Non-coherent Multi-layer Constellations for Unequal Error Protection

Kareem M. Attiah*, Karim Seddik†, Ramy H. Gohary‡ and Halim Yanikomeroglu‡
*Department of Electrical Engineering, Alexandria University, Alexandria 21544, Egypt
†Electronics Engineering Department, American University in Cairo, AUC Avenue, New Cairo 11835, Egypt
‡Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada

***Abstract*—In this paper, we consider the design of multi-resolution non-coherent multiple-input multiple-output (MIMO) systems that enable Unequal Error Protection (UEP). A method for designing multi-layer non-coherent Grassmannian constellations is introduced. Specifically, the proposed method yields multi-layer constellations that are amenable to a natural set partitioning strategy. The resulting subsets from such partitioning are used to encode the more protected symbols. On the other hand, the less protected symbols are mapped to points within these subsets. Furthermore, we present two methods to establish the link between the gain of the more protected layer and the design parameters of this construction. Finally, we exploit the underlying structure to develop a sequential decoding approach. Numerical results suggest that employing such decoding scheme leads to computational savings with respect to the optimal decoding while maintaining comparable performance.**

I. Introduction

MIMO communication systems are multi-antenna systems in which data streams are transmitted simultaneously over a common channel. For these systems, two modes of operations can be identified. In a coherent MIMO, the information is communicated over a channel whose fading coefficients are known at the receiver, whereas in non-coherent MIMO [1], [2] it is assumed that channel coefficients are unavailable at both the transmitter and the receiver. Despite the losses in achievable rates relative to the coherent case, non-coherent signaling is particularly attractive in situations where estimating the channel coefficients is either inefficient or infeasible.

In a standard communication setting, all transmitted bits of information are considered to be equally valuable. This symmetric view encourages uniform error protection to segments of information against channel imperfections. Despite the benefits of Equal Error Protection (EEP) schemes, in many applications, the nature of information being conveyed over the channel motivates the need to provide a non-uniform protection against channel perturbations. For instance, multi-resolution systems enable the receiver to reconstruct transmitted information at different levels of precision depending on the state of the channel. If the channel conditions are convenient, the receiver will be able to recover all transmitted information and reconstruct the message with high accuracy. On the other hand, if the channel conditions are less favorable, it is adequate for the receiver to reconstruct a coarse view by only retrieving the most basic or important parts of the message. In the previous example, it is clear that a uniform treatment of bits of information being transmitted is inappropriate. Instead, it is more natural to support transmission of information by a scheme that allows different parts of information to be protected at varying degrees based on their level of importance [3].

UEP has been previously studied in the literature. Many existing techniques utilize coded [4] as well as uncoded [5] coherent constellations to provide more protection to certain parts of information. While there have been several efforts in the literature to design multi-resolution UEP-based coherent constellations, the authors are not aware of any such attempts in the context of non-coherent signaling.

Motivated by the above discussion, this paper investigates the potential of utilizing a non-coherent multi-layer technique to achieve UEP. This scheme was previously proposed by the authors in [6]. In particular, we exploit the underlying structure of this construction to develop a natural set partitioning scheme. In other words, the non-coherent constellation representing points on the Grassmann manifold is broken down into smaller disjoint subsets. Furthermore, the information symbols are decomposed into two segments of smaller symbols and the more protected symbols are allowed to choose one of possible subsets whereas the less protected symbols will be conveyed by points within the chosen subset. By varying the distance between subsets we demonstrate that coding gains can be realized over the EEP case for the more protected symbols. Additionally, we provide two methods of characterizing the gain of the more protected layer in terms of a specific design parameter. These characterizations simplify the problem of choosing the design parameter to yield a desired nominal gain. Finally, we make use of the multi-layer structure to simplify the decoding task by utilizing a sequential decoding scheme. Numerical simulations reveal that the proposed decoder exhibits similar or comparable performance to the exhaustive optimal decoder while demanding much less computational effort than that of the optimal decoder.

II. Channel Model and Preliminaries

A. Channel Model

Consider the block Rayleigh fading model in [1]. This model assumes constant channel coefficients during channel coherence interval of T symbols duration, which then take on a statistically-independent realization in each of the subsequent coherence intervals. Using M and N to denote the numbers of transmit and receive antennas, respectively, the $T \times N$ received signal $\mathbf{Y}$ at a given coherence interval can be expressed as

$$\mathbf{Y} = \mathbf{X}\mathbf{H} + \sqrt{M/\gamma T}\,\mathbf{W}, \tag{1}$$

where $\mathbf{X}$ is the complex $T \times M$ transmitted matrix, $\mathbf{H}$ is the $M \times N$ fading matrix whose coefficients are independent, identically distributed (i.i.d.) drawn from the standard complex Gaussian distribution $\mathcal{CN}(0,1)$ and are assumed to be unknown at both the transmitter and the receiver, $\mathbf{W}$ is the $T \times N$ noise matrix which is also modeled as i.i.d. complex Gaussian whose elements are $\mathcal{CN}(0,1)$ and γ is the SNR.

