Constructing Packings in Grassmannian Manifolds via Alternating Projection

I. S. Dhillon, R. W. Heath Jr., T. Strohmer, and J. A. Tropp

CONTENTS

This paper describes a numerical method for finding good packings in Grassmannian manifolds equipped with various metrics. This investigation also encompasses packing in projective spaces. In each case, producing a good packing is equivalent to constructing a matrix that has certain structural and spectral properties. By alternately enforcing the structural condition and then the spectral condition, it is often possible to reach a matrix that satisfies both. One may then extract a packing from this matrix.

This approach is both powerful and versatile. In cases in which experiments have been performed, the alternating projection method yields packings that compete with the best packings recorded. It also extends to problems that have not been studied numerically. For example, it can be used to produce packings of subspaces in real and complex Grassmannian spaces equipped with the Fubini–Study distance; these packings are valuable in wireless communications. One can prove that some of the novel configurations constructed by the algorithm have packing diameters that are nearly optimal.

1. INTRODUCTION

Let us begin with the standard facetious example. Imagine that several mutually inimical nations build their capital cities on the surface of a featureless globe. Being concerned about missile strikes, they wish to locate the closest pair of cities as far apart as possible. In other words, what is the best way to pack points on the surface of a two-dimensional sphere?

This question, first discussed by the Dutch biologist Tammes [Tammes 30], is the prototypical example of packing in a compact metric space. It has been studied in detail for the last 75 years. More recently, researchers have started to ask about packings in other compact spaces. In particular, several communities have investigated how to arrange subspaces in a Euclidean

2000 AMS Subject Classification: Primary 51N15, 52C17

Keywords: Combinatorial optimization, packing, projective spaces, Grassmannian spaces, Tammes' Problem

1058-6458/2008 $ 0.50 per page

space so that they are as distinct as possible. An equivalent formulation is to find the best packings of points in a Grassmannian manifold. This problem has applications in quantum computing and wireless communications. There has been theoretical interest in subspace packing since the 1960s [Tóth 65], but the first detailed numerical study appears in a 1996 paper of Conway, Hardin, and Sloane [Conway et al. 96].

The aim of this paper is to describe a flexible numerical method that can be used to construct packings in Grassmannian manifolds equipped with several different metrics. The rest of this introduction provides a formal statement of abstract packing problems, and it offers an overview of our approach to solving them.

1.1 Abstract Packing Problems

Although we will be working with Grassmannian manifolds, it is more instructive to introduce packing problems in an abstract setting. Let $\mathbb{M}$ be a compact metric space endowed with the distance function $\text{dist}_{\mathbb{M}}$. The *packing diameter* of a finite subset $\mathcal{X}$ is the minimum distance between some pair of distinct points drawn from $\mathcal{X}$. That is,

$$\text{pack}_{\mathbb{M}}(\mathcal{X}) \stackrel{\text{def}}{=} \min_{m \neq n} \text{dist}_{\mathbb{M}}(x_m, x_n).$$

In other words, the packing diameter of a set is the diameter of the largest open ball that can be centered at each point of the set without encompassing any other point. (It is also common to study the *packing radius*, which is half the diameter of this ball.) An optimal packing of N points is an ensemble $\mathcal{X}$ that solves the mathematical program

$$\max_{|\mathcal{X}|=N} \text{pack}_{\mathbb{M}}(\mathcal{X}),$$

where $|\cdot|$ returns the cardinality of a finite set. The optimal packing problem is guaranteed to have a solution because the metric space is compact and the objective is a continuous function of the ensemble $\mathcal{X}$.

This article focuses on a *feasibility problem* closely connected with optimal packing. Given a number ρ, the goal is to produce a set of N points for which

$$\text{pack}_{\mathbb{M}}(\mathcal{X}) \geq \rho. \tag{1–1}$$

This problem is notoriously difficult to solve because it is highly nonconvex, and it is even more difficult to determine the maximum value of ρ for which the feasibility problem is soluble. This maximum value of ρ corresponds to the diameter of an optimal packing.

1.2 Alternating Projection

We will attempt to solve the feasibility problem (1–1) in Grassmannian manifolds equipped with a number of different metrics, but the same basic algorithm applies in each case. Here is a high-level description of our approach.

First, we show that each configuration of subspaces is associated with a block Gram matrix whose blocks control the distances between pairs of subspaces. Then we prove that a configuration solves the feasibility problem (1–1) if and only if its Gram matrix possesses both a structural property and a spectral property. The overall algorithm consists of the following steps.

1. Choose an initial configuration and construct its matrix.
2. Alternately enforce the structural condition and the spectral condition in hope of reaching a matrix that satisfies both.
3. Extract a configuration of subspaces from the output matrix.

In our work, we choose the initial configuration randomly and then remove similar subspaces from it with a simple algorithm. One can imagine more sophisticated approaches to constructing the initial configuration.

Flexibility and ease of implementation are the major advantages of alternating projection. This article demonstrates that appropriate modifications of this basic technique allow us to construct solutions to the feasibility problem in Grassmannian manifolds equipped with various metrics. Some of these problems have never been studied numerically, and the experiments point toward intriguing phenomena that deserve theoretical attention. Moreover, we believe that the possibilities of this method have not been exhausted and that it will see other applications in the future.

Alternating projection does have several drawbacks. It may converge very slowly, and it does not always yield a high level of numerical precision. In addition, it may not deliver good packings when the ambient dimension or the number of subspaces in the configuration is large.

1.3 Motivation and Related Work

This work was motivated by applications in electrical engineering. In particular, subspace packings solve certain extremal problems that arise in multiple-antenna communication systems [Zheng and Tse 02, Hochwald et al. 00, Love et al. 04]. This application requires complex

Grassmannian packings that consist of a small number of subspaces in an ambient space of low dimension. Our algorithm is quite effective in this parameter regime. The resulting packings fill a significant gap in the literature, since existing tables consider only the real case [Sloane 04a]. See Section 6.1 for additional discussion of the wireless application.

The approach to packing via alternating projection was discussed in a previous publication [Tropp et al. 05], but the experiments were limited to a single case. We are aware of several other numerical methods that can be used to construct packings in Grassmannian manifolds [Conway et al. 96, Trosset 01, Agrawal et al. 01]. These techniques rely on ideas from nonlinear programming.

1.4 Historical Interlude

The problem of constructing optimal packings in various metric spaces has a long and lovely history. The most famous example may be Kepler's conjecture that an optimal packing of spheres in three-dimensional Euclidean space[1] locates them at the points of a face-centered cubic lattice. For millennia, greengrocers have applied this theorem when stacking oranges, but it has only been established rigorously within the last few years [Hales 04]. Packing problems play a major role in modern communications because error-correcting codes may be interpreted as packings in the Hamming space of binary strings [Cover and Thomas 91]. The standard reference on packing is the magnum opus of Conway and Sloane [Conway and Sloane 98]. Classical monographs on the subject were written by L. Fejes Tóth [Tóth 64] and C. A. Rogers [Rogers 64].

The idea of applying alternating projection to feasibility problems first appeared in the work of von Neumann [von Neumann 50]. He proved that an alternating projection between two closed subspaces of a Hilbert space converges to the orthogonal projection of the initial iterate onto the intersection of the two subspaces. Cheney and Goldstein subsequently showed that an alternating projection between two closed, convex subsets of a Hilbert space always converges to a point in their intersection (provided that the intersection is nonempty) [Cheney and Goldstein 59]. This result does not apply in our setting because one of the constraint sets we define is not convex.

[1]The infinite extent of a Euclidean space necessitates a more subtle definition of an optimal packing.

1.5 Outline of Article

Here is a brief overview of this article. In Section 2, we develop a basic description of Grassmannian manifolds and present some natural metrics. Section 3 explains why alternating projection is a natural algorithm for producing Grassmannian packings, and it outlines how to apply this algorithm for one specific metric. Section 4 gives some theoretical upper bounds on the optimal diameter of packings in Grassmannian manifolds. Section 5 describes the outcomes of an extensive set of numerical experiments and explains how to apply the algorithm to other metrics. Section 6 offers some discussion and conclusions. Appendix 7 explores how our methodology applies to Tammes' problem of packing on the surface of a sphere.

Our experiments resulted in tables of packing diameters. We did not store the configurations produced by the algorithm. The Matlab code that produced these data is available on request from jtropp@acm.caltech.edu.

These tables and figures are intended only to describe the results of our experiments; it is likely that many of the packing diameters could be improved with additional effort. In all cases, we present the results of calculations for the stated problem, even if we obtained a better packing by solving a different problem. For example, a complex packing should always improve on the corresponding real packing. If the numbers indicate otherwise, it just means that the complex experiment yielded an inferior result. As a second example, the optimal packing diameter must not decrease as the number of points increases. When the numbers indicate otherwise, it means that running the algorithm with more points yielded a better result than running it with fewer. These failures may reflect the difficulty of various packing problems.

2. PACKING IN GRASSMANNIAN MANIFOLDS

This section introduces our notation and a simple description of the Grassmannian manifold. It presents several natural metrics on the manifold, and it shows how to represent a configuration of subspaces in matrix form.

2.1 Preliminaries

We work in the vector space $\mathbb{C}^d$. The symbol * denotes the complex-conjugate transpose of a vector (or matrix). We equip the vector space with its usual inner product $\langle \boldsymbol{x},\ \boldsymbol{y} \rangle = \boldsymbol{y}^* \boldsymbol{x}$.

This inner product generates the ℓ_2 norm via the formula $\|\boldsymbol{x}\|_2^2 = \langle \boldsymbol{x},\ \boldsymbol{x} \rangle$.

The d-dimensional identity matrix is $\mathbf{I}_d$; we sometimes omit the subscript if it is unnecessary. A square matrix is *positive semidefinite* when its eigenvalues are all nonnegative. We write $\boldsymbol{X} \succcurlyeq \mathbf{0}$ to indicate that $\boldsymbol{X}$ is positive semidefinite.

A square complex matrix $\boldsymbol{U}$ is *unitary* if it satisfies $\boldsymbol{U}^*\boldsymbol{U} = \mathbf{I}$. If in addition, the entries of $\boldsymbol{U}$ are real, the matrix is *orthogonal*. The unitary group $\mathbb{U}(d)$ can be presented as the collection of all $d \times d$ unitary matrices with ordinary matrix multiplication. The real orthogonal group $\mathbb{O}(d)$ can be presented as the collection of all $d \times d$ real orthogonal matrices with the usual matrix multiplication.

Suppose that $\boldsymbol{X}$ is a general matrix. The Frobenius norm is calculated as $\|\boldsymbol{X}\|_{\mathrm{F}}^2 = \operatorname{trace} \boldsymbol{X}^*\boldsymbol{X}$, where the trace operator sums the diagonal entries of the matrix. The spectral norm is denoted by $\|\boldsymbol{X}\|_{2,2}$; it returns the largest singular value of $\boldsymbol{X}$. Both these norms are unitarily invariant, which means that $\|\boldsymbol{U}\boldsymbol{X}\boldsymbol{V}^*\| = \|\boldsymbol{X}\|$ whenever $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary.

2.2 Grassmannian Manifolds

The (complex) Grassmannian manifold $\mathbb{G}(K, \mathbb{C}^d)$ is the collection of all K-dimensional subspaces of $\mathbb{C}^d$. This space is isomorphic to a quotient of unitary groups:

$$\mathbb{G}(K, \mathbb{C}^d) \cong \frac{\mathbb{U}(d)}{\mathbb{U}(K) \times \mathbb{U}(d-K)}.$$

To understand the equivalence, note that each orthonormal basis from $\mathbb{C}^d$ can be split into K vectors that span a K-dimensional subspace and $d-K$ vectors that span the orthogonal complement of that subspace. To obtain a unique representation for the subspace, it is necessary to divide by isometries that fix the subspace and by isometries that fix its complement. It is evident that $\mathbb{G}(K, \mathbb{C}^d)$ is always isomorphic to $\mathbb{G}(d-K, \mathbb{C}^d)$.

Similarly, the real Grassmannian manifold $\mathbb{G}(K, \mathbb{R}^d)$ is the collection of all K-dimensional subspaces of $\mathbb{R}^d$. This space is isomorphic to a quotient of orthogonal groups:

$$\mathbb{G}(K, \mathbb{R}^d) \cong \frac{\mathbb{O}(d)}{\mathbb{O}(K) \times \mathbb{O}(d-K)}.$$

If we need to refer to the real and complex Grassmannians simultaneously, we write $\mathbb{G}(K, \mathbb{F}^d)$.

In the theoretical development, we concentrate on complex Grassmannians, since the development for the real case is identical, except that all the matrices are real-valued instead of complex-valued. A second reason for focusing on the complex case is that complex packings arise naturally in wireless communications [Love et al. 03].

When each subspace has dimension $K = 1$, the Grassmannian manifold reduces to a simpler object called a *projective space*. The elements of a projective space can be viewed as lines through the origin of a Euclidean space. The standard notation is $\mathbb{P}^{d-1}(\mathbb{F}) \stackrel{\text{def}}{=} \mathbb{G}(1, \mathbb{F}^d)$. We will spend a significant amount of attention on packings of this manifold.

2.3 Principal Angles

Suppose that $\mathcal{S}$ and $\mathcal{T}$ are two subspaces in $\mathbb{G}(K, \mathbb{C}^d)$. These subspaces are inclined against each other by K different *principal angles*. The smallest principal angle θ_1 is the minimum angle formed by a pair of unit vectors $(\boldsymbol{s}_1, \boldsymbol{t}_1)$ drawn from $\mathcal{S} \times \mathcal{T}$. That is,

$$\theta_1 = \min_{(\boldsymbol{s}_1, \boldsymbol{t}_1) \in \mathcal{S} \times \mathcal{T}} \arccos \langle \boldsymbol{s}_1,\ \boldsymbol{t}_1 \rangle$$

subject to

$$\|\boldsymbol{s}_1\|_2 = 1 \quad \text{and} \quad \|\boldsymbol{t}_1\|_2 = 1.$$

The second principal angle θ_2 is defined as the smallest angle attained by a pair of unit vectors $(\boldsymbol{s}_2, \boldsymbol{t}_2)$ that is orthogonal to the first pair, i.e.,

$$\theta_2 = \min_{(\boldsymbol{s}_2, \boldsymbol{t}_2) \in \mathcal{S} \times \mathcal{T}} \arccos \langle \boldsymbol{s}_2,\ \boldsymbol{t}_2 \rangle$$

subject to

$$\begin{aligned} \|\boldsymbol{s}_2\|_2 = 1 \quad \text{and} \quad \|\boldsymbol{t}_2\|_2 = 1, \\ \langle \boldsymbol{s}_1,\ \boldsymbol{s}_2 \rangle = 0 \quad \text{and} \quad \langle \boldsymbol{t}_1,\ \boldsymbol{t}_2 \rangle = 0. \end{aligned}$$

The remaining principal angles are defined analogously. The sequence of principal angles is nondecreasing, and it is contained in the range $[0, \pi/2]$. We consider only metrics that are functions of the principal angles between two subspaces.

