Constellation Design with Unequal Priors and New Distance Criteria for the Low SNR Noncoherent Rayleigh Fading Channel

Shivratna Giri Srinivasan and Mahesh K. Varanasi

e-mail: {srinivsg, varanasi}@dsp.colorado.edu
Electrical & Computer Engineering Department
University of Colorado, Boulder, CO 80309

Abstract — **Code design for the noncoherent MIMO Rayleigh fading channel is considered for constellations with unequal prior probabilities. The motivation for such constellations with unequal priors arises from recent information theoretic results at low Signal-to-noise ratio (SNR). Different distance measures like the Bhattacharya distance and J-divergence are related to the average pairwise error probability (APEP) and are employed to design single antenna codes of general dimension both for equal and unequal priors. For equal priors, such codes are observed to perform at least as well, and often better than codes designed based on the Kullback-Liebler (KL) distance in the low SNR regime. Substantial performance gains are observed for codes designed with unequal prior probabilities in the low SNR regime as compared to codes designed with equal priors.**

I. Introduction

In this paper, we study code design for the noncoherent independent and identically distributed (i.i.d.) Rayleigh fading MIMO channel with an average power constraint, where the transmitter and the receiver have no knowledge of the channel. Recent papers, (c.f. [1, 2]) study the mutual information of i.i.d. Rayleigh fading channels at low SNR and provide some insights into the differences in code design paradigms from the high SNR case. For instance, results in [1, 2] indicate that it is better to avoid spreading power across more antennas in the low SNR regime, and at sufficiently low SNR, the mutual information is maximized by keeping only one antenna on.

From a coding modulation viewpoint, while there have been a number of papers on unitary designs and training codes, which are primarily suited for the high SNR regime, the low SNR scenario is relatively new. Recently, [3] presented methods to design codes based on the KL distance criteria for the spatially i.i.d. Rayleigh fading MIMO channel, which perform better than some existing noncoherent codes at low SNR. Through many examples, it is shown that the KL distance between constellation points is a more suitable distance measure than the Euclidean distance measure. For their single antenna designs, the equiprobable constellation points were constrained to lie on concentric spheres and after making an approximation on the allowed arrangement of points, the KL distance between the closest points was maximized. Even though the designs are for any specified SNR, they are different from the unitary designs only in the low SNR regime, where the constellation points occupy multiple levels with a point typically at the origin. Single antenna codes designed in this paper improve upon those in [3] in the following ways:

Unequal transmission probabilities: We derive upper and lower bounds on the average word error probability in terms of the APEPs for unequal priors and the MAP detector. At low SNRs, allowing unequal priors often results in the assigning of higher probability to the origin and lower probabilities to points away from the origin, and this lowers the average codeword error probability substantially. This is expected based on the recent information theoretic results [2, 4] which indicate that at sufficiently low SNR, On-Off signaling achieves capacity, and the capacity achieving distribution gets peaky. Also, for systems with a single transmit antenna, the capacity achieving signal given in [5], is of the form $\boldsymbol{\phi} v$ where $\boldsymbol{\phi}$ is an isotropically distributed unit vector, while v is a real scalar random variable, whose density is not necessarily uniform for a general SNR. These results indicate that allowing for unequal priors in the designs may be helpful especially at low SNR. Moreover, the general analysis and code design technique subsumes the case of equal priors.

More suitable distance measures: In Section II, we explain why the APEP leads to a more intuitive symmetric distance measure between constellation points as compared to the pairwise error probability (PEP) alone for the noncoherent channel. This is especially needed when the design aim is to maximize the minimum distance in a constellation, which is the case here. We evaluate upper and lower bounds on the APEP, which are related to the Bhattacharya distance and the J-divergence. These give rise to distance measures for the noncoherent fading channel and can be shown to serve as better candidates for signal design over the KL distance. We offer explanations as to why these distance measures are superior to the KL distance for the design method adopted. In confirmation of this, the codes designed using these distance measures are observed to perform at least as well, and often better than those designed in [3] using the KL distance, in the low SNR regime.

Also, the examples of single transmit antenna codes in [3] involved only two dimensional real constellations. We incorporate the single antenna unitary designs of [6], for a general dimension in our designs and for comparisons. This helps in making the rate of transmission flexible too.

This work was supported in part by NSF Grant CCR-0112977 and by ARO Grant DAAD 19-03-1-0248.

