}$. Note that similar fixed points arose in Section IV.

Theorem 10: A sufficient condition for a full-rank matrix S from $\mathscr{S}$ to be a fixed point of the alternating projection between $\mathscr{S}$ and $\mathscr{X}_\alpha$ is that the columns of S are all eigenvectors of SS^*. That is, $SS^*S = S\Lambda$, where $\Lambda \in \mathbb{C}^{N\times N}$ is diagonal and positive, with no more than d distinct entries.

Proof: Refer to Appendix II-E. ■

C. Numerical Examples

Let us demonstrate that alternating projection can indeed produce tight frames whose columns have specified PAR and specified norm. We shall produce complex tight frames because, in the real case, PAR constraints can lead to a discrete optimization problem. The experiments all begin with the initial 3×6 matrix

$$\left[\begin{array}{rrrrrr} .0748+.3609\mathrm{i} & .0392+.4558\mathrm{i} & .5648+.3635\mathrm{i} & -.2567+.4463\mathrm{i} & .7064+.6193\mathrm{i} & .1586+.6825\mathrm{i} \\ .5861-.0570\mathrm{i} & -.2029+.8024\mathrm{i} & -.5240+.4759\mathrm{i} & -.1806-.1015\mathrm{i} & -.1946-.1889\mathrm{i} & .5080+.0226\mathrm{i} \\ -.7112+.1076\mathrm{i} & -.2622-.1921\mathrm{i} & -.1662+.1416\mathrm{i} & .0202+.8316\mathrm{i} & .0393-.2060\mathrm{i} & .2819+.4135\mathrm{i} \end{array}\right].$$

The respective PAR values of its columns are 1.5521, 2.0551, 1.5034, 2.0760, 2.6475 and 1.4730.

Unit-PAR tight frames are probably the most interesting example. In each column of a unit-PAR tight frame, the entries share an identical modulus, which depends on the norm of the column. Let us apply our algorithm to calculate a unit-PAR, unit-norm tight frame:

$$\left[\begin{array}{rrrrrr} .1345+.5615\mathrm{i} & .1672+.5526\mathrm{i} & .4439+.3692\mathrm{i} & -.3358+.4696\mathrm{i} & .4737+.3300\mathrm{i} & .0944+.5696\mathrm{i} \\ .5410-.2017\mathrm{i} & -.0303+.5766\mathrm{i} & -.5115+.2679\mathrm{i} & -.5432-.1956\mathrm{i} & -.3689-.4442\mathrm{i} & .5747+.0554\mathrm{i} \\ -.5768+.0252\mathrm{i} & -.2777-.5062\mathrm{i} & -.2303+.5294\mathrm{i} & .1258+.5635\mathrm{i} & -.0088-.5773\mathrm{i} & .4132+.4033\mathrm{i} \end{array}\right].$$

Indeed, each of the columns has unit PAR and unit norm. The singular values of the matrix are identical to eight decimal places. The calculation required 78 iterations lasting 0.1902 seconds.

Alternating projection can also compute tight frames whose columns have unit PAR but different norms. For example, if we request the column norms 0.75, 0.75, 1, 1, 1.25 and 1.25, the algorithm yields

$$\left[\begin{array}{rrrrrr} .3054+.3070\mathrm{i} & .1445+.4082\mathrm{i} & .3583+.4527\mathrm{i} & -.4264+.3893\mathrm{i} & .4252+.5831\mathrm{i} & .3622+.6242\mathrm{i} \\ .4295-.0549\mathrm{i} & .1235+.4150\mathrm{i} & -.5597+.1418\mathrm{i} & -.5393-.2060\mathrm{i} & -.4425-.5701\mathrm{i} & .7165-.0863\mathrm{i} \\ -.4228-.0936\mathrm{i} & -.0484-.4303\mathrm{i} & .0200+.5770\mathrm{i} & .2585+.5162\mathrm{i} & -.2894-.6611\mathrm{i} & .1291+.7101\mathrm{i} \end{array}\right].$$

One can check that the column norms, PAR and singular values all satisfy the design requirements to eight or more decimal places. The computation took 84 iterations over 0.1973 seconds.

Less stringent constraints on the PAR pose even less trouble. For example, we might like to construct a tight frame whose PAR is bounded by two and whose columns have norms 0.75, 0.75, 1, 1, 1.25 and 1.25. It is

$$\left[\begin{array}{rrrrrr} .0617+.1320\mathrm{i} & .0184+.2764\mathrm{i} & .4299+.3593\mathrm{i} & -.1382+.2511\mathrm{i} & .6847+.7436\mathrm{i} & .2933+.6939\mathrm{i} \\ .4256-.1031\mathrm{i} & -.0558+.5938\mathrm{i} & -.5920+.4974\mathrm{i} & -.4306-.2650\mathrm{i} & -.2095-.3072\mathrm{i} & .7317+.0928\mathrm{i} \\ -.5912+.0025\mathrm{i} & -.1304-.3363\mathrm{i} & -.0807+.2857\mathrm{i} & .0852+.8093\mathrm{i} & -.3504-.5289\mathrm{i} & .2918+.6048\mathrm{i} \end{array}\right].$$

The computer worked for 0.0886 seconds, during which it performed 49 iterations. As usual, the singular values match to eight decimal places. It is interesting to observe that the frame exceeds the design specifications. The respective PAR values of its columns are 1.8640, 1.8971, 1.7939, 1.9867, 1.9618 and 1.0897.

VII. Discussion

As advertised, we have developed an alternating projection method for solving frame design problems, and we have provided ample evidence that it succeeds. In this section, we discuss some implementation issues and some of the limitations of the algorithm. We conclude with a collection of related problems that one can also solve with alternating projection.

A. The Starting Point

For alternating projection to succeed, it is essential to choose a good starting point. Here are a few general strategies that may be useful.

The simplest method is to select N vectors uniformly at random from the surface of the unit sphere in $\mathbb{C}^d$ and form them into an initial matrix. Although this technique sometimes works, it is highly probable that there will be pairs of strongly correlated vectors, and it is usually preferable for the frame to contain dissimilar vectors. Nevertheless, a collection of random vectors converges almost surely to a tight frame as more vectors are added [55].

A more practical idea is to select many vectors, say $2\,dN$, and then use a clustering algorithm—such as Lloyd-Max [63], spherical k-means [64] or diametrical clustering [65]—to separate these vectors into N clusters. The cluster representatives will usually be much more diverse than vectors chosen at random. A related approach would select many random vectors and then greedily remove vectors that are highly correlated with the remaining vectors. This method seems to furnish excellent starting points for constructing equiangular tight frames. One might also build up a collection of random vectors by allowing a new vector to join only if it is weakly correlated with the current members.

Another technique is to start with a tight frame that has been developed for another application. By rotating the frame at random, it is possible to obtain many different starting points that retain some of the qualities of the original frame. In particular, equiangular tight frames make excellent initializers.

It is also possible to choose a collection of N vectors from a larger frame for $\mathbb{C}^d$. Similarly, one might truncate some coordinates from a frame in a higher-dimensional space. In particular, one might truncate an orthonormal basis for $\mathbb{C}^N$ to retain only d coordinates. See [66], for example, which uses the Fourier transform matrix in this manner.

B. Limitations

Alternating projection is not a panacea that can alleviate all the pain of frame design. While preparing this report, we encountered several difficulties.

A theoretical irritation is the lack of a proof that alternating projection converges in norm. No general proof is possible, as the counterexample in [67] makes clear. Nevertheless, it would be comforting to develop sufficient conditions that guarantee the convergence of alternating projections between non-convex sets. The results of [67] are the best that we know of. We would also like to develop conditions that can ensure convergence to a pair of points at minimal distance. Here, the most general results are probably due to Csiszár and Tusnády [68].

Another major inconvenience is that alternating projection converges at a geometric rate (or worse) [49]–[52]. For large problems, it can be painful to wait on the solution. A valuable topic for future research would be a method of acceleration.

A more specific disappointment was the inability of alternating projection to construct tight frames over small finite alphabets. It is straightforward to solve the matrix nearness problem associated with a finite alphabet, and it can be shown that the algorithm always converges in norm to a fixed point. But the algorithm never once yielded a tight frame. This failure is hardly surprising; discrete constraints are some of the most difficult to deal with in optimization. It may be possible to use annealing to improve the performance of the algorithm. This would be a valuable topic for future research.

C. Related Problems

We have permitted a great deal of freedom in the selection of the structural constraint set, but we only considered the spectral constraints that arise naturally in connection with tight frames. Nevertheless, alternating projection offers a straightforward method for addressing other inverse eigenvalue problems. For example, one might try to construct general frames with prescribed lower and upper frame bounds, α and β. Instead of forcing the Gram matrix to be a rank-d orthogonal projector, one might impose only a rank constraint or a constraint on its condition number. To implement the algorithm, it would only be necessary to solve the matrix nearness problem associated with these spectral constraints.