Let us present a more computational definition of the principal angles [Björck and Golub 73]. Suppose that the columns of $\boldsymbol{S}$ and $\boldsymbol{T}$ form orthonormal bases for the subspaces $\mathcal{S}$ and $\mathcal{T}$. More rigorously, $\boldsymbol{S}$ is a $d \times K$ matrix that satisfies $\boldsymbol{S}^*\boldsymbol{S} = \mathbf{I}_K$ and $\operatorname{range} \boldsymbol{S} = \mathcal{S}$. The matrix $\boldsymbol{T}$ has an analogous definition. Next we compute a singular value decomposition of the product $\boldsymbol{S}^*\boldsymbol{T}$:

$$\boldsymbol{S}^*\boldsymbol{T} = \boldsymbol{U}\boldsymbol{C}\boldsymbol{V}^*,$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are $K \times K$ unitary matrices and $\boldsymbol{C}$ is a nonnegative diagonal matrix with nonincreasing entries. The matrix $\boldsymbol{C}$ of singular values is uniquely determined, and its entries are the cosines of the principal angles between $\mathcal{S}$ and $\mathcal{T}$:

$$c_{kk} = \cos \theta_k \qquad k = 1, 2, \ldots, K.$$

This definition of the principal angles is most convenient numerically because singular value decompositions can be computed efficiently with standard software. We also note that this definition of the principal angles does not depend on the choice of matrices $\boldsymbol{S}$ and $\boldsymbol{T}$ that represent the two subspaces.

2.4 Metrics on Grassmannian Manifolds

Grassmannian manifolds admit many interesting metrics, which lead to different packing problems. This section describes some of these metrics.

1. The *chordal distance* between two K-dimensional subspaces $\mathcal{S}$ and $\mathcal{T}$ is given by

$$\begin{aligned}\operatorname{dist}_{\text{chord}}(\mathcal{S},\mathcal{T}) &\stackrel{\text{def}}{=} \sqrt{\sin^2\theta_1 + \cdots + \sin^2\theta_K} \\ &= \left[K - \|\boldsymbol{S}^*\boldsymbol{T}\|_{\text{F}}^2\right]^{1/2}. \qquad (2\text{–}1)\end{aligned}$$

   The values of this metric range between zero and $\sqrt{K}$. The chordal distance is the easiest to work with, and it also yields the most symmetric packings [Conway et al. 96].

2. The *spectral distance* is

$$\begin{aligned}\operatorname{dist}_{\text{spec}}(\mathcal{S},\mathcal{T}) &\stackrel{\text{def}}{=} \min_k \sin\theta_k \\ &= \left[1 - \|\boldsymbol{S}^*\boldsymbol{T}\|_{2,2}^2\right]^{1/2}. \qquad (2\text{–}2)\end{aligned}$$

   The values of this metric range between zero and one. As we will see, this metric promotes a special type of packing called an *equi-isoclinic* configuration of subspaces.

3. The *Fubini–Study distance* is

$$\begin{aligned}\operatorname{dist}_{\text{FS}}(\mathcal{S},\mathcal{T}) &\stackrel{\text{def}}{=} \arccos\left(\prod\nolimits_k \cos\theta_k\right) \\ &= \arccos\left|\det \boldsymbol{S}^*\boldsymbol{T}\right|. \qquad (2\text{–}3)\end{aligned}$$

   This metric takes values between zero and $\pi/2$. It plays an important role in wireless communications [Love and Heath 05a, Love and Heath 05b].

4. The *geodesic distance* is

$$\operatorname{dist}_{\text{geo}}(\mathcal{S},\mathcal{T}) \stackrel{\text{def}}{=} \sqrt{\theta_1^2 + \cdots + \theta_K^2}.$$

   This metric takes values between zero and $\pi\sqrt{K}/2$. From the point of view of differential geometry, the geodesic distance is very natural, but it does not seem to lead to very interesting packings [Conway et al. 96], so we will not discuss it any further.

Grassmannian manifolds support several other interesting metrics, some of which are listed in [Barg and Nogin 02]. In case we are working in a projective space, i.e., $K = 1$, all of these metrics reduce to the acute angle between two lines or the sine thereof. Therefore, the metrics are equivalent up to a monotonically increasing transformation, and they promote identical packings.

2.5 Representing Configurations of Subspaces

Suppose that $\mathcal{X} = \{\mathcal{S}_1, \ldots, \mathcal{S}_N\}$ is a collection of N subspaces in $\mathbb{G}(K, \mathbb{C}^d)$. Let us develop a method for representing this configuration numerically. To each subspace $\mathcal{S}_n$, we associate a (nonunique) $d \times K$ matrix $\boldsymbol{X}_n$ whose columns form an orthonormal basis for that subspace, i.e., $\boldsymbol{X}_n^*\boldsymbol{X}_n = \mathbf{I}_K$ and range $\boldsymbol{X}_n = \mathcal{S}_n$. Now collate these N matrices into a $d \times KN$ configuration matrix

$$\boldsymbol{X} \stackrel{\text{def}}{=} \begin{bmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2 & \ldots & \boldsymbol{X}_N\end{bmatrix}.$$

In the sequel, we do not distinguish between the configuration $\mathcal{X}$ and the matrix $\boldsymbol{X}$.

The *Gram matrix* of $\boldsymbol{X}$ is defined as the $KN \times KN$ matrix $\boldsymbol{G} = \boldsymbol{X}^*\boldsymbol{X}$. By construction, the Gram matrix is positive semidefinite, and its rank does not exceed d. It is best to regard the Gram matrix as an $N \times N$ block matrix composed of $K \times K$ blocks, and we index it as such. Observe that each block satisfies

$$\boldsymbol{G}_{mn} = \boldsymbol{X}_m^*\boldsymbol{X}_n.$$

In particular, each diagonal block $\boldsymbol{G}_{nn}$ is an identity matrix. Meanwhile, the singular values of the off-diagonal block $\boldsymbol{G}_{mn}$ equal the cosines of the principal angles between the two subspaces range $\boldsymbol{X}_m$ and range $\boldsymbol{X}_n$.

Conversely, let $\boldsymbol{G}$ be an $N \times N$ block matrix with each block of size $K \times K$. Suppose that the matrix is positive semidefinite, that its rank does not exceed d, and that its diagonal blocks are identity matrices. Then we can factor $\boldsymbol{G} = \boldsymbol{X}^*\boldsymbol{X}$, where $\boldsymbol{X}$ is a $d \times KN$ configuration matrix. That is, the columns of $\boldsymbol{X}$ form orthogonal bases for N different K-dimensional subspaces of $\mathbb{C}^d$.

As we will see, each metric on the Grassmannian manifold leads to a measure of "magnitude" for the off-diagonal blocks on the Gram matrix $\boldsymbol{G}$. A configuration solves the feasibility problem (1–1) if and only if each off-diagonal block of its Gram matrix has sufficiently small magnitude. So solving the feasibility problem is equivalent to producing a Gram matrix with appropriate properties.

3. ALTERNATING PROJECTION FOR CHORDAL DISTANCE

In this section, we elaborate on the idea that solving the feasibility problem is equivalent to constructing a Gram matrix that meets certain conditions. These conditions fall into two categories: structural properties and spectral properties. This observation leads naturally to an alternating projection algorithm for solving the feasibility problem. The algorithm alternately enforces the structural properties and the spectral properties in hope of producing a Gram matrix that satisfies them all. This section illustrates how this approach unfolds when distances are measured with respect to the chordal metric. In Section 5, we describe adaptations for other metrics.

3.1 Packings with Chordal Distance

Suppose that we seek a packing of N subspaces in $\mathbb{G}(K, \mathbb{C}^d)$ equipped with the chordal distance. If $\boldsymbol{X}$ is a configuration of N subspaces, its packing diameter is

$$\begin{aligned}\operatorname{pack}_{\text{chord}}(\boldsymbol{X}) &\stackrel{\text{def}}{=} \min_{m \neq n} \operatorname{dist}_{\text{chord}}(\boldsymbol{X}_m, \boldsymbol{X}_n) \\ &= \min_{m \neq n} \left[K - \|\boldsymbol{X}_m^* \boldsymbol{X}_n\|_{\text{F}}^2 \right]^{1/2}.\end{aligned}$$

Given a parameter ρ, the feasibility problem elicits a configuration $\boldsymbol{X}$ that satisfies

$$\min_{m \neq n} \left[K - \|\boldsymbol{X}_m^* \boldsymbol{X}_n\|_{\text{F}}^2 \right]^{1/2} \geq \rho.$$

We may rearrange this inequality to obtain a simpler condition:

$$\max_{m \neq n} \|\boldsymbol{X}_m^* \boldsymbol{X}_n\|_{\text{F}} \leq \mu, \tag{3–1}$$

where

$$\mu = \sqrt{K - \rho^2}. \tag{3–2}$$

In fact, we may formulate the feasibility problem purely in terms of the Gram matrix. Suppose that the configuration $\boldsymbol{X}$ satisfies (3–1) with parameter μ. Then its Gram matrix $\boldsymbol{G}$ must have the following six properties:

1. $\boldsymbol{G}$ is Hermitian.
2. Each diagonal block of $\boldsymbol{G}$ is an identity matrix.
3. $\|\boldsymbol{G}_{mn}\|_{\text{F}} \leq \mu$ for each $m \neq n$.
4. $\boldsymbol{G}$ is positive semidefinite.
5. $\boldsymbol{G}$ has rank d or less.
6. $\boldsymbol{G}$ has trace KN.

Some of these properties are redundant, but we have listed them separately for reasons soon to become apparent. Conversely, suppose that a matrix $\boldsymbol{G}$ satisfies properties 1–6. Then it is always possible to factor it to extract a configuration of N subspaces that solves (3–1). The factorization of $\boldsymbol{G} = \boldsymbol{X}^* \boldsymbol{X}$ can be obtained most easily from an eigenvalue decomposition of $\boldsymbol{G}$.

3.2 The Algorithm

Observe that properties 1–3 are *structural* properties. By this, we mean that they constrain the entries of the Gram matrix directly. Properties 4–6, on the other hand, are *spectral* properties. That is, they control the eigenvalues of the matrix. It is not easy to enforce structural and spectral properties simultaneously, so we must resort to half measures. Starting from an initial matrix, our algorithm will alternately enforce properties 1–3 and properties 4–6 in hope of reaching a matrix that satisfies all six properties at once.

To be more rigorous, let us define the structural constraint set

$$\begin{aligned}\mathcal{H}(\mu) \stackrel{\text{def}}{=} \big\{ \boldsymbol{H} \in \mathbb{C}^{KN \times KN} : \boldsymbol{H} = \boldsymbol{H}^*, \boldsymbol{H}_{nn} = \mathbf{I}_K \\ \text{for } n = 1, 2, \dots, N, \\ \text{and } \|\boldsymbol{H}_{mn}\|_{\text{F}} \leq \mu \text{ for all } m \neq n \big\}.\end{aligned} \tag{3–3}$$

Although the structural constraint set evidently depends on the parameter μ, we will usually eliminate μ from the notation for simplicity. We also define the spectral constraint set

$$\begin{aligned}\mathcal{G} \stackrel{\text{def}}{=} \big\{ \boldsymbol{G} \in \mathbb{C}^{KN \times KN} : \boldsymbol{G} \succcurlyeq \mathbf{0}, \operatorname{rank} \boldsymbol{G} \leq d, \\ \text{and } \operatorname{trace} \boldsymbol{G} = KN \big\}.\end{aligned} \tag{3–4}$$

Both constraint sets are closed and bounded, hence compact. The structural constraint set $\mathcal{H}$ is convex, but the spectral constraint set is not.

To solve the feasibility problem (3–1), we must find a matrix that lies in the intersection of $\mathcal{G}$ and $\mathcal{H}$. This section states the algorithm, and the succeeding two sections provide some implementation details.

Algorithm 3.1. (Alternating Projection.)
INPUT:

- A $KN \times KN$ Hermitian matrix $\boldsymbol{G}^{(0)}$
- The maximum number of iterations T

OUPUT:

- A $KN \times KN$ matrix $\boldsymbol{G}_{\text{out}}$ that belongs to $\mathcal{G}$ and whose diagonal blocks are identity matrices

PROCEDURE:

1. Initialize $t \leftarrow 0$.
2. Determine a matrix $\boldsymbol{H}^{(t)}$ that solves
$$\min_{\boldsymbol{H} \in \mathcal{H}} \left\| \boldsymbol{H} - \boldsymbol{G}^{(t)} \right\|_{\text{F}}.$$
3. Determine a matrix $\boldsymbol{G}^{(t+1)}$ that solves
$$\min_{\boldsymbol{G} \in \mathcal{G}} \left\| \boldsymbol{G} - \boldsymbol{H}^{(t)} \right\|_{\text{F}}.$$
4. Increment t.
5. If $t < T$, return to Step 2.
6. Define the block-diagonal matrix $\boldsymbol{D} = \operatorname{diag} \boldsymbol{G}^{(T)}$.
7. Return the matrix
$$\boldsymbol{G}_{\text{out}} = \boldsymbol{D}^{-1/2} \boldsymbol{G}^{(T)} \boldsymbol{D}^{-1/2}.$$

The iterates generated by this algorithm are not guaranteed to converge in norm. Therefore, we have chosen to halt the algorithm after a fixed number of steps instead of checking the behavior of the sequence of iterates. We discuss the convergence properties of the algorithm in the sequel.

The scaling in the last step normalizes the diagonal blocks of the matrix but preserves its inertia (i.e., numbers of negative, zero, and positive eigenvalues). Since $\boldsymbol{G}^{(T)}$ is a positive semidefinite matrix with rank d or less, the output matrix $\boldsymbol{G}_{\text{out}}$ shares these traits. It follows that the output matrix always admits a factorization $\boldsymbol{G}_{\text{out}} = \boldsymbol{X}^* \boldsymbol{X}$, where $\boldsymbol{X}$ is a $d \times KN$ configuration matrix. Property 3 is the only one of the six properties that may be violated.

3.3 The Matrix Nearness Problems

To implement Algorithm 3.1, we must solve the matrix nearness problems in steps 2 and 3. The first one is straightforward.

Proposition 3.2. *Let $\boldsymbol{G}$ be a Hermitian matrix. With respect to the Frobenius norm, the unique matrix in $\mathcal{H}(\mu)$ nearest to $\boldsymbol{G}$ has diagonal blocks equal to the identity and off-diagonal blocks that satisfy*

$$\boldsymbol{H}_{mn} = \begin{cases} \boldsymbol{G}_{mn} & \text{if } \|\boldsymbol{G}_{mn}\|_{\text{F}} \leq \mu, \\ \mu \boldsymbol{G}_{mn} / \|\boldsymbol{G}_{mn}\|_{\text{F}}, & \text{otherwise.} \end{cases}$$

It is rather more difficult to find a nearest matrix in the spectral constraint set. To state the result, we define the plus operator by the rule $(x)_+ = \max\{0, x\}$.

