Flexible Codebook Design for Limited Feedback Systems Via Sequential Smooth Optimization on the Grassmannian Manifold

Ahmed Medra and Timothy N. Davidson, *Member, IEEE*

***Abstract*—Grassmannian quantization codebooks play a central role in a number of limited feedback schemes for single and multi-user multiple-input multiple-output (MIMO) communication systems. In practice, it is often desirable that these codebooks possess additional properties that facilitate their implementation, beyond the provision of good quantization performance. Although some good codebooks exist, their design tends to be a rather intricate task. The goal of this paper is to suggest a flexible approach to the design of Grassmannian codebooks based on sequential smooth optimization on the Grassmannian manifold and the use of smooth penalty functions to obtain additional desirable properties. As one example, the proposed approach is used to design rank-2 codebooks that have a nested structure and elements from a phase-shift keying (PSK) alphabet. In some numerical comparisons, codebooks designed using the proposed methods have larger minimum distances than some existing codebooks, and provide tangible performance gains when applied to a simple MIMO downlink scenario with zero-forcing beamforming, per-user unitary beamforming and rate control (PU^{2}RC), and block diagonalization signalling. Furthermore, the proposed approach yields codebooks that attain desirable additional properties without incurring a substantial degradation in performance.**

***Index Terms*—Feedback, Grassmannian, codebook design, quantization, line, subspace packing, defined alphabet, nested.**

I. Introduction

The structure of effective schemes for communicating wirelessly between nodes with multiple antennas is dependent on the extent to which the transmitter can adapt its transmission to the current state of the communication environment. As an example, in multi-user multiple-input multiple-output (MIMO) systems, channel state information (CSI) can be used by the transmitter to expand the achievable rate region through prudent management (and exploitation) of interference. In particular, consider the case of downlink transmission over a richly-scattered environment from a base station with M antennas to K mobile stations, each with a single antenna, in the presence of additive white Gaussian noise at the receivers. In such a system, if the base station has perfect knowledge of the channels to the receivers, then the sum of the ergodic achievable rates of coherent communication to the mobile users grows as $\min\{M,K\}\log(\mathsf{SNR})$ as the SNR increases, where SNR is the signal-to-noise ratio [1]. When $M \geq K$ this rate of growth can be achieved by employing a simple linear zero-forcing transmission strategy [1], [2]. However, in the absence of any information about the state of the channel, the sum rate grows only as $\log(\mathsf{SNR})$ as the SNR increases [1], [3]. In other words, the availability of perfect CSI at the transmitter enables a $\min\{M,K\}$-fold increase in the rate of growth of the achievable sum rate over that of a system with a single antenna at the transmitter, whereas in the absence of CSI at the transmitter the rate of growth is the same as that of a system with a single transmit antenna; see also [4].

Seeing as coherent receivers obtain an estimate of the channel, typically through training, prior to processing their received signals, a natural approach to providing the transmitter with information regarding the channel state would be for the receivers to feed back these estimates (or some function thereof) to the transmitter. However, account needs to be taken of the communication resources that are allocated to the feedback. When the feedback is digital, the question of how the channel state information should be quantized arises naturally [4]–[6]. In essence, this is a source coding problem [7], in which the channel (or some function thereof) is regarded as a source that needs to be represented using a small number of bits while introducing only a small "distortion". In this application, the notion of distortion is associated with the degradation of a chosen measure of the performance of the communication system with respect to the perfect CSI case. Based on insight from the source coding literature, a number of feedback strategies have been developed [5], many of which are based on the principles of memoryless vector quantization [7]. In those schemes, the receiver compares its channel estimate (or some function thereof) to "codewords" in a codebook and the index of the codeword that is closest to the channel estimate, in some appropriate sense, is fed back to the transmitter. The transmitter has a copy of the codebook and uses the index it receives to reconstruct the codeword that the receiver selected.

Although generic techniques for the design of quantization codebooks (e.g., [7]) could be applied to the limited feedback problem, these techniques tend to be rather cumbersome and

Manuscript received June 06, 2013; revised October 10, 2013 and December 24, 2013; accepted December 28, 2013. Date of publication January 17, 2014; date of current version February 12, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Xiqi Gao. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada and the Canada Research Chairs Program. A preliminary version of a portion of this work appears in *Proceedings of theThirteenth IEEE International Workshop on Signal Processing Advances in Wireless Communications,* Çeşme, Turkey, June 2012.

The authors are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada(e-mail: medraam@grads.ece.mcmaster.ca; davidson@mcmaster.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2014.2301137

the resulting codebooks can be rather awkward to implement. As a result, in many scenarios, the quantization codebook is partitioned into a codebook that represents the directional information and a codebook that represents the gain [5]. In a number of important scenarios, the relevant communication performance metric depends on the subspaces spanned by the directions, rather than the directions themselves. Since subspaces of a given dimension in a given ambient space can be represented by points on the Grassmannian manifold of the corresponding dimensions, in these scenarios the natural setting for the design of the codebook that represents the directional information is the Grassmannian manifold [8]–[13].

Despite the insight provided by posing the codebook design problem on the manifold, codebook design remains a rather intricate task, even when the scenario is such that the codewords should be uniformly distributed on the manifold with respect to a certain distance metric; a problem that is typically referred to as Grassmannian Subspace Packing [14]. One reason for this intricacy is that, with the exception of a few special cases (e.g., [15]), the problem has proven quite resistant to analysis of the structure of the optimal codebook. Furthermore, numerical optimization procedures are complicated by the fact that the manifold itself is inherently non-convex, as are the objectives that are typically used to capture the desired performance properties of the codebook. Some numerical design methods have been proposed, including adaptations of Lloyd's algorithm to the manifold [11], [16], an alternating projection method [17], and an "expansion-compression" algorithm [18]. In addition, methods for the design of Grassmannian constellations for non-coherent point-to-point MIMO communication (e.g., [19]–[22]) can be employed.

Numerical methods have been quite successful in designing codebooks in which each codeword represents a subspace of dimension one; i.e., a single direction. Such codebooks are often referred to as being "rank one" codebooks, and the subspace packing problem is often referred to as the Grasssmannian Line Packing problem. Some of these numerical methods have been also used to design "rank two" codebooks of codewords representing two-dimensional subspaces for systems with small number of transmitter antennas. However, for larger problems, there are fewer results available, especially in the case in which distances other than the chordal distance are chosen as the metric on the manifold.

In addition to pure quantization performance, practical considerations suggest that codebooks ought to possess additional properties that facilitate their implementation and storage [23]. For example, (i) in the single-user case one might seek to constrain the elements of the codewords to have constant modulus, so as to avoid power imbalances at the transmitter; (ii) for reasons of storage and computational cost, it may be desirable to have the elements of the codewords come from a defined alphabet such as the phase-shift keying (PSK) alphabet [24]; and (iii) in applications in which there is the option of multiple signalling modes (e.g., [25]), it is desirable for the codebook to be nested, in the sense that structure is imposed on the higher rank codewords so that they can also be used to generate codewords for codebooks of lower rank. In that case, the transmitter and the receiver need to store only a single codebook, and calculations results for a lower rank codeword can be re-used for the calculation of a higher rank one. A nested codebook also facilitates the transmitter's use of "rank overriding", in which the transmitter overrides the receiver's decision on how many data streams to transmit and uses lower-rank transmission [23].

The desire to obtain codebooks that possess additional properties significantly complicates what is already quite a difficult problem. In the case of codebooks from a defined alphabet, the problem becomes a combinational one. For certain alphabets, codebooks of certain dimensions and sizes can be generated [26], [27] using the concept of mutually unbiased bases (MUBs) from the theory of quantum mechanics [28]. By construction, codebooks derived from MUBs exhibit large distances between each pair of codewords, but the inherent restrictions on the technique limit the scenarios in which it can be applied.

Given the rather specialized nature of the existing approaches to structured Grassmannian codebook design, the goal of this paper is to suggest a flexible design approach based on smooth optimization on the Grassmannian manifold [29], [30]. In particular, to optimize the quantization performance of otherwise unconstrained codebooks we use insight from the technique in [22] and construct a sequence of smooth approximations to the objective. Since each approximation has continuous first and second order derivatives, a codebook that (locally) optimizes the approximation can be obtained by using adaptations of the conventional smooth optimization techniques, such as gradient descent, conjugate gradient and Newton's method, to the manifold [29], [30]. The optimization with respect to the next approximation in the sequence is initialized with the codebook obtained by optimizing the current approximation. In order to tackle problems with constraints on the codebook, we suggest the use of a combination of smooth penalty functions, which enable the algorithms for unconstrained optimization on the manifold to be retained, rounding techniques and low-dimensional searches.

Since both the original design problem and our smooth unconstrained approximations are not convex, the proposed approach does not necessarily yield optimal codebooks. However, in a number of our examples for the unconstrained case the codebooks we have obtained achieve a known upper bound on the codebook quantization performance and hence are optimal. In the case of the Fubini-Study distance, for which a simple performance bound is not available, the unconstrained codebooks that we have obtained exhibit larger minimum distances than the existing packings [31]. Furthermore, the proposed approach has been able to generate codebooks with excellent properties for systems with a larger number of transmit antennas, larger subspace dimensions and larger size than those currently available. In the case of the chordal distance, we have been able to design unconstrained codebooks that have essentially the same minimum distances as the codebooks in [17]. In the case of constrained codebooks, we have been able to design some codebooks that achieve an upper bound on the performance and hence are optimal, and other codebooks with distance properties that are close to those of the best unconstrained codebooks that we have been able to generate.

In order to assess the impact of the obtained codebook designs on the performance of practical limited feedback systems, we evaluate the performance of existing and proposed codebooks under three transmission schemes for the downlink of a wireless system (from base station to mobile station) with limited feedback, namely Zero-Forcing Beam Forming (ZFBF) [4], the Per-User Unitary beamforming and Rate Control (PU^{2}RC) system [32], and zero-forcing block diagonalization (ZFBD) [33]–[35]. In all three schemes, the transmitter employs linear precoding of the symbols intended for each user. In the first two schemes, the base station transmits a single data stream to each mobile station and hence rank-one codebooks designed using Grassmannian line packing are employed. The difference between those schemes lies in the way that precoding matrix is constructed from the quantized feedback. In the third scheme, the base station transmits multiple streams to each (multiple-antenna) mobile station, and higher rank subspace codebooks are employed. Our results show that the achievable rate can be increased and the outage probability can be decreased by implementing well-designed codebooks. These results also show that the proposed approach yields structured codebooks with performance close to that of the unstructured codebooks.