II. System Model, the MAP detector and the average codeword error probability

We consider a communication system with N_t transmit antennas and N_r receive antennas. In our channel model, the channel matrix $\mathbf{H} \in \mathbb{C}^{N_t \times N_r}$ is assumed to be constant for a duration of T symbols after which it changes to an independent value. $\mathbf{H}$ has i.i.d. circularly symmetric $\mathcal{CN}(0,1)$ entries. The channel matrix is assumed to be unknown to the transmitter and the receiver. Assuming that the transmitted symbol is $\mathbf{X} \in \mathbb{C}^{T \times N_t}$, the output of the channel can be written as

$$\mathbf{Y} = \mathbf{X}\mathbf{H} + \mathbf{N} \, . \tag{1}$$

The entries of $\mathbf{N}$ are i.i.d. circularly symmetric $\mathcal{CN}(0,1)$. Here, the symbols $\{\mathbf{X}\}$ are normalized so that $\frac{1}{T}\mathrm{E}[\mathrm{tr}(\mathbf{X}\mathbf{X}^\dagger)] \leq P$. This would mean that the average received SNR is P. Performing a vec operation on (1), we get

$$\begin{aligned} \mathbf{y} &= (\mathbf{I}_{N_r} \otimes \mathbf{X})\mathbf{h} + \mathbf{n} \\ &= \boldsymbol{\mathcal{X}}\mathbf{h} + \mathbf{n} \, , \end{aligned} \tag{2}$$

where $\mathbf{y} = \mathrm{vec}(\mathbf{Y})$ and $\mathbf{n} = \mathrm{vec}(\mathbf{N})$.

We next describe the *maximum a-posteriori* (MAP) noncoherent detector for a code $\{\mathbf{X}_i\}_{i=1}^L$, with prior probabilities $\{\pi_i\}_{i=1}^L$. Since $\boldsymbol{\mathcal{X}}_i = \mathbf{I}_{N_r} \otimes \mathbf{X}_i$, it is understood that $\mathbf{X}_i$ and $\boldsymbol{\mathcal{X}}_i$ refer to the i^{th} codeword. The pdf of $\mathbf{y}$ conditioned on codeword $\boldsymbol{\mathcal{X}}_i$ being sent is given by

$$p_i(\mathbf{y}) = \frac{1}{\pi^{TN_r}|\mathbf{I} + \boldsymbol{\mathcal{X}}_i\boldsymbol{\mathcal{X}}_i^\dagger|} e^{-\mathbf{y}^\dagger(\mathbf{I}+\boldsymbol{\mathcal{X}}_i\boldsymbol{\mathcal{X}}_i^\dagger)^{-1}\mathbf{y}} \, .$$

The MAP decoding rule is given by

$$\hat{i} = \arg \max_{j \in (1 \ldots L)} \pi_j p_j(\mathbf{y}) \, . \tag{3}$$

Let ϵ be the event "MAP decoded codeword $\neq$ codeword transmitted" and $P(\epsilon)$ denote the average codeword error probability (WEP). The WEP for the code is given by

$$\begin{aligned} P(\epsilon) &= \Sigma_{i=1}^L \pi_i P(\epsilon/i) \\ &= \Sigma_{i=1}^L \pi_i \int_{\Gamma_i^c} p_i(\mathbf{y}) \, , \end{aligned} \tag{4}$$

where Γ_i^c is the complement of the decoding region of the i^{th} codeword and is given by

$$\Gamma_i^c = \{\mathbf{y} : i \neq \arg \max_{j \in (1,\ldots,L)} \pi_j p_j(\mathbf{y})\} \, . \tag{5}$$

It should be noted that the conditional pdf of the observation $p_i(\mathbf{y})$ depends on the signal transmitted $\mathbf{X}_i$, only through the term $\mathbf{X}_i\mathbf{X}_i^\dagger$. Therefore, based on (4), the WEP depends on the signals only through $\{\mathbf{X}_i\mathbf{X}_i^\dagger\}_{i=1}^L$ and their respective prior probabilities $\{\pi_i\}_{i=1}^L$.

The following proposition helps in restricting the domain over which signals are picked, without changing the WEP and may be proved for the unequal prior case similarly as done in [3] for equal priors.

Proposition 1: Given any code $\boldsymbol{\mathcal{C}} = \{\mathbf{X}_i\}_{i=1}^L$ which has a WEP $P(\epsilon)$, there exists a code $\boldsymbol{\mathcal{C}}' = \{\boldsymbol{\Phi}_i\mathbf{V}_i\}_{i=1}^L$ with the same WEP. Here $\boldsymbol{\Phi}_i \in \mathbb{C}^{T \times T}$ is a unitary matrix and $\mathbf{V}_i \in \mathbb{C}^{T \times N_t}$ is a diagonal matrix.