Proposition 3.3. *Let $\boldsymbol{H}$ be a Hermitian matrix whose eigenvalue decomposition is $\sum_{j=1}^{KN} \lambda_j \boldsymbol{u}_j \boldsymbol{u}_j^*$ with the eigenvalues arranged in nonincreasing order: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{KN}$. With respect to the Frobenius norm, a matrix in $\mathcal{G}$ closest to $\boldsymbol{H}$ is given by*

$$\sum\nolimits_{j=1}^{d} (\lambda_j - \gamma)_+ \boldsymbol{u}_j \boldsymbol{u}_j^*,$$

where the scalar γ is chosen such that

$$\sum\nolimits_{j=1}^{d} (\lambda_j - \gamma)_+ = KN.$$

This best approximation is unique, provided that $\lambda_d > \lambda_{d+1}$.

The nearest matrix described by this theorem can be computed efficiently from an eigenvalue decomposition of $\boldsymbol{H}$. (See [Golub and Van Loan 96] for computational details.) The value of γ is uniquely determined, but one must solve a small root-finding problem to find it. The bisection method is an appropriate technique, since the plus operator is nondifferentiable. We omit the details, which are routine.

Proof: Given a Hermitian matrix $\boldsymbol{A}$, denote by $\boldsymbol{\lambda}(\boldsymbol{A})$ the vector of eigenvalues arranged in nonincreasing order. Then we may decompose $\boldsymbol{A} = \boldsymbol{U}\{\operatorname{diag} \boldsymbol{\lambda}(\boldsymbol{A})\}\boldsymbol{U}^*$ for some unitary matrix $\boldsymbol{U}$.

Finding the matrix in $\mathcal{G}$ closest to $\boldsymbol{H}$ is equivalent to solving the optimization problem

$$\min_{\boldsymbol{G}} \|\boldsymbol{G} - \boldsymbol{H}\|_{\text{F}}^2$$

subject to

$$\begin{aligned} \lambda_j(\boldsymbol{G}) &\geq 0 \text{ for } j = 1, \ldots, d, \\ \lambda_j(\boldsymbol{G}) &= 0 \text{ for } j = d+1, \ldots, KN, \\ \sum\nolimits_{j=1}^{KN} \lambda_j(\boldsymbol{G}) &= KN. \end{aligned}$$

First, we fix the eigenvalues of $\boldsymbol{G}$ and minimize with respect to the unitary part of its eigenvalue decomposition. In consequence of the Hoffman–Wielandt theorem [Horn and Johnson 85], the objective function is bounded below:

$$\|\boldsymbol{G} - \boldsymbol{H}\|_{\text{F}}^2 \geq \|\boldsymbol{\lambda}(\boldsymbol{G}) - \boldsymbol{\lambda}(\boldsymbol{H})\|_2^2.$$

Equality holds if and only if $\boldsymbol{G}$ and $\boldsymbol{H}$ are simultaneously diagonalizable by a unitary matrix. Therefore,

if we decompose $\boldsymbol{H} = \boldsymbol{U}\{\operatorname{diag}\boldsymbol{\lambda}(\boldsymbol{H})\}\boldsymbol{U}^*$, the objective function attains its minimal value whenever $\boldsymbol{G} = \boldsymbol{U}\{\operatorname{diag}\boldsymbol{\lambda}(\boldsymbol{G})\}\boldsymbol{U}^*$. Note that the matrix $\boldsymbol{U}$ may not be uniquely determined.

We find the optimal vector of eigenvalues $\boldsymbol{\xi}$ for the matrix $\boldsymbol{G}$ by solving the (strictly) convex program

$$\min_{\boldsymbol{\xi}} \|\boldsymbol{\xi} - \boldsymbol{\lambda}(\boldsymbol{H})\|_2^2$$

subject to

$$\begin{aligned} \xi_j &\geq 0 \text{ for } j = 1, \dots, d, \\ \xi_j &= 0 \text{ for } j = d+1, \dots, KN, \\ \textstyle\sum_{j=1}^{KN} \xi_j &= KN. \end{aligned}$$

This minimization is accomplished by an application of Karush–Kuhn–Tucker theory [Rockafellar 70]. In short, the top d eigenvalues of $\boldsymbol{H}$ are translated an equal amount, and those that become negative are set to zero. The size of the translation is chosen to fulfill the third condition (which controls the trace of $\boldsymbol{G}$). The entries of the optimal $\boldsymbol{\xi}$ are nonincreasing on account of the ordering of $\boldsymbol{\lambda}(\boldsymbol{H})$.

Finally, the uniqueness claim follows from the fact that the eigenspace associated with the top d eigenvectors of $\boldsymbol{H}$ is uniquely determined if and only if $\lambda_d(\boldsymbol{H}) > \lambda_{d+1}(\boldsymbol{H})$. □

3.4 Choosing an Initial Configuration

The success of the algorithm depends on adequate selection of the input matrix $\boldsymbol{G}^{(0)}$. We have found that the following strategy is reasonably effective. It chooses random subspaces and adds them to the initial configuration only if they are sufficiently distant from the subspaces that have already been chosen.

Algorithm 3.4. (Initial Configuration.)
INPUT:

- The ambient dimension d, the subspace dimension K, and the number N of subspaces
- An upper bound τ on the similarity between subspaces
- The maximum number T of random selections

OUTPUT:

- A $KN \times KN$ matrix $\boldsymbol{G}$ from $\mathcal{G}$ whose off-diagonal blocks also satisfy $\|\boldsymbol{G}_{mn}\|_{\mathrm{F}} \leq \tau$

PROCEDURE:

1. Initialize $t \leftarrow 0$ and $n \leftarrow 1$.
2. Increment t. If $t > T$, print a failure notice and stop.
3. Pick a $d \times K$ matrix $\boldsymbol{X}_n$ whose range is a uniformly random subspace in $\mathbb{G}(K, \mathbb{C}^d)$.
4. If $\|\boldsymbol{X}_m^* \boldsymbol{X}_n\|_{\mathrm{F}} \leq \tau$ for each $m = 1, \dots, n-1$, then increment n.
5. If $n \leq N$, return to step 2.
6. Form the matrix $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 & \dots & \boldsymbol{X}_N \end{bmatrix}$.
7. Return the Gram matrix $\boldsymbol{G} = \boldsymbol{X}^* \boldsymbol{X}$.

To implement step 3, we use the method developed in [Stewart 80]. Draw a $d \times K$ matrix whose entries are iid complex, standard normal random variables, and perform a $\boldsymbol{QR}$ decomposition. The first K columns of the unitary part of the $\boldsymbol{QR}$ decomposition form an orthonormal basis for a random K-dimensional subspace.

The purpose of the parameter τ is to prevent the starting configuration $\boldsymbol{X}$ from containing blocks that are nearly identical. The extreme case $\tau = \sqrt{K}$ places no restriction on the similarity between blocks. If τ is chosen too small (or if we are unlucky in our random choices), then this selection procedure may fail. For this reason, we add an iteration counter to prevent the algorithm from entering an infinite loop. We typically choose values of τ very close to the maximum value.

3.5 Theoretical Behavior of the Algorithm

It is important to be aware that packing problems are typically difficult to solve. Therefore, we cannot expect that our algorithm will necessarily produce a point in the intersection of the constraint sets. One may ask whether we can make any guarantees about the behavior of Algorithm 3.1. This turns out to be difficult. Indeed, there is potential that an alternating projection algorithm will fail to generate a convergent sequence of iterates [Meyer 76]. Nevertheless, it can be shown that the sequence of iterates has accumulation points and that these accumulation points satisfy a weak structural property.

In practice, the alternating projection algorithm seems to converge, but a theoretical justification for this observation is lacking. A more serious problem is that the algorithm frequently requires as many as five thousand iterations before the iterates settle down. This is one of the major weaknesses of our approach.

For reference, we offer the best theoretical convergence result that we know. The distance between a matrix and a compact collection of matrices is defined as

$$\operatorname{dist}(\boldsymbol{M}, \mathcal{C}) \stackrel{\text{def}}{=} \min_{\boldsymbol{C} \in \mathcal{C}} \|\boldsymbol{M} - \boldsymbol{C}\|_{\mathrm{F}}.$$

It can be shown that the distance function is Lipschitz, hence continuous.

Theorem 3.5. (Global Convergence.) *Suppose that Algorithm 3.1 generates an infinite sequence of iterates $\{(\boldsymbol{G}^{(t)}, \boldsymbol{H}^{(t)})\}$. This sequence has at least one accumulation point:*

- *Every accumulation point lies in $\mathcal{G} \times \mathcal{H}$.*
- *Every accumulation point $(\overline{\boldsymbol{G}}, \overline{\boldsymbol{H}})$ satisfies*

$$\left\|\overline{\boldsymbol{G}} - \overline{\boldsymbol{H}}\right\|_{\mathrm{F}} = \lim_{t \to \infty} \left\|\boldsymbol{G}^{(t)} - \boldsymbol{H}^{(t)}\right\|_{\mathrm{F}}.$$

- *Every accumulation point $(\overline{\boldsymbol{G}}, \overline{\boldsymbol{H}})$ satisfies*

$$\left\|\overline{\boldsymbol{G}} - \overline{\boldsymbol{H}}\right\|_{\mathrm{F}} = \operatorname{dist}(\overline{\boldsymbol{G}}, \mathcal{H}) = \operatorname{dist}(\overline{\boldsymbol{H}}, \mathcal{G}).$$

Proof sketch: The existence of an accumulation point follows from the compactness of the constraint sets. The algorithm does not increase the distance between successive iterates, which is bounded below by zero. Therefore, this distance must converge. The distance functions are continuous, so we can take limits to obtain the remaining assertions. □

A more detailed treatment requires the machinery of point-to-set maps, and it would not enhance our main discussion. Please see the appendices of [Tropp et al. 05] for additional information.

4. BOUNDS ON THE PACKING DIAMETER

To assay the quality of the packings that we produce, it helps to have some upper bounds on the packing diameter. If a configuration of subspaces has a packing diameter close to the upper bound, that configuration must be a nearly optimal packing. This approach allows us to establish that many of the packings we construct numerically have packing diameters that are essentially optimal.

Theorem 4.1. [Conway et al. 96] *The packing diameter of N subspaces in the Grassmannian manifold $\mathbb{G}(K, \mathbb{F}^d)$ equipped with chordal distance is bounded above as*

$$\operatorname{pack}_{\text{chord}}(\mathcal{X})^2 \leq \frac{K(d-K)}{d} \frac{N}{N-1}. \tag{4–1}$$

If the bound is met, all pairs of subspaces are equidistant. When $\mathbb{F} = \mathbb{R}$, the bound is attainable only if

$$N \leq \frac{1}{2} d(d+1).$$

When $\mathbb{F} = \mathbb{C}$, the bound is attainable only if $N \leq d^2$.

The complex case is not stated in [Conway et al. 96], but it follows from an identical argument. We refer to (4–1) as the *Rankin bound* for subspace packings with respect to the chordal distance. The reason for the nomenclature is that the result is established by embedding the chordal Grassmannian manifold into a Euclidean sphere and applying the classical Rankin bound for sphere packing [Rankin 47].

It is also possible to draw a corollary on packing with respect to the spectral distance; this result is novel. A subspace packing is said to be *equi-isoclinic* if all the principal angles between all pairs of subspaces are identical [Lemmens and Seidel 73].

Corollary 4.2. *We have the following bound on the packing diameter of N subspaces in the Grassmannian manifold $\mathbb{G}(K, \mathbb{F}^d)$ equipped with the spectral distance:*

$$\operatorname{pack}_{\text{spec}}(\mathcal{X})^2 \leq \frac{d-K}{d} \frac{N}{N-1}. \tag{4–2}$$

If the bound is met, the packing is equi-isoclinic.

We refer to (4–2) as the Rankin bound for subspace packings with respect to spectral distance.

Proof: The power mean inequality (equivalently, Hölder's inequality) yields

$$\min_k \sin \theta_k \leq \left[K^{-1} \sum\nolimits_{k=1}^{K} \sin^2 \theta_k \right]^{1/2}.$$

For angles between zero and $\pi/2$, equality holds if and only if $\theta_1 = \cdots = \theta_K$. It follows that

$$\operatorname{pack}_{\text{spec}}(\mathcal{X})^2 \leq K^{-1} \operatorname{pack}_{\text{chord}}(\mathcal{X})^2 \leq \frac{d-K}{d} \frac{N}{N-1}.$$

If the second inequality is met, then all pairs of subspaces are equidistant with respect to the chordal metric. Moreover, if the first inequality is met, then the principal angles between each pair of subspaces are constant. Together, these two conditions imply that the packing is equi-isoclinic. □

An upper bound on the maximum number of equi-isoclinic subspaces is available. Its authors do not believe that it is sharp.

Theorem 4.3. [Lemmens and Seidel 73] *The maximum number of equi-isoclinic K-dimensional subspaces of $\mathbb{R}^d$ is no greater than*

$$\frac{1}{2}d(d+1) - \frac{1}{2}K(K+1) + 1.$$

Similarly, the maximum number of equi-isoclinic K-dimensional subspaces of $\mathbb{C}^d$ does not exceed

$$d^2 - K^2 + 1.$$

5. EXPERIMENTS

Our approach to packing is experimental rather than theoretical, so the real question is how Algorithm 3.1 performs in practice. In principle, this question is difficult to resolve because the optimal packing diameter is unknown for almost all combinations of d and N. Whenever possible, we compared our results with the Rankin bound and with the "world record" packings tabulated by N. J. A. Sloane and his colleagues [Sloane 04a]. In many cases, the algorithm was able to identify a nearly optimal packing. Moreover, it yields interesting results for packing problems that have not received numerical attention.

In the next subsection, we describe detailed experiments on packing in real and complex projective spaces. Then, we move on to packings of subspaces with respect to the chordal distance. Afterward, we study the spectral distance and the Fubini–Study distance.

5.1 Projective Packings

Line packings are the simplest type of Grassmannian packing, so they offer a natural starting point. Our goal is to produce the best packing of N lines in $\mathbb{P}^{d-1}(\mathbb{F})$. In the real case, Sloane's tables allow us to determine how much our packings fall short of the world record. In the complex setting, there is no comparable resource, so we must rely on the Rankin bound to gauge how well the algorithm performs.

Let us begin with packing in real projective spaces. We attempted to construct configurations of real lines whose maximum absolute inner product μ fell within 10^{-5} of the best value tabulated in [Sloane 04a]. For pairs (d, N) with $d = 3, 4, 5$ and $N = 4, 5, \ldots, 25$, we computed the putatively optimal value of the feasibility parameter μ from Sloane's data and equation (3–2). In each of ten trials, we constructed a starting matrix using Algorithm 3.4 with parameters $\tau = 0.9$ and $T = 10{,}000$. (Recall that the value of T determines the maximum number of random subspaces that are drawn when one is trying to construct the initial configuration.) We applied alternating projection, Algorithm 3.1, with the computed value of μ and the maximum number of iterations $T = 5000$. (Our numerical experience indicates that increasing the maximum number of iterations beyond 5000 does not confer a significant benefit.) We halted the iteration in step 4 if the iterate $\boldsymbol{G}^{(t)}$ exhibited no off-diagonal entry with absolute value greater than $\mu + 10^{-5}$. After ten trials, we recorded the largest packing diameter attained, as well as the average value of the packing diameter. We also recorded the average number of iterations the alternating projection required per trial.