One can also use alternating projection to construct positive semi-definite (PSD) matrices that have certain structural properties. Higham, for example, has used a corrected alternating projection to produce the correlation matrix nearest to an input matrix [37]. (A *correlation matrix* is a PSD matrix with a unit diagonal.) Since the PSD matrices form a closed, convex set, it is possible to prove much more about the behavior of alternating algorithms.

We have also had good success using alternating projection to compute sphere packings in real and complex projective spaces. These methods can be extended to produce sphere packings in real and complex Grassmannian manifolds [60]. It seems clear that alternating projection has a promising future for a new generation of problems.

Acknowledgments

I. S. Dhillon acknowledges support by NSF CAREER Award No. ACI-0093404. R. W. Heath acknowledges support from the Texas Advanced Technology Program under Grant

No. 003658-0614-2001, the Samsung Advanced Institute of Technology, and the National Instruments Foundation. T. Strohmer acknowledges partial support by NSF grant DMS 0208568.

APPENDIX I
POINT-TO-SET MAPS

To understand the convergence of the algorithms, we rely on some basic results from the theory of point-to-set maps. Zangwill's book [69] is a good basic reference with applications to mathematical programming. More advanced surveys include [70], [71]. de Leeuw presents statistical applications in [51]. We have drawn from all these sources here.

A. Point-to-Set Maps

Let $\mathscr{Y}$ and $\mathscr{Z}$ be arbitrary sets. The *power set* of $\mathscr{Z}$ is the collection of all subsets of $\mathscr{Z}$, and it is denoted by $2^{\mathscr{Z}}$. A *point-to-set map* from $\mathscr{Y}$ to $\mathscr{Z}$ is a function $\Omega : \mathscr{Y} \to 2^{\mathscr{Z}}$. In words, Ω maps each point of $\mathscr{Y}$ to a subset of $\mathscr{Z}$.

There are several different ways of combining point-to-set maps. Take two maps $\Omega_{yz} : \mathscr{Y} \to 2^{\mathscr{Z}}$ and $\Omega_{zw} : \mathscr{Z} \to 2^{\mathscr{W}}$. The composition of these maps carries a point y to a subset of $\mathscr{W}$ via the rule

$$(\Omega_{zw} \circ \Omega_{yz})(y) = \bigcup_{z \in \Omega_{yz}(y)} \Omega_{zw}(z).$$

This definition can be extended in the obvious way to a longer composition of maps. Now, suppose Ω_{uv} maps $\mathscr{U}$ to $2^{\mathscr{V}}$. The Cartesian product of Ω_{uv} and Ω_{yz} is the point-to-set map from $\mathscr{U} \times \mathscr{Y}$ to $\mathscr{V} \times \mathscr{Z}$ given by

$$(\Omega_{uv} \times \Omega_{yz})(u, y) = \Omega_{uv}(u) \times \Omega_{yz}(y).$$

B. Topological Properties

Suppose that the underlying sets are endowed with topologies so that we may speak of convergence. A map $\Omega : \mathscr{Y} \to 2^{\mathscr{Z}}$ is *closed* at the point $\bar{y}$ in $\mathscr{Y}$ whenever the statements $y_j \to \bar{y}$, $z_j \in \Omega(y_j)$ and $z_j \to \bar{z}$ together imply that $\bar{z} \in \Omega(\bar{y})$. One may interpret this definition as saying that the set $\Omega(\bar{y})$ is "bigger" than the sets in the sequence $\{\Omega(y_j)\}$. On the other hand, the map Ω is *open* at $\bar{y}$ in $\mathscr{Y}$ whenever the statements $y_j \to \bar{y}$ and $\bar{z} \in \Omega(\bar{y})$ together imply the existence of a number J and a sequence of points $\{z_j\}$ such that $z_j \to \bar{z}$ and $z_j \in \Omega(y_j)$ for all $j \geq J$. More or less, this statement means that the set $\Omega(\bar{y})$ is "smaller" than the sets in the sequence $\{\Omega(y_j)\}$. A map which is both open and closed at $\bar{y}$ is said to be *continuous* at $\bar{y}$. We call Ω an *open map*, *closed map* or *continuous map* whenever it has the corresponding property for every point in $\mathscr{Y}$.

Finite Cartesian products and finite compositions of open maps are open. Finite Cartesian products of closed maps are closed. If $\Omega_{yz} : \mathscr{Y} \to 2^{\mathscr{Z}}$ and $\Omega_{zw} : \mathscr{Z} \to 2^{\mathscr{W}}$ are closed and $\mathscr{Z}$ is compact, then the composition $(\Omega_{zw} \circ \Omega_{yz})$ is closed.

C. Fixed Points

Suppose that Ω is a point-to-set map from $\mathscr{Y}$ to itself. Let y be a point of $\mathscr{Y}$ for which $\Omega(y) = \{y\}$. Then y is called a *fixed point* of the map Ω. In contrast, a *generalized fixed point* of Ω is a point for which $y \in \Omega(y)$. When we wish to emphasize the distinction, we may refer to a regular fixed point as a *strong* or *classical* fixed point.

D. Infimal Maps

Minimizing functions leads to a special type of point-to-set map. Suppose that $f : \mathscr{Y} \times \mathscr{Z} \to \mathbb{R}_+$ is a real-valued function of two variables, and let Ω be a point-to-set map from $\mathscr{Y}$ to $\mathscr{Z}$. Associated with f and Ω is an *infimal map* defined by

$$M^z(y) \stackrel{\text{def}}{=} \arg \min_{z \in \Omega(y)} f(y, z).$$

If $f(y, \cdot)$ attains no minimal value on $\Omega(y)$, then $M^z(y) = \emptyset$, the empty set. Under mild conditions, infimal maps are closed.

Theorem 11 (Dantzig-Folkman-Shapiro [72]): If Ω is continuous at $\bar{y}$ and $f(\bar{y}, \cdot)$ is continuous on $\Omega(\bar{y})$, then M^z is closed at $\bar{y}$.

In particular, the constant map $\Omega : y \mapsto \mathscr{Z}$ is continuous whenever $\mathscr{Z}$ is closed. So minimizing a continuous function over a fixed, closed set always yields a closed infimal map.

E. Iterative Algorithms

Zangwill was apparently the first to recognize that many procedures in mathematical programming find their most natural expression in the language of point-to-set maps [69]. An *algorithmic map* or *algorithm* is simply a function $\Omega : \mathscr{Y} \to 2^{\mathscr{Y}}$. Given an initial point y_0 of $\mathscr{Y}$, an algorithmic map generates a sequence of iterates according to the rule

$$y_{j+1} \in \Omega(y_j).$$

Suppose that $f : \mathscr{Y} \to \mathbb{R}_+$ is a continuous, non-negative function. We say that the algorithm Ω is *monotonic* with respect to f when

$$z \in \Omega(y) \qquad \text{implies} \qquad f(z) \leq f(y).$$

An algorithm *strictly monotonic* with respect to f is a monotonic algorithm for which

$$z \in \Omega(y) \text{ and } f(z) = f(y) \qquad \text{imply} \qquad z = y.$$

Zangwill showed that a closed, monotonic algorithm converges in a weak sense to a generalized fixed point. We present a streamlined version of his result.

Theorem 12 (Zangwill [69]): Let Ω be a closed algorithmic map on a compact set $\mathscr{Y}$, and assume that Ω is monotonic with respect to a continuous, non-negative function f. Suppose that the algorithm generates a sequence of iterates $\{y_j\}$.

- The sequence has at least one accumulation point in $\mathscr{Y}$.
- Each accumulation point $\bar{y}$ satisfies $f(\bar{y}) = \lim_j f(y_j)$.
- Each accumulation point $\bar{y}$ is a generalized fixed point of the algorithm.

R. R. Meyer subsequently extended Zangwill's Theorem to provide a more satisfactory convergence result for strictly monotonic algorithms. One version of his result follows. For reference, a sequence $\{y_j\}$ in a normed space is called *asymptotically regular* when $\|y_{j+1} - y_j\| \to 0$.

Theorem 13 (Meyer [67]): Let $\mathscr{Y}$ be a compact subset of a normed space, and assume that Ω is a closed algorithm on $\mathscr{Y}$ that it is strictly monotonic with respect to the continuous, non-negative function f. Suppose that Ω generates a sequence of iterates $\{y_j\}$. In addition to the conclusions of Zangwill's Theorem, the following statements hold.