Under the assumptions that $N \geq M$ and $T \geq N+M$ and by constraining the average transmitted power over all M transmitted antennas to be kept constant, it was shown in [2] that the asymptotic (in the SNR) capacity achieving signaling scheme is comprised of unitary signals ($\mathbf{X}_i^\dagger \mathbf{X}_i = \mathbf{I}_M$) that take on the form of isotropically distributed points on the Grassmann manifold $\mathbb{G}_{T,M}(\mathbb{C})$, i.e., the set of M dimensional subspaces in $\mathbb{C}^T$. In this paper, we only consider signaling constellations of this class, that is, unitary signals representing distinct points on $\mathbb{G}_{T,M}(\mathbb{C})$. In what follows, we do not differentiate between Grassmannian points and the unitary transmitted matrices $\mathbf{X}_i$.

B. The Optimal Detector

At the receiver, the optimal scheme for detecting the transmitted matrix is referred to as the Generalized Likelihood Ratio Test (GLRT) [7]

$$\hat{\mathbf{X}} = \arg\max_{\mathbf{X}\in\mathcal{C}} \mathrm{Tr}(\mathbf{Y}^\dagger \mathbf{X}\mathbf{X}^\dagger \mathbf{Y}), \tag{2}$$

which is shown in [7] to have the same performance as the Maximum Likelihood (ML) detector in i.i.d. fading environment and unitary signaling.

C. Pairwise Error Probability

For any two distinct transmitted matrices $\mathbf{X}_i, \mathbf{X}_j$ and for high SNR, the asymptotic Pairwise Error Probability (PEP) of the GLRT is approximated by [7]

$$P_{ij} \approx \frac{(\gamma T)^{-MN} M^{MN} \binom{2MN-1}{MN}}{\left|\mathbf{R}_{ii} - \mathbf{R}_{ij}\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji}\right|^N},$$

where

$$\begin{bmatrix} \mathbf{R}_{ii} & \mathbf{R}_{ij} \\ \mathbf{R}_{ji} & \mathbf{R}_{jj} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_i^\dagger \\ \mathbf{X}_j^\dagger \end{bmatrix} \begin{bmatrix} \mathbf{X}_i & \mathbf{X}_j \end{bmatrix}.$$

Under the assumption that the transmitted matrices are unitary, the asymptotic expression for the PEP becomes [8]

$$P_{ij} \approx (d_{cp}(\mathbf{X}_i,\mathbf{X}_j)\gamma T)^{-MN} M^{MN} \binom{2MN-1}{MN}, \tag{3}$$

where

$$d_{cp}(\mathbf{X}_i,\mathbf{X}_j) := \left(\prod_{k=1}^{M} \left(1-\sigma_k^2\right)\right)^{\frac{1}{M}},$$

where $\{\sigma_k\}_{1\leq k\leq M}$ are the singular values of $\mathbf{X}_i^\dagger \mathbf{X}_j$. Following along the lines in [8], we refer to this quantity as the product chordal distance. This is, in fact, a slight inaccuracy since $d_{cp}(\mathbf{X}_i,\mathbf{X}_j)$ is not a true distance.

A closely-related metric is the chordal distance [9]

$$d_c^2(\mathbf{X}_i,\mathbf{X}_j) = M - \|\mathbf{X}_i^\dagger \mathbf{X}_j\|_F^2.$$

Both metrics may be employed to aid in the design of Grassmannian constellations (e.g., see [10] for the chordal distance). Unlike the product chordal distance however, the chordal distance is a true distance. This handy feature makes it more appealing (and intuitive) to guide constellations design than the product chordal distance. Our multi-layer construction described in Section III is also based on the chordal distance.

Denoting the principal angles between the subspaces spanned by the columns of $\mathbf{X}_i$ and $\mathbf{X}_j$ by $\{\theta_k\}_{1\leq k\leq M}$, where $\theta_k \in \left[0, \frac{\pi}{2}\right]$, alternative expressions for both metrics can be obtained

$$d_{cp}(\mathbf{X}_i,\mathbf{X}_j) = \left(\prod_{k=1}^{M} \sin^2\theta_k\right)^{\frac{1}{M}}, \tag{4}$$

and

$$d_c^2(\mathbf{X}_i,\mathbf{X}_j) = \sum_{k=1}^{M} \sin^2\theta_k. \tag{5}$$

We find expressions (4) and (5) quite revealing. In particular, the connection between the two metrics can now be well-understood in terms of the Arithmetic Mean - Geometric Mean (AM-GM) inequality

$$d_{cp}(\mathbf{X}_i,\mathbf{X}_j) \leq \frac{1}{M} d_c^2(\mathbf{X}_i,\mathbf{X}_j). \tag{6}$$

With this inequality in mind, constellations designed by packing spheres on $\mathbb{G}_{T,M}(\mathbb{C})$ using the chordal distance will necessarily be suboptimal in the sense of minimizing the asymptotic union bound on the probability of error of the GLRT. Both of these metrics play an important role in this paper. Quite often we omit the distinction and simply speak of "distance" when either the context is clear or the statement applies to both measures.