Table 1 delivers the results of this experiment. Following Sloane, we have reported the degrees of arc subtended by the closest pair of lines. We believe that it is easiest to interpret the results geometrically when they are stated in this fashion. For collections of N points in the real projective space $\mathbb{P}^{d-1}(\mathbb{R})$, this table lists the best packing diameter (in degrees) and the average packing diameter (in degrees) obtained during ten random trials of the alternating projection algorithm. The error columns record how far our results deviate from the putative optimal packings (NJAS) reported in [Sloane 04a]. The last column gives the average number of iterations of alternating projection per trial before the termination condition is met.

According to the table, the best configurations produced by alternating projection consistently attain packing diameters tenths or hundredths of a degree away from the best configurations known. The average configurations returned by alternating projection are slightly worse, but they usually fall within a degree of the putative optima. Moreover, the algorithm finds certain configurations with ease. For the pair $(5, 16)$, fewer than one thousand iterations are required on average to achieve a packing within 0.001 degrees of optimal.

A second observation is that the alternating projection algorithm typically performs better when the number N of points is small. The largest errors are all clustered at larger values of N. A corollary observation is that the average number of iterations per trial tends to increase with the number of points.

There are several anomalies that we would like to point out. The most interesting pathology occurs at the pair $(d, N) = (5, 19)$. The best packing diameter calculated by alternating projection is about 1.76° worse than the optimal configuration, and it is also 1.76° worse than the best packing diameter computed for the pair $(5, 20)$. From Sloane's tables, we can see that the (putative) optimal packing of 19 lines in $\mathbb{P}^4(\mathbb{R})$ is actually a subset of

		Packing diameters (Degrees)					Iterations
d	N	NJAS	Best of 10	Error	Avg. of 10	Error	Avg. of 10
3	4	70.529	70.528	0.001	70.528	0.001	54
3	5	63.435	63.434	0.001	63.434	0.001	171
3	6	63.435	63.435	0.000	59.834	3.601	545
3	7	54.736	54.735	0.001	54.735	0.001	341
3	8	49.640	49.639	0.001	49.094	0.546	4333
3	9	47.982	47.981	0.001	47.981	0.001	2265
3	10	46.675	46.674	0.001	46.674	0.001	2657
3	11	44.403	44.402	0.001	44.402	0.001	2173
3	12	41.882	41.881	0.001	41.425	0.457	2941
3	13	39.813	39.812	0.001	39.522	0.291	4870
3	14	38.682	38.462	0.221	38.378	0.305	5000
3	15	38.135	37.934	0.201	37.881	0.254	5000
3	16	37.377	37.211	0.166	37.073	0.304	5000
3	17	35.235	35.078	0.157	34.821	0.414	5000
3	18	34.409	34.403	0.005	34.200	0.209	5000
3	19	33.211	33.107	0.104	32.909	0.303	5000
3	20	32.707	32.580	0.127	32.273	0.434	5000
3	21	32.216	32.036	0.180	31.865	0.351	5000
3	22	31.896	31.853	0.044	31.777	0.119	5000
3	23	30.506	30.390	0.116	30.188	0.319	5000
3	24	30.163	30.089	0.074	29.694	0.469	5000
3	25	29.249	29.024	0.224	28.541	0.707	5000
4	5	75.522	75.522	0.001	73.410	2.113	4071
4	6	70.529	70.528	0.001	70.528	0.001	91
4	7	67.021	67.021	0.001	67.021	0.001	325
4	8	65.530	65.530	0.001	64.688	0.842	3134
4	9	64.262	64.261	0.001	64.261	0.001	1843
4	10	64.262	64.261	0.001	64.261	0.001	803
4	11	60.000	59.999	0.001	59.999	0.001	577
4	12	60.000	59.999	0.001	59.999	0.001	146
4	13	55.465	55.464	0.001	54.390	1.074	4629
4	14	53.838	53.833	0.005	53.405	0.433	5000
4	15	52.502	52.493	0.009	51.916	0.585	5000
4	16	51.827	51.714	0.113	50.931	0.896	5000
4	17	50.887	50.834	0.053	50.286	0.601	5000
4	18	50.458	50.364	0.094	49.915	0.542	5000
4	19	49.711	49.669	0.041	49.304	0.406	5000
4	20	49.233	49.191	0.042	48.903	0.330	5000
4	21	48.548	48.464	0.084	48.374	0.174	5000
4	22	47.760	47.708	0.052	47.508	0.251	5000
4	23	46.510	46.202	0.308	45.789	0.722	5000
4	24	46.048	45.938	0.110	45.725	0.322	5000
4	25	44.947	44.739	0.208	44.409	0.538	5000
5	6	78.463	78.463	0.001	77.359	1.104	3246
5	7	73.369	73.368	0.001	73.368	0.001	1013
5	8	70.804	70.803	0.001	70.604	0.200	5000
5	9	70.529	70.528	0.001	69.576	0.953	2116
5	10	70.529	70.528	0.001	67.033	3.496	3029
5	11	67.254	67.254	0.001	66.015	1.239	4615
5	12	67.021	66.486	0.535	65.661	1.361	5000
5	13	65.732	65.720	0.012	65.435	0.297	5000
5	14	65.724	65.723	0.001	65.637	0.087	3559
5	15	65.530	65.492	0.038	65.443	0.088	5000
5	16	63.435	63.434	0.001	63.434	0.001	940
5	17	61.255	61.238	0.017	60.969	0.287	5000
5	18	61.053	61.048	0.005	60.946	0.107	5000
5	19	60.000	58.238	1.762	57.526	2.474	5000
5	20	60.000	59.999	0.001	56.183	3.817	3290
5	21	57.202	57.134	0.068	56.159	1.043	5000
5	22	56.356	55.819	0.536	55.173	1.183	5000
5	23	55.588	55.113	0.475	54.535	1.053	5000
5	24	55.228	54.488	0.740	53.926	1.302	5000
5	25	54.889	54.165	0.724	52.990	1.899	5000

TABLE 1. Packing in real projective spaces.

		Packing diameters (Degrees)		
d	N	DHST	Rankin	Difference
2	3	60.00	60.00	0.00
2	4	54.74	54.74	0.00
2	5	45.00	52.24	7.24
2	6	45.00	50.77	5.77
2	7	38.93	49.80	10.86
2	8	37.41	49.11	11.69
3	4	70.53	70.53	0.00
3	5	64.00	65.91	1.90
3	6	63.44	63.43	0.00
3	7	61.87	61.87	0.00
3	8	60.00	60.79	0.79
3	9	60.00	60.00	0.00
3	10	54.73	59.39	4.66
3	11	54.73	58.91	4.18
3	12	54.73	58.52	3.79
3	13	51.32	58.19	6.88
3	14	50.13	57.92	7.79
3	15	49.53	57.69	8.15
3	16	49.53	57.49	7.95
3	17	49.10	57.31	8.21
3	18	48.07	57.16	9.09
3	19	47.02	57.02	10.00
3	20	46.58	56.90	10.32
4	5	75.52	75.52	0.00
4	6	70.88	71.57	0.68
4	7	69.29	69.30	0.01
4	8	67.78	67.79	0.01
4	9	66.21	66.72	0.51
4	10	65.71	65.91	0.19
4	11	64.64	65.27	0.63
4	12	64.24	64.76	0.52
4	13	64.34	64.34	0.00
4	14	63.43	63.99	0.56
4	15	63.43	63.69	0.26
4	16	63.43	63.43	0.00
4	17	59.84	63.21	3.37
4	18	59.89	63.02	3.12
4	19	60.00	62.84	2.84
4	20	57.76	62.69	4.93
5	6	78.46	78.46	0.00
5	7	74.52	75.04	0.51
5	8	72.81	72.98	0.16
5	9	71.24	71.57	0.33
5	10	70.51	70.53	0.02
5	11	69.71	69.73	0.02
5	12	68.89	69.10	0.21
5	13	68.19	68.58	0.39
5	14	67.66	68.15	0.50
5	15	67.37	67.79	0.43
5	16	66.68	67.48	0.80
5	17	66.53	67.21	0.68
5	18	65.87	66.98	1.11
5	19	65.75	66.77	1.02
5	20	65.77	66.59	0.82
5	21	65.83	66.42	0.60
5	22	65.87	66.27	0.40
5	23	65.90	66.14	0.23
5	24	65.91	66.02	0.11
5	25	65.91	65.91	0.00

TABLE 2. Packing in complex projective spaces.

the best packing of 20 lines. Perhaps the fact that this packing is degenerate makes it difficult to construct. A similar event occurs (less dramatically) at the pair $(5, 13)$. The table also shows that the algorithm performs less effectively when the number of lines exceeds 20.

In complex projective spaces, this methodology does not apply because there are no tables available. In fact, we know of only one paper that contains numerical work on packing in complex projective spaces, [Agrawal et al. 01], but it gives very few examples of good packings. The only method we know for gauging the quality of a complex line packing is to compare it against an upper bound. The Rankin bound for projective packings, which is derived in Section 4, states that every configuration $\mathcal{X}$ of N lines in either $\mathbb{P}^{d-1}(\mathbb{R})$ or $\mathbb{P}^{d-1}(\mathbb{C})$ satisfies the inequality

$$\operatorname{pack}_{\mathbb{P}}(\mathcal{X})^2 \leq \frac{(d-1)N}{d(N-1)}.$$

This bound is attainable only for rare combinations of d and N. In particular, the bound can be met in $\mathbb{P}^{d-1}(\mathbb{R})$ only if $N \leq \frac{1}{2}d(d+1)$. In the space $\mathbb{P}^{d-1}(\mathbb{C})$, attainment requires that $N \leq d^2$. Any arrangement of lines that meets the Rankin bound must be equiangular. These optimal configurations are called *equiangular tight frames*. See [Strohmer and Heath 03, Holmes and Paulsen 04, Tropp et al. 05, Sustik et al. 07] for more details.

We performed some ad hoc experiments to produce configurations of complex lines with large packing diameters. For each pair (d, N), we used the Rankin bound to determine a lower limit on the feasibility parameter μ. Starting matrices were constructed with Algorithm 3.4 using values of τ ranging between 0.9 and 1.0. (Algorithm 3.4 typically fails for smaller values of τ.) For values of the feasibility parameter between the minimal value and twice the minimal value, we performed five thousand iterations of Algorithm 3.1, and we recorded the largest packing diameter attained during these trials.

Table 2 compares our results against the Rankin bound. We see that many of the complex line configurations have packing diameters much smaller than the Rankin bound, which is not surprising because the bound is usually not attainable. Some of our configurations fall within a thousandth of a degree of the bound, which is essentially optimal.

In the table, we compare our best configurations (DHST) of N points in the complex projective space $\mathbb{P}^{d-1}(\mathbb{C})$ against the Rankin bound (4–1). The packing diameter of an ensemble is measured as the acute angle (in degrees) between the closest pair of lines. The final column shows how far our configurations fall short of the bound.

Table 2 contains a few oddities. In $\mathbb{P}^4(\mathbb{C})$, the best packing diameter computed for $N = 18, 19, \ldots, 24$ is worse than the packing diameter for $N = 25$. This configuration of 25 lines is an equiangular tight frame, which means that it is an optimal packing [Tropp et al. 05, Table 1]. It seems likely that the optimal configurations for the preceding values of N are just subsets of the optimal arrangement of 25 lines. As before, it may be difficult to calculate this type of degenerate packing. A similar event occurs less dramatically at the pair $(d, N) = (4, 13)$ and at the pairs $(4, 17)$ and $(4, 18)$.

Figure 1 compares the quality of the best real projective packings from [Sloane 04a] with the best complex projective packings that we obtained. It is natural that the complex packings are better than the real packings because the real projective space can be embedded isometrically into the complex projective space. But it is remarkable how badly the real packings compare with the complex packings. The only cases in which the real and complex ensembles have the same packing diameter occur when the real configuration meets the Rankin bound.

5.2 The Chordal Distance

Emboldened by this success with projective packings, we move on to packings of subspaces with respect to the chordal distance. Once again, we are able to use Sloane's tables for guidance in the real case. In the complex case, we fall back on the Rankin bound.

For each triple (d, K, N), we determined a value for the feasibility parameter μ from the best packing diameter Sloane recorded for N subspaces in $\mathbb{G}(K, \mathbb{R}^d)$, along with equation (3–2). We constructed starting points using the modified version of Algorithm 3.4 with $\tau = \sqrt{K}$, which represents no constraint. (We found that the alternating projection performed no better with initial configurations generated from smaller values of τ.) Then we executed Algorithm 3.1 with the calculated value of μ for five thousand iterations.

Table 3 demonstrates how the best packings we obtained compare with Sloane's best packings. Many of our real configurations attained a squared packing diameter within 10^{-3} of the best value Sloane recorded. Our algorithm was especially successful for smaller numbers of subspaces, but its performance began to flag as the number of subspaces approached 20.