- Each accumulation point of the sequence is a (strong) fixed point of the algorithm.
- The sequence of iterates is asymptotically regular. In consequence, it has a continuum of accumulation points, or it converges in norm [73].
- In case that the fixed points of Ω on each isocontour of f form a discrete set, then the sequence of iterates converges in norm.

F. Alternating Projection

An alternating projection can be interpreted as a kind of monotonic algorithm. Suppose that $f : \mathscr{Y} \times \mathscr{Z} \to \mathbb{R}_+$ is a continuous function. Then f induces two natural infimal maps,

$$M_y(z) \stackrel{\text{def}}{=} \arg\min_{y \in \mathscr{Y}} f(y, z) \qquad \text{and}$$
$$M^z(y) \stackrel{\text{def}}{=} \arg\min_{z \in \mathscr{Z}} f(y, z).$$

If $\mathscr{Y}$ and $\mathscr{Z}$ are closed, then Theorem 11 shows that the maps M_y and M^z are both closed.

We interpret alternating projection as an algorithm on the *product space* $\mathscr{Y} \times \mathscr{Z}$ equipped with the usual product topology. Given an initial iterate y_0 from $\mathscr{Y}$, alternating projection generates a sequence of iterates $\{(y_j, z_j)\}$ via the rules

$$z_j \in M^z(y_j) \quad \text{and} \quad y_{j+1} \in M_y(z_j)$$

for each $j \geq 0$. Formally, this algorithm can be written as the composition of two sub-algorithms, Ω_{to} and Ω_{fro}, that are defined as

$$\Omega_{\text{to}} : (y, z) \mapsto \{y\} \times M^z(y) \qquad \text{and}$$
$$\Omega_{\text{fro}} : (y, z) \mapsto M_y(z) \times \{z\}.$$

It follows that $\Omega \stackrel{\text{def}}{=} \Omega_{\text{fro}} \circ \Omega_{\text{to}}$ is a closed algorithm whenever $\mathscr{Y}$ and $\mathscr{Z}$ are compact. Both sub-algorithms decrease the value of f, so it should also be clear that Ω is monotonic with respect to f. Zangwill's Theorem tenders a basic convergence result.

Corollary 14: Let $\mathscr{Y}$ and $\mathscr{Z}$ be compact. Suppose that the alternating projection between $\mathscr{Y}$ and $\mathscr{Z}$ generates a sequence of iterates $\{(y_j, z_j)\}$.

- The sequence has at least one accumulation point.
- Each accumulation point of the sequence lies in $\mathscr{Y} \times \mathscr{Z}$.
- Each accumulation point is a generalized fixed point of the algorithm.
- Each accumulation point $(\bar{y}, \bar{z})$ satisfies $f(\bar{y}, \bar{z}) = \lim_j f(y_j, z_j)$.

If the infimal maps M_y and M^z are single-valued, we can achieve a much more satisfactory result.

Corollary 15: Let $\mathscr{Y}$ and $\mathscr{Z}$ be compact subsets of a normed space, and assume that the infimal maps M_y and M^z are single-valued. Suppose that the alternating projection between $\mathscr{Y}$ and $\mathscr{Z}$ generates a sequence of iterates $\{(y_j, z_j)\}$. In addition to the conclusions of Corollary 14, we have the following.

- Each accumulation point is a classical fixed point of the alternating projection.
- The sequence of iterates is asymptotically regular.
- The sequence of iterates either converges in norm or it has a continuum of accumulation points.

Proof: We just need to show that the algorithm is strictly monotonic with respect to f. Suppose that $f(y, z) = f(\Omega(y, z))$. Since the infimal maps never increase the value of f, we have the equalities

$$\begin{aligned} f(y, z) &= f(y, M^z(y)) \\ &= f((M_y \circ M^z)(y), M^z(y)) = f(\Omega(y, z)). \end{aligned}$$

Since M^z yields the *unique* minimizer of f with its first argument fixed, the first equality implies that $M^z(y) = \{z\}$. Likewise, the second equality yields $(M_y \circ M^z)(y) = \{y\}$. That is, $\Omega(y, z) = \{(y, z)\}$. An application of Meyer's Theorem completes the argument. ∎

This result is a special case of a theorem of Fiorot and Huard [74]. In Appendix II, we shall translate the language of these corollaries into more familiar terms.

G. Literature on Alternating Projection

Like most good ideas, alternating projection has a long biography and several aliases, including *successive approximation*, *successive projection*, *alternating minimization* and *projection on convex sets*. This section offers a résumé of the research on alternating projection, but it makes no pretension to be comprehensive. Deutsch has written more detailed surveys, including [52], [75], [76].

According to Deutsch [75], alternating projection first appeared in a set of mimeographed lecture notes, written by John von Neumann in 1933. von Neumann proved that the alternating projection between two closed subspaces of a Hilbert space converges pointwise to the orthogonal projector onto their intersection [39]. Apparently, this theorem was not very well advertised, because many other authors have discovered it independently, including Aronszajn [49] and Wiener [77]. It was shown by Aronszajn [49] and Kayalar–Weinert [50] that both sequences of iterates converge geometrically with a rate exactly equal to the squared cosine of the (Friedrichs) principal angle between the two subspaces.

It is natural to extend the alternating projection between two subspaces by cyclically projecting onto several subspaces. Halperin demonstrated that, in a Hilbert space, the cyclic projection among a finite number of closed subspaces converges pointwise to the orthogonal projector onto their intersection [78]. The convergence is geometric [79]. Optimal bounds on the rate of convergence can be computed with techniques of Xu and Zikatonov [80]. Bauschke *et al.* study methods for accelerating cyclic projection in the recently minted paper [81].

It will come as no surprise that researchers have also studied alternating projection between subspaces of a Banach space. Unaware of von Neumann's work, Diliberto and Straus introduced an alternating method for computing the best sup-norm approximation of a bivariate continuous function as the sum of two univariate continuous functions, and they proved some weak convergence results [82]. The norm convergence of the sequence of iterates remained open until the work of Aumann [83]. M. Golomb extended the Diliberto–Straus algorithm to other best-approximation problems [84]. For more information on alternating algorithms in Banach spaces, see the monograph of Cheney and Light [85].

Another fruitful generalization is to consider projection onto convex subsets. The projector—or *proximity map*—onto a closed, convex subset of a Hilbert space is well-defined, because each point has a unique best approximation from that set. The basic result, due to Cheney and Goldstein, is that the alternating projection between two closed, convex subsets of a Hilbert space will converge to a pair of points at minimal distance from each other, so long as one set is compact [4]. Dykstra [86], [87] and Han [88] independently developed a cyclic projection technique that, given a point, can compute its best approximation from the intersection of a finite number of closed, convex sets in a Hilbert space. Their algorithm requires a correction to each projection. To date, the most extensive treatment of cyclic projection methods is the survey article by Bauschke and Borwein [89].

Most of the work on alternating projection has involved the Euclidean distance, but it is possible to develop results for other divergence measures. In particular, Csiszár and Tusnády have shown that alternate minimization of the Kullback-Leibler divergence can be used to find a pair of minimally distant points contained within two convex sets of probability measures [68].

There has been some research on alternating projection between non-convex sets, but the theoretical results so far are limited. Fiorot and Huard have applied the theorems of Zangwill and Meyer to obtain weak convergence results for a class of block relaxation schemes that include alternating and cyclic projection onto non-convex sets [74]. Combettes and Trussell have developed a technique which inflates the non-convex sets into convex sets; they offer some qualified convergence results [90]. Cadzow has also demonstrated empirically that cyclic projections among non-convex sets can effectively solve some signal enhancement problems [91]. More research in this direction would be valuable.

Alternating projection has found application to many different problems, of which we offer a (small) selection. The most famous example from these pages must be the Blahut–Arimoto algorithm for computing channel capacity and rate-distortion functions [92], [93]. In the field of signal restoration and recovery, we mention the work of Landau–Miranker [94], Gerchberg [95], Youla–Webb [96], Cadzow [91] and Donoho–Stark [97]. Çetin, Gerek and Yardimci show that projection on convex sets can compute multi-dimensional equiripple filters [98]. Xu and Zikatonov discuss how alternating projection can be used to solve the linear systems that arise in the discretization of partial differential equations [80]. In the matrix analysis community, alternating projection has been used as a computational method for solving inverse eigenvalue problems [36], [38] and for solving matrix nearness problems [37], [99]. In statistics, one may view the Expectation Maximization (EM) algorithm as an alternating projection [100]. de Leeuw has discussed other statistical applications in [51].

Appendix II
Convergence and Fixed Points

Armed with the theory of the last appendix, we are finally girded to attack the convergence of Algorithm 1. The results on point-to-set maps will allow us to dispatch this dragon quickly. Then we shall turn our attention to the convergence of the algorithm in the special case that the frame vectors have prescribed norms. This problem will require a longer siege, but it, too, will yield to our onslaught. The convergence results that we develop here are all novel.