II. System Model

Although the proposed approach to the design of structured Grassmannian quantization codebooks is quite generic and can be applied in many applications, we will position the approach in the context of a multi-user multiple-input multiple-output (MU-MIMO) downlink communication system in which a base station with M_t antennas communicates over narrow-band channels to K receivers, the kth of which is equipped with $M_{r_k} < M_t$ antennas. The case of a point-to-point MIMO link arises as the special case in which $K = 1$. At each channel use, the transmitter communicates simultaneously with all K receivers, sending $Q_k \leq \min\{M_t, M_{r_k}\}$ symbols to the kth receiver. The transmitter synthesizes the signal $\mathbf{x} = \sum_i \mathbf{P}_i \mathbf{s}_i \in \mathbb{C}^{M_t \times 1}$ for transmission, where $\mathbf{s}_i \in \mathbb{C}^{Q_i \times 1}$ contains the symbols intended for the ith receiver, normalized so that $E\{\mathbf{s}_i \mathbf{s}_i^H\} = \mathbf{I}_{Q_i}$, and $\mathbf{P}_i \in \mathbb{C}^{M_t \times Q_i}$ is a linear precoding matrix. The average transmitted power per symbol for the ith receiver is $\mathrm{trace}(\mathbf{P}_i \mathbf{P}_i^H)/Q_i$. The signal observed at the kth receiver is

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k = \mathbf{H}_k \mathbf{P}_k \mathbf{s}_k + \mathbf{H}_k \sum_{i \neq k} \mathbf{P}_i \mathbf{s}_i + \mathbf{n}_k, \qquad (1)$$

where $\mathbf{H}_k \in \mathbb{C}^{M_{r_k} \times M_t}$ is the matrix of complex channel gains from the antenna inputs at the transmitter to the antenna outputs at the kth receiver. The first term on the right hand side of (1) denotes the desired signal term, the second denotes the interference from signals transmitted to other receivers and the third term represents the additive noise at the kth receiver. A more compact form of the model in (1) can be obtained by concatenating the $Q = \sum_i Q_i$ symbols sent to the receivers into a single vector $\mathbf{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \ldots, \mathbf{s}_K^T]^T$, and concatenating the precoding matrices $\mathbf{P}_i$ to form $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_K] \in \mathbb{C}^{M_t \times Q}$. In that case, the expression in (1) can be rewritten as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{P} \mathbf{s} + \mathbf{n}_k. \qquad (2)$$

We consider scenarios in which the channels can be modelled as remaining (approximately) constant for long enough that it is viable for the receiver(s) to feed back information regarding the channel state to the transmitter over a dedicated channel of limited capacity, and for the transmitter to adapt its linear precoding matrix $\mathbf{P}$ (and possibly the coding and modulation schemes that produce $\mathbf{s}$) to the information that it receives. In point-to-point MIMO links, the information that is fed back to the transmitter is typically the precoder $\mathbf{P}$ that the transmitter is to use. One prominent class of such schemes is that in which unitary precoding is employed; i.e., $\mathbf{P}$ is a scaled semi-unitary matrix that satisfies $\mathbf{P}^H \mathbf{P} \propto \mathbf{I}_Q$, [8], [9], [12], [25]. For point-to-point links, unitary precoding is optimal at high SNRs (e.g., [36]) and for the case in which the receiver employs zero-forcing decision-feedback spatial equalization [12], provided that the channel matrix has full column rank. In such a system, communication performance metrics such as the achievable rate and the probability of error depend on the subspace spanned by the precoding matrix $\mathbf{P}$, rather than $\mathbf{P}$ itself, and hence the Grassmann manifold of representatives of these subspaces is the natural setting for the quantization problem [8]–[10]. Under the i.i.d. Rayleigh fading model for the channel gains, the quantization codebook should be uniformly distributed on the manifold with respect to a distance metric that is dependent on the chosen communication performance metric; i.e., a Grassmannian (subspace) packing [14].

In the case of downlink transmission from a base station to multiple receivers, the precoding matrix $\mathbf{P}$ must be adapted at the transmitter, as each receiver only has knowledge of its own channel. Although the receivers could quantize their channel matrices directly, the quantization scheme is typically partitioned into a quantizer that captures the "directions" of the channel (i.e., the right singular vectors of $\mathbf{H}_k$) and a quantizer that captures the "quality" of the channel in those directions (typically a measure of the ratio of the corresponding singular value and the noise variance at receiver k). The nature of the appropriate quantization scheme for the receiver to employ is dependent on how the transmitter adapts $\mathbf{P}$ to the information that it receives, but in some important cases (e.g., [4], [32], [35]) the desirable quantization scheme once again involves construction of Grassmannian codebooks that are uniformly distributed with respect to a chosen distance metric on the manifold.

III. Grassmannian Packings

Despite the prominent role that Grassmannian subspace packings play in point-to-point and downlink communication schemes with limited feedback, the design of good packings is widely regarded as a rather difficult problem; e.g., [9], [14], [17]. The goal of this paper is to propose effective algorithms for finding good packings and for finding good packings that are constrained to posses a structure that facilitates implementation. To formalize our discussion, we note that for a given dimension $M \leq M_t$, the complex Grassmannian manifold $\mathbb{G}_{M_t,M}$ is the set of all subspaces of dimension M in $\mathbb{C}^{M_t}$. In order to perform quantization on the manifold, we observe that a subspace $\mathcal{S}$ of dimension M in $\mathbb{C}^{M_t}$ can be expressed as the linear span of an orthonormal basis; i.e., $\mathcal{S} = \{\mathbf{F}\mathbf{x} | \mathbf{x} \in \mathbb{C}^M\}$, where each column of the matrix $\mathbf{F} \in \mathbb{C}^{M_t \times M}$ is an element of the orthonormal basis. Since the columns of $\tilde{\mathbf{F}} = \mathbf{F}\mathbf{Q}$, where $\mathbf{Q}$ is an $M \times M$ unitary matrix, also form an orthonormal basis for $\mathcal{S}$, all such $\tilde{\mathbf{F}}$ can be deemed to be equivalent in terms of the description of $\mathcal{S}$. If for each subspace $\mathcal{S}$ we select one such $\tilde{\mathbf{F}}$ as the representative of the equivalence class, then each of these representatives specifies a point on $\mathbb{G}_{M_t,M}$. For the ease of exposition, with a mild abuse of terminology we will also refer to the set of representatives as a Grassmannian manifold, with the context making the distinction; see also [29]. The Grassmannian packing problem is the problem of finding a codebook $\mathcal{F} = \{\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_N\}$ of representatives of N subspaces of dimension M in $\mathbb{C}^{M_t}$. The codewords $\mathbf{F}_i \in \mathbb{C}^{M_t \times M}$ have orthonormal columns, and hence such a codebook is often said to be a rank-M codebook. For the case of an unconstrained uniform codebook, the codebook design problem can be posed in terms of maximizing the minimum distance between codeword pairs. That is, find a codebook $\mathcal{F} = \{\mathbf{F}_i\}_{i=1}^N$ that solves:

$$\max_{\{\mathbf{F}_i\},\, \mathbf{F}_i \in \mathbb{G}_{M_t,M}} \min_{i \neq j} d(\mathbf{F}_i, \mathbf{F}_j), \tag{3}$$

where $d(\mathbf{F}_i, \mathbf{F}_j)$ is a measure of the distance between the subspaces spanned by its arguments. There are many valid distance metrics on the Grassmannian manifold that have been used in the design of packings for a variety of limited feedback applications; e.g., [8]–[13], [22]. Two of the more prominent metrics are the chordal distance,

$$\begin{aligned} d_{\mathrm{ch}}(\mathbf{F}_i, \mathbf{F}_j) &= \frac{1}{\sqrt{2}} \left\| \mathbf{F}_i \mathbf{F}_i^H - \mathbf{F}_j \mathbf{F}_j^H \right\|_F \\ &= \left(M - \|\mathbf{F}_j^H \mathbf{F}_i\|_F^2 \right)^{1/2} \end{aligned} \tag{4}$$

and the Fubini Study distance,

$$d_{\mathrm{FS}}(\mathbf{F}_i, \mathbf{F}_j) = \arccos \left| \det \left(\mathbf{F}_j^H \mathbf{F}_i \right) \right|. \tag{5}$$

(Consistent with the definition of the manifold, these metrics are invariant to the right multiplication of any codeword matrix by a unitary matrix.) The chordal distance has been used to quantify the effect of quantization on the achievable rate in limited feedback systems and to derive the bounds on the rate-distortion trade-off [13]. It has also been used to asses the impact of quantization in point-to-point links that employ unitary precoding of orthogonal space-time block codes [10]. The Fubini-Study distance has been used to bound the degradation of the achievable rate over an i.i.d. Rayleigh channel caused by quantization [9].

In the case in which the subspaces are of dimension $M = 1$, the subspace packing problem in (3) collapses to the problem of packing Grassmannian lines and, from the perspective of that problem all the distances become equivalent. For that case, we will use the distance

$$d_{\mathrm{line}}(\mathbf{f}_i, \mathbf{f}_j) = \sqrt{\left(1 - |\mathbf{f}_j^H \mathbf{f}_i|^2\right)}, \tag{6}$$

where $\mathbf{f}_i$ denotes the ith element of a rank-1 codebook.

There are three aspects of the codebook design problem in (3) that make it difficult to solve. First, the constraint that the codewords $\mathbf{F}_i$ lie on the manifold is non-convex. Second, the $\min$ operator is not differentiable, and third, many distance metrics for the manifold are both non-convex and non-differentiable. These aspects suggest that the basic problem in (3) may be difficult to solve, even before we seek to modify it to obtain codebooks with additional features. Given these difficulties, it can be useful to have an upper bound on the optimal value of the problem in (3) so that the quality of potentially suboptimal solutions can be assessed.

If we let $\delta(M_t, M, N)$ denote the optimal value of the problem in (3) for a particular distance metric, then for the case of the chordal distance in (4) we can use the Rankin bounds, i.e., the simplex and the orthoplex bounds, to evaluate the quality of a codebook [14], [37]–[39]. The simplex bound is defined as

$$\delta_{\mathrm{ch}}(M_t, M, N) \leq \sqrt{\frac{M(M_t - M)N}{M_t(N-1)}}. \tag{7}$$

This bound is attainable only if $N \leq M_t^2$. For $N > M_t^2$, the orthoplex bound

$$\delta_{\mathrm{ch}}(M_t, M, N) \leq \sqrt{\frac{M(M_t - M)}{M_t}} \tag{8}$$

is a tighter bound, and this bound is attainable only if $N \leq 2(M_t^2 - 1)$. These bounds are also valid in the case of line packing; with mild abuse of notation, $\delta_{\mathrm{line}}(M_t, N) \leq \sqrt{\frac{(M_t-1)N}{M_t(N-1)}}$ for $N \leq M_t^2$ and $\delta_{\mathrm{line}}(M_t, N) \leq \sqrt{\frac{(M_t-1)}{M_t}}$ for $N > M_t^2$. In this paper, we will also consider codebooks that are constrained so that each element of each codeword has the same modulus; i.e., for $i = 1, 2, \ldots, N$, $|[\mathbf{F}_i]_{\ell m}|^2 = \frac{1}{M_t}$ for $\ell = 1, 2, \ldots, M_t, m = 1, 2, \ldots, M$. In this case, the ranges of the bounds can be refined, as outlined in [39]. In particular for such constrained codebooks, the simplex bound is attainable only if $N \leq M_t^2 - M_t + 1$, while the orthoplex bound is attainable only if $N \leq 2(M_t^2 - M_t)$.