Analogous to the coherent case, the coding gain[1] of a non-coherent constellation can be readily deduced from (3) and is essentially equivalent to $\min\limits_{i,j,i\neq j} d_{cp}(\mathbf{X}_i,\mathbf{X}_j)$ [8].

III. Multi-layer Construction of Non-coherent Constellations

In [6], we propose a multi-layer approach to the design of Grassmannian constellations. In this section we summarize this strategy. Let L be the number of constellation layers and let $\xi_1,\ldots,\xi_L$ be these layers. We insist that the i^{th} layer ξ_i for $2 \leq i \leq L$ be constructed from $\xi_{1\leq j\leq i-1}$. Particularly, to find ξ_i, first find the i^{th} *parent* layer $\mathcal{C}_i = \bigcup_{j=1}^{i-1} \xi_j$. Then, for each *parent* point $\mathbf{X}_p \in \mathcal{C}_i$, find K *children* points by transitioning on $\mathbb{G}_{T,M}(\mathbb{C})$ along K geodesics emanating from $\mathbf{X}_p$. Finally, the newly found children points of all $\mathbf{X}_p \in \mathcal{C}_i$ collectively define the layer ξ_i.

Mathematically speaking, a geodesic on $\mathbb{G}_{T,M}(\mathbb{C})$ initially at some point $\mathbf{Z} = \mathbf{Z}(0)$ extending along some direction

[1] Although the chordal distance does not guarantee full diversity, we argue that when the constellation size is not very large the principal angles between the associated subspaces will take on nonzero values and hence $\min\limits_{i,j} d_{cp}(\mathbf{X}_i,\mathbf{X}_j) > 0$.

$\dot{\mathbf{Z}}(0) = \mathbf{Z}_\perp \mathbf{B}$ for some $(T-M)\times M$ complex matrix $\mathbf{B}$ is given by [11]

$$\mathbf{Z}(t) = \begin{bmatrix}\mathbf{Z} & \mathbf{Z}_\perp\end{bmatrix} \begin{bmatrix}\mathbf{V}\cos\mathbf{\Sigma}t \\ \mathbf{U}\sin\mathbf{\Sigma}t\end{bmatrix}, \tag{7}$$

where $\mathbf{B} = \mathbf{U\Sigma V}^\dagger$ is the compact SVD of $\mathbf{B}$ and t is a step parameter. Using (7), we can now formally define ξ_i

$$\xi_i = \left\{ \begin{bmatrix}\mathbf{X}_p & \mathbf{X}_{p\perp}\end{bmatrix} \begin{bmatrix}\mathbf{V}_k\cos\mathbf{\Sigma}_k t_i \\ \mathbf{U}_k\sin\mathbf{\Sigma}_k t_i\end{bmatrix} \middle| \forall \mathbf{X}_p \in \mathcal{C}_i, \forall k = 1,\ldots,K \right\}.$$

Observe that the above definition sets no restrictions on the directions defining children geodesics. We can impose some restrictions on this structure by further demanding that the pairwise chordal distances among children points associated with a common parent point be maximized. By doing so, we finally arrive at (cf. [6])

$$\xi_i = \left\{ \begin{bmatrix}\mathbf{X}_p & \mathbf{X}_{p\perp}\end{bmatrix} \begin{bmatrix}\alpha_i\tilde{\mathbf{V}}_k \\ \beta_i\mathbf{U}_i\end{bmatrix} \middle| \forall \mathbf{X}_p \in \mathcal{C}_i, \forall k = 1,\ldots,K \right\} \tag{8}$$

where $K \leq 4M^2$, $\alpha_i = \cos t_i$ and $\beta_i = \sin t_i$. The matrix $\mathbf{U}_i$ is an arbitrary $(T-M)\times M$ complex unitary matrix for which the natural choice is $\mathbf{I}_{T-M,M}$. However, as indicated by the subscript, it may also be possible to alter its value on a layer to layer basis in order to maximize the distances among children points of different parents. Typically, varying $\mathbf{U}_i$ amounts to altering the orientation of the children geodesics around the parent point without disrupting the distances among children points (of a common parent) as well as their distances from that parent. Finally, the K, $M\times M$, complex unitary matrices $\{\tilde{\mathbf{V}}_k\}_{1\leq k\leq K}$ are found by solving the optimization problem

$$\max_{\{\mathbf{V}_k^\dagger\mathbf{V}_k=\mathbf{I}_M\}_{1\leq k\leq K}} \min_{i,j,i\neq j} \|\mathbf{V}_i - \mathbf{V}_j\|_F^2.$$

