Constellation Design for the Noncoherent MIMO Rayleigh-Fading Channel at General SNR

Shivratna Giri Srinivasan, *Student Member, IEEE*, and Mahesh K. Varanasi, *Senior Member, IEEE*

***Abstract*—Constellation design for the noncoherent multiple-input–multiple-output (MIMO) block Rayleigh-fading channel is considered. For general signal-to-noise ratios (SNRs), starting from a given base unitary constellation of finite cardinality, and using the cutoff rate expression as the design criterion, input probabilities and per-antenna amplitudes for the constellation points are obtained via a *difference of convex programming* formulation. Using the mutual information as a performance metric, it is shown that the optimized constellations significantly outperform the base unitary designs from which they are obtained in the low-medium SNR regime, and indeed they also similarly outperform the mutual information achieved by isotropically distributed unitary inputs for the continuous input channel [i.e., the so-called unitary space–time capacity (USTC)]. At sufficiently high SNRs, the resulting mutual information coincides with that of the base unitary designs. Thus the optimum constellation design technique works over the entire range of SNRs. The bit energy/spectral efficiency tradeoff of the optimized constellations are also obtained, and these provide valuable insights on modulation and coding, which are especially useful for wideband channels where the SNR per degree of freedom is low.**

***Index Terms*—Capacity, constellation design, cutoff rate, difference of convex programming, global optimization, low signal-to-noise ratio (SNR), multiple-antenna systems, noncoherent channels, Rayleigh fading, unitary constellations, wideband channels.**

I. Introduction

The noncoherent multiple-input–multiple-output (MIMO) channel was studied from an information theoretic point of view by Marzetta and Hochwald [1] where a characterization, albeit partial, of the optimal input distribution is given for general signal-to-noise ratios (SNRs). The asymptotic regimes of high and low SNRs have since been studied extensively and can be found for instance in [2]–[6]. For example, Zheng and Tse study the high SNR behavior in [2] and conclude that isotropically distributed unitary inputs are optimal in that regime. For such inputs, Hassibi and Marzetta in [7] provide the mutual information of the MIMO channel for a general SNR and we refer to this as the unitary space–time capacity (USTC). The USTC serves as a lower bound on the true capacity which is tight in the high SNR regime. Recent results (cf. [3]–[6], [8]) shed some light on coding at low SNR and how it differs from the high SNR case. One major difference is that at low SNR, a peaky signaling scheme approaches capacity, while at high SNR isotropically distributed unitary inputs achieve capacity.

The characterization of the capacity at general SNR however remains an open problem. The authors obtain new results on coding modulation for the noncoherent MIMO channel at low SNR in [6].

Constellation design for the noncoherent MIMO channel has also received much attention. The case of designing high rate constellations is already well studied through the approaches of Unitary designs (cf. [9], [10]) and Training codes (cf. [11]). These constellations are eminently suited for high SNRs or long coherence lengths T, where high rates become feasible since the channel capacity is sufficiently large. Training codes also have efficient decoders, since they leverage state-of-the-art coherent codes along with channel estimation at the receiver. However, the performance of Unitary designs as well as Training codes deteriorate considerably in the low/medium SNR range. Therefore, primarily in the low/medium SNR range and when the coherence blocklength is small, the problem of constellation design is still not well understood. One approach to this problem is in [12], where the authors propose a method for designing SNR-dependent constellations aimed at maximizing the minimum Kullback-Liebler distance. The constellation points are however, restricted to have equal probabilities which severely limits their effectiveness at low SNR, where constellations with relatively larger peak-to-average power ratios are needed to approach capacity. Another major drawback of that work is an assumption that between any two concentric shells on which the constellation points are constrained to lie, there is at least one pair that is collinear with the origin. This assumption renders the constellations strictly suboptimal and not even locally optimal in most cases, since a mild perturbation of the points along a shell typically improves the minimum distance. These drawbacks offer potential for significant improvements to the designs in [12]. Furthermore, most existing constellations that use the mutual information as the performance measure, are designed according to criteria specific to either the low or the high SNR asymptotic regimes, and not for general SNRs.

To bridge this gap, we propose a constellation design technique using the cutoff rate expression as the design criterion. The cutoff rate expression is a lower bound on the mutual information and has been used before as a design criterion in [13] and the references contained therein. One major difference between [13] and our work is that [13] imposes only a peak power constraint on the constellations, whereas we impose an additional average power constraint which significantly changes the nature of the problem. The constellations that we propose have unequal transmission probabilities and unequal per-antenna amplitudes in general, and are obtained via a global optimization formulation by leveraging existing unitary constellations. A potential drawback is that such optimizations are tractable when the cardinality of the constellations is relatively small. However, information theoretic results suggest that unitary constellations which are good packings in the complex Grassmanian space are nearly optimal in the high SNR regime so that the low-medium SNR regime is of most interest. In this regime, constellations with relatively small cardinalities suffice since the channel capacities are also relatively small. With a fixed cardinality, the optimal constellations indicated over the whole range of SNRs make a smooth transition from constellations with large peak-to-average power ratio in the low SNR regime, to a constellation with no point at zero and equiprobable points at the same amplitude at sufficiently high SNRs. We show that such optimized constellations exhibit significant to moderate improvements in mutual information over the underlying unitary designs as well as the USTC in the low-medium SNR range while yielding the same mutual information as the underlying unitary designs at high SNR, thus providing a unified constellation design technique for the entire range of SNRs.

Results in [14] and [15] show that the capacity achieving input amplitude is discrete with a finite number of levels for the single-input–single-output (SISO) Rayleigh fast-fading channel and the SISO Rician fast-fading channel, respectively. The authors in [16] prove that the capacity achieving input is discrete under certain conditions for a conditionally Gaussian channel. The authors also enumerate several well

Manuscript received June 6, 2006; revised December 23, 2006. This work was supported in part by the National Science Foundation under Grant CCF-0431170. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Seattle, WA, July 2006.

The authors are with the Electrical and Computer Engineering Department, University of Colorado at Boulder, CO 80309 USA (e-mail: srinivsg@colorado.edu; varanasi@colorado.edu).

Communicated by R. R. Müller, Associate Editor for Communications.