In general however, the calculation of the WEP in closed form seems intractable. We hence relate the WEP of the system to the APEP via upper and lower bounds, in order to obtain code design criteria. We first present an upper bound on the WEP based on the union bound argument. Let $P(A/i)$ denote the probability of the event A conditioned on the i^{th} codeword being sent. Denote the pairwise decoding region for codeword j given that i is sent as $D_j^i = \{\mathbf{y} : \pi_j p_j(\mathbf{y}) > \pi_i p_i(\mathbf{y})\}$. This makes the pairwise error probability (PEP) $P(D_j^i/i) = \int_{D_j^i} p_i(\mathbf{y})$. We refer to $P_e^{(i,j)} = \pi_i P(D_j^i/i) + \pi_j P(D_i^j/j)$ as the average pairwise error probability (APEP) between codewords i and j. The following bounds may be derived using standard arguments

$$P(\epsilon) \leq \Sigma_i \Sigma_{j>i} P_e^{(i,j)} \leq \frac{L(L-1)}{2} \max_{i,j} \; P_e^{(i,j)} \tag{6}$$

$$P(\epsilon) \geq \max_{i,j} \; P_e^{(i,j)} \, . \tag{7}$$

Since the upper bound is a constant times the lower bound, a reasonable criterion for code design aiming to minimize the WEP would be

$$\min_{\boldsymbol{\mathcal{C}}} \max_{i,j \in \{1,\ldots,L\}} P_e^{(i,j)} \, . \tag{8}$$

Using APEPs to derive code design criteria is more suitable than using merely the PEPs as is commonly done, for two reasons. The first is due to the assumption of unequal priors, and the other is the asymmetry of the PEP for a noncoherent channel, for which $P(D_j^i/i) \neq P(D_i^j/j)$. For a discussion on the asymmetry of the noncoherent Rayleigh fading channel, see [3]. The notion of distance between two points in a constellation is typically inferred from the pairwise error probability expression, which is symmetric when the channel is symmetric. However, the PEPs in the noncoherent Rayleigh fading channel are asymmetric, as a result of which distance measures inferred from the PEP may be inadequate. This is especially true when the design approach is to maximize the minimum distance between points in a constellation, which is the case here. In such situations, since $P_e^{(i,j)} = P_e^{(j,i)}$, the APEP leads to a more suitable and intuitive symmetric distance measure between two points in a constellation.

We will subsequently use the criterion in (8) to design codes for the noncoherent channel in Section IV.

III. Upper and lower bounds on the APEP

For the general noncoherent i.i.d. Rayleigh fading channel, calculating the APEPs in closed form seems hard, and hence we resort to bounds on the APEP. The Bhattacharya bound on the APEP is a special case of the Chernoff bound, where the parameter α is assigned the value $\frac{1}{2}$. It is a useful upperbound when it is hard to find the minimizing α for the Chernoff bound. The Bhattacharya coefficient and the Bhattacharya distance are defined [7] to be

$$\rho = \int \sqrt{p_i(\mathbf{y})p_j(\mathbf{y})}d\mathbf{y} \quad (9)$$

$$B(i,j) = -\log \rho \quad (10)$$

The following upper and lower bounds for the APEP are known [8].

$$\frac{1}{2}\min(\pi_i,\pi_j)\rho^2 \leq P_e^{(i,j)} \leq \sqrt{\pi_i \pi_j}\,\rho \quad (11)$$

It is shown in [9], that the Bhattacharya upper bound is indeed the tightest Chernoff bound when $\pi_i = \pi_j$ and at sufficiently low SNR. This suggests that it may serve as a good design criterion for the noncoherent channel in the low SNR regime for equal priors. In the case of equal priors, minimizing the largest $P_e^{(i,j)}$ translates to the same criterion for the upper and lower bounds, namely maximizing the smallest $B(i,j)$ among all pairs i, j. In Section IV, we will compare codes designed using the Bhattacharya design criterion with those designed using the KL distance criterion, as in [3]. In the context of theoretically relating these two measures, a different measure known as the J-divergence is encountered. This is also known as the symmetric Kullback-Leibler distance and is given by $J = D(p_i||p_j) + D(p_j||p_i)$. We first state some known results relating the APEP, the Bhattacharya, J-divergence and KL distance measures :

$$\frac{1}{2}\min(\pi_i,\pi_j)e^{-D(p_i||p_j)} \leq \frac{1}{2}\min(\pi_i,\pi_j)\rho^2 \leq P_e^{(i,j)}\ , \quad (12)$$

$$\frac{1}{2}\min(\pi_i,\pi_j)e^{-J/2} \leq \frac{1}{2}\min(\pi_i,\pi_j)\rho^2 \leq P_e^{(i,j)}\ . \quad (13)$$

The proofs of the left inequalities in (12,13) follow by applying the Jensen's inequality on ρ and may be found in [7]. These inequalities are true for any channel model, and when used in (11), give lower bounds on $P_e^{(i,j)}$. However, for the channel model presented in Section II, the following bounds can also be derived.