The initial layer ξ_1 is of small size N' and is typically constructed by packing spheres on $\mathbb{G}_{T,M}(\mathbb{C})$ using the chordal distance as a packing metric. A number of good numerical algorithms have been proposed in the literature to attain this packing. One example would be [12] which is proven to be reliable when the packing size is small. It was found that this algorithm is appropriate for typical sizes of ξ_1. Finally, the overall non-coherent constellation is formed by taking the union of these L layers. Alternatively, it can be regarded as the $(L+1)^{th}$ parent layer $\mathcal{C} = \mathcal{C}_{L+1} = \{\mathbf{X}|\mathbf{X} \in \bigcup_{i=1}^L \xi_i\}$. The size of this constellation is $|\mathcal{C}| = N'(K+1)^{L-1}$.

Before we move on to the next section, it is insightful to draw a number of remarks. First, since the step parameter t_i is essentially the same for all members of the layer ξ_i, they all lie at an identical distance from their respective parents. To show this, let $\mathbf{X}_p$ be some parent point in $\mathcal{C}_i$ and let $\mathbf{X}_c$ be its child in ξ_i, the expression for the chordal distance between $\mathbf{X}_p$ and $\mathbf{X}_c$ can be readily derived using expressions (5) and (8)

$$d_c^2(\mathbf{X}_p, \mathbf{X}_c) = M\beta_i^2, \tag{9}$$

which is independent of both $\mathbf{X}_p$ and $\mathbf{X}_c$.

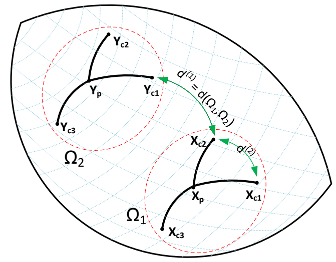

Fig. 1. This figure provides a pictorial visualization of the parent layer points X_p and Y_p, children layer points $X_{c1}, X_{c2}, X_{c3}, Y_{c1}, Y_{c2}$ and Y_{c3} and the set partitioning scheme. The minimum inter-subset distance, as well as the constellation minimum distance, is also shown.

Second, the evident hierarchical structure of this construction allows us to conveniently define a natural set partitioning scheme. To elaborate, the i^{th} level partition, $1 \leq i \leq L$, of $\mathcal{C}$ consists of the $|\mathcal{C}_{i+1}| = N'(K+1)^{i-1}$ subsets (i.e. subconstellations) $\Omega_1^{(i)}, \ldots, \Omega_{|\mathcal{C}_{i+1}|}^{(i)}$, where the k^{th} subset $\Omega_k^{(i)}$ contains exactly one element from $\mathcal{C}_{i+1}$ as well as all of its *descendants* (and not just its children) in $\xi_{i+1}, \ldots, \xi_L$. For instance, the 1^{st} level partition includes N' subsets $\Omega_1^{(1)}, \ldots, \Omega_{N'}^{(1)}$, where $\Omega_k^{(1)}$ contains a single element from $\mathcal{C}_2 = \xi_1$ and the *descendants* (and not only the children) of that element in $\xi_2, \ldots, \xi_L$. Similarly, the subsets $\Omega_1^{(2)}, \ldots, \Omega_{N'(K+1)}^{(2)}$ comprising the 2^{nd} level partition, each contains a single element from $\xi_1 \cup \xi_2$ and its descendants in $\xi_3, \ldots, \xi_L$.

Finally, since the step parameters control the intra-subset distances, we expect that inter-subset distances are also influenced by the same parameters. Guided by this observation, we conclude that with careful selection of $\{t_i\}_{2\leq i\leq L}$ it can be ensured that a certain distance criterion (be that local or global) holds. However, the lack of explicit formulas such as (9) in the global case makes it more challenging to accurately estimate these parameters.

IV. Unequal Error Protection

Consider a MIMO non-coherent communication system in which the source transmits one of possible $\mathcal{N}$ symbols from the set $\mathcal{S} = \{s_1, \ldots, s_\mathcal{N}\}$ by mapping distinct elements of this set onto distinct constellation points on the Grassmannian. For this system, assume that $\mathcal{N} = 2^m$ so that the rate $r = \frac{m}{T}$ bits per channel use (bpcu). Although the more general case can be treated in a similar manner, we focus on the case where it is desired that the data sequence be broken down into two segments to be transmitted with different reliability. Let f be the fraction of the transmitted bits such that $m_1 = fm$ is the length of the sequence of the more important bits to be reliably transmitted and $m_2 = m - m_1$ is the length of the sequence of the less important bits. Define the sets $\mathcal{S}^{(1)} = \{s_1^{(1)}, \ldots, s_{N'}^{(1)}\}$ and $\mathcal{S}^{(2)} = \{s_1^{(2)}, \ldots, s_{(K+1)}^{(2)}\}$ where $N' = 2^{m_1}$ and $K = 2^{m_2} - 1$. Moreover, define the bijective function $\mathcal{F}: \mathcal{S} \to \mathcal{S}^{(1)} \times \mathcal{S}^{(2)}$ that maps source symbols onto distinct pairs of symbols $\left(s_i^{(1)}, s_j^{(2)}\right)$. Clearly, $\mathcal{N} = N'(K+1)$ and $\mathcal{S}^{(1)}$ now includes the more important data symbols whereas $\mathcal{S}^{(2)}$ incorporates the less important data symbols.