Color version of Fig. 6 in this correspondence is available on line at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2007.892783

known channel models with constraints on the input, for which the capacity achieving distribution is known to be discrete. While this is not yet proved for the noncoherent MIMO channel, it is not unreasonable to expect well designed discrete constellations to approach capacity. In recent work, [17] proposes a technique to construct discrete distributions for a general class of SISO channels by maximizing the error exponent. It is also shown that simple discrete inputs can nearly achieve capacity, even when the capacity achieving distribution is known to be continuous. The mutual information of optimal constellations designed through the cutoff rate expression may, hence, serve as a good benchmark and provide a constructive lower bound on the capacity of noncoherent MIMO channels in the low-medium SNR range. This constellation design technique is the main contribution of this correspondence.

In his insightful paper, Verdu [8] points out that the bit-energy versus spectral-efficiency curve provides crucial insights into the tradeoffs between bandwidth, power and rate. Obtaining this tradeoff also sheds light on the minimum bit energy that is needed for reliable communication, and the ideal region of spectral efficiencies to operate in. The low SNR regime of operation is of great interest since the power-efficiency, which is closely related to the energy per bit, is highest in this regime. Due to the power limitations of battery operated devices, this regime of operation is likely to be encountered in practice in several applications. Also, the UWB spectrum of 7.5 GHz in the 3.1 to 10.6 GHz frequency band was allocated for unlicensed use by the Federal Communications Commission (FCC), and identifying the most efficient communication technique is an ongoing effort. Systems operating in the unlicensed bands can experience rapidly changing, uncontrolled interference from other colocated systems. This rapid variation of the channel can make phase estimation, and, hence, coherent detection, extremely difficult if not impossible. Emerging wireless systems like impulse radio also operate at high bandwidth, and in many instances noncoherently. Therefore, insights gained from the bit-energy/spectral efficiency tradeoff lead to efficient use of resources in such scenarios, where the SNR per degree of freedom is low. To obtain this tradeoff, the capacity of the channel as a function of SNR is required. Since the capacity of the general MIMO noncoherent channel is not known either in closed form or numerically, this approach poses a challenge. We show that constellations designed to maximize the cutoff rate expression at general SNRs can be used to generate benchmark bit-energy versus spectral-efficiency curves, and to obtain insights on optimal constellations. From these curves, a good upper bound on the minimum energy per bit required for reliable communication can also be obtained. This is a useful application of the constellation design technique that we propose.

We finally note here that the ideas of this correspondence can be applied to a more general setting than the independent, identically distributed (i.i.d.) Rayleigh fading channel. For example, in Section IV, we briefly outline how they may be extended to more general spatially correlated Rayleigh fading channels.

II. System Model

We begin by describing some notational conventions. For an integer N, $\mathbf{I}_N$ is an $N \times N$ identity matrix. Matrices are denoted by the boldfaced capital letters, and vectors by bold faced small letters. The matrices $\mathbf{A}^T$, $\bar{\mathbf{A}}$, $\mathbf{A}^*$ and $\mathbf{A}^\dagger$ denote the transpose, complex-conjugate, conjugate transpose of $\mathbf{A}$, and the pseudoinverse, respectively. We denote the block diagonal matrix with the matrices $\mathbf{A}_1, \ldots, \mathbf{A}_N$ along the block diagonal as $\mathrm{blockdiag}(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_N)$. $\mathrm{E}[\cdot]$ and tr denote the expectation and trace operators, respectively. We use the notation $o(\rho)$ to mean that $\lim_{\rho \to 0} \frac{o(\rho)}{\rho} = 0$. The vec operation on a matrix results in a column vector formed by stacking the columns of the matrix in the order in which they appear in the matrix, one on top of the other. The determinant of a matrix $\mathbf{A}$ is denoted as $|\mathbf{A}|$. The (i, j)th element of $\mathbf{A}$ is denoted as $[\mathbf{A}]_{ij}$. The Kronecker product of two matrices $\mathbf{A}$ and $\mathbf{B}$ is denoted as $\mathbf{A} \otimes \mathbf{B}$. A zero-mean, unit-variance, circularly symmetric, complex normal random variable is denoted as $\mathcal{CN}(0, 1)$.

We consider a MIMO channel with N_t transmit and N_r receive antennas. The random 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. It has independent and identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$ entries (spatial correlations are considered in Section IV) and the knowledge that the entries have this distribution is available to the transmitter and receiver. The realizations of $\mathbf{H}$, however, are unknown at both ends. 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. $\mathcal{CN}(0, 1)$ random variables. The symbol $\mathbf{X}$ is drawn from a finite constellation $\boldsymbol{\mathcal{C}}$ which in turn is normalized so that $\frac{1}{T}\mathrm{E}[\mathrm{tr}(\mathbf{X}\mathbf{X}^*)] \leq P$ so that the average received SNR is constrained to be P. An additional peak-power constraint $\frac{1}{TN_t}\mathrm{tr}(\mathbf{X}\mathbf{X}^*) \leq \mathcal{K}$ on each $\mathbf{X} \in \boldsymbol{\mathcal{C}}$ is also imposed.

Performing a vec operation on (1) and denoting $\mathbf{y} = \mathrm{vec}(\mathbf{Y})$, $\mathbf{n} = \mathrm{vec}(\mathbf{N})$, and $\mathbf{h} = \mathrm{vec}(\mathbf{H})$, we have

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

The probability density function (pdf) of $\mathbf{y}$ conditioned on $\boldsymbol{\mathcal{X}}$ being sent is given by

$$\begin{aligned} p(\mathbf{y}/\boldsymbol{\mathcal{X}}) &= \frac{1}{\pi^{TN_r}|\mathbf{I}_{TN_r} + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{X}}^*|} e^{-\mathbf{y}^*(\mathbf{I}_{TN_r} + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{X}}^*)^{-1}\mathbf{y}} \\ &= \frac{1}{\pi^{TN_r}|\mathbf{I}_{TN_r} + \mathbf{I}_{N_r} \otimes \mathbf{X}\mathbf{X}^*|} e^{-\mathbf{y}^*(\mathbf{I} + \mathbf{I}_{N_r} \otimes \mathbf{X}\mathbf{X}^*)^{-1}\mathbf{y}}. \end{aligned}$$

III. Constellation Design at General SNR

At a general SNR, the mutual information is not known in closed form. We therefore adopt the cutoff rate expression as our design criterion.