IV. Subspace Packing Design via Sequential Smooth Optimization

The proposed strategy for constructing an effective technique for finding good solutions to the codebook design problem in (3) is to use algorithms developed for optimization of smooth functions on the Grassmannian manifold [29], [30]. These algorithms are adaptations to the manifold of conventional algorithms for unconstrained optimization of smooth functions, such as gradient descent, conjugate gradient and Newton's method, and so in order to use them we need to construct a smooth approximation of the objective. Since we are dealing with Grassmannian manifolds in a complex-valued ambient space, we need to clarify that when we speak of a smooth approximation, we mean an approximation that has continuous derivatives with respect to the real and imaginary parts of the variable; cf. [40]. In the following discussions, we will focus on the Fubini-Study distance and the chordal distance as they often arise in limited feedback systems. That said, the principles that underlie the suggested procedures can be applied to any valid subspace distance or smoothed version thereof.

As a first step in the proposed approach we define $\tilde{d}_{\mathrm{FS}}(\mathbf{F}_i, \mathbf{F}_j) = |\det(\mathbf{F}_j^H \mathbf{F}_i)|$ and $\tilde{d}_{\mathrm{ch}}(\mathbf{F}_i, \mathbf{F}_j) = \|\mathbf{F}_j^H \mathbf{F}_i\|_F^2$. With these metrics, the problem in (3) can be rewritten as:

$$\min_{\{\mathbf{F}_i\}, \mathbf{F}_i \in \mathbb{G}_{M_t,M}} \max_{i \neq j} \tilde{d}(\mathbf{F}_i, \mathbf{F}_j). \tag{9}$$

In the case of the chordal distance, $\tilde{d}_{\mathrm{ch}}(\mathbf{F}_i, \mathbf{F}_j)$ is smooth and we can move directly to dealing with the non-smoothness of the $\max$ operation in (9). In the Fubini-Study case, the magnitude operator in $\tilde{d}_{\mathrm{FS}}(\mathbf{F}_i, \mathbf{F}_j)$ is not smooth, and hence we seek a smooth approximation. By employing some of the existing smooth approximations for the magnitude operator (e.g., ([41], Section 5.4.5), [42]), and by using $\mathbf{X}$ to denote $\mathbf{F}_j^H \mathbf{F}_i$, we obtain the following approximations of $|\det(\mathbf{X})|$:

$$\sqrt{\mu^2 + \det(\mathbf{X}^H \mathbf{X})} - \mu, \tag{10}$$

$$\mu \log\left(2\cosh\left(\sqrt{\det(\mathbf{X}^H \mathbf{X})}/\mu\right)\right), \tag{11}$$

$$\frac{2}{\pi}\sqrt{\det(\mathbf{X}^H \mathbf{X})}\, \tan^{-1}\left(\mu\sqrt{\det(\mathbf{X}^H \mathbf{X})}\right), \tag{12}$$

where μ is an appropriate constant, and for a non-negative real number x the derivative of $\sqrt{x}$ at $x = 0$ is defined as the limit of that derivative as x approaches zero from above. Other approximations include the Huber penalty function (e.g., ([43], p. 299)) and

$$(1 + \det(\mathbf{X}^H \mathbf{X}))\log(1 + \det(\mathbf{X}^H \mathbf{X})). \tag{13}$$

The choice among these approximations does not affect the principles of the proposed approach, and hence we will simply let $\hat{\tilde{d}}_{\mathrm{FS}}(\cdot,\cdot)$ denote a chosen smooth approximation of $\tilde{d}_{\mathrm{FS}}(\cdot,\cdot)$. For notational simplicity, we will also set $\hat{\tilde{d}}_{\mathrm{ch}}(\cdot,\cdot) = \tilde{d}_{\mathrm{ch}}(\cdot,\cdot)$. Having said that, the different shapes of the approximations do have an impact in terms of the convergence behavior of the proposed design approach. For instance, the approximation in (13) is quite flat around the origin and quite steep away from the origin. As a result, in our numerical experiments, the approximation in (13) tended to generate good solutions in a few iterations, but the refinement of those good solutions tended to be slow. The approximation in (11) has a reasonably consistent slope and in our experiments, it tended to refine good codebooks more quickly than the approximation in (13). However in the absence of a good initialization that approximation can become numerically unwieldy. Since our goal is to develop techniques that yield good codebooks with a modest effort, we will use the approximation in (13) in our designs, but techniques that switch between approximations can be implemented in a straightforward manner.

There are number of ways in which the expression $\max_{i\neq j} \hat{\tilde{d}}(\mathbf{F}_i, \mathbf{F}_j)$ in (9) can be smoothly approximated. One is to use the approximation $\max\{a, b\} \approx \log(e^a + e^b)$. (The reverse approximation is similar to the "max-log" approximation that is often used in soft decoding algorithms.) This approximation was successfully used in [22] for the construction of Grassmannian constellations for non-coherent MIMO communication and is also applicable here. An alternative is to use the approximation $\max_{i\neq j} \hat{\tilde{d}}(\mathbf{F}_i, \mathbf{F}_j) \approx \mu \log(\frac{1}{N(N-1)}\sum_i \sum_{j\neq i} \cosh(\hat{\tilde{d}}(\mathbf{F}_i, \mathbf{F}_j)/\mu))$, for some $\mu > 0$, [42]. As in the case of the distance metric, the proposed approach can incorporate a number of different smooth approximations, and in this paper we will employ approximations of $\max_{i\neq j} \hat{\tilde{d}}(\mathbf{F}_i, \mathbf{F}_j)$ of the form

$$J_1(\{\mathbf{F}_i\}) = \left(\sum_{i\neq j} \hat{\tilde{d}}(\mathbf{F}_i, \mathbf{F}_j)^{\beta}\right)^{1/\beta}, \tag{14}$$

where $\beta \geq 1$. That is, we approximate the ∞-norm of the vector of pair-wise metrics by its β-norm. (For some distances, the construction and optimization of $J_1(\cdot)$ can be simplified by restricting β to be even.)

Even though the problem

$$\min_{\{\mathbf{F}_i\},\, \mathbf{F}_i \in \mathbb{G}_{M_t,M}} J_1(\{\mathbf{F}_i\}) \tag{15}$$

is smooth, it remains non-convex, due, in part, to the nature of the manifold. Nevertheless, a straightforward approach for using (15) to obtain good solutions to (9), and to the original problem in (3), would be to select a value for β and an initial codebook, and seek a locally optimal solution by applying an adaptation of a conventional smooth optimization technique, such as gradient descent, conjugate gradient or Newton's method, to the manifold [29], [30]; see also [44]. The suggested approach enhances that approach by solving a sequence of problems with increasing values of β. Once a good codebook for the current problem has been found, that codebook is used is used to initialize the next problem. This sequential procedure takes advantage of the better conditioning of $J_1(\cdot)$ for smaller values of β, and the improved approximation for larger values of β.

In particular, our basic procedure is as follows:

Basic Sequential Optimization Procedure

1) Choose the initial value of β, denoted β_0, and the increment δ_β in β at each iteration.
2) Set $\beta = \beta_0$ and randomly select an initial codebook of matrices with orthonormal columns.
3) Starting from the codebook obtained in the previous iteration, obtain a good solution to (15) using an algorithm for smooth unconstrained optimization on the manifold [29], [30].
4) Evaluate the quality of the codebook against a known bound (if any), and/or evaluate the progress of the algorithm in terms of the increase in the minimum distance.
5) Terminate if desired, else $\beta \leftarrow \beta + \delta_\beta$ and return to step 3.

In our implementations we chose $\beta_0 = 2$, $\delta_\beta = 1$ or 2, and in step 3 we employed the gradient-based algorithm in [30]. Our numerical experiments have provided good empirical evidence for monotonic convergence of the procedure, but the question of whether convergence is guaranteed remains open.

There are several ways in which the cost function $J_1(\cdot)$ can be modified to improve the basic procedure. In the case that an upper bound on the minimum distance is known, such as the Rankin bound on the chordal distance, we can replace $J_1(\cdot)$ by

$$J_2(\{\mathbf{F}_i\}) = \left(\sum_{i\neq j} \left(\hat{\tilde{d}}(\mathbf{F}_i, \mathbf{F}_j) - \alpha\right)^{\beta}\right)^{1/\beta}, \tag{16}$$

where the constant α is set to the known bound. For the case of the Fubini-Study distance, and other metrics for which non-trivial bounds are not known, the value of the constant α can be adapted to the outcome of the previous iteration of the sequential procedure. For example, the value of α can be set to the average of the pair-wise distances obtained at the previous iteration. Our numerical experiments have suggested that employing $J_2(\cdot)$ with α set to an known bound or adapted to the previous iteration, tends to improve the convergence properties

TABLE I
MINIMUM FUBINI-STUDY DISTANCES OF UNCONSTRAINED CODEBOOKS

$M_t \times M$	N	Our Codebook	Codebook in [31]
4×2	4	1.5708	1.2451
4×2	8	1.3444	1.0414
4×2	16	1.2309	0.8654
4×2	64	0.9613	0.6059
6×3	8	1.5594	N/A
6×3	16	1.5261	1.1936
6×3	32	1.4187	1.0724
6×3	64	1.3514	0.9722
8×2	8	1.5707	N/A
8×2	16	1.5414	N/A
8×2	32	1.4738	1.3153
8×2	64	1.4084	N/A

and the quality of the generated packings. This idea is similar to idea presented in [14] for constructing packings in 4-dimensional space.

An important property of the Grassmannian manifold that simplifies the design process for codebooks of different ranks is the isomorphism between the Grassmannian manifolds $\mathbb{G}_{M_t,M}$ and $\mathbb{G}_{M_t,M_t-M}$. Accordingly, an optimal codebook of rank M can be used to synthesize an optimal codebook of rank $M_t - M$ as follows: Given a codebook $\mathcal{F} = \{\mathbf{F}_i\}_{i=1}^N, \mathbf{F}_i \in \mathbb{G}_{M_t,M}$ that is optimal for a given distance metric

1) For $i = 1, 2, \ldots N$, find a matrix $\mathbf{K}_i$ whose columns form an orthonormal basis for the null space of the codeword $\mathbf{F}_i$.
2) Each $\mathbf{K}_i$ is an element of $\mathbb{G}_{M_t,M_t-M}$, and the codebook $\bar{\mathcal{F}} = \{\mathbf{K}_i\}_{i=1}^N$ is optimal in the same distance metric sense.