Proposition 2: For the noncoherent i.i.d Rayleigh fading MIMO channel,

$$\frac{1}{2}\min(\pi_i,\pi_j)e^{-J/4} \leq \frac{1}{2}\min(\pi_i,\pi_j)\rho^2\ , \quad (14)$$

$$\sqrt{\pi_i\pi_j}\rho \leq \sqrt{\pi_i\pi_j}\frac{2}{\sqrt{J}}\ . \quad (15)$$

Proof: The proof for this is on similar lines as given in [8], though there is a mistake in the inequalities mentioned in [8] .

The inequalities in (14) and (15), when used with (11) give rise to bounds on $P_e^{(i,j)}$. It should be noted that the lower bound on $P_e^{(i,j)}$ involving J, obtained through (11) and (14), is tighter than that obtained using (11) and (13). For general prior probabilities, the lower bound involving the Bhattacharya distance is closer to $P_e^{(i,j)}$ than the lower bounds derived using the KL or the J-divergence measures. By simulations, we find the Bhattacharya lower bound to be more suitable for code design when compared with the Bhattacharya upper bound in (11). This makes the Bhattacharya lower bound a natural choice for code design with unequal priors.

Since the codes that are designed here are single antenna codes, these are most suitable for use on the low SNR channel. Moreover, we allow for unequal priors in our designs, for which the Bhattacharya lower bound is most suitable. Due to these reasons, we adopt the Bhattacharya design criterion for signal design in this work.

For the noncoherent Rayleigh fading channel, we next state the pertinent distance criteria next.

Proposition 3: The Bhattacharya coefficient between the pair of codewords $\mathbf{X}_i$ and $\mathbf{X}_j$ is

$$\rho = \frac{1}{\left(\frac{\left|\mathbf{I}+\frac{1}{2}(\mathbf{X}_j\mathbf{X}_j^\dagger+\mathbf{X}_i\mathbf{X}_i^\dagger)\right|}{\left|\mathbf{I}+\mathbf{X}_i\mathbf{X}_i^\dagger\right|^{1/2}\left|\mathbf{I}+\mathbf{X}_j\mathbf{X}_j^\dagger\right|^{1/2}}\right)^{N_r}} \quad (16)$$

The KL-distance for the channel is derived in [3] and is given by

$$\begin{aligned} D(p_i||p_j) &= \int_\Gamma p_i(\mathbf{y})\log\frac{p_i(\mathbf{y})}{p_j(\mathbf{y})}d\mathbf{y} \\ &= \{\mathrm{tr}(\mathbf{I}+\mathbf{X}_j\mathbf{X}_j^\dagger)^{-1}(\mathbf{I}+\mathbf{X}_i\mathbf{X}_i^\dagger) \\ &\quad -\log\det(\mathbf{I}+\mathbf{X}_j\mathbf{X}_j^\dagger)^{-1}(\mathbf{I}+\mathbf{X}_i\mathbf{X}_i^\dagger) \\ &\quad -T\}\times N_r\ . \end{aligned} \quad (17)$$

By the definition of J-divergence, we get that

$$\begin{aligned} J &= D(p_i||p_j) + D(p_j||p_i) \\ &= N_r \times \{\mathrm{tr}(\mathbf{I}+\mathbf{X}_j\mathbf{X}_j^\dagger)^{-1}(\mathbf{I}+\mathbf{X}_i\mathbf{X}_i^\dagger) \\ &\quad +\mathrm{tr}(\mathbf{I}+\mathbf{X}_i\mathbf{X}_i^\dagger)^{-1}(\mathbf{I}+\mathbf{X}_j\mathbf{X}_j^\dagger) - 2T\}\ . \end{aligned} \quad (18)$$

IV. Code design using the Bhattacharya bound

A. Code construction

At sufficiently low SNR, results in [1, 2] indicate that the mutual information is maximized by keeping only a single transmit antenna on. In this subsection, we show how single transmit antenna codes may be designed using the Bhattacharya lower bound. Let a rate R bits/s/Hz be fixed for the code as a design parameter. Another design parameter that we will fix is the cardinality of the constellation L. Since the algorithm we will describe allows for unequal priors, which are obtained as the output of an optimization, the choice of L has to necessarily satisfy $L \geq 2^{RT}$. For the case when $N_t = 1$, the signal constellation may be taken to be of the form $\mathbf{x}_i = \boldsymbol{\phi}_i r_i, i = 1, ..., L$, where $\{\boldsymbol{\phi}_i\}_{i=1}^L$ are unit norm vectors, while $\{r_i\}_{i=1}^L$ are real scalars by Proposition 1. We may use this in the lower bound in (11) to get after some simplification,

$$Bh_{LB}(i,j) = \frac{1}{\frac{2}{\min(\pi_i,\pi_j)}\left(\boldsymbol{\mathcal{B}}_1(i,j)+\boldsymbol{\mathcal{B}}_2(i,j)\right)^{2N_r}}, \quad (19)$$

where $\mathcal{B}_1(i,j) = \frac{1}{2}\sqrt{\frac{1+r_j^2}{1+r_i^2}} + \frac{1}{2}\sqrt{\frac{1+r_i^2}{1+r_j^2}}$ depends only on the magnitudes of the signals r_i and r_j, and $\mathcal{B}_2(i,j) = \frac{1}{4}\frac{r_i^2 r_j^2}{\sqrt{1+r_i^2}\sqrt{1+r_j^2}}\left(1-|\phi_j^\dagger\phi_i|^2\right)$ is a scaled function of the angle between the two signals, the scaling being a function of the radius. The function $\log(\mathcal{B}_1(i,j)+\mathcal{B}_2(i,j))$ is the Bhattacharya distance by definition.