A. Basic Convergence Proof

In this section, we establish the convergence of the basic alternating projection algorithm that appears in Section III-D. Our main burden is to translate the language of point-to-set maps into more familiar terms.

Theorem 16 (Global Convergence): Let $\mathscr{Y}$ and $\mathscr{Z}$ be closed sets, one of which is bounded. Suppose that alternating projection generates a sequence of iterates $\{(Y_j, Z_j)\}$. This sequence possesses at least one accumulation point, say $(\overline{Y}, \overline{Z})$.

- The accumulation point lies in $\mathscr{Y} \times \mathscr{Z}$.
- The accumulation point satisfies

$$\left\|\overline{Y} - \overline{Z}\right\|_{\mathrm{F}} = \lim_{j \to \infty} \|Y_j - Z_j\|_{\mathrm{F}}.$$

- The accumulation point satisfies

$$\left\|\overline{Y} - \overline{Z}\right\|_{\mathrm{F}} = \mathrm{dist}(\overline{Y}, \mathscr{Z}) = \mathrm{dist}(\overline{Z}, \mathscr{Y}).$$

Proof: Assume without loss of generality that $\mathscr{Y}$ is the compact set, while $\mathscr{Z}$ is merely closed. We must establish that we have all the compactness necessary to apply Corollary 14.

Without loss of generality, assume that $Y_0 \in \mathscr{Y}$. If $\delta = \|Y_0 - Z_0\|_{\mathrm{F}}$, then subsequent iterates always satisfy

$$\|Y_j - Z_j\|_{\mathrm{F}} \le \delta \qquad \text{and}$$
$$\|Y_{j+1} - Z_j\|_{\mathrm{F}} \le \delta.$$

Thus, we may restrict our attention to the sets

$$\mathscr{Y}_1 = \{Y \in \mathscr{Y} : \mathrm{dist}(Y, \mathscr{Z}) \le \delta\} \qquad \text{and}$$
$$\mathscr{Z}_1 = \{Z \in \mathscr{Z} : \mathrm{dist}(Z, \mathscr{Y}) \le \delta\}.$$

Since $\mathscr{Y}$ is compact, $\mathscr{Y}_1$ is compact because it is a closed subset of a compact set. On the other hand, $\mathscr{Z}_1$ is compact because it is the intersection of the closed set $\mathscr{Z}$ with a compact set, namely the collection of matrices within a fixed distance of $\mathscr{Y}$.

We may apply Corollary 14. Each of the conclusions of the corollary has a straightforward analogue among the conclusions of the present theorem. The only question that may remain is what it means for a pair of matrices $(\overline{Y}, \overline{Z})$ to be a generalized fixed point of the alternating projection.

A generalized fixed point of an algorithm is a point which is a possible successor of itself. In the present case, a pair of matrices can succeed itself if and only if the second component is a potential successor of the first *and* the first component is a potential successor of the second. The matrix $\overline{Z}$ can succeed the matrix $\overline{Y}$ if and only if

$$\left\|\overline{Z} - \overline{Y}\right\|_{\mathrm{F}} = \mathrm{dist}(\overline{Y}, \mathscr{Z}).$$

Likewise, $\overline{Y}$ can succeed $\overline{Z}$ if and only if

$$\left\|\overline{Y} - \overline{Z}\right\|_{\mathrm{F}} = \mathrm{dist}(\overline{Z}, \mathscr{Y}).$$

This observation completes the proof. ∎

Since the collection of α-tight frames and the collection of their Gram matrices are both compact, the theorem has two immediate corollaries.

Corollary 17: If $\mathscr{X}_\alpha$ is the collection of α-tight frames, and $\mathscr{S}$ is a closed set of matrices, then Theorem 16 applies with $\mathscr{Y} \stackrel{\text{def}}{=} \mathscr{S}$ and $\mathscr{Z} \stackrel{\text{def}}{=} \mathscr{X}_\alpha$.

Corollary 18: If $\mathscr{G}_\alpha$ contains the Gram matrices of all α-tight frames, and $\mathscr{H}$ is a closed set of Hermitian matrices, then Theorem 16 applies with $\mathscr{Y} \stackrel{\text{def}}{=} \mathscr{G}_\alpha$ and $\mathscr{Z} \stackrel{\text{def}}{=} \mathscr{H}$.

B. Stronger Convergence Results

Meyer's Theorem suggests that it might be possible to provide a stronger convergence result for Algorithm 1 if we can ensure that the matrix nearness problems have unique solutions. In many cases, the nearness problems are uniquely soluble whenever the iterates get sufficiently close together. This provides a local convergence result that is much stronger than Zangwill's Theorem allows. First, we prove a general version of this result. Afterward, we show that it applies to an alternating projection that involves one of the spectral constraint sets $\mathscr{X}_\alpha$ or $\mathscr{G}_\alpha$.

Recall that the distance between a matrix M and a set $\mathscr{Y}$ is defined as

$$\mathrm{dist}(M, \mathscr{Y}) \stackrel{\text{def}}{=} \inf_{Y \in \mathscr{Y}} \left\|M - Y\right\|_{\mathrm{F}}.$$

Theorem 19: Let $\mathscr{Y}$ and $\mathscr{Z}$ be closed sets of matrices, one of which is compact. Suppose that the alternating projection between $\mathscr{Y}$ and $\mathscr{Z}$ generates a sequence of iterates $\{(Y_j, Z_j)\}$, and assume that the matrix nearness problems

$$\min_{Y \in \mathscr{Y}} \left\|Y - M\right\|_{\mathrm{F}}$$
$$\min_{Z \in \mathscr{Z}} \left\|Z - M\right\|_{\mathrm{F}}$$

have unique solutions for any matrix M in the sequence of iterates. Then we reach the following conclusions.

- The sequence of iterates possesses at least one accumulation point, say $(\overline{Y}, \overline{Z})$.
- The accumulation point lies in $\mathscr{Y} \times \mathscr{Z}$.
- The pair $(\overline{Y}, \overline{Z})$ is a fixed point of the alternating projection. In other words, if we applied the algorithm to $\overline{Y}$ or to $\overline{Z}$ every iterate would equal $(\overline{Y}, \overline{Z})$.
- The accumulation point satisfies
$$\left\|\overline{Y} - \overline{Z}\right\|_{\mathrm{F}} = \lim_{j \to \infty} \left\|Y_j - Z_j\right\|_{\mathrm{F}}.$$
- The component sequences are asymptotically regular, i.e.
$$\left\|Y_{j+1} - Y_j\right\|_{\mathrm{F}} \to 0 \quad \text{and} \quad \left\|Z_{j+1} - Z_j\right\|_{\mathrm{F}} \to 0.$$
- *Either* the component sequences both converge in norm,
$$\left\|Y_j - \overline{Y}\right\|_{\mathrm{F}} \to 0 \quad \text{and} \quad \left\|Z_j - \overline{Z}\right\|_{\mathrm{F}} \to 0,$$
or the set of accumulation points forms a continuum.

Proof: The argument in the proof of Theorem 16 shows that we are performing an alternating minimization between two compact sets. The hypotheses of the theorem guarantee that each iterate is uniquely determined by the previous iterate. Corollaries 14 and 15 furnish the stated conclusions.

The only point that may require clarification is what it takes for a pair of matrices $(\overline{Y}, \overline{Z})$ to be a classical fixed point of the alternating projection. A classical fixed point of an algorithm is the *only* possible successor of itself. In the case of alternating projection, the matrix $\overline{Z}$ must be the unique successor of the $\overline{Y}$, and the matrix $\overline{Y}$ must be the unique successor of $\overline{Z}$. This observation completes the argument. ∎

Due to the peculiar structure of the spectral constraint sets $\mathscr{X}_\alpha$ and $\mathscr{G}_\alpha$, the solutions to the associated matrix nearness problems are often unique. Therefore, the alternating projection algorithms that we have considered in this paper sometimes have better performance than the basic convergence result, Theorem 16, would predict.

We remind the reader that

$$\mathscr{X}_\alpha \stackrel{\text{def}}{=} \{X \in \mathbb{C}^{d \times N} : XX^* = \alpha\, \mathrm{I}_d\}, \quad \text{and}$$
$$\mathscr{G}_\alpha \stackrel{\text{def}}{=} \{G \in \mathbb{C}^{N \times N} : G = G^*, \text{ and } G \text{ has eigenvalues } (\underbrace{\alpha, \dots, \alpha}_{d}, 0, \dots, 0)\}.$$

The uniqueness of the matrix nearness problems will follow from the Wielandt–Hoffman Theorem, a powerful result from matrix analysis.