A. The Cutoff Rate Expression

Consider a constellation $\{\mathbf{X}_i\}_{i=1}^L$ with corresponding transmission probabilities $\{P_i\}_{i=1}^L$. The cutoff rate for the discrete input (of cardinality L) and continuous output channel is given by

$$R_0 = \max_{\{P_i\}_{i=1}^L, \{\mathbf{X}_i\}_{i=1}^L} -\log\left\{\sum_i \sum_j P_i P_j \int \sqrt{p(\mathbf{y}/i)p(\mathbf{y}/j)}d\mathbf{y}\right\}. \tag{3}$$

For the noncoherent MIMO channel, the argument of $\max(\cdot)$ in (3) is easily shown to be (cf. [13])

$$-\log\left\{\sum_i \sum_j P_i P_j \frac{|\mathbf{I} + \mathbf{X}_i\mathbf{X}_i^*|^{N_r/2}|\mathbf{I} + \mathbf{X}_j\mathbf{X}_j^*|^{N_r/2}}{\left|\mathbf{I} + \frac{1}{2}\left(\mathbf{X}_i\mathbf{X}_i^* + \mathbf{X}_j\mathbf{X}_j^*\right)\right|^{N_r}}\right\}. \tag{4}$$

We refer to (4) as the *cutoff rate expression* (denoted as CR). The cutoff rate expression is a lower bound on the mutual information at any SNR. It is proved in Appendix-A that CR is an increasing function of SNR and N_r for any nontrivial constellation.

We first state a proposition which plays an important role in constellation design and is a slight modification of that in [13].

Proposition 1: Consider a constellation $\boldsymbol{\mathcal{C}} = \{\mathbf{X}_i\}_{i=1}^L$ with priors $\{P_i\}_{i=1}^L$ satisfying the average power constraint $\frac{1}{T}\sum_i P_i \mathrm{tr}(\mathbf{X}_i\mathbf{X}_i^*) \leq P$ and the peak power constraint $\frac{1}{TN_t}\mathrm{tr}(\mathbf{X}_i\mathbf{X}_i^*) \leq \mathcal{K}$. Then there exists a constellation with the structure $\boldsymbol{\mathcal{C}}' = \{\boldsymbol{\Phi}_i\mathbf{V}_i\}_{i=1}^L$ having the same cutoff rate expression value as $\boldsymbol{\mathcal{C}}$ and satisfying the same average and peak power constraints, where $\mathbf{\Phi}_i \in \mathbb{C}^{T\times N_t}$ is a matrix with orthonormal columns (i.e., $\mathbf{\Phi}_i^*\mathbf{\Phi}_i = \mathbf{I}$), and $\mathbf{V}_i \in \Re^{N_t\times N_t}$ is a diagonal matrix with nonnegative entries and $N_t \le T$.

The implication of Proposition 1 is that, without loss of generality, we may restrict attention to constellations of unitary matrices with variedly scaled columns. Note that even though we use the term "unitary" in line with popular usage, we only mean that $\mathbf{\Phi}_i^*\mathbf{\Phi}_i = \mathbf{I}$ and not $\mathbf{\Phi}_i\mathbf{\Phi}_i^* = \mathbf{I}$. The peak power constraint is active in the low SNR regime, since here the optimal signals under an average power constraint alone get peaky (and thus are hard to implement in practice). In the medium and high SNR regime, the peak-power constraint is usually not active, and only the average power constraint may be considered to hold.

In the low SNR regime, the cutoff rate expression upto second order in P is derived in [13] as

$$\mathrm{CR}_{\mathrm{low}} = \frac{N_r}{8}\sum_i\sum_j P_iP_j\,\mathrm{tr}\left\{\left(\mathbf{X}_i\mathbf{X}_i^* - \mathbf{X}_j\mathbf{X}_j^*\right)^2\right\} + o(P^2). \quad (5)$$

The key idea used in obtaining $\mathrm{CR}_{\mathrm{low}}$ from (4) is the approximation $\log|\mathbf{I}+\mathbf{A}| \approx \mathrm{tr}(\mathbf{A}) - \frac{1}{2}\mathrm{tr}(\mathbf{A}^2)$ which is valid when $\mathbf{A}$ is Hermitian and has small eigenvalues.

At asymptotically low SNR and when the fourth and sixth-order moments of the input signal are bounded, [4] obtains the mutual information up to the second order in P for the continuous input and continuous output channel. A similar analysis tailored to the discrete input and continuous output case yields the same expression and is given as

$$\mathrm{I}_{\mathrm{low}} = \frac{N_r}{2}\{\mathrm{E}[\mathrm{tr}\{(\mathbf{X}\mathbf{X}^*)^2\}] - \mathrm{tr}\{(\mathrm{E}[\mathbf{X}\mathbf{X}^*])^2\}\} + o(P^2) \quad (6)$$

with the expectations in (6) involving a discrete input and probability mass function (pmf), while those in [4] involve a continuous input and pdf.

The following simple proposition shows that the cutoff rate expression behaves *identically* to the mutual information at low SNRs.

Proposition 2: In the limit of low SNR, the cutoff rate approaches a value that is equal to half the mutual information, i.e.

$$\lim_{P\to 0}\frac{\mathrm{CR}}{\mathrm{I}} = \frac{1}{2}$$

Proof:

$$\begin{aligned}\mathrm{CR}_{\mathrm{low}} &= \frac{N_r}{8}\sum_{i,j}P_iP_j\,\mathrm{tr}\{(\mathbf{X}_i\mathbf{X}_i^* - \mathbf{X}_j\mathbf{X}_j^*)^2\} + o(P^2)\\ &= \frac{N_r}{8}\Bigg\{2\sum_i P_i\,\mathrm{tr}(\mathbf{X}_i\mathbf{X}_i^*\mathbf{X}_i\mathbf{X}_i^*)\\ &\quad -2\,\mathrm{tr}(\sum_i P_i\mathbf{X}_i\mathbf{X}_i^*\sum_j P_j\mathbf{X}_j\mathbf{X}_j^*)\Bigg\} + o(P^2)\\ &= \frac{N_r}{4}\{\mathrm{E}[\mathrm{tr}\{(\mathbf{X}\mathbf{X}^*)^2\}]\\ &\quad -\mathrm{tr}\{(\mathrm{E}[\mathbf{X}\mathbf{X}^*])^2\}\} + o(P^2)\\ &= \frac{1}{2}\mathrm{I}_{\mathrm{low}} + o(P^2). \qquad (7)\end{aligned}$$