The denominator of (19) may be viewed as a distance measure between codewords i and j. This distance measure is a monotonically increasing function of the Bhattacharya distance, scaled by $\frac{2}{\min(\pi_i,\pi_j)}$.

We seek a code which has the minimum $\max_{i,j} P_e^{(i,j)}$, and using the lower bound instead of $P_e^{(i,j)}$ gives us the following design criterion :

$$\min_{\substack{1\le K\le L,\ \Sigma_{k=1}^K \pi_k l_k r_k^2 \le PT,\ \Sigma_{k=1}^K l_k = L,\\ 0<\{\pi_k\}_{k=1}^K<1,\ -\Sigma_{k=1}^K l_k\pi_k\log\pi_k = R,\\ 0\le r_1<\ldots r_K,\ \Sigma_{k=1}^K l_k\pi_k=1}} \max_{i,j} Bh_{LB}(i,j)\,. \quad (20)$$

Solving this problem seems hard in general, and we adopt a suboptimal procedure instead. We perform this optimization over the set of all constellations such that signal points lie in concentric spheres $C_1, C_2, ..., C_K$, with the probabilities of signals lying in the same sphere being equal. We assume that the spheres have radii $r_1, r_2, ..., r_K$, and the spheres have $l_1, l_2, ..., l_K$ points with probabilities $\pi_1, \pi_2, ...\pi_K$. While this assumption need not hold with the optimal code, it is reasonable, as the capacity achieving distribution is known [5] to be of the form ϕv, where ϕ is an isotropically distributed unit vector, while v is a scalar whose distribution is not known in closed form. This suggests a coding-modulation analogue for the structure of the constellation as assumed, with probabilities of symbols on the same sphere to be the same. The other simplifying assumption made is that whenever two concentric spheres C_i and C_j have non-zero number of codewords lying on each, there is at least one codeword in C_i that lies along the same unit norm vector as a codeword in C_j. This is a property that need not be true of the optimal code. However, this assumption simplifies the code design greatly.

Notice that for any two points on the same sphere with radius r_k,

$$\begin{aligned}\mathcal{B}_1(i,j) &= 1\\ \text{and}\quad \mathcal{B}_2(i,j) &= \frac{1}{4}\frac{r_k^4}{1+r_k^2}\left(1-|\phi_j^\dagger\phi_i|^2\right)\,,\end{aligned}$$

where $1-|\phi_j^\dagger\phi_i|^2$ is a measure of the angular distance between ϕ_i and ϕ_j. For the pair of codewords with $\phi_i=\phi_j$ between any two shells at radii r_i and r_j,

$$\begin{aligned}\mathcal{B}_1(i,j) &= \frac{1}{2}\sqrt{\frac{1+r_j^2}{1+r_i^2}} + \frac{1}{2}\sqrt{\frac{1+r_i^2}{1+r_j^2}}\,.\\ \text{and}\quad \mathcal{B}_2(i,j) &= 0\,.\end{aligned}$$

We define the inter-sphere distance between codewords i,j to be $\mathcal{B}_{inter}(i,j)$

$$= \frac{2}{\min(\pi_i,\pi_j)}\left(\frac{1}{2}\sqrt{\frac{1+r_j^2}{1+r_i^2}} + \frac{1}{2}\sqrt{\frac{1+r_i^2}{1+r_j^2}}\right)^{2N_r}\,. \quad (21)$$

We define the minimum distance within sphere k to be the intra-sphere distance, given by $\mathcal{B}_{intra}(k)$

$$\begin{aligned}&= \frac{2}{\pi_k}\min_{\mathbf{x}_i,\mathbf{x}_j\in C_k}\left(1+\frac{1}{4}\frac{r_k^4}{1+r_k^2}\left(1-|\phi_j^\dagger\phi_i|^2\right)\right)^{2N_r}\\ &= \frac{2}{\pi_k}\left(1+\frac{1}{4}\frac{r_k^4}{1+r_k^2}\min_{\mathbf{x}_i,\mathbf{x}_j\in C_k}\left(1-|\phi_j^\dagger\phi_i|^2\right)\right)^{2N_r}\,. \quad (22)\end{aligned}$$