One conclusion that can be inferred from this property is that all the optimal codebooks on $\mathbb{G}_{3,M}$ are equivalent under any distance metric. That is, an optimal codebook on $\mathbb{G}_{3,1}$, which is the solution of a line packing problem, can be used to generate an optimal codebook on $\mathbb{G}_{3,2}$. Hence optimal chordal distance and Fubini Study distance based designs on $\mathbb{G}_{3,2}$ will result in codebooks with the same minimum distance. More generally, this conclusion can be applied to codebook pairs on $\mathbb{G}_{M_t,M}$ and $\mathbb{G}_{M_t,M_t-M}$.

In order to assess the basic design procedure, in Table I we compare the minimum Fubini-Study distances of some of the codebooks that we have designed to those of the corresponding codebooks in [31]. That comparison shows that our codebooks have significantly larger minimum distances than those in [31]. Indeed, in the case of the 4×2 codebook of size 4, our basic procedure achieves the upper bound on the Fubini-Study distance of $\pi/2$; cf. (5). The results reported in the first two rows of Table I can also be compared to the results for codebooks designed using the method in [17]; see Table VI in [17]. In those cases, our basic procedure produces codebooks with the same minimum distances as in [17]. However, as evidenced by Table VIII in Appendix I, our basic procedure has enabled us to design codebooks for systems with larger dimensions. We have also used the basic procedure to design codebooks for the chordal distance, and in that case the minimum distances of our codebooks are essentially the same as those in [17]. The codebooks whose distances are reported in this paper, and codebooks that we have designed for a number of other dimensions and sizes are available online at: http://www.ece.mcmaster.ca/~davidson/pubs/Flexible_codebook_design.html

V. LINE PACKINGS WITH (ESSENTIALLY) CONSTANT MODULUS

In scenarios in which at most one data stream is sent to each user, the generic codebook design problem in (3) reduces to the line packing problem of finding a codebook of vectors $\mathbf{f}_i \in \mathbb{G}_{M_t,1}$, that are maximally separated with respect to the distance metric $d_{\text{line}}(\mathbf{f}_i, \mathbf{f}_j) = (1 - |\mathbf{f}_j^H \mathbf{f}_i|^2)^{1/2}$. As in the previous section, the principle of sequential smooth approximation of the problem in (9) can be invoked in a number of different ways. For reasons analogous to those that led to the selection of the approximation in (13), one way that we have found to be particularly effective is to choose $\tilde{\hat{d}}_{\text{line}}(\mathbf{f}_i, \mathbf{f}_j) = (1 + |\mathbf{f}_j^H \mathbf{f}_i|^2) \log(1 + |\mathbf{f}_j^H \mathbf{f}_i|^2)$, and to apply the basic sequential design procedure to the problem of minimizing $J_2(\{\mathbf{f}_i\})$ in (16), where α is set to the value of the corresponding Rankin bound. That procedure produces codebooks that exhibit essentially the same distance properties as the best of the existing codebooks; e.g., [17]. In this section we leverage the basic sequential procedure to obtain codebooks with similar distance properties and the additional property that the elements of the codewords have (essentially) constant modulus. In feedback schemes for point-to-point links, this property ensures that the power transmitted from each antenna is (almost) the same and hence mitigates the effects of power imbalances at the transmitter.

For the line packing case, the constant modulus constraint is that the ℓth element of each vector $\mathbf{f}_i$ has modulus $1/\sqrt{M_t}$; i.e., $|[\mathbf{f}_k]_\ell| = 1/\sqrt{M_t}$. However, our basic design procedure is based on unconstrained optimization on the manifold. In order to use that procedure we define the smooth penalty term

$$P_{\text{cm}}(\{\mathbf{f}_i\}) = \left(\sum_{k,\ell} \left(|[\mathbf{f}_k]_\ell|^2 - \frac{1}{M_t} \right)^\beta \right)^{1/\beta}, \qquad (17)$$

where $\beta \geq 1$, and apply the basic sequential optimization procedure with the cost function

$$J_3(\{\mathbf{f}_i\}) = w_1 J_2(\{\mathbf{f}_i\}) + w_2 P_{\text{cm}}(\{\mathbf{f}_i\}) \qquad (18)$$

for appropriately chosen weights w_1 and w_2. Since we have only two weights, a variety of line search strategies (e.g., [41]) can be used to select an appropriate value for the ratio of the weights. Perhaps the simplest strategy would be to augment the basic sequential procedure with an additional loop in which w_2 is initially set to zero and then increased sequentially until a codebook whose elements are sufficiently close to being constant modulus is obtained. As in the basic procedure, the optimization at the current iteration would be initialized by the codebook obtained from the previous iteration.

Table II compares the minimum distances of the unconstrained and constant-modulus Grassmannian line packings that we have obtained with those of the unconstrained line packings presented in [17]. For the unconstrained case, our technique generates line packings with approximately the same minimum distances as those reported in [17]. In the constrained case, the constructed packings have elements with almost constant modulus, and have minimum distances that are very close to the unconstrained case.

TABLE II
MINIMUM DISTANCES OF UNCONSTRAINED AND CONSTANT MODULUS RANK-1 CODEBOOKS, ALONG WITH THE RANKIN BOUND. FOR THE CONSTANT MODULUS PACKINGS $\max_{k,\ell} \|[\mathbf{f}_k]_\ell - 1/M_t|^2| \leq 0.01/M_t$

M_t	N	Unconstrained Our Codebook	Unconstrained Codebook in [17]	Unconstrained Rankin Bound	Const. Mod. Our Codebook
4	5	0.9682	0.9682	0.9682	0.9680
4	6	0.9448	0.9448	0.9486	0.9441
4	7	0.9350	0.9353	0.9354	0.9188
4	8	0.9255	0.9257	0.9258	0.9114
4	9	0.9155	0.9150	0.9186	0.8969
4	10	0.9115	0.9114	0.9129	0.8977
4	16	0.8943	0.8943	0.8943	0.8656
4	20	0.8659	0.8458	0.8660	0.8240
5	6	0.9798	0.9798	0.9798	0.9797
5	7	0.9638	0.9637	0.9661	0.9522
5	8	0.9553	0.9553	0.9562	0.9444
5	9	0.9466	0.9468	0.9487	0.9350
5	10	0.9427	0.9427	0.9428	0.9354
5	16	0.9190	0.9183	0.9238	0.9145

VI. LINE PACKINGS WITH PSK ALPHABET

In this section, we seek line packings with codeword elements that are not only of constant modulus, but are also from a finite alphabet, such as a scaled phase-shift keying (PSK) alphabet, $\mathcal{A} = 1/\sqrt{M_t}\,\{e^{j(\phi_0+2\pi(\ell-1)/L)}\}_{\ell=1}^{L}$. This property significantly reduces the storage requirements of the codebook and may reduce the computational effort required to select the codeword at the receiver. In particular, in the case of the scaled 4-PSK alphabet, $1/\sqrt{M_t}\,\{1,-1,j,-j\}$, the scaling can be absorbed and the complex multiplications that are inherent in the selection process are reduced to sign changes and swaps of the real and imaginary parts.

The restriction to a defined alphabet $\mathcal{A}$ offers the possibility to design codebooks by exhaustively evaluating each admissible codebook. However, taking into account the rotational invariance of the codewords into account, there are $\binom{|\mathcal{A}|^{M_t-1}}{N}$ admissible codebooks and even for modestly sized codebooks the computational cost of this approach exceeds the computational resources that could reasonably be applied to the problem. In this section we describe two simple approaches to obtaining good codebooks with elements from a (scaled) PSK alphabet. In the first approach, the method of the previous section is used to generate partial codebooks with (scaled) PSK elements and good properties. These partial codebooks are then completed by exhaustive search, but that exhaustive search is typically over a much smaller number of codewords than the size of the codebook. In the second approach, we employ an incremental construction in which the size of the codebook is doubled at each step. The new codewords are Hadamard products of the existing vectors and a single vector that is obtained using a smooth optimization and rounding procedure similar to that used in the first approach. This incremental construction was motivated, in part, by some of the principles used in the generation of mutually unbiased bases (MUB) with elements from the (scaled) 4-PSK alphabet [26]–[28].

A. Relax-Round-Expurgate-Replace Approach

The first of the proposed approaches begins with the application of a variant of the method in Section V to design an initial codebook of codewords with constant modulus elements. Those elements are then rounded to the given alphabet. The "bad" codewords in this rounded codebook are expurgated and are replaced through an exhaustive search procedure. The selection of the number of codewords to be expurgated, and hence the size of the subsequent search is based on the notion of a satisfactory codebook.

Given that we are considering Grassmannian line packings, one way in which the quality of a finite alphabet codebook could be assessed would be to compare the achieved minimum distance to the Rankin bound. That approach can be refined by observing that the distance spectrum of a finite alphabet codebook is discrete and the set of distance values depends on the alphabet being used. A refined upper bound on the minimum distance is the largest of these discrete values that is no larger than the Rankin bound. As an example, the set of possible distances between pairs of codewords of a codebook of size 8 that resides on $\mathbb{G}_{4,1}$ and whose elements are selected from the (scaled) 4-PSK alphabet can be shown to be $\{1, 0.9354, 0.8660, 0.7071, 0.6123, 0\}$. The Rankin bound for a codebook of size 8 on $\mathbb{G}_{4,1}$ is 0.92582. Comparing the Rankin bound with the set of admissible distances, it can be deduced that the minimum distance of the (scaled) 4-PSK codebook is bounded above by 0.8660. A satisfactory codebook would achieve a large fraction of this "quantized" Rankin bound.

A more detailed description of the proposed approach is as follows: We begin by choosing the first codeword of the codebook to be a randomly generated vector with elements from the scaled PSK alphabet. Then we relax the finite alphabet constraint on the remaining $N-1$ codewords and use the procedure in Section V to design those $(N-1)$ codewords so that we obtain a good codebook of size N with (essentially) constant modulus elements. The elements of the $(N-1)$ designed codewords in that codebook are then rounded to the nearest point in the alphabet. This rounding process could be done in a more sophisticated way, but element-wise rounding of each codeword is simple to implement and has sufficed in our numerical examples. The quantized codebook is then analyzed to determine whether its minimum distance achieves a sufficiently large fraction of the quantized Rankin bound to be deemed satisfactory. If not, the codeword that appears in the largest number of pairwise distances that are deemed unsatisfactory is removed, and the resulting codebook is analyzed again. This process is repeated until a satisfactory codebook is found. If $\bar{N}$ denotes the number of codewords that are expurgated in this way, the satisfactory partial codebook, which is of size $N-\bar{N}$, is completed by adding $\bar{N}$ codewords via an exhaustive search. Typically, $\bar{N} \ll N$ and this reduced-dimension exhaustive search is often viable. In cases in which $\bar{N}$ is deemed to be too large to attempt an exhaustive search, the basic relaxation and rounding aspects of the proposed approach can be recapitulated, but with the existing $N-\bar{N}$ codewords being fixed, rather than the single codeword that was fixed in the first stage.