□

B. The General SNR Case

Without loss of generality, we assume that the optimal constellation is of the form $\{\mathbf{\Phi}_i\mathbf{V}_i\}_{i=1}^L$ with respective probabilities $\{P_i\}_{i=1}^L$. It is a solution to the following optimization problem:

$$\mathcal{C} = \arg\max_{\{P_i\}_{i=1}^L,\{\mathbf{\Phi}_i\}_{i=1}^L,\{\mathbf{V}_i\}_{i=1}^L}\mathrm{CR}. \quad (8)$$

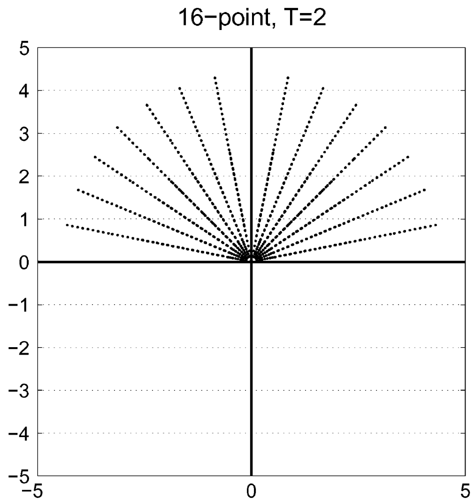

Fig. 1. Two-dimensional real space divided by 16 equally spaced rays.

In Section III-C, we show that for low rate constellations with $L \le \frac{T}{N_t}$, the jointly global optimal solution can be found numerically. The optimization of CR jointly over $\{P_i\}_{i=1}^L$, $\{\mathbf{\Phi}_i\}_{i=1}^L$ and $\{\mathbf{V}_i\}_{i=1}^L$ seems hard in the general case. Moreover, it involves $L + LN_t + LTN_t$ variables which is usually too large a number.

In this section, we address the general SNR case and we adopt an alternative approach which is very reasonable and complexity-wise tractable in many interesting cases. We motivate this approach by considering a two dimensional real constellation example ($T = 2, N_t = 1$). Consider the two dimensional real space of points partitioned by L rays spaced equi-angularly and passing through the origin, as shown in Fig. 1. It should be noted that only two quadrants are used in plotting the constellation points since $\mathbf{X}$ and $-\mathbf{X}$ are indistinguishable using the conditional pdf of the received signal. If the constellation points are constrained to lie on these rays (one on each ray), then only the amplitudes and prior probabilities need to be determined. In a sense, the rays quantize the angular space and the resulting optimization has a lot fewer parameters. This intuition may be extended to the $T\times N_t$ dimensional complex space and its being divided into L "rays" indicated by $\{\mathbf{\Phi}_i\}_{i=1}^L$ spaced apart so as to maximize the minimum "angular" distance between them. Optimizing over the diagonal matrices $\{\mathbf{V}_i\}_{i=1}^L$ is similar to optimizing over amplitudes in the real case.

We choose any existing constellation of matrices with orthonormal columns that is a good packing in complex Grassmanian space and use it for $\{\mathbf{\Phi}_i\}_{i=1}^L$. Such constellations are popularly referred to as *unitary constellations*, examples of which are given in [9], [10]. We then optimize the cutoff rate expression jointly over $\{P_i\}_{i=1}^L$ and $\{\mathbf{V}_i\}_{i=1}^L$ to obtain our constellation. The optimization is now over only $L + LN_t$ variables.

A long-standing open problem in noncoherent MIMO communications is to find the distribution of $\mathbf{V}$ so that the signal $\mathbf{\Phi}\mathbf{V}$ achieves capacity at a general SNR, where $\mathbf{\Phi}$ is an isotropically distributed unitary matrix [1]. The capacity of the MIMO channel under isotropically distributed inputs for a general SNR is derived in [7] and we refer to this as the unitary space time capacity (USTC). Numerical results show that the systematic unitary designs achieve the USTC in the low SNR regime and have mutual informations quite close to the USTC in the medium SNR regime, indicating that they form a good packing in complex Grasmannian space and they exhibit nearly the same mutual information as isotropically distributed unitary matrices in this regime. Our problem of maximizing the cutoff rate expression over $\{P_i\}_{i=1}^L$ and $\{\mathbf{V}_i\}_{i=1}^L$, by *leveraging* $\{\mathbf{\Phi}_i\}_{i=1}^L$ that are designed to achieve the

USTC, resembles a discrete version of the problem of finding the distribution of $\mathbf{V}$, while using the cutoff rate expression as the design criterion, which in turn, by Proposition 1, is equivalent to optimizing the mutual information in the critical low SNR regime. Moreover, for small T and in the low-medium SNR range, the capacity of the noncoherent channel is small enough to enable constellations of small cardinality to approach it. This renders the optimizations tractable in many interesting cases. Indeed, it is in this regime that most gains are expected over unitary designs. The joint optimization over the amplitudes and probabilities can also be viewed as giving rise to power-control and "probability-control" strategies that must be applied to the base unitary constellations to boost their performance. This is particularly relevant from a practical standpoint since algebraic unitary designs utilizing the QPSK alphabet are available (c.f. [10]) that achieve mutual information close to the USTC bound in the low SNR regime. Finally, the mutual information obtained this way can be seen to provide a constructive lower bound to the noncoherent MIMO capacity at general SNRs and one that we will see is much tighter than the USTC bound in the low-medium SNR regime. It can be seen that the cutoff rate expression is nonconvex in general with respect to $\{P_i\}_{i=1}^L$ and $\{\mathbf{V}_i\}_{i=1}^L$, and, hence, this problem comes under the realm of *deterministic global optimization*[18], [19]. In order to maximize the cutoff rate expression at a general SNR, a globally optimal solution can be obtained by formulating it as a *difference of convex programming* problem (dc programming). We give some definitions from [18], [19] in this regard.