The simplified max-min problem corresponding to (20) may be expressed as

$$\max_{\substack{1\le K\le L,\ \Sigma_{k=1}^K \pi_k l_k r_k^2 \le PT,\ \Sigma_{k=1}^K l_k = L,\\ 0<\{\pi_k\}_{k=1}^K<1,\ -\Sigma_{k=1}^K l_k\pi_k\log\pi_k = R,\\ 0\le r_1<\ldots<r_K,\ \Sigma_{k=1}^K l_k\pi_k=1}} \min_{i,j}\ \mathcal{S}\,, \quad (23)$$

where

$$\mathcal{S} = \left\{\{\mathcal{B}_{inter}(i,j)\}_{i>j}, \{\mathcal{B}_{intra}(k)\}_{k=1}^K\right\}\,. \quad (24)$$

For a given $\{l_k\}_{k=1}^K$ such that $\Sigma_{k=1}^K l_k = L$, this optimization may be decoupled into the problem of finding the best constellation within each sphere, and using this constellation in (23), to get $\{\pi_k\}_{k=1}^K$ and $\{r_k\}_{k=1}^k$ using numerical methods. This is because for a given r_k, π_k and l_k corresponding to a sphere, maximizing $\mathcal{B}_{intra}(k)$ is equivalent to finding max $\min_{\mathbf{x}_i,\mathbf{x}_j\in C_k}(1-|\phi_j^\dagger\phi_i|^2)$ from (22). This is precisely the problem of finding the best l_k point constellation a sphere, and is studied for instance in [6, 10, 11], and so we may use these constellations and their minimum distances for this stage of the optimization. This procedure may be adopted for each iteration corresponding to a set of $\{l_k\}_{k=1}^K$ with the minimum distances, $\{r_k\}_{k=1}^K$ and $\{\pi_k\}_{k=1}^K$ of each resulting constellation stored. Then the candidate that has the largest minimum distance between pairs is selected as the best constellation.

For the special case of equal priors, this algorithm is similar to that used in [3] to design codes using the KL distance as the design criterion. The KL-distance between two codewords i,j also has a term which depends only on the radii r_i, r_j through $\frac{1+r_j^2}{1+r_i^2}$, and a term which is a scaled function of the angle between the two codewords. After some simplification to (18), it may be seen that the J-divergence also has a similar form, and can also be used to design codes using the same algorithm for equal priors. Based on our discussion in Section II, we saw that choosing a symmetric distance measure related to the APEP gives rise to a more intuitive distance measure between two points in a constellation. Since the KL distance is asymmetric while the Bhattacharya distance is symmetric, and from the discussion in Section III, it would seem like the Bhattacharya distance is a better candidate. For equal priors, this is corroborated by our observation that codes designed using the Bhattacharya criterion seem to do at least as well and often better than those using the KL distance criterion. However, the comparison between the Bhattacharya distance and the J-divergence is not as straightforward for equal priors. Both distance measures are symmetric, and in the low SNR regime, the codes indicated as optimal by both criteria match in most, if not all of the examples that we examined. However, it may help to compare codes designed independently according to both criteria and choose the better performing code for higher SNR regimes.

B. Simulation results

For the simple case of a 2-dimensional real constellation with $\boldsymbol{\phi}_i = \begin{bmatrix} \cos\theta_i \\ \sin\theta_i \end{bmatrix}$ and $\boldsymbol{\phi}_j = \begin{bmatrix} \cos\theta_j \\ \sin\theta_j \end{bmatrix}$, then $1 - |\boldsymbol{\phi}_j^\dagger \boldsymbol{\phi}_i|^2 = \sin^2(\theta_j - \theta_i)$. Here, $(\theta_j - \theta_i)$ is the angle between the two vectors $\boldsymbol{\phi}_j$ and $\boldsymbol{\phi}_i$. The solution for the best constellation of l_k points within a circle would hence be given by $\max\min \quad \sin^2(\theta_j - \theta_i)$, which would have the solution $\left\{ \begin{bmatrix} \cos((l-1)\pi/L) \\ \sin((l-1)\pi/L \end{bmatrix} \right\}_{l=1}^{L}$. Since this constellation admits easy representation in the cartesian plane, we plot signals designed for $R = 2, P = 1, T = 2$ using the various design criteria in the Figures 1,2 and 3. In Figure 4, we compare the WEPs of these constellations using the MAP detector.

One way of implementing a system with unequal prior probabilities is given in [12]. To implement the code in Figure 3, the i.i.d. input binary sequence should be parsed using a prefix free code $\{0, 100, 101, 110, 111\}$. Since the code is prefix free, the parsing is unique. Thereafter one can map these words of the prefix code to the codewords used on the channel according to the priors. Here, the respective probabilities of occurrences of $\{0, 100, 101, 110, 111\}$ in a sequence of i.i.d. bits would be $\{\frac{1}{2}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}\}$, which are the probabilities of the codewords of the required code designed according to the Bhattacharya lower bound. Clearly, this kind of mapping works only because the prior probabilities involved were dyadic. In the case when the probabilities are not dyadic, this method of implementation would necessitate an approximation.