Table 3 contains several anomalies. For example, our configurations of $N = 11, 12, \ldots, 16$ subspaces in $\mathbb{R}^4$ yield

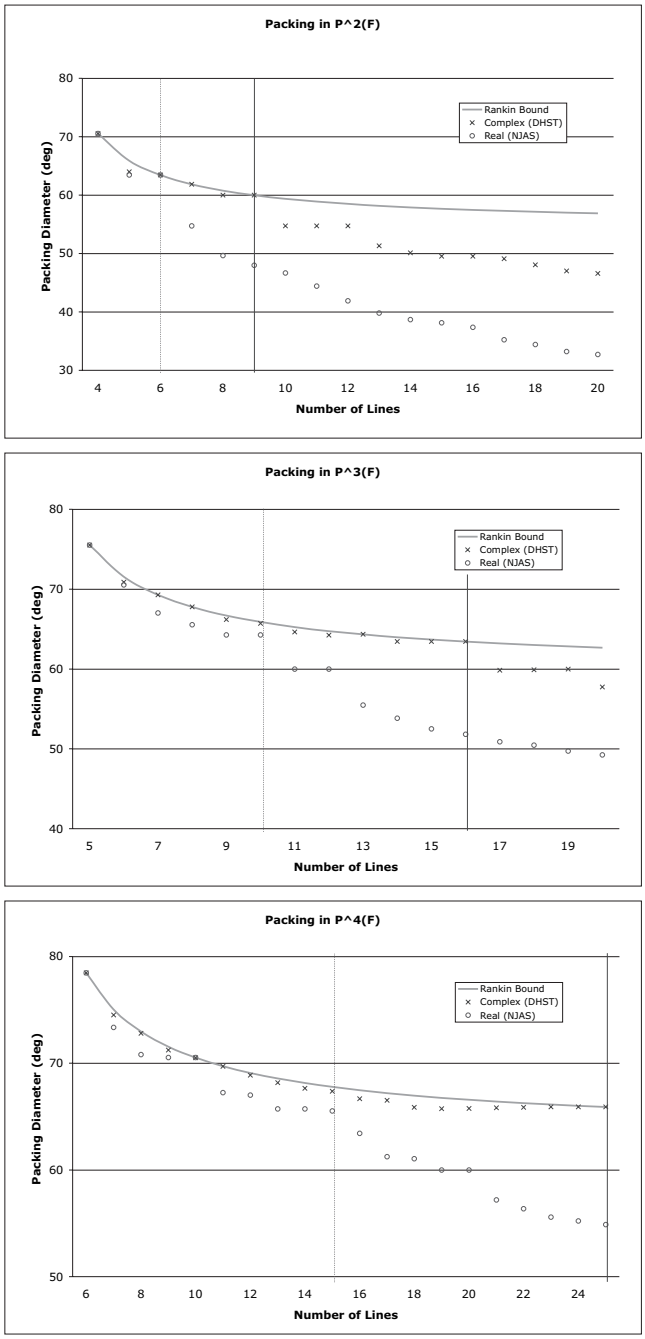

FIGURE 1. Real and Complex Projective Packings. These three graphs compare the packing diameters attained by configurations in real and complex projective spaces with $d = 3, 4, 5$. The circles indicate the best real packings obtained by Sloane and his colleagues [Sloane 04a]. The crosses indicate the best complex packings produced by the authors. Rankin's upper bound (4–1) is depicted in gray. The dashed vertical line marks the largest number of real lines for which the Rankin bound is attainable, while the solid vertical line marks the maximum number of complex lines for which the Rankin bound is attainable.

worse packing diameters than the configuration of 17 subspaces. It turns out that this configuration of 17 subspaces is optimal, and Sloane's data show that the (putative) optimal arrangements of 11 to 16 subspaces are all subsets of this configuration. This is the same problem that occurred in some of our earlier experiments, and it suggests again that our algorithm has difficulty locating these degenerate configurations precisely.

			Squared Packing diameters		
K	d	N	DHST	NJAS	Difference
2	4	3	1.5000	1.5000	0.0000
2	4	4	1.3333	1.3333	0.0000
2	4	5	1.2500	1.2500	0.0000
2	4	6	1.2000	1.2000	0.0000
2	4	7	1.1656	1.1667	0.0011
2	4	8	1.1423	1.1429	0.0005
2	4	9	1.1226	1.1231	0.0004
2	4	10	1.1111	1.1111	0.0000
2	4	11	0.9981	1.0000	0.0019
2	4	12	0.9990	1.0000	0.0010
2	4	13	0.9996	1.0000	0.0004
2	4	14	1.0000	1.0000	0.0000
2	4	15	1.0000	1.0000	0.0000
2	4	16	0.9999	1.0000	0.0001
2	4	17	1.0000	1.0000	0.0000
2	4	18	0.9992	1.0000	0.0008
2	4	19	0.8873	0.9091	0.0218
2	4	20	0.8225	0.9091	0.0866
2	5	3	1.7500	1.7500	0.0000
2	5	4	1.6000	1.6000	0.0000
2	5	5	1.5000	1.5000	0.0000
2	5	6	1.4400	1.4400	0.0000
2	5	7	1.4000	1.4000	0.0000
2	5	8	1.3712	1.3714	0.0002
2	5	9	1.3464	1.3500	0.0036
2	5	10	1.3307	1.3333	0.0026
2	5	11	1.3069	1.3200	0.0131
2	5	12	1.2973	1.3064	0.0091
2	5	13	1.2850	1.2942	0.0092
2	5	14	1.2734	1.2790	0.0056
2	5	15	1.2632	1.2707	0.0075
2	5	16	1.1838	1.2000	0.0162
2	5	17	1.1620	1.2000	0.0380
2	5	18	1.1589	1.1909	0.0319
2	5	19	1.1290	1.1761	0.0472
2	5	20	1.0845	1.1619	0.0775

TABLE 3. Packing in Real Grassmannians with Chordal Distance. We compare our best configurations (DHST) of N points in $\mathbb{G}(K, \mathbb{R}^d)$ against the best packings (NJAS) reported in [Sloane 04a]. The squared packing diameter is the squared chordal distance (2–1) between the closest pair of subspaces. The last column lists the difference between the columns (NJAS) and (DHST).

The literature contains very few experimental results on packing in complex Grassmannian manifolds equipped with chordal distance. To our knowledge, the only numerical work appears in two short tables from [Agrawal et al. 01]. Therefore, we found it valuable to compare our results against the Rankin bound for subspace packings, which is derived in Section 4. For reference, this bound requires that every configuration $\mathcal{X}$ of N subspaces in $\mathbb{G}(K, \mathbb{F}^d)$ satisfy the inequality

$$\operatorname{pack}_{\text{chord}}(\mathcal{X})^2 \leq \frac{K(d-K)}{d} \frac{N}{N-1}.$$

This bound cannot always be met. In particular, the bound is attainable in the complex setting only if $N \leq d^2$. In the real setting, the bound requires $N \leq \frac{1}{2}d(d+1)$. When the bound is attained, each pair of subspaces in $\mathcal{X}$ is equidistant.

We performed some ad hoc experiments to construct a table of packings in $\mathbb{G}(K, \mathbb{C}^d)$ equipped with the chordal distance. For each triple (d, K, N), we constructed random starting points using Algorithm 3.4 with $\tau = \sqrt{K}$ (which represents no constraint). Then we used the Rankin bound to calculate a lower limit on the feasi-

			Squared Packing diameters		
K	d	N	DHST	Rankin	Difference
2	4	3	1.5000	1.5000	0.0000
2	4	4	1.3333	1.3333	0.0000
2	4	5	1.2500	1.2500	0.0000
2	4	6	1.2000	1.2000	0.0000
2	4	7	1.1667	1.1667	0.0000
2	4	8	1.1429	1.1429	0.0000
2	4	9	1.1250	1.1250	0.0000
2	4	10	1.1111	1.1111	0.0000
2	4	11	1.0999	1.1000	0.0001
2	4	12	1.0906	1.0909	0.0003
2	4	13	1.0758	1.0833	0.0076
2	4	14	1.0741	1.0769	0.0029
2	4	15	1.0698	1.0714	0.0016
2	4	16	1.0658	1.0667	0.0009
2	4	17	0.9975	1.0625	0.0650
2	4	18	0.9934	1.0588	0.0654
2	4	19	0.9868	1.0556	0.0688
2	4	20	0.9956	1.0526	0.0571
2	5	3	1.7500	1.8000	0.0500
2	5	4	1.6000	1.6000	0.0000
2	5	5	1.5000	1.5000	0.0000
2	5	6	1.4400	1.4400	0.0000
2	5	7	1.4000	1.4000	0.0000
2	5	8	1.3714	1.3714	0.0000
2	5	9	1.3500	1.3500	0.0000
2	5	10	1.3333	1.3333	0.0000
2	5	11	1.3200	1.3200	0.0000
2	5	12	1.3090	1.3091	0.0001
2	5	13	1.3000	1.3000	0.0000
2	5	14	1.2923	1.2923	0.0000
2	5	15	1.2857	1.2857	0.0000
2	5	16	1.2799	1.2800	0.0001
2	5	17	1.2744	1.2750	0.0006
2	5	18	1.2686	1.2706	0.0020
2	5	19	1.2630	1.2667	0.0037
2	5	20	1.2576	1.2632	0.0056
2	6	4	1.7778	1.7778	0.0000
2	6	5	1.6667	1.6667	0.0000
2	6	6	1.6000	1.6000	0.0000
2	6	7	1.5556	1.5556	0.0000
2	6	8	1.5238	1.5238	0.0000
2	6	9	1.5000	1.5000	0.0000
2	6	10	1.4815	1.4815	0.0000
2	6	11	1.4667	1.4667	0.0000
2	6	12	1.4545	1.4545	0.0000
2	6	13	1.4444	1.4444	0.0000
2	6	14	1.4359	1.4359	0.0000
2	6	15	1.4286	1.4286	0.0000
2	6	16	1.4221	1.4222	0.0001
2	6	17	1.4166	1.4167	0.0000
2	6	18	1.4118	1.4118	0.0000
2	6	19	1.4074	1.4074	0.0000
2	6	20	1.4034	1.4035	0.0001
2	6	21	1.3999	1.4000	0.0001
2	6	22	1.3968	1.3968	0.0001
2	6	23	1.3923	1.3939	0.0017
2	6	24	1.3886	1.3913	0.0028
2	6	25	1.3862	1.3889	0.0027
3	6	3	2.2500	2.2500	0.0000
3	6	4	2.0000	2.0000	0.0000
3	6	5	1.8750	1.8750	0.0000
3	6	6	1.8000	1.8000	0.0000
3	6	7	1.7500	1.7500	0.0000
3	6	8	1.7143	1.7143	0.0000
3	6	9	1.6875	1.6875	0.0000
3	6	10	1.6667	1.6667	0.0000
3	6	11	1.6500	1.6500	0.0000
3	6	12	1.6363	1.6364	0.0001
3	6	13	1.6249	1.6250	0.0001
3	6	14	1.6153	1.6154	0.0000
3	6	15	1.6071	1.6071	0.0000
3	6	16	1.5999	1.6000	0.0001
3	6	17	1.5936	1.5938	0.0001
3	6	18	1.5879	1.5882	0.0003
3	6	19	1.5829	1.5833	0.0004
3	6	20	1.5786	1.5789	0.0004
3	6	21	1.5738	1.5750	0.0012
3	6	22	1.5687	1.5714	0.0028
3	6	23	1.5611	1.5682	0.0070
3	6	24	1.5599	1.5652	0.0053
3	6	25	1.5558	1.5625	0.0067
3	6	26	1.5542	1.5600	0.0058
3	6	27	1.5507	1.5577	0.0070
3	6	28	1.5502	1.5556	0.0054
3	6	29	1.5443	1.5536	0.0092
3	6	30	1.5316	1.5517	0.0201
3	6	31	1.5283	1.5500	0.0217
3	6	32	1.5247	1.5484	0.0237
3	6	33	1.5162	1.5469	0.0307
3	6	34	1.5180	1.5455	0.0274
3	6	35	1.5141	1.5441	0.0300
3	6	36	1.5091	1.5429	0.0338

TABLE 4. Packing in Complex Grassmannians with Chordal Distance. We compare our best configurations (DHST) of N points in $\mathbb{G}(K, \mathbb{C}^d)$ against the Rankin bound, equation (4–1). The squared packing diameter is calculated as the squared chordal distance (2–1) between the closest pair of subspaces. The final column shows how much the computed ensemble deviates from the Rankin bound. When the bound is met, all pairs of subspaces are equidistant.

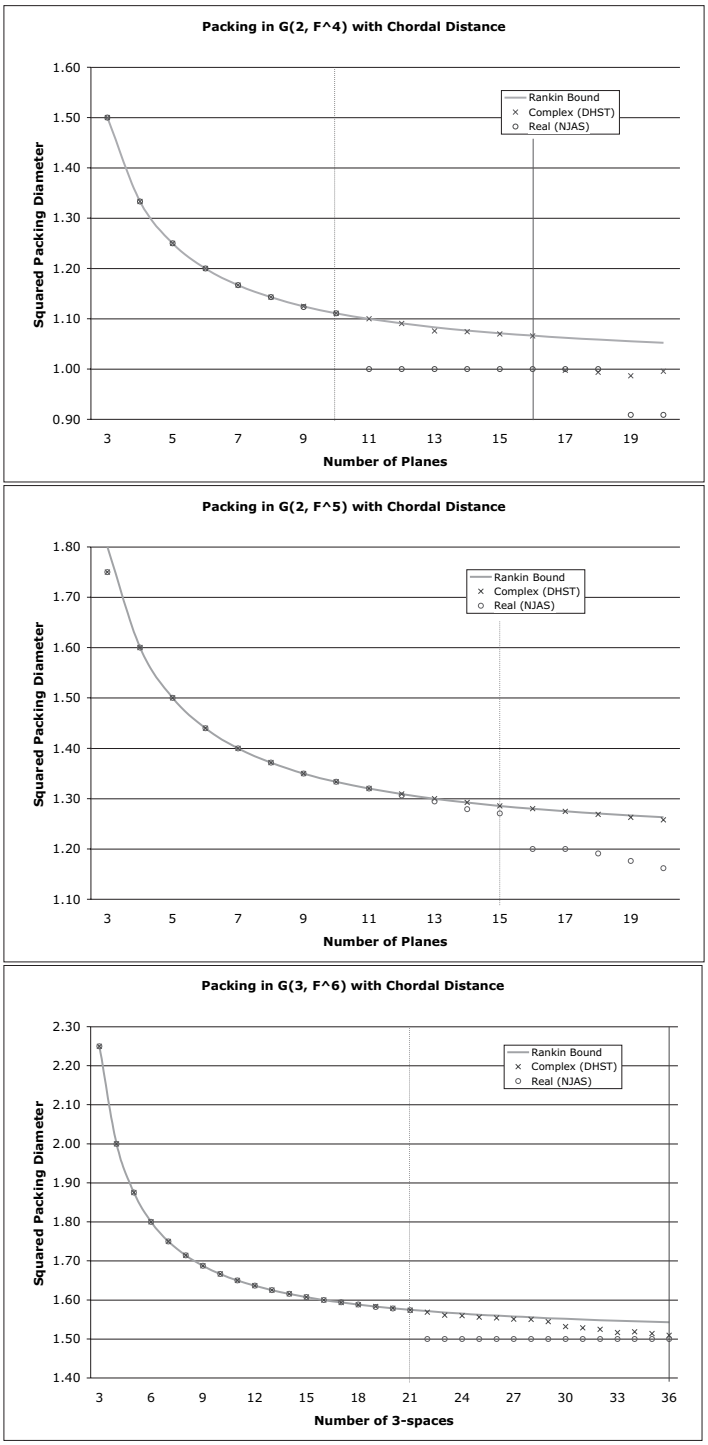

FIGURE 2. Packing in Grassmannians with Chordal Distance. This figure shows the packing diameters of N points in the Grassmannian $\mathbb{G}(K, \mathbb{F}^d)$ equipped with the chordal distance. The circles indicate the best real packings ($\mathbb{F} = \mathbb{R}$) obtained by Sloane and his colleagues [Sloane 04a]. The crosses indicate the best complex packings ($\mathbb{F} = \mathbb{C}$) produced by the authors. Rankin's upper bound (4–1) appears in gray. The dashed vertical line marks the largest number of real subspaces for which the Rankin bound is attainable, while the solid vertical line marks the maximum number of complex subspaces for which the Rankin bound is attainable.

bility parameter μ. For this value of μ, we executed the alternating projection, Algorithm 3.1, for five thousand iterations.