Definition 1: A real valued function f defined on a convex set $\boldsymbol{\mathcal{A}} \subseteq \Re^n$ is called dc (difference of convex) on $\boldsymbol{\mathcal{A}}$ if, for all $\mathbf{x} \in \boldsymbol{\mathcal{A}}$, f can be expressed in the form

$$f(\mathbf{x}) = p(\mathbf{x}) - q(\mathbf{x}) \tag{9}$$

where p, q are convex functions on $\boldsymbol{\mathcal{A}}$. The representation (9) is said to be a dc decomposition of f.

It should be noted that no dc decomposition is unique. For instance, $(p(\mathbf{x}) + \|\mathbf{x}\|^2) - (q(\mathbf{x}) + \|\mathbf{x}\|^2)$ is also a dc decomposition of $f(\mathbf{x})$ in (9), where $\|\mathbf{x}\|$ is the l_2-norm of $\mathbf{x}$.

Definition 2: A global optimization problem is called a dc programming problem or a dc program if it has the form

$$\begin{aligned} \min \quad & f_0(\mathbf{x}) \\ \text{s.t} \quad & \mathbf{x} \in \boldsymbol{\mathcal{A}} \\ & f_i(\mathbf{x}) \le 0 \quad (i = 1, \ldots, m) \end{aligned} \tag{10}$$

where $\boldsymbol{\mathcal{A}}$ is a closed convex subset of $\Re^n$ and all functions $f_i, (i = 0, 1, \ldots, m)$ are d.c on $\boldsymbol{\mathcal{A}}$. If the set of all constraints form a polytope, then the problem is called a dc program over a polytope.

There are a number of algorithms given in [18], [19] to find the global minimum of a dc program if the dc decomposition is known. The next lemma is needed to prove the ensuing theorem.

Lemma 1: The function $f(\mu, \mathbf{D}_1, \mathbf{D}_2) = |(\mathbf{I} + \mu(\mathbf{A}\mathbf{D}_1\mathbf{A}^* + \mathbf{B}\mathbf{D}_2\mathbf{B}^*))^{-1}|$ defined over positive semidefinite diagonal matrices $\mathbf{D}_1, \mathbf{D}_2$ and positive number μ is a jointly log-convex function of $\mathbf{D}_1$ and $\mathbf{D}_2$ for fixed μ.

Proof: The functions indicated are all compositions of the function $h(\mathbf{C}) = -\log|\mathbf{I} + \mathbf{C}|$ and linear functions of the form $g(\mathbf{D}_1, \mathbf{D}_2) = \mathbf{A}\mathbf{D}_1\mathbf{A}^* + \mathbf{B}\mathbf{D}_2\mathbf{B}^*$. Since $h(\mathbf{C})$ is convex over positive semidefinite $\mathbf{C}$, the composition $f = h \circ g$ is also convex [20]. □

A lemma that is useful in obtaining dc decompositions of complicated functions is proved next.

Lemma 2: Let $h_i(\mathbf{x})$ and $g_i(\mathbf{x})$ be log-convex functions $\forall i = 1 \ldots, L$ over $\Re^n$ and c_i be nonnegative constants. Then $f(\mathbf{x}) = \log(\sum_i c_i \frac{g_i(\mathbf{x})}{h_i(\mathbf{x})})$ is dc and has a dc decomposition

$$\log\left(\sum_i c_i g_i(\mathbf{x}) \prod_{j \ne i} h_j(\mathbf{x})\right) - \sum_i \log h_i(\mathbf{x}). \tag{11}$$

Proof:

$$\begin{aligned} f(\mathbf{x}) &= \log\left(\sum_i c_i \frac{g_i(\mathbf{x})}{h_i(\mathbf{x})}\right) && (12) \\ &= \log\left(\frac{\sum_i c_i g_i(\mathbf{x}) \prod_{j \ne i} h_j(\mathbf{x})}{\prod_i h_i(\mathbf{x})}\right). && (13) \end{aligned}$$

The product of log-convex functions is log-convex, the sum of log-convex functions is log-convex and a positive constant times a log-convex function is log-convex [20]. Hence the argument of $\log(\cdot)$ in (13) is the ratio of log-convex functions. Therefore, a dc decomposition for $f(\mathbf{x})$ is $\log(\sum_i c_i g_i(\mathbf{x}) \prod_{j \ne i} h_j(\mathbf{x})) - \sum_i \log h_i(\mathbf{x})$. □

We consider a constellation $\boldsymbol{\mathcal{C}} = \{\boldsymbol{\Phi}_i \mathbf{V}_i\}_{i=1}^L$, which is in the form indicated in Proposition 1. If we define the positive semidefinite diagonal matrix $\mathbf{D}_i = \mathbf{V}_i^2$, the average power constraint over $\boldsymbol{\mathcal{C}}$ can be expressed as $\frac{1}{T}\sum_i P_i \operatorname{tr}(\mathbf{D}_i) \le P$, and the peak power constraint is $\frac{1}{N_t T}\operatorname{tr}(\mathbf{D}_i) \le \mathcal{K}, \forall i$. We define the constraint set over which we find the optimal constellations as

$$\boldsymbol{\mathcal{G}} = \left\{ \{P_i, \mathbf{V}_i\}_{i=1}^L : P_i \ge \epsilon, \sum_i P_i = 1, \mathbf{D}_i = \mathbf{V}_i^2 \ge \mathbf{0}, \right. \\ \left. \frac{1}{N_t T}\operatorname{tr}(\mathbf{D}_i) \le \mathcal{K} \;\; \forall i, \frac{1}{T}\sum_i P_i \operatorname{tr}(\mathbf{D}_i) \le P \right\}.$$

The number $\epsilon > 0$ may be chosen to be small, so that the optimal constellations over $\boldsymbol{\mathcal{G}}$ are as close as we please to the optimal constellations without the constraints $P_i \ge \epsilon \; \forall i$.

Theorem 1: Consider a constellation $\boldsymbol{\mathcal{C}} = \{\boldsymbol{\Phi}_i \mathbf{V}_i\}_{i=1}^L$ with corresponding transmission probabilities $\{P_i\}_{i=1}^L$. Given $\{\boldsymbol{\Phi}_i\}_{i=1}^L$, the optimal set of $\{P_i\}_{i=1}^L$ and $\{\mathbf{V}_i\}_{i=1}^L$ in $\boldsymbol{\mathcal{G}}$ may be obtained through a dc program.