We also plot constellations designed for a coherence time of $T = 6$ and decoded using the MAP detector. For the maximization of $\boldsymbol{\mathcal{B}}_{intra}(k)$, we use a single antenna unitary design of dimension $T = 6$ in [6]. The comparison between WEPs of codes indicated as optimal by the various criteria is shown in Figure 5. For $R = 3, P = 0.25$, 8-point constellation having equal priors, code design according to either the Bhattacharya distance or the J-divergence led to the 8-point single antenna unitary design with all points at a radius of $\frac{\sqrt{3}}{\sqrt{2}}$. However, with unequal priors and a 17-point constellation with the same rate, the Bhattacharya lower bound indicated a constellation that has a point at the origin with probability $\frac{1}{2}$, and 16 points at a radius of $\sqrt{3}$ forming a single antenna unitary design each with probability $\frac{1}{32}$. This code may also be implemented as described in the previous paragraph using the prefix code $\{0, 1xxxx\}$, where $1xxxx$ denotes any possible length 5 sequence beginning with 1.

V. CONCLUSION

We described a method to design single antenna codes, allowing for the constellation points to have unequal priors. We showed that a lower bound on the APEP between a pair of codewords which is related to the Bhattacharya distance is most suitable for the method adopted, as compared to lower bounds involving the KL-distance or the J-divergence. The constellation that minimized the maximum of the Bhattacharya lower bound on the APEP over all pairs is sought. Since it is best to have only one transmit antenna on in the low SNR regime based on recent information theoretic results, the codes designed here are well suited for low SNR i.i.d. Rayleigh fading MIMO channels. These codes were shown to perform substantially better than codes designed for equal priors using any of the known criteria and similar design techniques in the low SNR regime. In particular, at received SNR = 0 dB, $T = 2$, simulations showed that a code designed with unequal priors needed 4 receive antennas fewer than a code designed with equal priors, to exhibit the same WEP $\approx 1.5 \times 10^{-2}$. At received SNR = -6 dB, $T = 6$, a code designed with unequal priors needed 4 receive antennas fewer than a code designed with equal priors, for the same WEP $\approx 7 \times 10^{-2}$.

REFERENCES

[1] B. Hassibi and T. L. Marzetta, "Multiple-antennas and isotropically-random unitary inputs: The received signal density in closed-form," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1473–1484, June 2002, Special Issue on Shannon Theory: Perspective, Trends, and Applications.

[2] C. Rao and B. Hassibi, "Analysis of multiple-antenna wireless links at low SNR," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 2123 – 2130, Sept. 2004.

[3] M. J. Borran, A. Sabharwal, and B. Aazhang, "On design criteria and construction of noncoherent space-time constellations," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2332–2351, Oct. 2003.

[4] I. C. Abou-Faycal, M. D. Trott, and S. Shamai (Shitz), "The Capacity of discrete-time memoryless Rayleigh fading channels," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.

[5] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.

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

[7] T. Kailath, "The Divergence and Bhattacharyya distance measures in signal selection," *IEEE Trans. Commun.*, vol. 15, no. 1, pp. 52–60, Feb. 1967.

[8] T. T. Kadota and L. A. Shepp, "On the best finite set of linear observables for discriminating two Gaussian signals," *IEEE Trans. Inform. Theory*, vol. 13, no. 2, pp. 278–284, Apr. 1967.

[9] A. Dogandzic, "Chernoff bounds on pairwise error probabilities of space-time codes," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1327–1336, May 2003.

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

[11] M. L. McCloud and M. K. Varanasi, "Modulation and coding for noncoherent communications," *Journ. VLSI Signal Processing*, vol. 30, no. 1–3, pp. 35–54, Jan.-Feb.-Mar. 2002.

[12] G. D. Forney, R. G. Gallager, G. R. Lang, F. M. Longstaff, and S. U. Quereshi, "Efficient-modulation for bandlimited channels," *IEEE J. Select. Areas Commun.*, vol. 2, no. 5, pp. 632–647, Sept. 1984.