Now consider the signaling scheme described in Section III with $L = 2$. Clearly, the overall constellation is comprised of two layers ξ_1 and ξ_2 which will be referred to as the parent and children layers, respectively. Furthermore, the set of step parameters is now reduced down to a single step parameter that will be called t. Now let us induce the level one partition described in the previous section and let $\Omega_1, \ldots, \Omega_{N'}$ be the resulting subsets. One way to achieve UEP is to map the information symbols to constellation points by letting data symbols in $\mathcal{S}^{(1)}$ choose one of the possible subsets while letting data symbols in $\mathcal{S}^{(2)}$ be conveyed by points within the subset (cf. [4] for the coherent case). Fig. 1 illustrates the concept.

Define the inter-subset distance (in the product chordal sense)

$$d_{cp}(\Omega_i, \Omega_j) = \min_{\mathbf{X}\in\Omega_i, \mathbf{Y}\in\Omega_j} d_{cp}(\mathbf{X}, \mathbf{Y}).$$

Let us use $d^{(1)}$ and $d^{(2)}$ to denote the minimum inter-subset distance and the minimum constellation distance, then

$$d^{(1)} = \min_{i,j,i\neq j} d_{cp}(\Omega_i, \Omega_j), \tag{10}$$

and

$$d^{(2)} = \min_{\mathbf{X},\mathbf{Y}\in\mathcal{C}} d_{cp}(\mathbf{X}, \mathbf{Y}). \tag{11}$$

It is clear that $d^{(1)}$ ($d^{(2)}$) characterizes the probability of symbol error of the more (less) important data symbols. Typically, in EEP scenarios, the constellation is designed in such a way that $d^{(1)}\big|_{EEP} = d^{(2)}\big|_{EEP}$. In UEP however, we dictate that $d^{(1)}\big|_{UEP} > d^{(2)}\big|_{UEP}$. In the UEP case, it is often advantageous to measure the coding gain of the parent/children layer with respect to that of the EEP case. We define the relative gain of layer i as

$$\Gamma_i = 10\log_{10}\frac{d^{(i)}\big|_{UEP}}{d^{(i)}\big|_{EEP}}, \qquad i\in\{1,2\}. \tag{12}$$

As was previously stated in Section III, selecting t to ensure some nominal coding gain is not an easy task, due to the complicated dependence of global metrics in (10) and (11) on several other factors besides t (e.g., the directions of children geodesics and the location of parent points on the manifold). Proposed next are two methods for estimating the sine of this parameter, namely $\beta = \sin t$.

A. Geometric Mean Approximation (GMA)

In this section, a simple method for computing β in terms of a lower bound on $d^{(1)}$ is described. This method is suitable for $M \leq 2$. We first introduce the following lemmas.

Lemma 1: Let $\omega_1, \ldots, \omega_M$ be a sequence of non-negative numbers with variance σ_ω^2 and fixed mean μ, then

$$\sigma_\omega^2 \leq (M-1)\mu^2.$$

Proof: $\sigma_\omega^2 = \frac{1}{M}\sum_{i=1}^M \omega_i^2 - \mu^2 \leq \frac{1}{M}\left(\sum_{i=1}^M \omega_i\right)^2 - \mu^2 = M\mu^2 - \mu^2$. ∎

Lemma 2: Let $\mathbf{X}$ and $\mathbf{Y}$ be two children points such that $\mathbf{X}\in\Omega_i$ and $\mathbf{Y}\in\Omega_{j\neq i}$ and let $\mathbf{X}_p$ and $\mathbf{Y}_p$ be their parents. If $N' \leq T^2$, then the following lower bound holds on $d_c(\mathbf{X}, \mathbf{Y})$

$$d_c(\mathbf{X}, \mathbf{Y}) \geq \frac{M(T-M)}{T}\sqrt{\frac{N'}{N'-1}} - 2\sqrt{M}\beta. \tag{13}$$

Proof: By the triangle inequality, $d_c(\mathbf{X}_p, \mathbf{Y}_p) \leq d_c(\mathbf{X}_p, \mathbf{X}) + d_c(\mathbf{X}, \mathbf{Y}) + d_c(\mathbf{Y}, \mathbf{Y}_p)$. Rearranging terms and denoting the packing radius of the parent layer by d_{pack} we arrive at $d_c(\mathbf{X}, \mathbf{Y}) \geq d_{pack} - 2\sqrt{M}\beta$. Where the second term on the right hand side follows from (9). The proof is complete by noting that the first term in (13) is the packing distance when $N' \leq T^2$ [12]. ∎