Proof: Substituting $\mathbf{X}_i = \boldsymbol{\Phi}_i \mathbf{V}_i \; \forall i$ in (4), and using the identity $|\mathbf{I} + \mathbf{A}\mathbf{B}| = |\mathbf{I} + \mathbf{B}\mathbf{A}|$, the resulting expression for CR may be rewritten as

$$-\log\left(\sum_i \frac{1}{P_i^{-2}} + \sum_{\substack{i,j \\ j \ne i}} \frac{\left|\mathbf{I} + \frac{1}{2}(\boldsymbol{\Phi}_i \mathbf{D}_i \boldsymbol{\Phi}_i^* + \boldsymbol{\Phi}_j \mathbf{D}_j \boldsymbol{\Phi}_j^*)\right|^{-N_r}}{P_i^{-1} P_j^{-1} \left(|\mathbf{I} + \mathbf{D}_i|\,|\mathbf{I} + \mathbf{D}_j|\right)^{-\frac{N_r}{2}}}\right). \tag{14}$$

By Lemma 1, $|\mathbf{I} + \frac{1}{2}(\boldsymbol{\Phi}_i \mathbf{D}_i \boldsymbol{\Phi}_i^* + \boldsymbol{\Phi}_j \mathbf{D}_j \boldsymbol{\Phi}_j^*)|^{-1}$, $|\mathbf{I} + \mathbf{D}_i|^{-1}$ and $|\mathbf{I} + \mathbf{D}_j|^{-1}$ are log-convex functions of $(\mathbf{D}_i, \mathbf{D}_j)$. Also, P_i^{-2} and $P_i^{-1} P_j^{-1}$ are log-convex functions of (P_i, P_j) in $\boldsymbol{\mathcal{G}}$. Using the fact that the product of log-convex functions is log-convex and that a log-convex function raised to a positive index is log-convex, it can be seen that CR is now in the form given in Lemma 2 and, hence, a dc decomposition of the objective function is obtained.

Let D_{ik} denote the kth diagonal element of $\mathbf{D}_i \; \forall i$. The power constraint is $\frac{1}{T}\sum_i P_i \operatorname{tr}(\mathbf{D}_i) = \frac{1}{T}\sum_i P_i \sum_k D_{ik} \le P$. It can be shown that the set of $\{P_i\}_{i=1}^L$ and $\{D_{ik}\}_{i,k}$ satisfying this constraint is a nonconvex set in general. This constraint can be written equivalently as

$$\left(\sum_i P_i \sum_k D_{ik} + \sqrt{N_t}\left(\sum_i P_i^2 + \sum_{i,k} D_{ik}^2\right) - PT\right) \\ -\sqrt{N_t}\left(\sum_i P_i^2 + \sum_{i,k} D_{ik}^2\right) \le 0 \quad (15)$$

and it is shown in Appendix-B that the left-hand side is a difference of two convex functions. Therefore, it conforms to a dc constraint. The other constraints in $\boldsymbol{\mathcal{G}}$ are $\sum_i P_i = 1$ and for each i, $P_i \geq \epsilon$, $\mathbf{D}_i \geq \mathbf{0}$ and $\frac{1}{N_t T}\mathrm{tr}(\mathbf{D}_i) \leq \mathcal{K}$, which are all convex constraints and, hence, dc as well. Therefore, the maximization of the cutoff rate expression in the set $\boldsymbol{\mathcal{G}}$ can be solved by a dc program. Once the optimal $\mathbf{D}_i$ are obtained, the optimal $\mathbf{V}_i$ are obtained using $\mathbf{V}_i = \mathbf{D}_i^{1/2}$. □

An example of an algorithm that solves the dc program is the Simplical Branch and Bound algorithm (Section 4.6) [18]. Alternatively, one could convert the problem into a canonical dc program and then use the Edge Following Algorithm (Section 4.5) [18]. More algorithms may be found in [18] and [21].

The following definitions are needed for the theorem that follows.

Definition 3: A *concave minimization* problem is an optimization problem in the following form:

$$\min_{\mathbf{x} \in X} \quad f(\mathbf{x}), \tag{16}$$

where $f(\mathbf{x})$ is a concave function and $X \subset \Re^n$ is a convex set.

Definition 4: A separable function $f(\mathbf{x})$ is one which can be expressed as $f(\mathbf{x}) = \sum_i f_i(x_i)$ where $\mathbf{x} = [x_1 \; x_2 \; \cdots \; x_n]^T$.

Definition 5: A separable concave minimization program is a concave minimization program with a separable and concave objective function $f(\mathbf{x})$, where $\mathbf{x} = [x_1 \; x_2 \; \cdots \; x_n]^T$.

Definition 6: A reverse convex set is a set whose complement is an open convex set. A separable reverse convex set is a set of the form $\{\mathbf{x} : q(\mathbf{x}) \geq 0\}$, where $q(\mathbf{x})$ is a separable convex function of $\mathbf{x} = [x_1 \; x_2 \; \cdots \; x_n]^T$.

Since CR is expressed as a difference of convex functions over $\{P_i\}_{i=1}^L$ and $\{\mathbf{D}_i\}_{i=1}^L$ in Theorem 1, the same dc decomposition holds even when considered as a function of $\{\mathbf{D}_i\}_{i=1}^L$ and for fixed $\{P_i\}_{i=1}^L$. Since the power constraint $\frac{1}{T}\sum_i P_i \mathrm{tr}(\mathbf{D}_i) \leq P$ is a polytope over $\{\mathbf{D}_i\}_{i=1}^L$ for a fixed $\{P_i\}_{i=1}^L$, and the other constraint $D_{ik} \geq 0$ is linear, the optimal $\{\mathbf{D}_i\}_{i=1}^L$ may be obtained through a dc program over a polytope. We show however, that the optimization problem in this case may be simplified and transformed into other standard global optimization problems with more structure and, hence, state the following theorem separately. We first define the constraint set over which we find the optimal constellations as

$$\boldsymbol{\mathcal{H}} = \left\{ \{\mathbf{V}_i\}_{i=1}^L : \mathbf{D}_i = \mathbf{V}_i^2 \geq \mathbf{0}, \frac{1}{N_t T}\mathrm{tr}(\mathbf{D}_i) \leq \mathcal{K} \; \forall i, \right. \\ \left. \frac{1}{T}\sum_i P_i \; \mathrm{tr}(\mathbf{D}_i) \leq P \right\}.$$

Theorem 2: Consider a constellation with the structure $\{\boldsymbol{\Phi}_i\mathbf{V}_i\}_{i=1}^L$ as given in Proposition 1. Let the matrices be transmitted with probabilities $\{P_i\}_{i=1}^L$. Then, given $\{\boldsymbol{\Phi}_i\}_{i=1}^L$ and $\{P_i\}_{i=1}^L$, the optimal set of $\{\mathbf{V}_i\}_{i=1}^L$ in $\boldsymbol{\mathcal{H}}$ may be obtained through either

(i) a dc program over a polytope, or
(ii) a separable concave minimization program, or
(iii) a convex minimization program with an additional separable reverse convex constraint.