We used the above procedure to generate 16 lines in the 4 dimensional complex space using the (scaled) 4-PSK alphabet. The generated packings have a minimum distance of 0.8660, while the corresponding codebook reported in [23] has a minimum distance of 0.7071. The codebooks generated incrementally using the notion of mutually unbiased bases [26], [27] also achieve a minimum distance of 0.8660. As noted above, the quantized Rankin bound in this case is 0.8660. Hence the codebooks that we obtained, and those in [26], [27], are essentially optimal.

B. Incremental Construction

The good performance of the incremental construction based on mutually unbiased bases (MUBs) [26]–[28] in the previous example suggests that the MUB construction might also be of interest in other scenarios. While that is the case to some degree, in its raw form the incremental MUB construction is only applicable to cases in which the number of transmitter antennas, M_t, is a power of two. In this section, we exploit our earlier work to develop an incremental construction that is based on some of the ideas that underlie the MUB construction, but is more flexible in that it can be applied to systems with an arbitrary number of antennas. Furthermore, at each step the proposed approach doubles the size of the codebook, whereas the MUB construction only adds M_t codewords. Since the computational cost of each step is typically lower in the proposed approach, this facilitates the construction of sizeable codebooks with elements from (scaled) PSK alphabets.

In order to place the proposed approach in context, we first briefly describe the existing MUB-based construction [26], which can be applied to systems for which M_t is a power of two: To design a codebook of size $N = (\check{n}+1)M_t$ in $\check{n}$ steps,

1) Let $\mathbf{A}^{(0)}$ to be the normalized Hadamard matrix of size $M_t \times M_t$, and let $\mathbf{a}_i^{(0)}$ denote its ith column. Set $n = 0$.
2) While $n < \check{n}$, increment n and find a length-M_t vector $\mathbf{u}^{(n)}$ with elements $\mathcal{A} = 1/\sqrt{M_t}\,\{\pm 1, \pm j\}$ such that all the elements of the vectors ${\mathbf{A}^{(n-1)}}^H\mathbf{u}^{(n)}$, ${\mathbf{A}^{(n-2)}}^H\mathbf{u}^{(n)}, \ldots, {\mathbf{A}^{(0)}}^H\mathbf{u}^{(n)}$ lie in $\mathcal{A}$.
3) Construct $\mathbf{A}^{(n)} = \sqrt{M_t}\,[\mathbf{u}^{(n)} \odot \mathbf{a}_1^{(0)}, \mathbf{u}^{(n)} \odot \mathbf{a}_2^{(0)}, \ldots, \mathbf{u}^{(n)} \odot \mathbf{a}_{M_t}^{(0)}]$, where $\odot$ denotes the element-wize (Hadamard) product, and return to 2.
4) The desired codebook is the columns of the matrix $[\mathbf{A}^{(0)}, \mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(\check{n})}]$.

In conventional implementations, the vector $\mathbf{u}^{(n)}$ in step 2) is found using an exhaustive search.

In the proposed approach, we design codebook of size $N = 2^{\check{q}}N_0$ in $\check{q}$ steps, where N_0 is the number of codewords in the initial codebook. To do so, the above steps are modified in a number of ways:

- The proposed approach is initialized by a matrix $\mathbf{B}^{(0)}$ of size $M_t \times N_0$ with elements from the scaled PSK alphabet whose columns yield a satisfactory codebook; cf. Section VI.A. In our implementations, when M_t is a power of two we have chosen $\mathbf{B}^{(0)}$ to be the normalized Hadamard matrix of the appropriate size, in which case, $N_0 = M_t$. In other cases we have chosen $N_0 = 1$ and $\mathbf{B}^{(0)} = (1/\sqrt{M_t})\,\mathbf{1}$, where $\mathbf{1}$ denotes the length-M_t vector with all elements equal to 1.
- Similar to step 2) in the existing method, at each iteration, starting from $q = 1$, the proposed approach seeks a vector $\mathbf{v}^{(q)}$ with elements in $\mathcal{A}$ such that the vector ${\mathbf{B}^{(q-1)}}^H\mathbf{v}^{(q)}$ has all its elements in $\mathcal{A}$. We do so using a variant of the procedure in Section VI.A in which we first relax the finite alphabet constraint and apply the basic sequential optimization procedure in Section IV to
$$J_4(\mathbf{v}) = \omega_1 P_{\mathrm{cm}}\left({\mathbf{B}^{(q-1)}}^H\mathbf{v}\right) + \omega_2 P_{\mathrm{cm}}(\mathbf{v}), \tag{19}$$
where $P_{\mathrm{cm}}(\,\cdot\,)$ was defined in (17). Once a good solution to that problem has been obtained, the elements of the resulting $\mathbf{v}$ are rounded to the nearest member of the alphabet to form a candidate for $\mathbf{v}^{(q)}$. That candidate is then used to generate a candidate for the incremented codebook, as described below.
- Distinct from the existing method, the candidate vector $\mathbf{v}^{(q)}$ is used to generate the vectors $\mathbf{w}_i = \sqrt{M_t}\mathbf{v}^{(q)} \odot \mathbf{b}_i^{(q-1)}$, $i = 1, 2, \ldots, 2^{(q-1)}N_0$, where $\mathbf{b}_i^{(q-1)}$ is the ith column of $\mathbf{B}^{(q-1)}$. A candidate codebook consisting of the union of these $2^{(q-1)}N_0$ new vectors and the existing codebook is evaluated. If it is deemed to be satisfactory (cf. Section VI.A), the candidate codebook becomes the incremented codebook and is represented by the columns of the matrix $\mathbf{B}^{(q)} = [\mathbf{B}^{(q-1)}, \mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{2^{(q-1)}N_0}]$. Otherwise, we seek a new candidate for $\mathbf{v}^{(q)}$. One can be found by picking a vector with elements from $\mathcal{A}$ that lies in the neighborhood of the existing candidate, or by returning to the previous step and optimizing the expression in (19) from a different starting point. (Obviously, the search for new candidates for $\mathbf{v}^{(q)}$ must be controlled so that all possibilities are eventually explored.)

In applications in which the space required to store the codewords is limited, incremental constructions such as the MUB-based construction and that proposed above have the advantage that when $\mathbf{A}^{(0)}$, respectively $\mathbf{B}^{(0)}$, can be generated deterministically, one need only store the sequence of vectors $\mathbf{u}^{(n)}$, $n = 1, 2, \ldots, \check{n}$, respectively $\mathbf{v}^{(q)}$, $q = 1, 2, \ldots, \check{q}$. For a codebook of size $N = (\check{n}+1)M_t$, the MUB approach needs to store $\check{n}$ vectors, whereas for a codebook of size $N = 2^{\check{q}}N_0$, the proposed approach need only store $\check{q}$ vectors. Furthermore, since the size of the codebook we obtain doubles at each iteration, and since our relax-round approach typically finds an appropriate $\mathbf{v}^{(q)}$ quite quickly, for large codebooks our approach had a significant computational advantage in our numerical experiments. As such we have been able to design significantly larger codebooks than those that have been previously obtained using the MUB approach.

Table III presents the minimum distances of some (scaled) 4-PSK line packings designed with the proposed incremental approach. Many of these packings are optimal, in the sense that they achieve the "quantized" Rankin bound. In the case of $M_t = 4$ and $N = 32$ the designed codebook does not achieve the quantized Rankin bound, but it can be shown to be optimal in the minimum distance sense via an exhaustive search. In the other three cases in Table III in which the designed codebook does not achieve the "quantized" Rankin bound, the optimality, or otherwise, of the obtained codebooks remains to be determined.

Interestingly, when M_t is a power of two, the codewords in the packings generated by our technique can be arranged into sets of mutually unbiased bases. In particular, the union of our packing with $M_t = 16$ and $N = 256$ and an identity matrix generates a complete set of mutually unbiased bases for dimension 16; a result that may be of independent interest.

TABLE III
MINIMUM DISTANCES OF RANK-1 CODEBOOKS WITH (SCALED) 4-PSK ALPHABET, ALONG WITH THE RANKIN BOUND FOR CONSTANT MODULUS CODEBOOKS AND AN INDICATION OF WHETHER OUR CODEBOOK IS KNOWN TO BE OPTIMAL

M_t	N	Our PSK Codebook	Rankin Bound	Optimal?
4	8	0.8660	0.9258	Yes
4	16	0.8660	0.8660	Yes
4	32	0.7071	0.8660	Yes
5	8	0.8944	0.9562	Yes
5	16	0.8944	0.9238	Yes
5	32	0.8000	0.8944	?
6	8	0.9428	0.976	Yes
6	16	0.9428	0.9428	Yes
6	32	0.8819	0.9129	Yes
7	8	0.9897	0.9897	Yes
7	16	0.9035	0.9562	?
7	32	0.9035	0.9406	Yes
8	16	0.9354	0.9661	Yes
8	32	0.9354	0.9504	Yes
8	64	0.9354	0.9354	Yes
16	32	0.9682	0.9837	?
16	64	0.9682	0.9759	Yes
16	128	0.9682	0.9720	Yes
16	256	0.9682	0.9682	Yes

TABLE IV
MINIMUM DISTANCES ACHIEVED BY DESIGNED NESTED CODEBOOKS WITH FUBINI-STUDY DISTANCE AS THE RANK-2 DISTANCE. THE RANKIN BOUND OF THE RANK-1 CASE AND THE MINIMUM FUBINI-STUDY DISTANCE OF OUR UNSTRUCTURED RANK-2 CODEBOOKS ARE PROVIDED FOR COMPARISON

$M_t \times M$	N	Rank 1		Rank 2, Fubini-Study dist.	
		Our Nested	Rankin Bound	Our Nested	Our Unstructured
4×2	8	0.9118	0.9258	1.3069	1.3444
4×2	16	0.8594	0.8944	1.1659	1.2309
4×2	32	0.7687	0.8660	0.9689	1.1159
6×2	8	0.9740	0.9759	1.5245	1.5584
6×2	16	0.9358	0.9428	1.3855	1.4812
6×2	32	0.8980	0.9275	1.3258	1.3636
8×2	8	1	1	1.5282	1.5642
8×2	16	0.9629	0.9661	1.4829	1.5414
8×2	32	0.9388	0.9504	1.4203	1.4738