The best packing diameters we obtained are listed in Table 4. We see that there is a remarkable correspondence between the squared packing diameters of our configurations and the Rankin bound. Indeed, many of our packings are within 10^{-4} of the bound, which means that these configurations are essentially optimal. The algorithm was less successful as N approached d^2, which is an upper bound on the number N of subspaces for which the Rankin bound is attainable.

Figure 2 compares the packing diameters of the best configurations in real and complex Grassmannian spaces equipped with chordal distance. It is remarkable that both real and complex packings almost meet the Rankin bound for all N where it is attainable. Notice how the real packing diameters fall off as soon as N exceeds $\frac{1}{2}d(d+1)$. In theory, a complex configuration should always attain a better packing diameter than the corresponding real configuration because the real Grassmannian space can be embedded isometrically into the complex Grassmannian space. The figure shows that our best arrangements of 17 and 18 subspaces in $\mathbb{G}(2, \mathbb{C}^4)$ are actually slightly worse than the real arrangements calculated by Sloane. This indicates a failure of the alternating projection algorithm.

5.3 The Spectral Distance

Next, we consider how to compute Grassmannian packings with respect to the spectral distance. This investigation requires some small modifications to the algorithm, which are described in the next subsection. Afterward, we provide the results of some numerical experiments.

5.3.1 Modifications to the Algorithm. To construct packings with respect to the spectral distance, we tread a familiar path. Suppose that we wish to produce a configuration of N subspaces in $\mathbb{G}(K, \mathbb{C}^d)$ with a packing diameter ρ. The feasibility problem requires that

$$\max_{m \neq n} \|\boldsymbol{X}_m^* \boldsymbol{X}_n\|_{2,2} \leq \mu, \tag{5–1}$$

where $\mu = \sqrt{1-\rho^2}$. This leads to the convex structural constraint set

$$\begin{aligned}\mathcal{H}(\mu) \stackrel{\text{def}}{=} \{\boldsymbol{H} \in \mathbb{C}^{KN \times KN} : \boldsymbol{H} = \boldsymbol{H}^*,\\ \boldsymbol{H}_{nn} = \mathbf{I} \text{ for } n = 1, 2, \dots, N,\\ \text{and} \quad \|\boldsymbol{H}_{mn}\|_{2,2} \leq \mu \text{ for all } m \neq n\}.\end{aligned}$$

The spectral constraint set is the same as before. The next proposition shows how to find the matrix in $\mathcal{H}$ closest to an initial matrix. In preparation, define the truncation operator $[x]_\mu = \min\{x, \mu\}$ for numbers, and extend it to matrices by applying it to each component.

Proposition 5.1. *Let $\boldsymbol{G}$ be a Hermitian matrix. With respect to the Frobenius norm, the unique matrix in $\mathcal{H}(\mu)$ nearest to $\boldsymbol{G}$ has a block identity diagonal. If the off-diagonal block $\boldsymbol{G}_{mn}$ has a singular value decomposition $\boldsymbol{U}_{mn}\boldsymbol{C}_{mn}\boldsymbol{V}_{mn}^*$, then*

$$\boldsymbol{H}_{mn} = \begin{cases} \boldsymbol{G}_{mn} & \text{if } \|\boldsymbol{G}_{mn}\|_{2,2} \leq \mu, \\ \boldsymbol{U}_{mn}[\boldsymbol{C}_{mn}]_\mu \boldsymbol{V}_{mn}^*, & \text{otherwise.} \end{cases}$$

Proof: To determine the (m, n) off-diagonal block of the solution matrix $\boldsymbol{H}$, we must solve the optimization problem

$$\min_{\boldsymbol{A}} \frac{1}{2}\|\boldsymbol{A} - \boldsymbol{G}_{mn}\|_{\mathrm{F}}^2 \quad \text{subject to} \quad \|\boldsymbol{A}\|_{2,2} \leq \mu.$$

The Frobenius norm is strictly convex and the spectral norm is convex, so this problem has a unique solution.

Let $\boldsymbol{\sigma}(\cdot)$ return the vector of decreasingly ordered singular values of a matrix. Suppose that $\boldsymbol{G}_{mn}$ has the singular value decomposition $\boldsymbol{G}_{mn} = \boldsymbol{U}\{\operatorname{diag}\boldsymbol{\sigma}(\boldsymbol{G}_{mn})\}\boldsymbol{V}^*$. The constraint in the optimization problem depends only on the singular values of $\boldsymbol{A}$, and so the Hoffman–Wielandt theorem for singular values [Horn and Johnson 85] allows us to check that the solution has the form $\boldsymbol{A} = \boldsymbol{U}\{\operatorname{diag}\boldsymbol{\sigma}(\boldsymbol{A})\}\boldsymbol{V}^*$.

To determine the singular values $\boldsymbol{\xi} = \boldsymbol{\sigma}(\boldsymbol{A})$ of the solution, we must solve the (strictly) convex program

$$\min_{\boldsymbol{\xi}} \frac{1}{2}\big\|\boldsymbol{\xi} - \boldsymbol{\sigma}(\boldsymbol{G}_{mn})\big\| \quad \text{subject to} \quad \xi_k \leq \mu.$$

An easy application of Karush–Kuhn–Tucker theory [Rockafellar 70] proves that the solution is obtained by truncating the singular values of $\boldsymbol{G}_{mn}$ that exceed μ. □

5.3.2 Numerical Results. To our knowledge, there are no numerical studies of packing in Grassmannian spaces equipped with spectral distance. To gauge the quality of our results, we compare them against the upper bound of Corollary 4.2. In the real or complex setting, a configuration $\mathcal{X}$ of N subspaces in $\mathbb{G}(K, \mathbb{F}^d)$ with respect to the spectral distance must satisfy the bound

$$\operatorname{pack}_{\text{spec}}(\mathcal{X})^2 \leq \frac{d-K}{d}\frac{N}{N-1}.$$

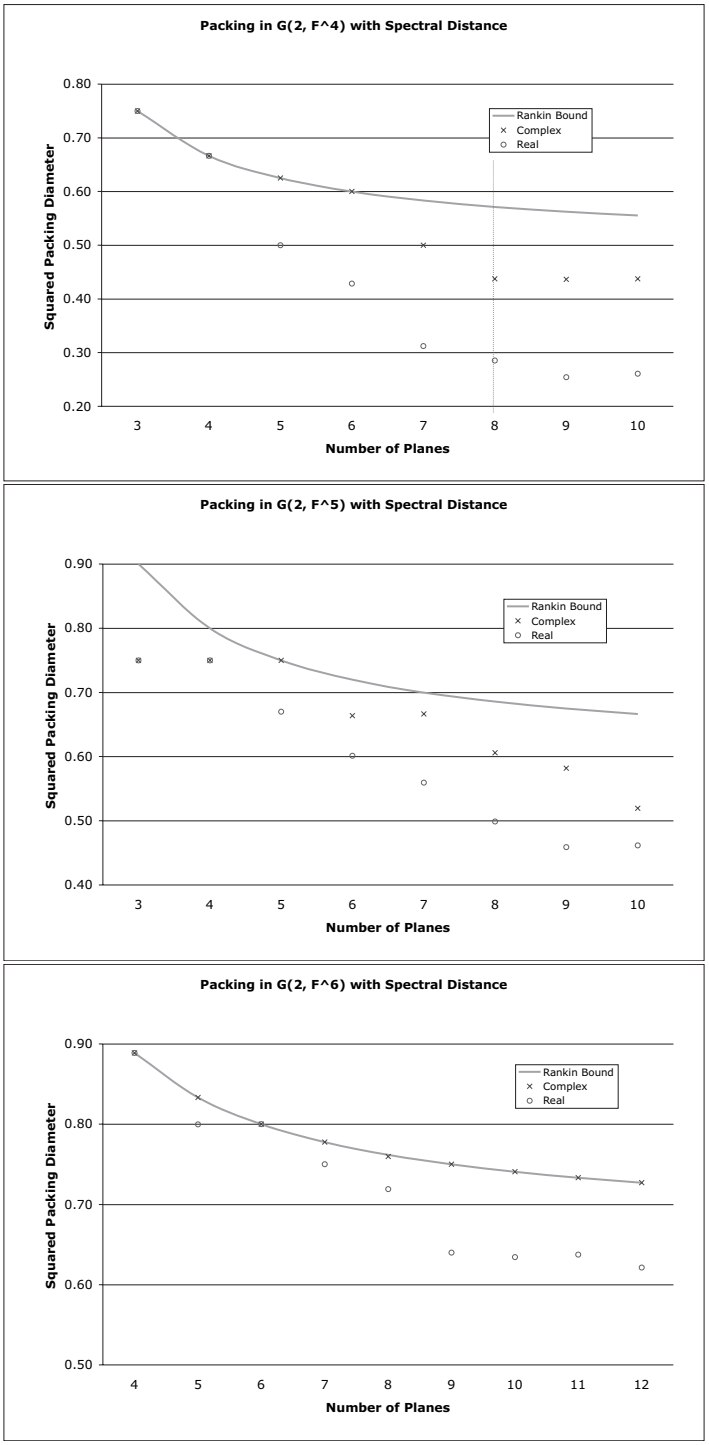

FIGURE 3. Packing in Grassmannians with spectral distance: This figure shows the packing diameters of N points in the Grassmannian $\mathbb{G}(K, \mathbb{F}^d)$ equipped with the spectral distance. The circles indicate the best real packings ($\mathbb{F} = \mathbb{R}$) obtained by the authors, while the crosses indicate the best complex packings ($\mathbb{F} = \mathbb{C}$) obtained. The Rankin bound (4–2) is depicted in gray. The dashed vertical line marks an upper bound on largest number of real subspaces for which the Rankin bound is attainable according to Theorem 4.3.

			Squared Packing diameters				
d	K	N	Rankin	$\mathbb{R}$	Difference	$\mathbb{C}$	Difference
4	2	3	0.7500	0.7500	0.0000	0.7500	0.0000
4	2	4	0.6667	0.6667	0.0000	0.6667	0.0000
4	2	5	0.6250	0.5000	0.1250	0.6250	0.0000
4	2	6	0.6000	0.4286	0.1714	0.6000	0.0000
4	2	7	0.5833	0.3122	0.2712	0.5000	0.0833
4	2	8	0.5714	0.2851	0.2863	0.4374	0.1340
4	2	9	0.5625	0.2544	0.3081	0.4363	0.1262
4	2	10	0.5556	0.2606	0.2950	0.4375	0.1181
5	2	3	0.9000	0.7500	0.1500	0.7500	0.1500
5	2	4	0.8000	0.7500	0.0500	0.7500	0.0500
5	2	5	0.7500	0.6700	0.0800	0.7497	0.0003
5	2	6	0.7200	0.6014	0.1186	0.6637	0.0563
5	2	7	0.7000	0.5596	0.1404	0.6667	0.0333
5	2	8	0.6857	0.4991	0.1867	0.6060	0.0798
5	2	9	0.6750	0.4590	0.2160	0.5821	0.0929
5	2	10	0.6667	0.4615	0.2052	0.5196	0.1470
6	2	4	0.8889	0.8889	0.0000	0.8889	0.0000
6	2	5	0.8333	0.7999	0.0335	0.8333	0.0000
6	2	6	0.8000	0.8000	0.0000	0.8000	0.0000
6	2	7	0.7778	0.7500	0.0278	0.7778	0.0000
6	2	8	0.7619	0.7191	0.0428	0.7597	0.0022
6	2	9	0.7500	0.6399	0.1101	0.7500	0.0000
6	2	10	0.7407	0.6344	0.1064	0.7407	0.0000
6	2	11	0.7333	0.6376	0.0958	0.7333	0.0000
6	2	12	0.7273	0.6214	0.1059	0.7273	0.0000
6	3	3	0.7500	0.7500	0.0000	0.7500	0.0000
6	3	4	0.6667	0.5000	0.1667	0.6667	0.0000
6	3	5	0.6250	0.4618	0.1632	0.4999	0.1251
6	3	6	0.6000	0.4238	0.1762	0.5000	0.1000
6	3	7	0.5833	0.3590	0.2244	0.4408	0.1426
6	3	8	0.5714	—	—	0.4413	0.1301
6	3	9	0.5625	—	—	0.3258	0.2367

TABLE 5. Packing in Grassmannians with spectral distance: We compare our best real ($\mathbb{F} = \mathbb{R}$) and complex ($\mathbb{F} = \mathbb{C}$) packings in $\mathbb{G}(K, \mathbb{F}^d)$ against the Rankin bound, equation (4–2). The squared packing diameter of a configuration is the squared spectral distance (2–2) between the closest pair of subspaces. When the Rankin bound is met, all pairs of subspaces are equi-isoclinic. The algorithm failed to produce any configurations of eight or nine subspaces in $\mathbb{G}(3, \mathbb{R}^6)$ with nontrivial packing diameters.

In the real case, the bound is attainable only if $N \leq \frac{1}{2}d(d+1) - \frac{1}{2}K(K+1) + 1$, while attainment in the complex case requires that $N \leq d^2 - K^2 + 1$ [Lemmens and Seidel 73]. When a configuration meets the bound, the subspaces are not only equidistant but also *equi-isoclinic*. That is, all principal angles between all pairs of subspaces are identical.

We performed some limited ad hoc experiments in an effort to produce good configurations of subspaces with respect to the spectral distance. We constructed random starting points using the modified version of Algorithm 3.4 with $\tau = 1$, which represents no constraint. (Again, we did not find that smaller values of τ improved the performance of the alternating projection.) From the Rankin bound, we calculated the smallest possible value of the feasibility parameter μ. For values of μ ranging from the minimal value to twice the minimal value, we ran the alternating projection, Algorithm 3.1, for five thousand iterations, and we recorded the best packing diameters that we obtained.

Table 5 displays the results of our calculations. We see that some of our configurations essentially meet the Rankin bound, which means that they are equi-isoclinic. It is clear that alternating projection also succeeds reasonably well for this packing problem.

The most notable pathology in the table occurs for configurations of eight and nine subspaces in $\mathbb{G}(3, \mathbb{R}^6)$. In these cases, the algorithm always yielded arrangements of subspaces with a zero packing diameter, which implies that two of the subspaces intersect nontrivially. Nevertheless, we were able to construct random starting points with a nonzero packing diameter, which means that the algorithm is making the initial configuration worse. We do not understand the reason for this failure.

Figure 3 makes a graphical comparison between the real and complex subspace packings. On the whole, the

complex packings are much better than the real packings. For example, every configuration of subspaces in $\mathbb{G}(2, \mathbb{C}^6)$ nearly meets the Rankin bound, while just two of the real configurations achieve the same distinction. In comparison, it is curious how few arrangements in $\mathbb{G}(2, \mathbb{C}^5)$ come anywhere near the Rankin bound.