Proof: We need to maximize CR in (14) over the set $\boldsymbol{\mathcal{H}}$. This is equivalent to maximizing the following expression due to the monotonicity of $\log(1+x)$ and $\log(x)$

$$-\log\left\{2\sum_{\substack{i,j\\j>i}} P_iP_j \frac{\left|\mathbf{I} + \frac{1}{2}(\boldsymbol{\Phi}_i\mathbf{D}_i\boldsymbol{\Phi}_i^* + \boldsymbol{\Phi}_j\mathbf{D}_j\boldsymbol{\Phi}_j^*)\right|^{-N_r}}{|\mathbf{I}+\mathbf{D}_i|^{-\frac{N_r}{2}}\,|\mathbf{I}+\mathbf{D}_j|^{-\frac{N_r}{2}}}\right\} \tag{17}$$

where the argument of the $-\log(\cdot)$ may be expressed as shown at the bottom of the page. By Lemma 1, $|\mathbf{I} + \frac{1}{2}(\boldsymbol{\Phi}_i\mathbf{D}_i\boldsymbol{\Phi}_i^* + \boldsymbol{\Phi}_j\mathbf{D}_j\boldsymbol{\Phi}_j^*)|^{-1}$, $|\mathbf{I}+\mathbf{D}_i|^{-1}$ and $|\mathbf{I}+\mathbf{D}_j|^{-1}$ are log convex functions of $(\mathbf{D}_i, \mathbf{D}_j)$. A log-convex function raised to a positive index is still log-convex. Also, a positive constant times a log-convex function is log-convex. The sum and product of log-convex functions is still log convex. This implies that the last expression is the ratio of two log-convex functions. Taking $-\log(\cdot)$ therefore gives a dc decomposition for CR, which is

$$\mathrm{CR} = \log\Big(\prod_{\substack{i\\j>i}} |(\mathbf{I}+\mathbf{D}_i)^{-1}||(\mathbf{I}+\mathbf{D}_j)^{-1}|\Big)^{N_r/2} - q(\mathbf{D}_1,\ldots,\mathbf{D}_L)$$

where $q(\mathbf{D}_1,\ldots,\mathbf{D}_L)$ is jointly convex over $\{\mathbf{D}_i\}_{i=1}^L$. Let D_{ik} denote the kth diagonal element of $\mathbf{D}_i \; \forall i$. The constraint set is the intersection of $\frac{1}{T}\sum_i P_i \; \mathrm{tr}(\mathbf{D}_i) \leq P$ and the linear constraints $\{D_{ik} \geq 0\}_{i=1,k=1}^{i=L,k=N_t}$, and is, hence, a polytope. The optimal $\{\mathbf{D}_i\}_{i=1}^L$ can hence be obtained through a dc program over a polytope.

We will now express the maximization of (17) as a separable concave minimization program. We need to maximize the $-\log(\cdot)$ of the last expression, which is

$$= \max_{\substack{\frac{1}{T}\sum_i P_i\mathrm{tr}(\mathbf{D}_i)\leq P\\ \mathbf{D}_i\geq\mathbf{0}\;\forall i}} \log\Bigg(\prod_{\substack{i\\j>i}} |\mathbf{I}+\mathbf{D}_i|\,|\mathbf{I}+\mathbf{D}_j|\Bigg)^{-N_r/2} - q(\mathbf{D}_1,\ldots,\mathbf{D}_L) \tag{18}$$

$$= \max_{\substack{\frac{1}{T}\sum_i P_i\mathrm{tr}(\mathbf{D}_i)\leq P\\ \mathbf{D}_i\geq\mathbf{0}\;\forall i}} \sum_{\substack{i\\j>i}}\sum_k \left\{-\frac{N_r}{2}\log(1+D_{ik}) - \frac{N_r}{2}\log(1+D_{jk})\right\} - q(\mathbf{D}_1,\ldots,\mathbf{D}_L). \tag{19}$$

An additional variable t is now introduced to obtain the equivalent optimization problem

$$\max_{\substack{\frac{1}{T}\sum_i P_i\mathrm{tr}(\mathbf{D}_i)\leq P\\ q(\mathbf{D}_1,\ldots,\mathbf{D}_L)\leq t}} \sum_{\substack{i\\j>i}}\sum_k \left\{-\frac{N_r}{2}\log(1+D_{ik}) - \frac{N_r}{2}\log(1+D_{jk})\right\} - t \tag{20}$$

$$= \min_{\substack{\frac{1}{T}\sum_i P_i\mathrm{tr}(\mathbf{D}_i)\leq P\\ q(\mathbf{D}_1,\ldots,\mathbf{D}_L)\leq t}} t + \sum_{\substack{i\\j>i}}\sum_k \left\{\frac{N_r}{2}\log(1+D_{ik}) + \frac{N_r}{2}\log(1+D_{jk})\right\}. \tag{21}$$

Since $q(\mathbf{D}_1,\ldots,\mathbf{D}_L) - t$ is a convex function, $q(\mathbf{D}_1,\ldots,\mathbf{D}_L) - t \leq 0$ is a convex set. Since the intersection of convex sets is convex,

$$\frac{2\sum_{i,j>i} P_iP_j \left|\mathbf{I} + \frac{1}{2}(\boldsymbol{\Phi}_i\mathbf{D}_i\boldsymbol{\Phi}_i^* + \boldsymbol{\Phi}_j\mathbf{D}_j\boldsymbol{\Phi}_j^*)\right|^{-N_r} \prod_{\substack{k\neq i\\ l\neq j, l>k}} (|\mathbf{I}+\mathbf{D}_k|\,|\mathbf{I}+\mathbf{D}_l|)^{-N_r/2}}{\prod_{i,j>i}(|\mathbf{I}+\mathbf{D}_i|\,|\mathbf{I}+\mathbf{D}_j|)^{-N_r/2}}.$$

the constraint set is convex. Hence the problem in (21) is a separable concave minimization problem over $\{\mathbf{D}_i\}_{i=1}^L$ and t, since the objective function is separable concave, and the constraint set is convex.