TABLE V
MINIMUM DISTANCES ACHIEVED BY DESIGNED NESTED CODEBOOKS WITH CHORDAL DISTANCE AS THE RANK-2 DISTANCE. THE RANKIN BOUND OF THE RANK-1 CASE AND THE MINIMUM CHORDAL DISTANCE OF THE UNSTRUCTURED RANK-2 CODEBOOKS REPORTED IN [17] ARE PROVIDED FOR COMPARISON

$M_t \times M$	N	Rank 1		Rank 2, chordal dist.	
		Our Nested	Rankin Bound	Our Nested	Unstructured in [17]
4×2	8	0.9242	0.9258	1.0519	1.0690
4×2	16	0.8660	0.8944	1	1.0323
4×2	32	0.7512	0.8660	0.8925	N/A
6×2	8	0.9744	0.9759	1.2311	1.2344
6×2	16	0.9372	0.9428	1.1563	1.1925
6×2	32	0.8895	0.9275	1.1073	N/A
8×2	8	1	1	1.3093	N/A
8×2	16	0.9645	0.9661	1.2352	N/A
8×2	32	0.9372	0.9504	1.2024	N/A

VII. SUBSPACE PACKINGS WITH NESTED STRUCTURE

In a number of scenarios, substantial performance gains can be made by enabling "multi-mode" precoding [25], in which the number of data streams to be transmitted to a given receiver is adapted to the state of the channel, in addition to the directions in which those streams are sent. To facilitate the implementation of a multi-mode system, it is desirable that the codebook have a nested structure, in which a codebook of higher rank contains codebooks of lower rank. For example, a nested rank-2 codebook, $\mathcal{F}_2 = \{\mathbf{F}_i\}_{i=1}^N = \{[\mathbf{f}_i, \check{\mathbf{f}}_i]\}_{i=1}^N$ contains the rank-1 codebook $\mathcal{F}_1 = \{\mathbf{f}_i\}_{i=1}^N$. The design of a nested codebook intrinsically involves making trade-offs between the quality of the codebook of each rank. To capture that trade-off without adding constraints, we formulate the design of a nested codebook using a weighted sum of metrics. For the case of rank-2 nested codebooks, that design problem can be written as

$$\max_{\{\mathbf{F}_i\},\, \mathbf{F}_i \in \mathbb{G}_{M_t,2}} \min_{i \neq j} \omega_2 d_2(\mathbf{F}_i, \mathbf{F}_j) + \omega_1 d_1(\mathbf{f}_i, \mathbf{f}_j), \tag{20}$$

where $\mathbf{f}_i$ denotes the first column of $\mathbf{F}_i$, $d_m(\cdot,\cdot)$ is any valid subspace distance on $\mathbb{G}_{M_t,m}$, and ω_m is the weight assigned to the distance metric for the codebook of rank m. (In practice, the distance metrics for the rank-1 and rank-2 codebooks would likely be chosen to be of the same type, but the approach can also handle metrics of different types.) The extension to nested codebooks of higher rank is conceptually straightforward by somewhat awkward from a notational perspective. By defining smooth functions $\hat{\tilde{d}}(\mathbf{F}_i, \mathbf{F}_j)$ and $\hat{\tilde{d}}_{\text{line}}(\mathbf{f}_i, \mathbf{f}_j)$ as in Section IV, good solutions to (20) can be obtained by applying the basic sequential optimization procedure. In particular, if we let the dimensions of the argument of $J_2(\,\cdot\,)$ in (16) implicitly define the appropriate metric and if we once again let $\mathbf{f}_i$ denote the first column of $\mathbf{F}_i$, then the basic sequential optimization procedure can be applied to

$$J_5(\{\mathbf{F}_i\}) = \omega_2 J_2(\{\mathbf{F}_i\}) + \omega_1 J_2(\{\mathbf{f}_i\}), \tag{21}$$

where $J_2(\{\mathbf{F}_i\}) = (\sum_{i\neq j}(\hat{\tilde{d}}(\mathbf{F}_i, \mathbf{F}_j) - \alpha_2)^\beta)^{1/\beta}$, $J_2(\{\mathbf{f}_i\}) = (\sum_{i\neq j}(\hat{\tilde{d}}_{\text{line}}(\mathbf{f}_i, \mathbf{f}_j) - \alpha_1)^\beta)^{1/\beta}$, with α_2 and α_1 set or adapted as described in Section IV and the Grassmannian optimization steps are over $\mathbb{G}_{M_t,2}$. For simplicity, we have used the same notation for the weights in (20) and (21), but they may be adjusted if desired.

Table IV presents some results from the application of this approach to the design of nested codebooks of rank two in which the Fubini Study distance is chosen as the subspace metric. The weights ω_1 and ω_2 were chosen so as to achieve a balance between the properties of the rank-1 and rank-2 codebooks. The third and fourth columns of the table contain the minimum distance achieved by the rank-1 component of the nested codebook and the (unconstrained) Rankin bound for rank-1 codebooks, respectively. The fifth column contains the minimum Fubini study distance of the designed nested codebook, and the sixth column contains the best minimum Fubini-Study distance that we have obtained for unstructured rank-2 codebooks. (Some of those distances appear in Table I.) Table V presents corresponding results for nested codebooks with the chordal distance as the subspace metric. In that table, the comparator for the quality of the rank-2 codebook is best of the corresponding minimum distances reported in [17], where corresponding results are available.

VIII. Subspace Packings With Nested Structure and PSK Alphabet

In this section we develop an effective design technique for codebooks that not only have the nested structure, but also have codeword elements from a (scaled) PSK alphabet. One approach to developing such a technique would be to tackle the problem directly, by applying discretization techniques to the method in the previous section, perhaps employing an appropriate constant-modulus penalty along the way. However, such techniques can become quite unwieldy as the rank of the codebook and the number of antennas grow. As an alternative, in this section we leverage the line packing technique developed in Section VI.A to develop a layered method in which the codebook is designed one rank at a time, in a greedy fashion. The greedy nature of this method is reasonable, because in many wireless systems lower rank codebooks will be employed more often.

To simplify the exposition of the proposed technique, we will, once again, focus on the case of nested codebooks of rank 2. We begin with a rank-1 codebook $\{\mathbf{f}_i\}_{i=1}^N$ with codeword elements from the (scaled) PSK alphabet, possibly designed using one of the techniques in Section VI. We then seek vectors $\{\check{\mathbf{f}}_i\}_{i=1}^N$ also with elements from the PSK alphabet so that $\mathcal{F} = \{[\mathbf{f}_i, \check{\mathbf{f}}_i]\}_{i=1}^N$ forms a good nested rank-2 codebook with elements from the alphabet.

The proposed technique for designing the vectors $\{\check{\mathbf{f}}_i\}_{i=1}^N$ is based on the principles of the relax-round-expurgate-replace approach in Section VI.A. To employ those principles, we need to capture the desired properties of $\{\check{\mathbf{f}}_i\}_{i=1}^N$ in the absence of the PSK alphabet constraint. To that end we observe that first we must ensure that the matrices $[\mathbf{f}_i, \check{\mathbf{f}}_i]$ have orthonormal columns. Second, the rank-2 codebook should have a large minimum distance in the chosen subspace distance metric. Third, to facilitate the rounding procedure, the relaxed codewords should have elements of almost constant modulus. We capture these properties in the smooth objective

$$J_6(\{\check{\mathbf{f}}_i\}) = \omega_1 \left(\sum_i \left| \mathbf{f}_i^H \check{\mathbf{f}}_i \right|^\beta \right)^{1/\beta} + \omega_2 J_2(\{[\mathbf{f}_i, \check{\mathbf{f}}_i]\}) + \omega_3 P_{\mathrm{cm}}(\{\check{\mathbf{f}}_i\}), \quad (22)$$

where β is constrained to be even and} the subspace metric is implicit in $J_2(\cdot)$. Then we apply the basic sequential optimization procedure using this objective. Next, we apply the principles of the approach in Section VI.A, and round the elements of the obtained vectors to the nearest point in the alphabet. Any rounded $\check{\mathbf{f}}_i$ that is not orthogonal to $\mathbf{f}_i$ is immediately expurgated, and the remaining set of rank-2 codewords, $\{[\mathbf{f}_i, \check{\mathbf{f}}_i]\}$ is examined. Any $\check{\mathbf{f}}_i$ that induces a distance, or distances, in the rank-2 codebook that is deemed unsatisfactory is also expurgated. The expurgated vectors can be replaced using an exhaustive search, or through a recapitulation of the relax-round-expurgate-replace method. However, in both cases the size of the new problem is the number of codewords to be replaced, which is typically much smaller than N.

We have summarized some of the distance results for codebooks designed using this technique in Table VI. In this case, the distance metric in the optimization of the rank-2 codebook is the chordal distance. This table shows that in spite of the nesting and alphabet constraints, our approach can find codebooks that come reasonably close to the Rankin bound at both ranks. The obtained minimum distances are also quite close to those of the otherwise unconstrained nested codebooks in Table V.

TABLE VI
Minimum Distances Achieved By Designed Nested Codebooks With (Scaled) 4-PSK Alphabet and the Chordal Distance As the Rank-2 Distance. The Rankin Bounds for Constant Modulus Codebooks for Each Rank Are Provided for Comparison

$M_t \times M$	N	Rank 1		Rank 2, chordal dist.	
		Our Nested PSK	Rankin Bound	Our Nested PSK	Rankin Bound
4×2	8	0.8660	0.9258	1	1.0690
4×2	16	0.8660	0.8660	1	1
4×2	32	0.7071	0.8660	0.8660	1.0000
6×2	8	0.9428	0.9759	1.0541	1.2344
6×2	16	0.9428	0.9428	1	1.1926
6×2	32	0.8819	0.9129	1	1.1547
8×2	8	1	1	1.1456	1.3093
8×2	16	0.9354	0.9661	1.0897	1.2649
8×2	32	0.9354	0.9504	1.0607	1.2443

IX. Performance Evaluation

We now evaluate the performance of communication systems based on the codebooks obtained using the suggested approaches in the case of the MIMO downlink with zero-forcing beamforming (ZFBF) [4], per-user unitary precoding and rate control (PU^{2}RC) [32], and zero-forcing block diagonalization (ZFBD) [35] signalling under a simple Rayleigh fading channel model. In that model, the channel gains are modelled as i.i.d. circular complex Gaussian random variables with zero mean and unit variance, and each receiver is assumed to know its channel realization. The kth receiver has a codebook $\mathcal{F}_k$ that is a random rotation of the transmitter's codebook $\mathcal{F}$, [45]. (The transmitter knows the rotation.) For rank-1 codebooks used in the beamforming cases, each receiver determines the codeword $\mathbf{f}_{i^\star} \in \mathcal{F}_k$ that solves $\max_{\mathbf{f}_i \in \mathcal{F}_k} |\mathbf{h}_k \mathbf{f}_i^*|$, [4], and for the higher-rank codebooks used in the ZFBD case, the receiver determines the codeword $\mathbf{F}_{i^\star} = \arg\max_{\mathbf{F}_i \in \mathcal{F}_k} \|\mathbf{H}\mathbf{F}_i\|_F$, [35]. The index $i^\star$ is fed back to the transmitter, which collects all the quantized versions of the channel estimates from the scheduled users in order to compute the beamforming/precoding matrix according to the system architecture.