Theorem 20 (Wielandt–Hoffman [25]): Suppose that A and B are $N \times N$ Hermitian matrices, and let the vectors $\boldsymbol{\lambda}(A)$ and $\boldsymbol{\lambda}(B)$ list the eigenvalues of A and B in algebraically non-increasing order. Then

$$\left\|\boldsymbol{\lambda}(A) - \boldsymbol{\lambda}(B)\right\|_2 \le \left\|A - B\right\|_{\mathrm{F}}.$$

Suppose instead that A and B are $d \times N$ rectangular matrices with $d \le N$, and let $\boldsymbol{\sigma}(A)$ and $\boldsymbol{\sigma}(B)$ list the largest d singular values of A and B in non-increasing order. Then

$$\left\|\boldsymbol{\sigma}(A) - \boldsymbol{\sigma}(B)\right\|_2 \le \left\|A - B\right\|_{\mathrm{F}}.$$

Note that if we solving matrix nearness problems with respect to the spectral norm, Weyl's Theorem would allow us to provide stronger bounds [25].

Corollary 21 (Local Convergence with Constraint $\mathscr{X}_\alpha$): Let $\mathscr{S}$ be a closed set of $d \times N$ matrices for which the associated matrix nearness problem

$$\min_{S \in \mathscr{S}} \left\|S - M\right\|_{\mathrm{F}}$$

has a unique solution whenever $\mathrm{dist}(M, \mathscr{S}) < \varepsilon$. Imagine that the alternating projection between $\mathscr{S}$ and $\mathscr{X}_\alpha$ generates a sequence of iterates $\{(S_j, X_j)\}$ in which

$$\left\|S_J - X_J\right\|_{\mathrm{F}} < \min\{\varepsilon, \alpha\} \quad \text{for some index } J.$$

Then the conclusions of Theorem 19 are in force.

Proof: According to Theorem 2, the matrix in $\mathscr{X}_\alpha$ nearest to a matrix M is unique so long as M has full rank. A $d \times N$ matrix is rank-deficient only if its d-th largest singular value is zero. Observe that the largest d singular values of each matrix in $\mathscr{X}_\alpha$ all equal $\alpha > 0$. According to the Wielandt–Hoffman Theorem, any matrix sufficiently close to $\mathscr{X}_\alpha$ cannot be rank-deficient. More precisely, $\operatorname{dist}(M, \mathscr{X}_\alpha) < \alpha$ implies that M has full rank, which in turn shows that M has a unique best approximation in $\mathscr{X}_\alpha$.

Define the constraint sets

$$\begin{aligned} \mathscr{Y} &\stackrel{\text{def}}{=} \mathscr{S} \cap \operatorname{closure}\{S_j : j \geq J\} \qquad \text{and} \\ \mathscr{Z} &\stackrel{\text{def}}{=} \mathscr{X}_\alpha \cap \operatorname{closure}\{X_j : j \geq J\}. \end{aligned}$$

Note that $\mathscr{Y}$ is closed and that $\mathscr{Z}$ is compact. We will apply Theorem 19 to the tail of the sequence of iterates, beginning with index J. For $j \geq J$, each matrix S_j is close enough to $\mathscr{Z}$ and each matrix X_j is close enough to $\mathscr{Y}$ that we can ensure the matrix nearness problems have unique solutions. ∎

Corollary 22 (Local Convergence with Constraint $\mathscr{G}_\alpha$): Let $\mathscr{H}$ be a closed set of $N \times N$ matrices for which the associated matrix nearness problem

$$\min_{H \in \mathscr{H}} \|H - M\|_{\mathrm{F}}$$

has a unique solution whenever $\operatorname{dist}(M, \mathscr{H}) < \varepsilon$. Imagine that the alternating projection between $\mathscr{G}_\alpha$ and $\mathscr{H}$ generates a sequence of iterates $\{(G_j, H_j)\}$ in which

$$\|G_j - H_j\|_{\mathrm{F}} < \min\{\varepsilon, \alpha/\sqrt{2}\} \quad \text{for some index } J.$$

The conclusions of Theorem 19 are in force.

Proof: Theorem 3 indicates that the matrix in $\mathscr{G}_\alpha$ nearest to a matrix M is unique so long as its d-th and $(d+1)$-st eigenvalues are distinct. Imagine that M is a matrix whose d-th and $(d+1)$-st eigenvalues both equal τ. Since the d-th and $(d+1)$-st eigenvalue of a matrix in $\mathscr{G}_\alpha$ are α and zero, the Wielandt–Hoffman Theorem shows that

$$\operatorname{dist}(M, \mathscr{G}_\alpha)^2 \geq (\alpha - \tau)^2 + \tau^2.$$

Varying τ, the minimum value of the right-hand side is $\alpha^2/2$. Therefore, $\operatorname{dist}(M, \mathscr{G}_\alpha) < \alpha/\sqrt{2}$ implies that the d-th and $(d+1)$-st eigenvalues of M are distinct. In consequence, M has a unique best approximation from $\mathscr{G}_\alpha$.

As before, define the constraint sets

$$\begin{aligned} \mathscr{Y} &\stackrel{\text{def}}{=} \mathscr{H} \cap \operatorname{closure}\{H_j : j \geq J\} \qquad \text{and} \\ \mathscr{Z} &\stackrel{\text{def}}{=} \mathscr{G}_\alpha \cap \operatorname{closure}\{G_j : j \geq J\}. \end{aligned}$$

The set $\mathscr{Y}$ is closed, and $\mathscr{Z}$ is compact. We will apply Theorem 19 to the tail of the sequence of iterates, beginning with index J. For $j \geq J$, each matrix H_j is close enough to $\mathscr{Z}$ and each matrix G_j is close enough to $\mathscr{Y}$ that we can ensure the matrix nearness problems have unique solutions. ∎

C. Specified Column Norms

This section offers a detailed analysis of the alternating projection between the set of α-tight frames and the collection of matrices with specified column norms.

Let $c_1, \dots, c_N$ be positive numbers that denote the squared column norms we desire in the frame. Without loss of generality, we assume that $\sum_n c_n/d = 1$ to streamline the proofs. Then the structural constraint set is

$$\mathscr{S} \stackrel{\text{def}}{=} \{S \in \mathbb{C}^{d \times N} : \|\boldsymbol{s}_n\|_2^2 = c_n\}.$$

The tightness parameter of the frame α must equal one, so we define the set of 1-tight frames as

$$\mathscr{X}_1 \stackrel{\text{def}}{=} \{X \in \mathbb{C}^{d \times N} : XX^* = \mathrm{I}_d\}.$$

Suppose that S_0 is a full-rank matrix drawn from $\mathscr{S}$, and perform an alternating projection between the sets $\mathscr{S}$ and $\mathscr{X}_1$ to obtain sequences $\{S_j\}$ and $\{X_j\}$. Proposition 23 of the sequel shows that the sequence $\{S_j\}$ lies in a compact subset of $\mathscr{S}$ whose elements have full rank, while the sequence $\{X_j\}$ lies in a compact subset of $\mathscr{X}_1$ whose elements have non-zero columns. By an appeal to the matrix nearness results, Theorem 2 and Proposition 5, we see that each iterate is uniquely determined by its predecessor. We may therefore apply Corollary 15.

In this subsection, we complete the foregoing argument by demonstrating that the iterates are well-behaved. In the next subsection, we classify the full-rank fixed points of the alternating projection between $\mathscr{S}$ and $\mathscr{X}_1$.

Set $c_{\min} = \min_n c_n$, and define the diagonal matrix C whose entries are $\sqrt{c_1}, \dots, \sqrt{c_N}$.

Proposition 23: Assume that the initial iterate S_0 is a full rank matrix from $\mathscr{S}$. For every positive index j,

1) the Euclidean norm of each column of X_j exceeds $\sqrt{c_{\min}}/\|C\|_{\mathrm{F}}$; and
2) the smallest singular value of S_j exceeds $\sqrt{c_{\min}}$.

The matrices that satisfy these bounds form compact subsets of $\mathscr{X}_1$ and $\mathscr{S}$.