We will use these lemmas to establish a lower bound on $d^{(1)}$. Let $\mathbf{X}\in\Omega_i$ and $\mathbf{Y}\in\Omega_{j\neq i}$ be two constellation points whose product chordal distance attains (10). Denote the squares of the sines of the principal angles between the subspaces associated with $\mathbf{X}$ and $\mathbf{Y}$ by $\{\omega_i\}_{1\leq i\leq M}$ such that $(d^{(1)})^M = \prod_{i=1}^M \omega_i$. The well-known approximation in [13] states that the difference between the AM and GM of a sequence of positive numbers can be well-approximated by one half the variance of this sequence if the numbers in the sequence are close to one. From this we deduce

$$\begin{aligned} d^{(1)} &\approx \frac{1}{M}d_c^2(\mathbf{X}, \mathbf{Y}) - \frac{\sigma_\omega^2}{2} \\ &\gtrapprox \frac{1}{M}d_c^2(\mathbf{X}, \mathbf{Y}) - \frac{M-1}{2M^2}d_c^4(\mathbf{X}, \mathbf{Y}) \end{aligned} \tag{14}$$

where in the second line, the result from Lemma 1 is used. Observe that the expression on the right hand side of (14) is increasing in the range $0 \leq d_c(\mathbf{X}, \mathbf{Y}) \leq \sqrt{\frac{M}{M-1}}$, from which we conclude $M \leq 2$. Using Lemma (2), we can further lower bound (14) in the range for which the chordal distance is increasing without violating the inequality

$$d^{(1)} \gtrapprox \frac{1}{M}(d_{pack} - 2\sqrt{M}\beta)^2 - \frac{M-1}{2M^2}(d_{pack} - 2\sqrt{M}\beta)^4 \tag{15}$$

where d_{pack} is the first term on the right hand side of (13). Equation (15) provides a remarkably simple method for selecting β: For a given minimum inter-subset distance d, find β that makes the right hand expression identical to d. The resulting constellation will have an actual minimum inter-subset distance (in the product chordal sense) that is bounded below by d. In other words, the actual coding gain Γ_1 is going to be at least as large as the nominal coding gain.

It is crucial to know the limitations associated with expression (15). Specifically, the GM approximation will not be adequately accurate when the squares of the sines of the principal angles are bounded away from one (i.e. $\mathbf{X}$ and $\mathbf{Y}$ are nearby). This typically occurs when either the overall constellation size or the parameter t (or β) is too large. Under such circumstances, the bound in (15) need not even hold.

B. Polynomial Regression (PR)

As was stated in the previous subsection, the GM approximation method is likely bound to fail under conditions where the principal angles between subspaces belonging to different partitioning subsets of the constellation are small. A different (and more reliable) method for establishing the relationship between minimum inter-subset distance and the step parameter is a one that is borrowed from statistical analysis and modeling. In essence, it is assumed that the $d^{(1)}$ is modeled as an n^{th} degree polynomial in β, where the

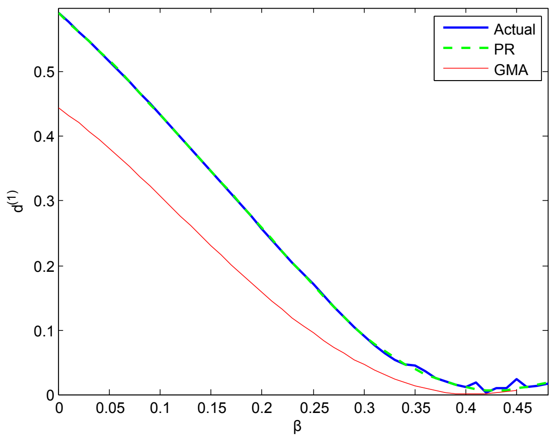

Fig. 2. Minimum inter-subset distance against β for 4×16 constellation.

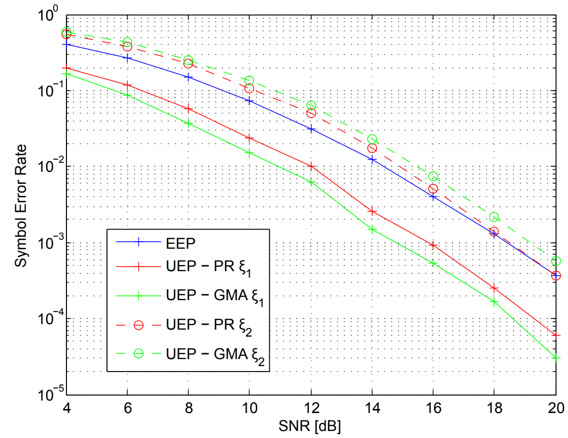

Fig. 3. SER of ξ_1 and ξ_2 when 33% of transmitted data is important

modeling coefficients are unknown parameters to be estimated. Given enough observations $l > n$, the modeling coefficients are evaluated via least-squares estimation.