First, we consider the ZFBF case for a system with $M_t = 4$ transmitter antennas, $K = 4$ receivers and rank-1 codebooks of size $N = 16$, and 64. In Fig. 1 we have plotted the cumulative distribution (cdf) of the sum of the rates that can be achieved using ZFBF and Gaussian signalling at an average SNR of 15 dB, where if σ^2 denotes the variance of the noise at each receiver, then the SNR is $\mathrm{trace}(\mathbf{P}^H\mathbf{P})/(K\sigma^2)$. We consider codebooks designed using our basic procedure, and Love's codebooks [31]. We also provide the average performance of a set of codebooks generated randomly using the uniform distribution on the manifold. For the codebooks of size $N = 16$, our basic procedure generated a codebook with a minimum distance of 0.8943, whereas the codebook in [31] has a minimum distance of 0.8670. For the codebooks of size $N = 64$, the corresponding minimum distances are 0.7175 and 0.6035, respectively. Fig. 1 demonstrates that the improved minimum distances of our codebooks generate improved sum rate statistics.

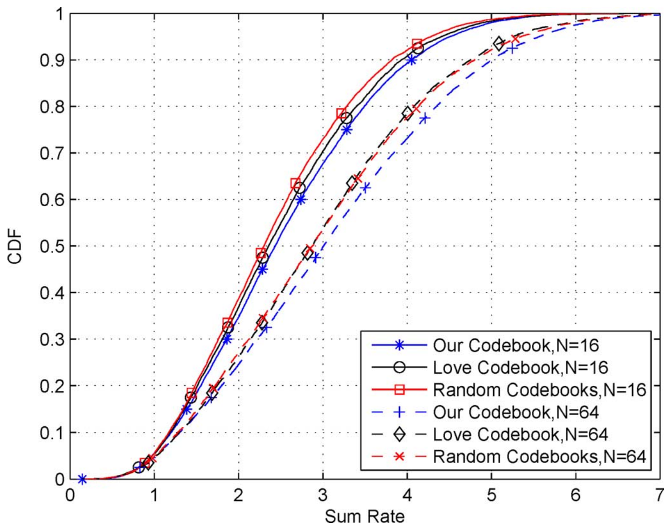

Fig. 1. CDF of the sum rate for a ZFBF system with various unconstrained codebooks.

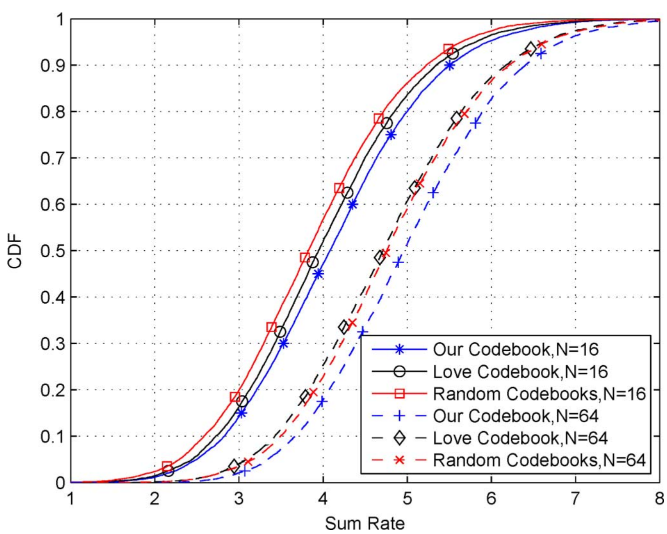

Fig. 2. CDF of the sum rate for a $\mathrm{PU^2RC}$ system with various unconstrained codebooks.

For a given target rate, our codebooks yield a lower outage probability, and for a given target outage probability, our codebooks yield a larger rate. In Fig. 2 we plot the corresponding results for $\mathrm{PU^2RC}$ signalling. In that case, the tangible performance advantages of our codebooks extend over a broader range of rates.

Fig. 3 examines the performance of rank-1 codebooks of size $N = 16$ and 32 with elements from the (scaled) 4-PSK defined alphabet in the $\mathrm{PU^2RC}$ scheme at an SNR of 15 dB. For the codebooks of size $N = 16$, our incremental construction generated a 4-PSK codebook with a minimum distance of 0.8660 (cf. Table III), whereas the unconstrained codebook has a minimum distance of 0.8943. For the codebooks of size $N = 32$, the corresponding distances are 0.7071 and 0.8115. As suggested by their good distance properties, our 4-PSK codebooks provide performance that is close to that of the unconstrained codebooks and better than the average performance of randomly generated codebooks (with unconstrained coefficients).

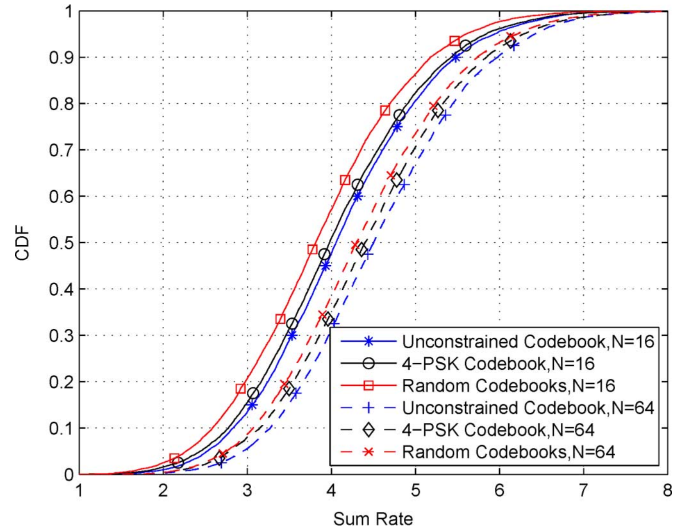

Fig. 3. CDF of the sum rate for a $\mathrm{PU^2RC}$ system with unconstrained, 4-PSK and random codebooks.

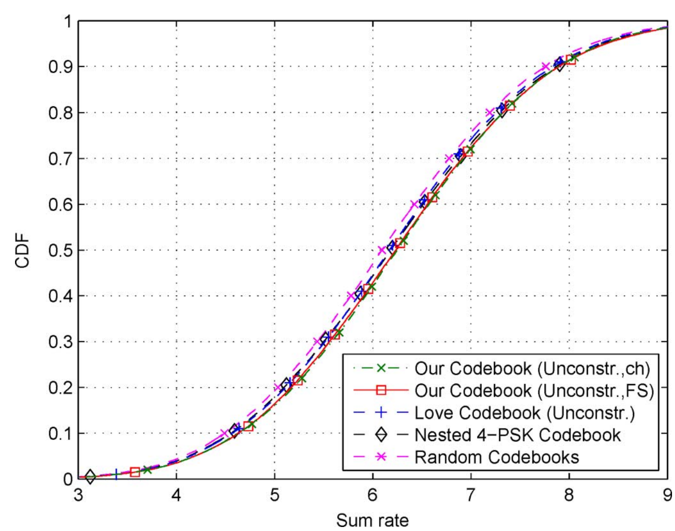

Fig. 4. CDF of the sum rate for a rank-2 ZFBD system with unconstrained, and nested 4-PSK codebooks.

Finally, we examine the performance of the proposed approach to the design of nested codebooks with PSK elements in the context of a ZFBD system with $M_t = 4$ transmit antennas, 2 users, and rank-2 codebooks of size $N = 16$. The CDF of the sum rate at an SNR of 15 dB for a nested codebook with 4-PSK elements designed using the approach in Section VIII with the chordal distance is plotted in Fig. 4. To provide benchmarks against which the performance can be assessed, we have also plotted the CDF for the average of a set of randomly generated unconstrained codebooks, for Love's codebook [31], and for unconstrained codebooks designed using the technique in Section IV with the Fubini-Study and the chordal distances. For reference, the minimum distances of the designed codebooks are provided in Table VII. The key feature of Fig. 4 is that in this setting, the memory and computational savings associated with the nested 4-PSK codebook that we have designed, and the convenience of the

TABLE VII
DISTANCE PROPERTIES OF THE RANK-2 CODEBOOKS USED IN FIG. 4

Type	Design Metric	Min. Ch. dist.	Min. FS dist.
Nested 4-PSK	Chordal dist.	1	1.0472
Unconstr. [31]	—	0.8188	0.8654
Unconstr.	Chordal dist.	1.0309	1.0863
Unconstr.	FS dist.	0.9428	1.2309

nested structure, are obtained without incurring a substantial reduction in performance. Fig. 4 also shows that both the unconstrained codebooks designed using the proposed technique provide improved performance over the codebook in [31]. In a similar setting with $M_t = 8$, 2 users and rank-4 codebooks (not shown here) the unconstrained codebook that was obtained using the proposed method with the chordal distance provided marginally better performance than the codebook that was obtained using the Fubini-Study distance.

X. CONCLUSION

In this paper we have developed a flexible approach to the design of structured Grassmannian codebooks (packings) for communication systems that employ limited feedback. The proposed approach is based on (otherwise) unconstrained optimization over the Grassmannian manifold of a sequence of smooth objective functions. In addition to generating unstructured subspace codebooks with better minimum Fubini-Study distances than some existing codebooks, the proposed approach was also used to generate structured codebooks with other desirable properties. In particular, one variant of the approach was used to generate rank-2 codebooks with a nested structure and elements from the (scaled) 4-PSK alphabet. An outcome of another variant was the complete set of mutually unbiased bases of dimension 16; a result that may be of independent interest. In the unstructured case, the designed codebooks were shown to provide tangible performance gains when applied to a simple multiple antenna downlink communication system with limited feedback. In was also shown that the structured codebooks obtained using the proposed approaches do not incur a substantial degradation in performance.