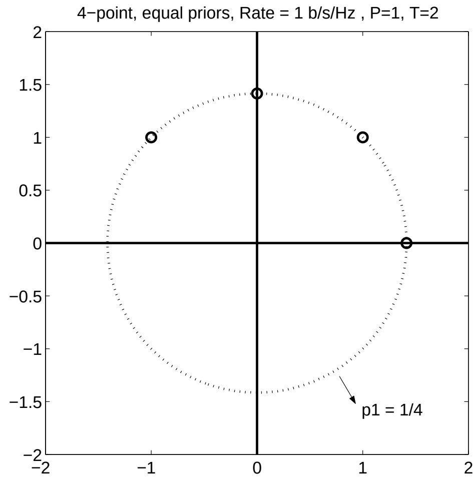

Fig. 1. 4-point equal prior constellation obtained using the KL distance at received SNR = 0 dB

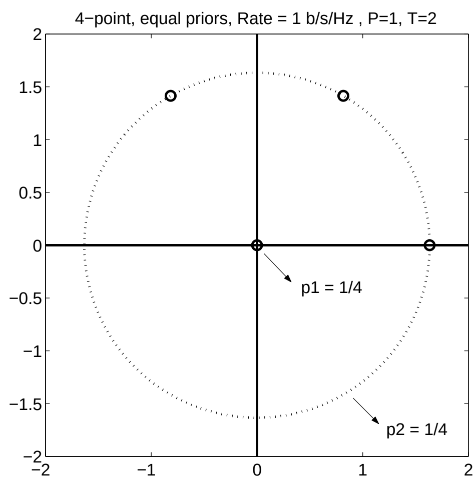

Fig. 2. 4-point equal prior constellation obtained using either the Bhattacharya distance or J-divergence at received SNR = 0 dB

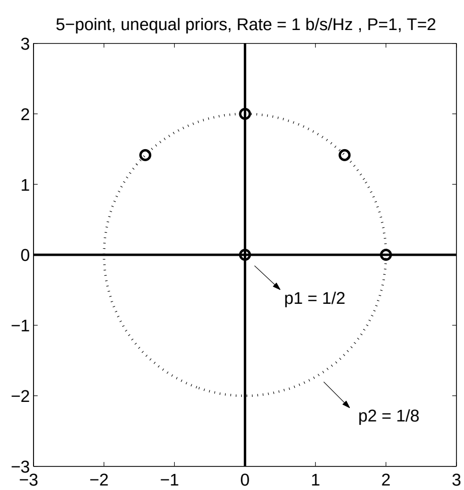

Fig. 3. 5-point unequal prior constellation designed using the Bhattacharya lower bound at received SNR = 0 dB . Probability of transmitting '0'=$\frac{1}{2}$ and the probability of the remaining symbols =$\frac{1}{8}$

.

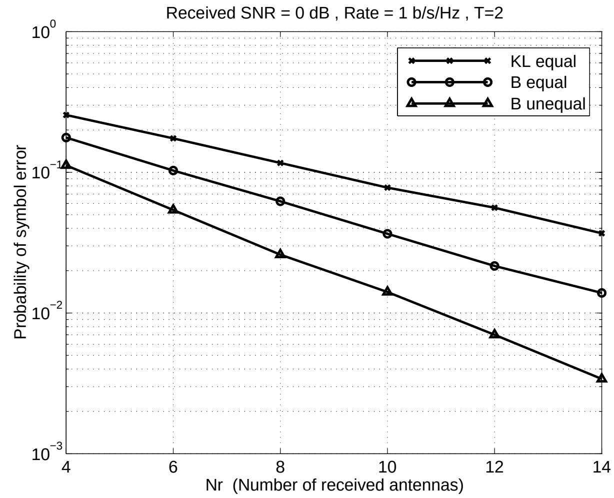

Fig. 4. Comparison of WEPs of constellations given in figures 1,2 and 3 over a spatially i.i.d. Rayleigh fading channel with received SNR = 0 dB.

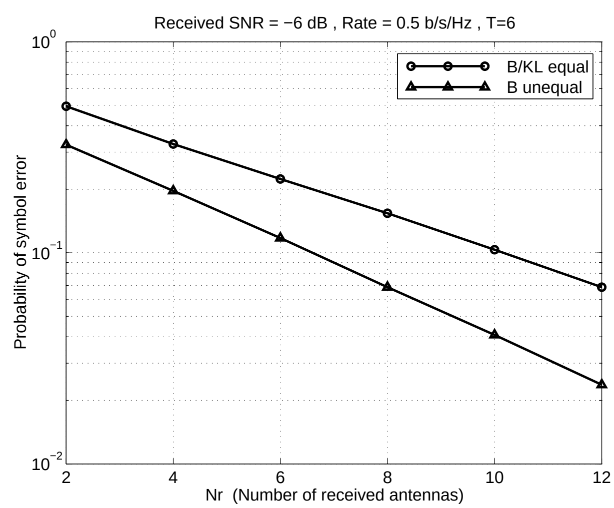

Fig. 5. Comparison of WEPs of codes designed according to the Bhattacharya distance measure with equal and unequal priors over a spatially i.i.d. Rayleigh fading channel with received SNR = -6 dB. The single antenna unitary design of dimension $T = 6$ in [6] is utilized for the intra-sphere code.