Proof: Assume that $j \geq 0$, and make the inductive assumption that S_j has full rank. First, we bound the top singular value of S_j by exploiting the relationship between the singular values of a matrix and its Frobenius norm. Since C lists the column norms of S_j, it follows that $\|S_j\|_{\mathrm{F}}^2 = \|C\|_{\mathrm{F}}^2$. The squared Frobenius norm also equals the sum of the squared singular values of S_j. It is immediate that the maximum singular value of S_j satisfies

$$\sigma_{\max}(S_j)^2 \leq \|C\|_{\mathrm{F}}^2 . \tag{9}$$

Next we use this relation to estimate the column norms of X_j. Let S_j have singular value decomposition $U\Sigma V^*$, and write the n-th columns of S_j and X_j as $\boldsymbol{s}_n$ and $\boldsymbol{x}_n$. On account of the fact that $X_j = (S_j {S_j}^*)^{-1/2} S_j$, we have

$$\begin{aligned} \|\boldsymbol{x}_n\|_2 &= \left\|(S_j {S_j}^*)^{-1/2}\, \boldsymbol{s}_n\right\|_2 \\ &= \left\|U \Sigma^{-1} U^*\, \boldsymbol{s}_n\right\|_2 \\ &\geq \sqrt{c_{\min}}/\sigma_{\max}(S_j), \end{aligned} \tag{10}$$

since the norm of $\boldsymbol{s}_n$ is at least $\sqrt{c_{\min}}$. Introducing the estimate (9) into (10) yields the first part of the proposition.

Now, we show that the smallest singular value of S_{j+1} remains well away from zero. The Courant–Fischer Theorem for singular values [25] states that one may calculate the k-th largest singular value of a matrix $\mathsf{B} \in \mathbb{C}^{d \times N}$ as

$$\sigma_k(\mathsf{B}) = \max_{\mathscr{Z}} \min_{\boldsymbol{z} \neq \boldsymbol{0}} \frac{\|\mathsf{B}\,\boldsymbol{z}\|_2}{\|\boldsymbol{z}\|_2},$$

where $\mathscr{Z}$ ranges over all k-dimensional subspaces of $\mathbb{C}^N$ and $\boldsymbol{z} \in \mathscr{Z}$ [25]. Define T_j to be the diagonal matrix that lists the column norms of X_j. Therefore, the nearest matrix in $\mathscr{S}$ can be written as $\mathsf{S}_{j+1} = \mathsf{X}_j\,\mathsf{T}_j^{-1}\mathsf{C}$. Then put

$$\mathscr{Z} \stackrel{\text{def}}{=} \{\mathsf{C}^{-1}\,\mathsf{T}_j\,\boldsymbol{x} : \boldsymbol{x} \in \operatorname{rowspan} \mathsf{X}_j\}.$$

Since X_j has full row-rank, $\mathscr{Z}$ forms a d-dimensional subspace. Select a unit vector $\boldsymbol{z}$ from $\mathscr{Z}$, and express it as $\boldsymbol{z} = \mathsf{C}^{-1}\,\mathsf{T}_j\,\boldsymbol{x}$ for some $\boldsymbol{x}$ in $\operatorname{rowspan} \mathsf{X}_j$. By construction, X_j has orthonormal rows, so we may compute

$$\begin{aligned} \|\mathsf{S}_{j+1}\,\boldsymbol{z}\|_2 &= \left\|\mathsf{X}_j\,\mathsf{T}_j^{-1}\,\mathsf{C}\,\boldsymbol{z}\right\|_2 \\ &= \|\mathsf{X}_j\,\boldsymbol{x}\|_2 \\ &= \|\boldsymbol{x}\|_2 \qquad\qquad (\text{since } \boldsymbol{x} \in \operatorname{rowspan} \mathsf{X}_j) \\ &= \left\|\mathsf{T}_j^{-1}\,\mathsf{C}\,\boldsymbol{z}\right\|_2 . \end{aligned}$$

The matrix X_j is a submatrix of a unitary matrix, so its column norms cannot exceed one. Thus every entry of T_j^{-1} must be at least one. It follows that

$$\|\mathsf{S}_{j+1}\,\boldsymbol{z}\|_2 \geq \sqrt{c_{\min}}.$$

Applying the Courant-Fischer Theorem yields

$$\sigma_{\min}(\mathsf{S}_{j+1}) \geq \sqrt{c_{\min}}.$$

The second part of the proposition is complete.

Finally, we must make the compactness argument. We have shown that the squared singular values of an iterate S_j must lie in the closed interval $[\sqrt{c_{\min}}, \|\mathsf{C}\|_{\mathrm{F}}]$. The minimum squared singular value of a matrix is a continuous function of the matrix entries, which follows from the Wielandt–Hoffman Theorem. Therefore, the matrices whose smallest singular value lies in this interval form a closed set. We conclude that the intersection of this set with the compact set $\mathscr{S}$ is compact. The same argument implies that the sequence $\{\mathsf{X}_j\}$ lies in a compact subset of $\mathscr{X}_1$ whose matrices have column norms bounded away from zero. ∎

D. Fixed Points I

It remains to characterize the fixed points of the alternating projection between the set of matrices with fixed column norms and the set of α-tight frames.

Proposition 24: The full-rank stationary points of an alternating projection between $\mathscr{S}$ and $\mathscr{X}_\alpha$ are precisely those full-rank matrices S from $\mathscr{S}$ whose columns are all eigenvectors of $\mathsf{S}\mathsf{S}^*$. That is, $\mathsf{S}\mathsf{S}^*\mathsf{S} = \mathsf{S}\Lambda$, where $\Lambda \in \mathbb{C}^{N \times N}$ is diagonal and positive with at most d distinct values.

Proof: Define the diagonal matrix $\mathsf{T} = \mathsf{T}(\mathsf{S})$ whose entries are the column norms of $\alpha\,(\mathsf{S}\mathsf{S}^*)^{-1/2}\,\mathsf{S}$.

Suppose that S is a full-rank fixed point of the algorithm. Thus projecting S onto $\mathscr{X}_\alpha$ and projecting back to $\mathscr{S}$ returns S. Symbolically,

$$\mathsf{S} = (\alpha\,(\mathsf{S}\mathsf{S}^*)^{-1/2}\,\mathsf{S})\,(\mathsf{T}^{-1}\,\mathsf{C}).$$

Define $\Lambda = \alpha\,\mathsf{T}^{-1}\mathsf{C}$. Then the equation becomes $(\mathsf{S}\mathsf{S}^*)^{-1/2}\,\mathsf{S} = \mathsf{S}\Lambda^{-1}$. Due to the joint eigenstructure of a positive-definite matrix and its positive-definite roots [25], it follows that $(\mathsf{S}\mathsf{S}^*)\,\mathsf{S} = \mathsf{S}\Lambda^2$.

Conversely, suppose that S has full rank and that $(\mathsf{S}\mathsf{S}^*)\,\mathsf{S} = \mathsf{S}\Lambda^2$ for some positive diagonal matrix Λ. Equivalently, $(\mathsf{S}\mathsf{S}^*)\,\boldsymbol{s}_n = \lambda_n^2\,\boldsymbol{s}_n$ for each n. It follows that

$$(\mathsf{S}\mathsf{S}^*)^{-1/2}\,\boldsymbol{s}_n = \lambda_n^{-1}\,\boldsymbol{s}_n \quad \text{for each } n.$$

Multiply by α, and take norms to see that $t_n = \alpha\,\lambda_n^{-1}\sqrt{c_n}$. Combine these equations into the matrix equation $\Lambda = \alpha\,\mathsf{T}^{-1}\,\mathsf{C}$. It follows that S is a fixed point of the algorithm. ∎

E. Fixed Points II

Proposition 24 allows us to provide a partial characterization of the fixed points of any alternating projection between the set of α-tight frames $\mathscr{X}_\alpha$ and any structural constraint set $\mathscr{Z}$ that contains matrices with fixed column norms. This result applies even if the matrices in $\mathscr{Z}$ have additional properties.

Proposition 25: Suppose that the column norms of matrices in $\mathscr{Z}$ are fixed. A sufficient condition for a full-rank matrix Z in $\mathscr{Z}$ to be a fixed point of the alternating projection between $\mathscr{Z}$ and $\mathscr{X}_\alpha$ is that the columns of Z are all eigenvectors of $\mathsf{Z}\mathsf{Z}^*$. That is, $\mathsf{Z}\mathsf{Z}^*\mathsf{Z} = \mathsf{Z}\Lambda$, where Λ is a positive, diagonal matrix with no more than d distinct entries.

Proof: Let $\mathscr{Z}$ be a closed subset of the closed set $\mathscr{S}$, a collection of matrices with prescribed column norms. Suppose that $P(\cdot)$ is a sufficient condition for S to be a fixed point of the alternating projection between $\mathscr{S}$ and $\mathscr{X}_\alpha$. Assume that $P(\mathsf{Z})$ for a matrix Z in $\mathscr{Z}$, and let X be the matrix in $\mathscr{X}_\alpha$ closest to Z. Since $P(\mathsf{Z})$ and $\mathsf{Z} \in \mathscr{S}$, it follows that Z is the matrix in $\mathscr{S}$ closest to X. Therefore, Z is also a fixed point of the alternating minimization between $\mathscr{Z}$ and $\mathscr{X}_\alpha$. An appeal to Proposition 24 completes the proof. ∎

References

[1] O. Christensen, *An introduction to frames and Riesz bases*. Boston: Birkhäuser, 2003.