5.4 The Fubini–Study Distance

When we approach the problem of packing in Grassmannian manifolds equipped with the Fubini–Study distance, we are truly out in the wilderness. To our knowledge, the literature contains neither experimental nor theoretical treatments of this question. Moreover, we are not presently aware of general upper bounds on the Fubini–Study packing diameter that we might use to assay the quality of a configuration of subspaces. Nevertheless, we attempted a few basic experiments. The investigation entails some more modifications to the algorithm, which are described below. Afterward, we go over our experimental results. We view this work as very preliminary.

5.4.1 Modifications to the Algorithm. Suppose that we wish to construct a configuration of N subspaces whose Fubini–Study packing diameter exceeds ρ. The feasibility condition is

$$\max_{m \neq n} |\det \boldsymbol{X}_m^* \boldsymbol{X}_n| \leq \mu, \tag{5–2}$$

where $\mu = \cos \rho$. This leads to the structural constraint set

$$\begin{aligned} \mathcal{H}(\mu) \stackrel{\text{def}}{=} \{ \boldsymbol{H} \in \mathbb{C}^{KN \times KN} : \boldsymbol{H} = \boldsymbol{H}^*, \\ \boldsymbol{H}_{nn} = \mathbf{I} \text{ for } n = 1, 2, \dots, N, \\ \text{and} \quad |\det \boldsymbol{H}_{mn}| \leq \mu \text{ for all } m \neq n \}. \end{aligned}$$

Unhappily, this set is no longer convex. To produce a nearest matrix in $\mathcal{H}$, we must solve a nonlinear programming problem. The following proposition describes a numerically favorable formulation.

Proposition 5.2. *Let $\boldsymbol{G}$ be a Hermitian matrix. Suppose that the off-diagonal block $\boldsymbol{G}_{mn}$ has singular value decomposition $\boldsymbol{U}_{mn} \boldsymbol{C}_{mn} \boldsymbol{V}_{mn}^*$. Let $\boldsymbol{c}_{mn} = \operatorname{diag} \boldsymbol{C}_{mn}$, and find a (real) vector $\boldsymbol{x}_{mn}$ that solves the optimization problem*

$$\min_{\boldsymbol{x}} \quad \frac{1}{2} \|\exp(\boldsymbol{x}) - \boldsymbol{c}_{mn}\|_2^2 \quad \textit{subject to} \quad \mathbf{e}^* \boldsymbol{x} \leq \log \mu.$$

In Frobenius norm, a matrix $\boldsymbol{H}$ from $\mathcal{H}(\mu)$ that is closest to $\boldsymbol{G}$ has a block-identity diagonal and off-diagonal blocks

$$\boldsymbol{H}_{mn} = \begin{cases} \boldsymbol{G}_{mn} & \textit{if } |\det \boldsymbol{G}_{mn}| \leq \mu, \\ \boldsymbol{U}_{mn} \{\operatorname{diag}(\exp \boldsymbol{x}_{mn})\} \boldsymbol{V}_{mn}^* & \textit{otherwise.} \end{cases}$$

We use $\exp(\cdot)$ to denote the componentwise exponential of a vector. One may establish that the optimization problem is not convex by calculating the Hessian of the objective function.

Proof: To determine the (m, n) off-diagonal block of the solution matrix $\boldsymbol{H}$, we must solve the optimization problem

$$\min_{\boldsymbol{A}} \frac{1}{2} \|\boldsymbol{A} - \boldsymbol{G}_{mn}\|_{\mathrm{F}}^2 \quad \text{subject to} \quad |\det \boldsymbol{A}| \leq \mu.$$

We may reformulate this problem as

$$\min_{\boldsymbol{A}} \frac{1}{2} \|\boldsymbol{A} - \boldsymbol{G}_{mn}\|_{\mathrm{F}}^2$$

subject to

$$\sum\nolimits_{k=1}^{K} \log \sigma_k(\boldsymbol{A}) \leq \log \mu.$$

A familiar argument proves that the solution matrix has the same left and right singular vectors as $\boldsymbol{G}_{mn}$. To obtain the singular values $\boldsymbol{\xi} = \boldsymbol{\sigma}(\boldsymbol{A})$ of the solution, we consider the mathematical program

$$\min_{\boldsymbol{\xi}} \frac{1}{2} \|\boldsymbol{\xi} - \boldsymbol{\sigma}(\boldsymbol{G}_{mn})\|_2^2$$

subject to

$$\sum\nolimits_{k=1}^{K} \log \xi_k \leq \log \mu.$$

Change variables to complete the proof. □

5.4.2 Numerical Experiments. We implemented the modified version of Algorithm 3.1 in Matlab, using the built-in nonlinear programming software to solve the optimization problem required by the proposition. For a few triples (d, K, N), we ran one hundred to five hundred iterations of the algorithm for various values of the feasibility parameter μ. (Given the exploratory nature of these experiments, we found that the implementation was too slow to increase the number of iterations.)

The results appear in Table 6. For small values of N, we find that the packings exhibit the maximum possible packing diameter $\pi/2$, which shows that the algorithm is succeeding in these cases. For larger values of N, we are unable to judge how close the packings might be from optimal.

Figure 4 compares the quality of our real packings against our complex packings. In each case, the complex packing is at least as good as the real packing, as we would expect. The smooth decline in the quality of the complex packings suggests that there is some underlying order to the packing diameters, but it remains to be discovered.

			Squared Packing diameters	
d	K	N	$\mathbb{R}$	$\mathbb{C}$
2	4	3	1.0000	1.0000
2	4	4	1.0000	1.0000
2	4	5	1.0000	1.0000
2	4	6	1.0000	1.0000
2	4	7	0.8933	0.8933
2	4	8	0.8447	0.8559
2	4	9	0.8196	0.8325
2	4	10	0.8176	0.8216
2	4	11	0.7818	0.8105
2	4	12	0.7770	0.8033
2	5	3	1.0000	1.0000
2	5	4	1.0000	1.0000
2	5	5	1.0000	1.0000
2	5	6	0.9999	1.0000
2	5	7	1.0000	0.9999
2	5	8	1.0000	0.9999
2	5	9	1.0000	1.0000
2	5	10	0.9998	1.0000
2	5	11	0.9359	0.9349
2	5	12	0.9027	0.9022

TABLE 6. Packing in Grassmannians with Fubini–Study distance: Our best real packings ($\mathbb{F} = \mathbb{R}$) compared with our best complex packings ($\mathbb{F} = \mathbb{C}$) in the space $\mathbb{G}(K, \mathbb{F}^d)$. The packing diameter of a configuration is the Fubini–Study distance (2–3) between the closest pair of subspaces. Note that we have scaled the distance by $2/\pi$ so that it ranges between zero and one.

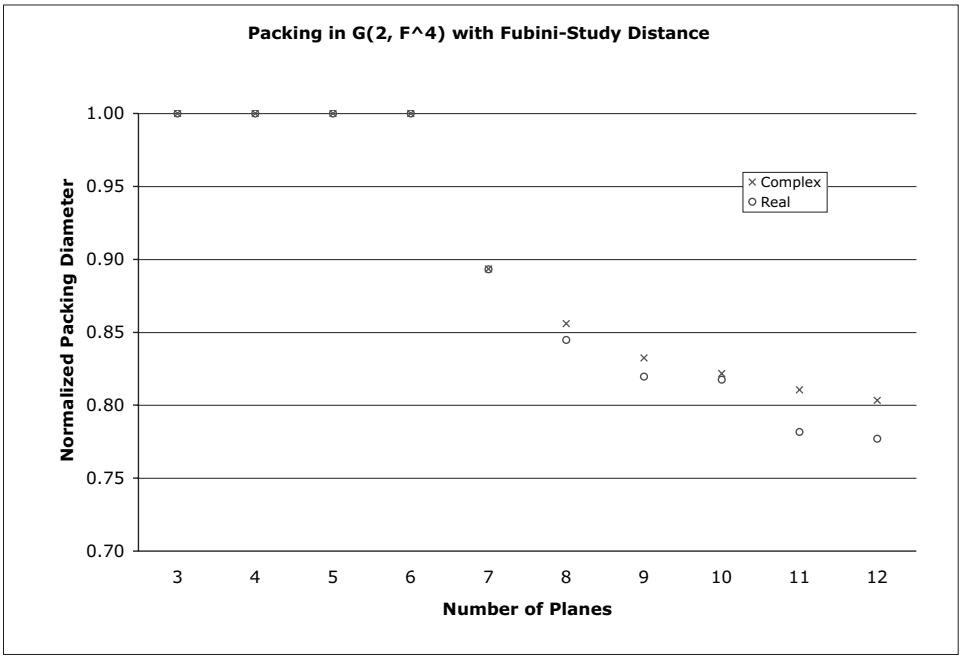

FIGURE 4. Packing in Grassmannians with Fubini–Study distance: This figure shows the packing diameters of N points in the Grassmannian $\mathbb{G}(K, \mathbb{F}^d)$ equipped with the Fubini–Study distance. The circles indicate the best real packings ($\mathbb{F} = \mathbb{R}$) obtained by the authors, while the crosses indicate the best complex packings ($\mathbb{F} = \mathbb{C}$) obtained.

To perform large-scale experiments, it will probably be necessary to tailor an algorithm that can solve the nonlinear programming problems more quickly. It may also be essential to implement the alternating projection in a programming environment more efficient than Matlab. Therefore, a detailed study of packing with respect to the Fubini–Study distance must remain a topic for future research.

6. DISCUSSION

6.1 Subspace Packing in Wireless Communications

Configurations of subspaces arise in several aspects of wireless communication, especially in systems with multiple transmit and receive antennas. The intuition behind this connection is that the transmitted and received signals in a multiple antenna system are connected by a matrix transformation, or *matrix channel*.

Subspace packings occur in two wireless applications: noncoherent communication and subspace quantization. The noncoherent application is primarily of theoretical interest, while subspace quantization has a strong impact on practical wireless systems. Grassmannian packings appear in these situations due to an assumption that the matrix channel should be modeled as a complex Gaussian random matrix.

In the noncoherent communication problem, it has been shown that from an information-theoretic perspective, under certain assumptions about the channel matrix, the optimum transmit signal corresponds to a packing in $\mathbb{G}(K, \mathbb{C}^d)$, where K corresponds to the minimum of the number of transmit and receive antennas and d corresponds to the number of consecutive samples over which the channel is constant [Zheng and Tse 02, Hochwald and Marzetta 00]. In other words, the number of subspaces K is determined by the system configuration, while d is determined by the carrier frequency and the degree of mobility in the propagation channel.

On account of this application, several papers have investigated the problem of finding packings in Grassmannian manifolds. One approach for the case of $K = 1$ is presented in [Hochwald and Marzetta 00]. This paper proposes a numerical algorithm for finding line packings, but it does not discuss its properties or connect it with the general subspace packing problem. Another approach, based on discrete Fourier transform matrices, appears in [Hochwald et al. 00]. This construction is both structured and flexible, but it does not lead to optimal packings. The paper [Agrawal et al. 01] studies Grassmannian packings in detail, and it contains an algorithm for finding packings in the complex Grassmannian manifold equipped with chordal distance. This algorithm is quite complex: it uses surrogate functionals to solve a sequence of relaxed nonlinear programs. The authors tabulate several excellent chordal packings, but it is not clear whether their method generalizes to other metrics.

The subspace quantization problem also leads to Grassmannian packings. In multiple-antenna wireless systems, one must quantize the dominant subspace in the matrix communication channel. Optimal quantizers can be viewed as packings in $\mathbb{G}(K, \mathbb{C}^d)$, where K is the dimension of the subspace and d is the number of transmit antennas. The chordal distance, the spectral distance, and the Fubini–Study distance are all useful in this connection [Love and Heath 05a, Love and Heath 05b]. This literature does not describe any new algorithms for constructing packings; it leverages results from the noncoherent communication literature. Communication strategies based on quantization via subspace packings have been incorporated into at least one recent standard [WirelessMAN 05].

6.2 Conclusions

We have shown that the alternating projection algorithm can be used to solve many different packing problems. The method is easy to understand and to implement, even while it is versatile and powerful. In cases in which experiments have been performed, we have often been able to match the best packings known. Moreover, we have extended the method to solve problems that have not been studied numerically. Using the Rankin bounds, we have been able to show that many of our packings are essentially optimal. It seems clear that alternating projection is an effective numerical algorithm for packing.

7. TAMMES' PROBLEM

The alternating projection method can also be used to study Tammes' problem of packing points on a sphere [Tammes 30]. This question has received an enormous amount of attention over the last 75 years, and extensive tables of putatively optimal packings are available [Sloane 04b]. This section offers a brief treatment of our work on this problem.

7.1 Modifications to the Algorithm

Suppose that we wish to produce a configuration of N points on the unit sphere $\mathbb{S}^{d-1}$ with a packing diameter ρ.

The feasibility problem requires that

$$\max_{m \neq n} \langle \boldsymbol{x}_m, \boldsymbol{x}_n \rangle \leq \mu, \tag{7–1}$$

where $\mu = \sqrt{1 - \rho^2}$. This leads to the convex structural constraint set

$$\begin{aligned} \mathcal{H}(\mu) \stackrel{\text{def}}{=} \{ \boldsymbol{H} \in \mathbb{R}^{N \times N} : \boldsymbol{H} = \boldsymbol{H}^*, \\ h_{nn} = 1 \text{ for } n = 1, 2, \dots, N, \\ \text{and} \quad -1 \leq h_{mn} \leq \mu \text{ for all } m \neq n \}. \end{aligned}$$

The spectral constraint set is the same as before. The associated matrix nearness problem is trivial to solve.

Proposition 7.1. *Let $\boldsymbol{G}$ be a real symmetric matrix. With respect to the Frobenius norm, the unique matrix in $\mathcal{H}(\mu)$ closest to $\boldsymbol{G}$ has a unit diagonal and off-diagonal entries that satisfy*

$$h_{mn} = \begin{cases} -1, & g_{mn} < -1, \\ g_{mn}, & -1 \leq g_{mn} \leq \mu, \\ \mu, & \mu < g_{mn}. \end{cases}$$

7.2 Numerical Results

Tammes' problem has been studied for 75 years, and many putatively optimal configurations are available. Therefore, we attempted to produce packings whose maximum inner product μ fell within 10^{-5} of the best value tabulated by N. J. A. Sloane and his colleagues [Sloane 04b]. This resource draws from all the experimental and theoretical work on Tammes' problem, and it should be considered the gold standard.

Our experimental setup echoes the setup for real projective packings. We implemented the algorithms in Matlab, and we performed the following experiment for pairs (d, N) with $d = 3, 4, 5$ and $N = 4, 5, \dots, 25$. First, we computed the putatively optimal maximum inner product μ using the data from [Sloane 04b]. In each of ten trials, we constructed a starting matrix using Algorithm 3.4 with parameters $\tau = 0.9$ and $T = 10{,}000$. Then, we executed the alternating projection, Algorithm 3.1, with the calculated value of μ and the maximum number of iterations set to $T = 5000$. We stopped the alternating projection in step 4 if the iterate $\boldsymbol{G}^{(t)}$ contained no off-diagonal entry greater than $\mu + 10^{-5}$ and proceeded with step 6. After ten trials, we recorded the largest packing diameter attained, as well as the average value of the packing diameter. We also recorded the average number of iterations the alternating projection required during each trial.