The design of unstructured Grassmannian codebooks is widely acknowledged as a difficult problem, and the kinds of structure that we have imposed on the structured codebooks do not make that problem easier. As a result, our focus has been on the development of sound heuristics that guide us towards good codebooks. Inherently, the development of heuristics involves choices, and the choices that we have made herein could certainly be debated. However, our numerical results have shown that with the choices that we have made we have been able to obtain codebooks of considerable size that possess the desirable structural properties and yet have a minimum distance that comes close to the bound on the minimum distance of unstructured codebooks. Furthermore, the basic principles of the approach that we have proposed can accommodate, in a straightforward way, a myriad of other choices for the smooth objectives used to capture the desirable properties of the codebook. In the interests of facilitating the refinement of these functions, we have made the codebooks that we have designed freely available online at: http://www.ece.mcmaster.ca/~davidson/pubs/Flexible_codebook_design.html.

In closing, we draw attention to a point that was implicit in step 3 of the basic sequential optimization procedure in Section IV, namely that the proposed approach can be extended in a straightforward way to certain other manifolds, including the Stiefel manifold. The extension to that manifold may be of interest in some related applications of limited feedback in wireless communications; e.g., [46].

TABLE VIII
MINIMUM FUBINI-STUDY DISTANCES OF UNCONSTRAINED CODEBOOKS (CONTINUED FROM TABLE I)

$M_t \times M$	N	Our Codebook
4×3	8	1.1820
4×3	16	1.1070
4×3	32	0.8805
4×3	64	0.7413
6×2	8	1.5691
6×2	16	1.4812
6×2	32	1.3636
6×2	64	1.2986
8×3	8	1.5652
8×3	16	1.5420
8×3	32	1.5080
8×3	64	1.4628
8×4	8	1.5679
8×4	16	1.5410
8×4	32	1.5109
8×4	64	1.4703
10×2	8	1.5707
10×2	16	1.5468
10×2	32	1.5277
10×2	64	1.4708
10×3	8	1.5706
10×3	16	1.5624
10×3	32	1.5230
10×3	64	1.5002
12×2	8	1.5708
12×2	16	1.5600
12×2	32	1.5247
12×2	64	1.4945
16×2	8	1.5708
16×2	16	1.5684
16×2	32	1.5254
16×2	64	1.4979
16×3	8	1.5708
16×3	16	1.5704
16×3	32	1.5580
16×3	64	1.5394
16×4	8	1.5708
16×4	16	1.5708
16×4	32	1.5652
16×4	64	1.5546

APPENDIX A
FURTHER RESULTS FOR THE MINIMUM FUBINI-STUDY DISTANCE OF UNCONSTRAINED CODEBOOKS

In Table VIII we provide the minimum Fubini-Study distances of unconstrained codebooks that we have designed for dimensions for which there is no corresponding codebook in [31].

REFERENCES

[1] G. Caire and S. Shamai, "On the achievable throughput of a multi-antenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[2] J. Lee and N. Jindal, "High SNR analysis for MIMO broadcast channels: Dirty paper coding versus linear precoding," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4787–4792, Dec. 2007.

[3] S. A. Jafar and A. J. Goldsmith, "Isotropic fading vector broadcast channels: The scalar upper bound and loss in degrees of freedom," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 848–857, Mar. 2005.

[4] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[5] D. J. Love, R. W. Heath, Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[6] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, June 2010.

[7] A. Gersho and R. M. Gray*, Vector Quantization and Signal Compression*. Norwell, MA, USA: Kluwer, 1991.

[8] D. J. Love, R. W. Heath, Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.

[9] D. J. Love and R. W. Heath, Jr., "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2967–2976, Oct. 2005.

[10] D. J. Love and R. W. Heath, Jr., "Limited feedback unitary precoding for orthogonal space-time block codes," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 64–73, Jan. 2005.

[11] J. C. Roh and B. D. Rao, "Design and analysis of MIMO spatial multiplexing systems with quantized feedback," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 2874–2886, Aug. 2006.

[12] M. B. Shenouda and T. N. Davidson, "A design framework for limited feedback MIMO systems with zero-forcing DFE," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1578–1587, Oct. 2008.

[13] W. Dai, Y. Liu, and B. Rider, "Quantization bounds on Grassmann manifolds and applications to MIMO communications," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1108–1123, 2008.

[14] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Exper. Math.*, vol. 5, pp. 139–159, 1996.

[15] R. A. Pitaval, H. L. Määttänen, K. Schober, O. Tirkkonen, and R. Wichman, "Beamforming codebooks for two transmit antenna systems based on optimum Grassmannian packings," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6591–6602, Oct. 2011.

[16] P. Xia and G. Giannakis, "Design and analysis of transmit-beamforming based on limited-rate feedback," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1853–1863, May 2006.

[17] I. S. Dhillon, R. W. Heath, T. Strohmer, and J. A. Tropp, "Constructing packings in Grassmannian manifolds via alternating projection," *Exper. Math*, vol. 17, no. 1, pp. 9–35, 2008.

[18] K. Schober, P. Jnis, and R. Wichman, "Geodesical codebook design for precoded MIMO systems," *IEEE Commun. Lett.*, vol. 13, no. 10, pp. 773–775, 2009.

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

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

[21] I. Kammoun, A. M. Cipriano, and J.-C. Belfiore, "Non-coherent codes over the Grassmannian," *IEEE Trans. Wireless Commun.*, vol. 6, no. 10, pp. 3657–3667, Oct. 2007.

[22] R. H. Gohary and T. N. Davidson, "Noncoherent MIMO communication: Grassmannian constellations and efficient detection," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1176–1205, Mar. 2009.

[23] B. Clerckx, Y. Zhou, and S. Kim, "Practical codebook design for limited feedback spatial multiplexing," in *Proc. IEEE Int. Conf. Commun.*, Beijing, China, 2008, pp. 3982–3987.

[24] D. J. Ryan, I. V. L. Clarkson, I. B. Collings, D. Guo, and M. L. Honig, "QAM and PSK codebooks for limited feedback MIMO beamforming," *IEEE Trans. Commun.*, vol. 57, pp. 1184–1196, Apr. 2009.

[25] D. J. Love and R. W. Heath, Jr., "Multimode precoding for MIMO wireless systems," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3674–3687, 2005.

[26] B. Mondal, T. A. Thomas, and M. Harrison, "Rank-independent codebook design from a quaternary alphabet," in *Conf. Rec. 41st Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, 2007, pp. 297–301.

[27] T. Inoue and R. W. Heath, Jr., "Kerdock codes for limited feedback precoded MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3711–3716, Sep. 2009.

[28] I. Tselniker, M. Nazarathy, and M. Orenstein, "Mutually unbiased bases in 4, 8, and 16 dimensions generated by means of controlled-phase gates with application to entangled-photon QKD protocols," *IEEE J. Sel. Topics Quantum Electron.*, vol. 15, no. 6, pp. 1713–1723, Nov./Dec. 2009.

[29] A. Edelman, T. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, 1998.

[30] J. H. Manton, "Optimization algorithms exploiting unitary constraints," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 635–650, Mar. 2002.

[31] D. J. Love, Grassmannian Subspace Packing webpage [Online]. Available: https://engineering.purdue.edu/~djlove/grass.html [Online]:

[32] J. S. Kim, H. Kim, C. S. Park, and K. B. Lee, "On the performance of multiuser MIMO systems in WCDMA/HSDPA: Beamforming, feedback and user diversity," *IEICE Trans. Commun.*, vol. E89-B, no. 8, pp. 2161–2169, Aug. 2006.

[33] L.-U. Choi and R. D. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 20–24, Jan. 2004.

[34] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[35] N. Ravindran and N. Jindal, "Limited feedback-based block diagonalization for the MIMO broadcast channel," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1473–1482, Oct. 2008.

[36] D. Tse and P. Viswanath*, Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[37] A. Barg and D. Y. Nogin, "Bounds on packings of spheres in the Grassmann manifold," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2450–2454, Sept. 2002.

[38] O. Henkel, "Sphere-packing and bounds in the Grassmann and Stiefel manifolds," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3445–3456, Oct. 2005.

[39] R.-A. Pitaval, O. Tirkkonen, and S. Blostein, "Low complexity MIMO precoding codebooks from Orthoplex packings," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2011, pp. 1–5.

[40] K. Kreutz-Delgado, "The complex gradient operator and the $\mathbb{CR}$-calculus," Electr. and Comput. Eng. Dept., Univ. of California, San Diego, CA, USA, Tech. Rep., Jun. 2009 [Online]. Available: http://arxiv.org/abs/0906.4835

[41] D. P. Bertsekas*, Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 1999.

[42] Y. Nesterov, "Smooth minimization of non-smooth functions," *Math. Program.*, vol. 103, pp. 127–152, 2005.

[43] S. Boyd and L. Vandenberghe*, Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.

[44] R. H. Gohary and T. N. Davidson, "An explicit expression for the Newton direction on the complex Grassmann manifold," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1303–1309, 2011.

[45] P. Ding, D. J. Love, and M. D. Zoltowski, "Multiple antenna broadcast channels with shape feedback and limited feedback," *IEEE Trans. Signal Process.*, vol. 55, no. 7–1, pp. 3417–3428, 2007.

[46] R. Pitaval and O. Tirkkonen, "Incorporating Stiefel geometry in codebook design and selection for improved base station cooperation," presented at the 75th IEEE Veh. Technol. Conf., Yokohama, Japan, May 2012.

Ahmed Medra received the B.Sc. degree (with honors.) in electrical engineering, in 2005, and the M.Sc. degree, in 2010, both from Alexandria University, Egypt. He is currently working toward the Ph.D. degree in electrical engineering from McMaster University, Canada. He eceived the Philomathia Foundation Fellowship in 2011 and the McMaster International Excellence Award in 2012. His research interest includes MIMO communications, Grassmannian codebook design, linear precoding, and Interference Alignment.

Timothy N. Davidson (M'96) received the B.Eng. (Hons. I) degree in electronic engineering from the University of Western Australia (UWA), Perth, in 1991 and the D.Phil. degree in engineering science from the University of Oxford, U.K., in 1995.

He is currently a Professor in the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he holds the (Tier II) Canada Research Chair in Communication Systems and is currently serving as Chair of the Department. His research interests lie in the general areas of communications, signal processing, and control.

Dr. Davidson received the 1991 J. A. Wood Memorial Prize from UWA, the 1991 Rhodes Scholarship for Western Australia, and a 2011 Best Paper Award from the IEEE Signal Processing Society. He has served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II. He has also served as a Guest Co-Editor of issues of the *IEEE Journal on Selected Areas in Communications*, the *IEEE Journal of Selected Topics in Signal Processing*, and *EURASIP Journal on Advances in Signal Processing*. Dr. Davidson is currently serving as the Chair of the IEEE Signal Processing Society's Technical Committee on Signal Processing for Communications and Networking, as a General Co-Chair for the 2014 IEEE International Workshop on Signal Processing Advances in Wireless Communications, and a Technical Program Co-Chair for the 2nd IEEE Global Conference on Signal and Information Processing. He is a Registered Professional Engineer in the Province of Ontario.