[2] J. A. Tropp, R. W. Heath, and T. Strohmer, "Optimal CDMA signature sequences, inverse eigenvalue problems and alternating projection," in *Proceedings of the 2003 IEEE International Symposium on Information Theory*, Yokohama, July 2003, p. 407.

[3] J. A. Tropp, I. S. Dhillon, R. W. Heath, and T. Strohmer, "CDMA signature sequences with low peak-to-average-power ratio via alternating projection," in *Proceedings of the 37th Annual Asilomar Conference on Signals, Systems and Computers*, Monterrey, Nov. 2003.

[4] E. W. Cheney and A. A. Goldstein, "Proximity maps for convex sets," *Proc. Amer. Math. Soc.*, vol. 10, no. 3, pp. 448–450, June 1959.

[5] L. M. Bregman, "Finding the common point of convex sets by the method of successive projection," *Dokl. Akad. Nauk. USSR*, vol. 162, pp. 487–490, 1965.

[6] M. Rupf and J. L. Massey, "Optimum sequence multisets for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 40, no. 4, pp. 1261–1266, July 1994.

[7] P. Viswanath and V. Anantharam, "Optimal sequences and sum capacity of synchronous CDMA systems," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 1984–1991, Sept. 1999.

[8] P. Viswanath and A. Anantharam, "Optimal sequences for CDMA under colored noise: A Schur-saddle function property," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1295–1318, June 2002.

[9] L. Zheng and D. N. Tse, "Communication on the Grassmann manifold: A geometric approach to the non-coherent multiple-antenna channel," *IEEE Trans. Inform. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.

[10] R. W. Heath, T. Strohmer, and A. J. Paulraj, "On quasi-orthogonal signatures for CDMA systems," Sept. 2002, submitted to IEEE Trans. Inform. Theory. [Online]. Available: http://www.ece.utexas.edu/~rheath/papers/2002/quasicdma/

[11] V. Tarokh and I.-M. Kim, "Existence and construction of noncoherent unitary space-time codes," *IEEE Trans. Inform. Theory*, vol. 48, no. 12, pp. 3112–3117, Dec. 2002.

[12] T. Strohmer and R. W. Heath, "Grassmannian frames with applications to coding and communication," *Appl. Comp. Harmonic Anal.*, vol. 14, no. 3, pp. 257–275, May 2003.

[13] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," Institut de Recherce en Informatique et Systèmes Aléatoires, Tech. Rep. 1499, Nov. 2002.

[14] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," University of Texas at Austin," ICES Report, Feb. 2003.

[15] D. L. Donoho and M. Elad, "Maximal sparsity representation by ℓ_1 minimization," *Proc. Natl. Acad. Sci.*, vol. 100, pp. 2197–2202, Mar. 2003.

[16] D. V. Sarwate, "Meeting the Welch Bound with equality," in *Sequences and their Applications*. London: Springer, 1998, pp. 79–102.

[17] P. G. Casazza and M. T. Leon, "Existence and construction of finite tight frames," 2002, unpublished.

[18] J. A. Tropp, I. S. Dhillon, and R. W. Heath, "Finite-step algorithms for constructing optimal CDMA signature sequences," The University of Texas at Austin, Wireless Networking and Communications Group Technical Report WNCG-TR-2003-05-08, May 2003, submitted. [Online]. Available: http://www.ece.utexas.edu/~rheath/papers/2003/finitestep/

[19] C. Rose, "CDMA codeword optimization: interference avoidance and class warfare," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2368–2382, Sept. 2001.

[20] S. Ulukus and R. D. Yates, "Iterative construction of optimum signature sequence sets in synchronous CDMA systems," *IEEE Trans. Inform. Theory*, vol. 47, no. 5, pp. 1989–1998, 2001.

[21] C. Rose, S. Ulukus, and R. D. Yates, "Wireless systems and interference avoidance," *IEEE Trans. Wireless Comm.*, vol. 1, no. 3, pp. 415–428, Jul. 2002.

[22] S. Ulukus and A. Yener, "Iterative joint optimization of CDMA signature sequence and receiver filters," in *2002 Conference on Information Sciences and Systems*, Princeton Univ., Mar. 2002.

[23] P. Anigstein and V. Anantharam, "Ensuring convergence of the MMSE iteration for interference avoidance to the global optimum," *IEEE Trans. Inform. Th.*, vol. 49, no. 4, pp. 873–885, Apr. 2003.

[24] L. R. Welch, "Lower bounds on the maximum cross-correlation of signals," *IEEE Trans. Inform. Theory*, vol. 20, pp. 397–399, 1974.

[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[26] A. W. Marshall and I. Olkin, *Inequalities: Theory of majorization and its applications*. Academic Press, 1979.

[27] P. G. Casazza, M. Leon, and J. C. Tremain, "Custom built tight frames," 2002, unpublished.

[28] W. M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, 2003.

[29] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Experimental Math.*, vol. 5, no. 2, pp. 139–159, 1996.

[30] M. T. Chu, "Inverse eigenvalue problems," *SIAM Rev.*, vol. 40, no. 1, pp. 1–39, 1998.

[31] M. T. Chu and G. H. Golub, "Structured inverse eigenvalue problems," *Acta Numerica*, 2001.

[32] S. Friedland, J. Nocedal, and M. L. Overton, "The formulation and analysis of numerical methods for inverse eigenvalue problems," *SIAM J. Numer. Anal.*, pp. 634–667, 1987.

[33] M. T. Chu, "Numerical methods for inverse singular value problems," *SIAM J. Numer. Anal.*, vol. 29, pp. 885–903, 1992.

[34] ——, "On the refinement of a Newton method for the inverse Toeplitz eigenvalue problem," North Carolina State University, Tech. Rep., 1994.

[35] M. T. Chu and K. R. Driessel, "The projected gradient method for least squares matrix approximations with spectral constraints," *SIAM J. Numer. Anal.*, vol. 27, no. 4, pp. 1050–1060, Aug. 1990.

[36] M. T. Chu, "Constructing a Hermitian matrix from its diagonal entries and eigenvalues," *SIAM J. Matrix Anal. Appl.*, vol. 16, pp. 207–217, 1995.

[37] N. J. Higham, "Computing the nearest correlation matrix—a problem from finance," *IMA J. Numer. Anal.*, vol. 22, pp. 329–343, 2002.

[38] M. T. Chu, R. Funderlic, and R. J. Plemmons, "Structured low-rank approximation," *Linear Algebra Appl.*, vol. 366, pp. 157–172, 2003.

[39] J. von Neumann, *Functional Operators, Vol. II*, ser. Annals of Mathematics Studies. Princeton Univ. Press, 1950, no. 22.

[40] G. H. Golub, A. Hoffman, and G. W. Stewart, "A generalization of the Eckart-Young-Mirsky matrix approximation theorem," *Linear Algebra Appl.*, vol. 88/89, pp. 317–327, 1987.

[41] A. Ruhe, "Closest normal matrix finally found!" *BIT*, vol. 27, pp. 585–598, 1987.

[42] N. J. Higham, "Computing a nearest symmetric positive semi-definite matrix," *Linear Algebra Appl.*, vol. 103, pp. 103–118, 1988.

[43] ——, *Applications of Matrix Theory*. Oxford Univ. Press, 1989, ch. Matrix nearness problems and applications, pp. 1–27.

[44] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Johns Hopkins University Press, 1996.

[45] B. F. Green, "The orthogonal approximation of an oblique structure in factor analysis," *Psychometrika*, vol. 17, pp. 429–440, 1952.

[46] P. H. Schönemann, "A generalized solution of the orthogonal Procrustes problem," *Psychometrika*, vol. 31, pp. 1–10, 1966.

[47] N. J. Higham, "Computing the polar decomposition—with applications," *SIAM J. Sci. Stat. Comput.*, no. 7, pp. 1160–1174, 1986.

[48] N. J. Higham and P. Papadimitriou, "A parallel algorithm for computing the polar decomposition," *Parallel Comput.*, vol. 20, no. 8, pp. 1161–1173, Aug. 1994.

[49] N. Aronszajn, *Introduction to the Theory of Hilbert Spaces, Vol. I*. Stillwater, OK: Research Foundation of Oklahoma A & M College, 1950.

[50] S. Kayalar and H. Weinert, "Error bounds for the method of alternating projections," *Math. Control Signal Systems*, pp. 43–59, 1988.

[51] J. de Leeuw, *Information Systems and Data Analysis*. Berlin: Springer, 1994, ch. Block relaxation algorithms in statistics, pp. 308–325.