Table 7 provides the results of this experiment. The most striking feature of the table is that the best configurations returned by alternating projection consistently attain packing diameters that fall hundredths or thousandths of a degree away from the best packing diameters recorded by Sloane. If we examine the maximum inner product in the configuration instead, the difference is usually on the order of 10^{-4} or 10^{-5}, which we expect based on our stopping criterion. The average-case results are somewhat worse. Nevertheless, the average configuration returned by alternating projection typically attains a packing diameter only several tenths of a degree away from optimal.

For collections of N points on the $(d-1)$-dimensional sphere, this table lists the best packing diameter and the average packing diameter obtained during ten random trials of the alternating projection algorithm. The error columns record how far our results diverge from the putative optimal packings (NJAS) reported in [Sloane 04b]. The last column gives the average number of iterations of alternating projection per trial.

A second observation is that the alternating projection algorithm typically performs better when the number of points N is small. The largest errors are all clustered at larger values of N. A corollary observation is that the average number of iterations per trial tends to increase with the number of points. We believe that the explanation for these phenomena is that Tammes' problem has a combinatorial regime, in which solutions have a great deal of symmetry and structure, and a random regime, in which the solutions have very little order. The algorithm typically seems to perform better in the combinatorial regime, although it fails for certain unusually structured ensembles.

This claim is supported somewhat by theoretical results for $d = 3$. Optimal configurations have been established only for $N = 1, 2, \dots, 12$ and $N = 24$. Of these, the cases $N = 1, 2, 3$ are trivial. The cases $N = 4, 6, 8, 12, 24$ fall from the vertices of various well-known polyhedra. The cases $N = 5, 11$ are degenerate, obtained by leaving a point out of the solutions for $N = 6, 12$. The remaining cases involve complicated constructions based on graphs [Ericson and Zinoviev 01]. The algorithm was able to calculate the known optimal configurations to a high order of accuracy, but it generally performed slightly better for the nondegenerate cases.

On the other hand, there is at least one case in which the algorithm failed to match the optimal packing diameter, even though the optimal configuration is highly symmetric. The best arrangement of 24 points on $\mathbb{S}^3$ lo-

		Packing diameters (Degrees)					Iterations
d	N	NJAS	Best of 10	Error	Avg. of 10	Error	Avg. of 10
3	4	109.471	109.471	0.001	109.471	0.001	45
3	5	90.000	90.000	0.000	89.999	0.001	130
3	6	90.000	90.000	0.000	90.000	0.000	41
3	7	77.870	77.869	0.001	77.869	0.001	613
3	8	74.858	74.858	0.001	74.858	0.001	328
3	9	70.529	70.528	0.001	70.528	0.001	814
3	10	66.147	66.140	0.007	66.010	0.137	5000
3	11	63.435	63.434	0.001	63.434	0.001	537
3	12	63.435	63.434	0.001	63.434	0.001	209
3	13	57.137	57.136	0.001	56.571	0.565	4876
3	14	55.671	55.670	0.001	55.439	0.232	3443
3	15	53.658	53.620	0.038	53.479	0.178	5000
3	16	52.244	52.243	0.001	51.665	0.579	4597
3	17	51.090	51.084	0.007	51.071	0.019	5000
3	18	49.557	49.548	0.008	49.506	0.050	5000
3	19	47.692	47.643	0.049	47.434	0.258	5000
3	20	47.431	47.429	0.002	47.254	0.177	5000
3	21	45.613	45.576	0.037	45.397	0.217	5000
3	22	44.740	44.677	0.063	44.123	0.617	5000
3	23	43.710	43.700	0.009	43.579	0.131	5000
3	24	43.691	43.690	0.001	43.689	0.002	3634
3	25	41.634	41.458	0.177	41.163	0.471	5000
4	5	104.478	104.478	0.000	104.267	0.211	2765
4	6	90.000	90.000	0.000	89.999	0.001	110
4	7	90.000	89.999	0.001	89.999	0.001	483
4	8	90.000	90.000	0.000	89.999	0.001	43
4	9	80.676	80.596	0.081	80.565	0.111	5000
4	10	80.406	80.405	0.001	77.974	2.432	2107
4	11	76.679	76.678	0.001	75.881	0.798	2386
4	12	75.522	75.522	0.001	74.775	0.748	3286
4	13	72.104	72.103	0.001	71.965	0.139	4832
4	14	71.366	71.240	0.126	71.184	0.182	5000
4	15	69.452	69.450	0.002	69.374	0.078	5000
4	16	67.193	67.095	0.098	66.265	0.928	5000
4	17	65.653	65.652	0.001	64.821	0.832	4769
4	18	64.987	64.987	0.001	64.400	0.587	4713
4	19	64.262	64.261	0.001	64.226	0.036	4444
4	20	64.262	64.261	0.001	64.254	0.008	3738
4	21	61.876	61.864	0.012	61.570	0.306	5000
4	22	60.140	60.084	0.055	59.655	0.485	5000
4	23	60.000	59.999	0.001	58.582	1.418	4679
4	24	60.000	58.209	1.791	57.253	2.747	5000
4	25	57.499	57.075	0.424	56.871	0.628	5000
5	6	101.537	101.536	0.001	95.585	5.952	4056
5	7	90.000	89.999	0.001	89.999	0.001	1540
5	8	90.000	89.999	0.001	89.999	0.001	846
5	9	90.000	89.999	0.001	89.999	0.001	388
5	10	90.000	90.000	0.000	89.999	0.001	44
5	11	82.365	82.300	0.065	81.937	0.429	5000
5	12	81.145	81.145	0.001	80.993	0.152	4695
5	13	79.207	79.129	0.078	78.858	0.349	5000
5	14	78.463	78.462	0.001	78.280	0.183	1541
5	15	78.463	78.462	0.001	77.477	0.986	1763
5	16	78.463	78.462	0.001	78.462	0.001	182
5	17	74.307	74.307	0.001	73.862	0.446	4147
5	18	74.008	74.007	0.001	73.363	0.645	3200
5	19	73.033	73.016	0.017	72.444	0.589	5000
5	20	72.579	72.579	0.001	72.476	0.104	4689
5	21	71.644	71.639	0.005	71.606	0.039	5000
5	22	69.207	68.683	0.524	68.026	1.181	5000
5	23	68.298	68.148	0.150	67.568	0.731	5000
5	24	68.023	68.018	0.006	67.127	0.896	5000
5	25	67.690	67.607	0.083	66.434	1.256	5000

TABLE 7. Packing on spheres.

cates them at vertices of a very special polytope called the 24-cell [Sloane 04b]. The best configuration produced by the algorithm has a packing diameter that is worse by 1.79°. It seems that this optimal configuration is very difficult for the algorithm to find. Less dramatic failures occurred at pairs $(d, N) = (3, 25)$, $(4, 14)$, $(4, 25)$, $(5, 22)$, and $(5, 23)$. But in each of these cases, our best packing diverged by more than a tenth of a degree from the best recorded.

ACKNOWLEDGMENTS

ISD was supported by NSF grant CCF-0431257, NSF Career Award ACI-0093404, and NSF-ITR award IIS-0325116. RWH was supported by NSF CCF Grant #514194. TS was supported by NSF DMS Grant #0511461. JAT was supported by an NSF Graduate Fellowship, a J. T. Oden Visiting Faculty Fellowship, and NSF DMS Grant #0503299.

REFERENCES

[Agrawal et al. 01] D. Agrawal, T. J. Richardson, and R. L. Urbanke. "Multiple-Antenna Signal Constellations for Fading Channels." *IEEE Trans. Inform. Theory* 6:47 (2001), 2618–2626.

[Barg and Nogin 02] A. Barg and D. Yu. Nogin. "Bounds on Packings of Spheres in the Grassmannian Manifold." *IEEE Trans. Inform. Theory* 48:9 (2002), 2450–2454.

[Björck and Golub 73] Å. Björck and G. Golub. "Numerical Methods for Computing Angles between Linear Subspaces." *Mathematics of Computation* 123:27 (1973), 579–594.

[Cheney and Goldstein 59] E. W. Cheney and A. A. Goldstein. "Proximity Maps for Convex Sets." *Proc. Amer. Math. Soc.* 10:3 (1959), 448–450.

[Conway and Sloane 98] J. H. Conway and N. J. A. Sloane. *Sphere Packing, Lattices and Groups*, 3rd edition, Grundlehren der mathematischen Wissenschaften, 290. New York: Springer-Verlag, 1998.

[Conway et al. 96] J. H. Conway, R. H. Hardin, and N. J. A. Sloane. "Packing Lines, Planes, etc.: Packings in Grassmannian Spaces." *Experimental Math.* 5:2 (1996), 139–159.

[Cover and Thomas 91] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. New York: John Wiley and Sons, 1991.

[Ericson and Zinoviev 01] T. Ericson and V. Zinoviev. *Codes on Euclidean Spheres*. Amsterdam: Elsevier, 2001.

[Golub and Van Loan 96] G. H. Golub and C. F. Van Loan. *Matrix Computations*, 3rd edition. Baltimore: Johns Hopkins University Press, 1996.

[Hales 04] T. C. Hales. "A Proof of the Kepler Conjecture (DCG Version)." Available online (http://www.math.pitt.edu/~thales/kepler04/fullkepler.pdf), 2004.

[Hochwald and Marzetta 00] B.M. Hochwald and T.L. Marzetta. "Unitary Space-Time Modulation for Multiple-Antenna Communications in Rayleigh Flat Fading." *IEEE Trans. Info. Theory* 46:2 (2000), 543–564.

[Hochwald et al. 00] 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* 46:6 (2000), 1962–1973.

[Holmes and Paulsen 04] R. B. Holmes and V. I. Paulsen. "Optimal Frames for Erasures." *Linear Algebra Appl.* 377 (2004), 31–51.

[Horn and Johnson 85] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge: Cambridge University Press, 1985.

[Lemmens and Seidel 73] P. W. H. Lemmens and J. J. Seidel. "Equi-isoclinic Subspaces of Euclidean Spaces." *Proc. Nederl. Akad. Wetensch. Series A* 76 (1973), 98–107.

[Love and Heath 05a] D. J. Love and R. W. Heath Jr. "Limited Feedback Unitary Precoding for Orthogonal Space-Time Block Codes." *IEEE Trans. Signal Processing* 53:1 (2005), 64–73.

[Love and Heath 05b] D. J. Love and R. W. Heath Jr. "Limited Feedback Unitary Precoding for Spatial Multiplexing Systems." *IEEE Trans. Info. Theory* 51:8 (2005), 2967–2976.

[Love et al. 03] D. J. Love, R. W. Heath Jr., and T. Strohmer. "Grassmannian Beamforming for Multiple-Input Multiple-Output Wireless Systems." *IEEE Trans. Info. Theory* 49:10 (2003), 2735–2747.

[Love et al. 04] D. J. Love, R. W. Heath Jr., W. Santipach, M. L. Honig. "What Is the Value of Limited Feedback for MIMO Channels?" *IEEE Comm. Mag.* 42:10 (2004), 54–59.

[Meyer 76] R. R. Meyer. "Sufficient Conditions for the Convergence of Monotonic Mathematical Programming Algorithms." *J. Comp. Sys. Sci.* 12 (1976), 108–121.

[von Neumann 50] J. von Neumann. *Functional Operators*, vol. II, Annals of Mathematics Studies, 22. Princeton: Princeton University Press, 1950.

[Rankin 47] R. A. Rankin. "On the Closest Packing of Spheres in n Dimensions." *Ann. Math.* 48 (1947), 1062–1081.

[Rockafellar 70] R. T. Rockafellar. *Convex Analysis*. Princeton: Princeton University Press, 1970.

[Rogers 64] C. A. Rogers. *Packing and Covering*. Cambridge: Cambridge University Press, 1964.

[Sloane 04a] N. J. A. Sloane. "Table of Best Grassmannian Packings." Available online (http://www.research.att.com/~njas/grass/grassTab.html), 2004.

[Sloane 04b] N. J. A. Sloane. "Tables of Spherical Codes." Available online (http://www.research.att.com/~njas/packings/), 2004.

[Stewart 80] G. W. Stewart "The Efficient Generation of Random Orthogonal Matrices with an Application to Condition Estimation." *SIAM J. Numer. Anal.* 17:30 (1980), 403–409.

[Strohmer and Heath 03] T. Strohmer and R. W. Heath Jr. "Grassmannian Frames with Applications to Coding and Communication." *Appl. Comp. Harmonic Anal.* 14:3 (2003), 257–275.

[Sustik et al. 07] M. A. Sustik, J. A. Tropp, I. S. Dhillon, and R. W. Heath Jr.. "On the Existence of Equiangular Tight Frames." *Linear Algebra Appl.* 426:2–3 (2007), 619–635.

[Tammes 30] P. M. L. Tammes. "On the Origin of Number and Arrangement of the Places of Exit on the Surface of Pollen Grains." *Rec. Trav. Bot. Neerl.* 27 (1930) 1–84.

[Tóth 64] L. Fejes Tóth. *Regular Figures*. New York: Macmillan, 1964.

[Tóth 65] L. Fejes Tóth. "Distribution of Points in the Elliptic Plane." *Acta Math. Acad. Hung. Sci.* 16 (1965), 437–440.

[Tropp et al. 05] J. A. Tropp, I. S. Dhillon, R. W. Heath Jr., T. Strohmer. "Designing Structured Tight Frames via Alternating Projection." *IEEE Trans. Info. Theory* 51:1 (2005), 188–209.

[Trosset 01] M. W. Trosset. "Approximate Maximin Distance Designs." In *Proceedings of the Section on Physical and Engineering Sciences*, pp. 223–227. Alexandria, VA: American Statistical Association, 2001.

[WirelessMAN 05] IEEE WirelessMAN. "Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems." *IEEE P802.16e/D8*, May 2005.

[Zheng and Tse 02] L. Zheng and D. N. C. Tse. "Communication on the Grassmann Manifold: A Geometric Approach to the Noncoherent Multiple-Antenna Channel." *IEEE Trans. Info. Theory* 48:2 (2002), 359–383.

I. S. Dhillon, Department of Computer Sciences, University of Texas at Austin, 1 University Station C0500, Austin, TX 78712 (inderjit@cs.utexas.edu).

R. W. Heath Jr., Department of Electrical and Computer Engineering, University of Texas at Austin, 1 University Station C0500, Austin, TX 78712 (rheath@ece.utexas.edu)

T. Strohmer, Department of Mathematics, University of California, Davis, Mathematical Sciences Building, 1 Shields Ave., Davis, CA 95616 (strohmer@ucdavis.edu)

J. Tropp, Department of Applied and Computational Mathematics, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125 (jtropp@acm.caltech.edu)

Received November 21, 2006; accepted February 21, 2007.