Unlike GM approximation, the practicality of this method depends upon the availability of values of β for which the corresponding values of $d^{(1)}$ are known. This generally means that this scheme must be preceded by an offline step in which these values are acquired. Once this is performed, and the modeling coefficients are acquired, they can be stored at the transmitter and reused as many times as needed.

C. Examples

We now assess the utility of our multi-layer scheme. Before we introduce the examples, it is necessary to point out that (12) provides a method for computing the nominal coding gain over the EEP case. The actual code gain incurs additional losses relative to the nominal value due to a number of reasons. Typically, the actual gain is reduced by a factor that depends on the number of nearest neighbors. Moreover, the assumption made in Section II-C that the coding gain is based on product chordal distance follows from an asymptotic union bound, which is known to be loose when either the constellation size is large or the SNR is small. Having said that, the purpose of the following examples is not to show how our scheme attains the exact numeric gains set by the design parameters, but rather to demonstrate it has the potential of realizing UEP in the context of non-coherent signaling, which, to our knowledge, has never been investigated in the literature. In what follows, we present two examples. In both cases, we compare the parent layer Symbol Error Rate (SER) of the UEP scheme to that of EEP. Also, in both cases we assume $T = 4$ and $M = N = 2$. The detection is performed according to the GLRT.

1) Rate 1.5 *with* 33% *of the data is important:* In this example, the size of the multi-layer constellation is $\mathcal{N} = 64$, corresponding to a rate of 1.5 bpcu. The size of the parent layer is $N' = 4$. Therefore, the fraction of bits that are important is 33%. In Fig. 2, we plot the actual minimum inter-subset distance $d^{(1)}$ as a function against $\beta = \sin t$. It is interesting to remark that despite the relationship in Fig. 2 represents one instance of the family of constellations described in Section III, the general features of this curve are common to all instances of this family. Specifically, two main regions are identified for this curve that are common to all constellations of this family. The first region indicates consistent decrease of the inter-subset distance in β. For this particular curve, this is the region extending up to ≈ 0.4. This consistency can be intuitively explained: Initially, when β is small, all children points lie in close proximity to their parents and the regions enclosing the partitioning subsets are all disjoint. As β increases, these regions continue to expand in size thus bringing children of different subsets closer. This consistency maintains until the partitioning subsets blend in and the associated enclosing regions are no longer disjoint. In which case we enter the second region where the inter-subset distances vary uncontrollably with the increase of β. Clearly, it is of interest to only select β within first region. The GM approximation lower bound as well as the polynomial modeling ($n = 9$ and $l = 25$) are also plotted.

In Fig. 3, we plot the SER of the parent layer in the EEP case as well as UEP case when β is estimated via the methods described in Sections IV-A and IV-B. The nominal coding gain relative to the EEP case is $\Gamma_1 = 3$ dB, which corresponds to a factor of 2 in terms of the minimum inter-subset distance relative to the EEP case. As high SNR, it can be seen that the actual gain is roughly about 2.5 dB in the PR case and 3.5 dB in the GMA case. We observe a loss of 0.5 dB relative to nominal value is incurred in the PR case. In the GMA case, however, this loss is compensated since the GMA selects β in such a way that overestimates the gain.

2) Rate 2 *with* 50% *of the data is important:* In this case, the size of the overall constellation is $\mathcal{N} = 256$. The size of the parent layer is $N' = 16$, which corresponds to a fraction of 50% important data. Again, we set $\Gamma_1 = 3$ dB. For this case, the lower bound in (15) does not hold and thus the GMA method fails. In Fig. 4, the SER of the UEP-PR is compared with the EEP for the first layer. We can see that the actual coding gain is roughly 3.5 dB.

V. Simplified Decoder

The optimal detector in the case of non-coherent signaling was introduced in Section II-B. Although optimal, the GLRT detector must examine every point of the constellation before