[52] F. Deutsch, *Best Approximation in Inner-Product Spaces*. New York: Springer-Verlag, 2001.

[53] J. J. Benedetto and M. Fickus, "Finite normalized tight frames," *Adv. Comp. Math.*, 2002, to appear.

[54] P. Casazza and J. Kovačević, "Equal-norm tight frames with erasures," *Adv. Comp. Math.*, 2002.

[55] V. K. Goyal, M. Vetterli, and N. T. Thao, "Quantized overcomplete expansions in $\mathbb{R}^n$: analysis, synthesis and algorithms," *IEEE Trans. Inform. Theory*, vol. 44, no. 1, pp. 16–31, 1998.

[56] M. Sustik, J. A. Tropp, I. S. Dhillon, and R. W. Heath, "Necessary conditions for existence of equiangular tight frames," University of Texas at Austin," Dept. of Computer Sciences, 2003, in preparation.

[57] A. R. Calderbank, P. J. Cameron, W. M. Kantor, and J. J. Seidel, "$\mathbb{Z}_4$-Kerdock codes, orthogonal spreads and extremal Euclidean line sets," *Proc. London Math. Soc.*, vol. 75, no. 2, pp. 436–480, 1997.

[58] N. J. A. Sloane, "Packing planes in four dimensions and other mysteries," in *Proceedings of the Conference on Algebraic Combinatorics and Related Topics*, Yamagata, Japan, Nov. 1997.

[59] D. Agrawal, T. J. Richardson, and R. L. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2618–2626, Sept. 2001.

[60] J. Tropp, I. S. Dhillon, R. W. H. Jr., and T. Strohmer, "Constructing Grassmannian packings via alternating projection," Dec. 2003, in preparation.

[61] V. K. N. Lau, "On the analysis of peak-to-average ratio (PAR) for IS95 and CDMA2000d systems," *IEEE Trans. Vehicular Tech.*, vol. 49, no. 6, pp. 2174–2188, Nov. 2000.

[62] R. T. Rockafellar, *Convex Analysis*. Princeton Univ. Press, 1970.

[63] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, 1992.

[64] I. S. Dhillon and D. S. Modha, "Concept decompositions for large sparse text data using clustering," *Machine Learning*, vol. 42, no. 1, pp. 143–175, 2001.

[65] I. S. Dhillon, E. Marcotte, and U. Roshan, "Diametrical clustering for identifying anti-correlated gene clusters," *Bioinformatics*, vol. 19, pp. 1612–1619, 2003.

[66] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 1962–1973, 2000.

[67] R. R. Meyer, "Sufficient conditions for the convergence of monotonic mathematical programming algorithms," *J. Comp. Sys. Sci.*, vol. 12, pp. 108–121, 1976.

[68] I. Csiszár and G. Tusnády, "Information geometry and alternating minimization procedures," *Statistics and Decisions*, vol. 1, pp. 205–237, 1984, supplemental issue.

[69] W. I. Zangwill, *Nonlinear Programming: A Unified Approach*. Englewood Cliffs: Prentice-Hall, 1969.

[70] W. W. Hogan, "Point-to-set maps in mathematical programming," *SIAM Rev.*, vol. 15, no. 3, pp. 591–603, July 1973.

[71] P. Huard, Ed., *Point-to-Set Maps and Mathematical Programming*, ser. Mathematical Programming Study. Amsterdam: North-Holland Publishing. Co., 1979, no. 10.

[72] G. B. Dantzig, J. Folkman, and N. Shapiro, "On the continuity of the minimum set of a continuous function," *J. Math. Anal. Appl.*, vol. 17, pp. 519–548, 1967.

[73] A. M. Ostrovsky, *Solution of Equations and Systems of Equations*. New York: Academic Press, 1966.

[74] J. C. Fiorot and P. Huard, "Composition and union of general algorithms of optimization," in *Point-to-Set Maps and Mathematical Programming*, ser. Mathematical Programming Studies. North Holland, 1979, vol. 10, pp. 69–85.

[75] F. Deutsch, *Multivariate Approximation Theory*. Basel: Birkhäuser Verlag, 1979, ch. The alternating method of von Neumann, pp. 83–96.

[76] ——, *Approximation Theory, Spline Functions and Applications*. The Netherlands: Kluwer Academic Publishers, 1992, ch. The method of alternating orthogonal projections, pp. 105–121.

[77] N. Wiener, "On the factorization of matrices," *Comment. Math. Helv.*, vol. 29, pp. 97–111, 1955.

[78] I. Halperin, "The product of projection operators," *Acta Sci. Math.*, vol. 23, pp. 96–99, 1962.

[79] K. T. Smith, D. C. Solmon, and S. L. Wagner, "Practical and mathematical aspects of the problem of reconstructing objects from radiographs," *Bull. Amer. Math. Soc.*, vol. 83, pp. 1227–1270, 1977.

[80] J. Xu and L. Zikatonov, "The method of alternating projections and the method of subspace corrections in Hilbert space," *J. Amer. Math. Soc.*, vol. 15, no. 3, pp. 573–597, Apr. 2002.

[81] H. H. Bauschke, F. Deutsch, H. Hundal, and S.-H. Park, "Accelerating the convergence of the method of alternating projections," *Trans. Amer. Math. Soc.*, vol. 355, no. 9, pp. 3433–3461, May 2003.

[82] S. P. Diliberto and E. G. Straus, "On the approximation of a function of several variables by the sum of functions of fewer variables," *Pacific J. Math.*, vol. 1, pp. 195–210, 1951.

[83] G. Aumann, *Über approximative Nomographie I, II and III*. Bayer. Akad. Wiss. Math.-Nat. Kl. S.B., 1958–1960.

[84] M. Golomb, *On Numerical Approximation*. Madison: Univ. of Wisconsin Press, 1959, ch. Approximation by functions of fewer variables, pp. 275–327.

[85] E. W. Cheney and W. A. Light, *Approximation Theory in Tensor Product Spaces*, ser. Lecture Notes in Mathematics. Berlin: Springer-Verlag, 1985, no. 1169.

[86] R. L. Dykstra, "An algorithm for restricted least squares regression," *J. Amer. Stat. Assoc.*, vol. 78, pp. 837–842, 1983.

[87] J. P. Boyle and R. L. Dykstra, *Advances in Order Restricted Inference 37*, ser. Lecture notes in statistics. Springer, 1985, ch. A method for finding projections onto the intersection of convex sets in Hilbert spaces, pp. 28–47.

[88] S.-P. Han, "A successive projection method," *Math. Prog.*, vol. 40, pp. 1–14, 1988.

[89] H. H. Bauschke and J. Borwein, "On projection algorithms for solving convex feasibility problems," *SIAM Rev.*, vol. 38, no. 3, pp. 367–426, 1996.

[90] P. L. Combettes and H. J. Trussell, "Method of successive projections for finding a common point of sets in metric spaces," *J. Optim. Theory and Appl.*, vol. 67, no. 3, pp. 487–507, 1990.

[91] J. A. Cadzow, "Signal enhancement—a composite property mapping algorithm," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 36, no. 1, pp. 49–62, Jan. 1988.

[92] R. Blahut, "Computation of channel capacity and rate distortion functions," *IEEE Trans. Inform. Theory*, vol. 18, pp. 460–473, 1972.

[93] S. Arimoto, "An algorithm for calculating the capacity of an arbitary discrete memoryless channel," *IEEE Trans. Inform. Theory*, vol. 18, pp. 14–20, 1972.

[94] H. J. Landau and W. L. Miranker, "The recovery of distorted band-limited signals," *J. Math Anal. Appl.*, vol. 2, pp. 97–104, 1961.

[95] R. W. Gerchberg, "Super-resolution through error energy reduction," *Optica Acta*, vol. 21, pp. 709–720, 1974.

[96] D. C. Youla and H. Webb, "Image restoration by the method of convex projections. I—theory," *IEEE Trans. Medical Imaging*, vol. 1, pp. 81–94, 1982.

[97] D. L. Donoho and P. B. Stark, "Uncertainty principles and signal recovery," *SIAM J. Appl. Math.*, vol. 49, no. 3, pp. 906–931, Jun. 1989.

[98] A. E. Çetin, Ö. N. Gerek, and Y. Yardimci, "Equiripple FIR filter design by the FFT algorithm," *IEEE Signal Process. Mag.*, Mar. 1997.

[99] W. Glunt and T. L. Hayden, "The nearest "doubly stochastic" matric to a real matrix with the same first moment," *Numer. Linear Algebra Appl.*, vol. 5, no. 6, Dec. 1998.

[100] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Stat. Soc.*, vol. 39, pp. 1–38, 1977.