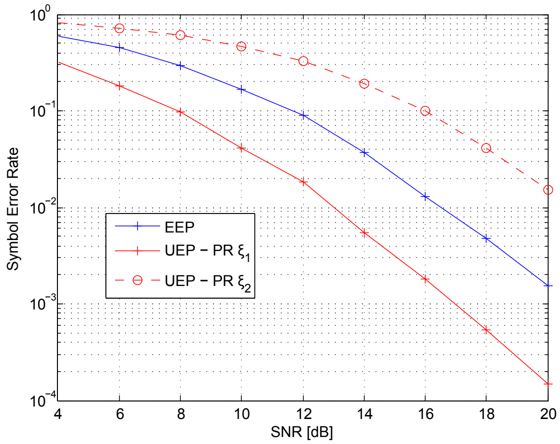

Fig. 4. SER of ξ_1 and ξ_2 when 50% of transmitted data is important

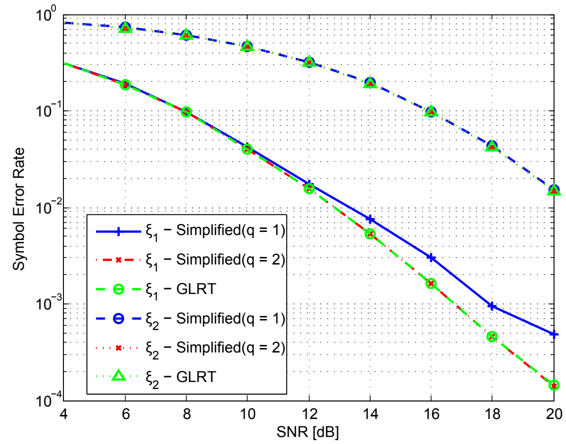

Fig. 5. Performance comparison of the simplified decoder with $q = 1, 2$ against the GLTR

making its decision. As a result, it may become computationally infeasible to employ the usual GLRT detector for moderate to large constellations. Additionally, for multi-resolution systems in which the channel conditions are unfavorable (e.g., low SNR), it is sometimes of interest to receive the more reliable information without having the need to decipher the less reliable portion of the received data. In this section, we propose a more appropriate decoding scheme that overcomes the exhaustive search of the GLRT. The new decoder makes use of the hierarchical structure of the design in Section III. Consider the constellations described in Section III with $L = 2$ (the structure and analysis of this decoder in the general L layer case are omitted for space limitations), the decoder initially examines all points in ξ_1 by computing the GLRT metric and identifying the q parents with the largest GLRT metric, where $q < N'$. Having done that, the decoder then investigates only the children points in ξ_2 associated with those parents while ignoring all children of other parents. Finally, a decision is made in favor of some $\hat{\mathbf{X}} \in \mathcal{C}$ if $\hat{\mathbf{X}}$ maximizes the GLR test over all examined points. With this setup, the number of GLRT operations needed in the children layer is reduced by a factor of q/K. Intuitively, the choice of the step parameter directly influences the performance of this detector. In particular, as the step parameter grows larger, the children points come closer to other parent points and further from their own parents, thus, leading to performance degradation relative to the usual GLRT detector. However, as our simulations suggest, this loss in performance can be overcome by employing more candidates in the parent layer. In Fig. 5, we plot the SER of both layers of the constellation described in IV-C2 for the GLRT decoder as well as the simplified decoder with $q = 1, 2$. We can see that, in the case $q = 1$, a performance loss in the parent layer SER is incurred relative to the GLRT detector. However, that loss is compensated in the $q = 2$ case, and the simplified decoder becomes essentially equivalent to the GLRT.

VI. Conclusion

In this paper, we present a non-coherent multi-layer approach suitable for UEP scenarios. This approach gives rise to a natural set partitioning scheme, which is utilized to provide more protection to the more important information. Also, we introduce two characterizations that simplify the design procedure by linking a local design parameter to the more global inter-subset distance. Finally, a simplified detector is proposed that enables simple detection while maintaining comparable performance to the optimal detector.

References

[1] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 543–564, 2000.

[2] L. Zheng and D. N. Tse, "Communication on the grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Transactions on Information Theory*, vol. 48, no. 2, pp. 359–383, 2002.

[3] S. Borade, B. Nakiboglu, and L. Zheng, "Some fundamental limits of unequal error protection," in *2008 IEEE International Symposium on Information Theory*. IEEE, 2008, pp. 2222–2226.

[4] A. R. Calderbank and N. Seshadri, "Multilevel codes for unequal error protection," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1234–1248, 1993.

[5] C.-E. Sundberg, W. Wong, and R. Steele, "Logarithmic pcm weighted qam transmission over gaussian and rayleigh fading channels," *IEE Proceedings F - Communications, Radar and Signal Processing*, vol. 134, no. 6, pp. 557–570, 1987.

[6] K. M. Attiah, K. Seddik, R. H. Gohary, and H. Yanikomeroglu, "A systematic design approach for non-coherent grassmannian constellations," in *2016 IEEE International Symposium on Information Theory (ISIT)*, July 2016, pp. 2948–2952.

[7] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receivers in rayleigh-fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2383–2399, 2001.

[8] I. Kammoun, A. M. Cipriano, and J. C. Belfiore, "Non-coherent codes over the grassmannian," *IEEE Transactions on Wireless Communications*, vol. 6, no. 10, pp. 3657–3667, 2007.

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

[10] D. Argawal, T. Richardson, and R. Urbanke, "Multiple-antenna signal constellations for fading channel," *IEEE Transactions on Informatiion Theory*, vol. 47, pp. 2618–2626, 2001.

[11] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM Journal on Matrix Analysis and Applications*, vol. 20, no. 2, pp. 303–353, 1998.

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

[13] R. A. Becker, "The variance drain and Jensen's inequality," *SSRN Working Paper Series*, 2012.