The additional variable t may also be introduced in place of the other convex function in (19), to obtain the equivalent problem

$$\max_{\substack{\frac{1}{T}\sum_i P_i \mathrm{tr}(\mathbf{D}_i)\le P\\ t\le r(\mathbf{D}_1,\dots,\mathbf{D}_L)}} t - q(\mathbf{D}_1,\dots,\mathbf{D}_L) = \min_{\substack{\frac{1}{T}\sum_i P_i \mathrm{tr}(\mathbf{D}_i)\le P\\ t\le r(\mathbf{D}_1,\dots,\mathbf{D}_L)}} q(\mathbf{D}_1,\dots,\mathbf{D}_L) - t \quad (22)$$

where $r(\mathbf{D}_1,\dots,\mathbf{D}_L) = \sum_{\substack{i\\ j>i}}\sum_{k,n}\{-\frac{1}{2}\log(1+D_{ik}) - \frac{1}{2}\log(1+D_{jk})\}$. Since $r(\mathbf{D}_1,\dots,\mathbf{D}_L)$ is a separable convex function of $\{\mathbf{D}_i\}_{i=1}^L$, $t \le r(\mathbf{D}_1,\dots,\mathbf{D}_L)$ is a separable reverse convex constraint. Since $q(\mathbf{D}_1,\dots,\mathbf{D}_L) - t$ is a convex function, this form of the optimization problem is a convex minimization problem with an additional separable reverse convex constraint (or separable concave constraint).

Once the optimal $\{\mathbf{D}_i\}_{i=1}^L$ are found using any of the optimization techniques, the optimal $\{\mathbf{V}_i\}_{i=1}^L$ are obtained through $\mathbf{V}_i = (\mathbf{D}_i)^{1/2}$. □

All the three forms indicated in Theorem 2 are more structured global optimization problems than the one encountered in Theorem 1. The concave minimization problem over a convex set is known to have a solution at an extreme point of the constraint region and this fact is exploited in algorithms to obtain the global optimum [18]. Branch and bound algorithms to solve this problem are given in Section 3.7 of [18] and in [21]. The problem that we encounter is a separable concave minimization problem, which is a concave minimization problem with additional structure. The separability in the objective function can be exploited to further simplify the search by using the rectangular branch and bound algorithm given in [21, Sec. 3.7.4 and 3.7.5]. Another algorithm which uses a partial outer approximation and branch and bound method to solve the separable concave minimization problem is described in [22]. The formulation involving a convex minimization problem with an additional reverse convex constraint can be solved by a branch and bound algorithm given in [23].

The number of optimization variables in the problem is $L + LN_t$, and is independent of N_r. With large values of T, one may need more constellation points L to approach capacity, and this results in higher complexity. Difference of convex programming problems and concave minimization problems come under the class of nonconvex optimization problems and therefore, are in general NP-hard. This would mean that the worst case number of iterations would grow exponentially with the number of optimization variables. However, in practice depending on the problem structure, there are algorithms proposed that require smaller iterations on average to either solve the problem or obtain a good approximate solution. For relatively small constellations and transmit antennas, the algorithms mentioned may be solved with tractable complexity.

An example where the formulation using fixed probabilities is useful, is when there is a need for equiprobable constellations. Even in this case, it can be seen that in the low SNR regime many points in the optimal solution are assigned the same zero amplitude during the optimization. This is expected as a constellation with higher peak-to-average power ratio is needed in this regime to maximize the mutual information. We observe via simulations that equiprobable constellations with a suitable constellation expansion (to offset the loss of source entropy due to the zero amplitude points coinciding at the origin) can have nearly the same mutual information and CR. While the performance loss is small, the constellations designed this way are usually more regular than constellations obtained by jointly optimizing over $\{P_i\}_{i=1}^L$ and $\{\mathbf{V}_i\}_{i=1}^L$. The optimization problem in this case can also be transformed into a separable concave minimization problem which has more structure compared to a general nonconvex problem, and can, hence, be exploited to decrease the complexity. Due to these reasons, we conclude that this formulation can be quite useful practically, even when unequal probabilities are allowed. Numerical examples and comparisons for this formulation can be found in Section V-B.

C. Low Rate Constellations: $L \le \frac{T}{N_t}$

When $L \le \frac{T}{N_t}$, which makes the constellation have low rate, it is possible to globally optimize the cutoff rate jointly over $\{P_i\}_{i=1}^L$, $\{\mathbf{\Phi}_i\}_{i=1}^L$ and $\{\mathbf{V}_i\}_{i=1}^L$. It is for this case, that [13] derives the cutoff rate in closed form by first showing the optimality of scaled orthogonal matrices under a peak power constraint and we take a similar approach here. Using [13, Lemma 6], we can write the cutoff rate in the following form:

$$\mathrm{CR} = -\log\left\{\sum_i\sum_j P_iP_j \times \frac{|\mathbf{I}+\mathbf{X}_i^*\mathbf{X}_i|^{N_r/2}\,|\mathbf{I}+\mathbf{X}_j^*\mathbf{X}_j|^{N_r/2}}{|\mathbf{I}+\frac{1}{2}\mathbf{X}_i^*\mathbf{X}_i|^{N_r}\,|\mathbf{I}+\frac{1}{2}\mathbf{X}_j^*\mathbf{X}_j|^{N_r}\,|\mathbf{I}-\boldsymbol{\mathcal{F}}_{ij}^*\boldsymbol{\mathcal{F}}_{ij}|^{N_r}}\right\} \quad (23)$$

where $\boldsymbol{\mathcal{F}}_{ij} = \tilde{\mathbf{X}}_j^*\tilde{\mathbf{X}}_i$, $\tilde{\mathbf{X}}_i = \mathbf{X}_i[\mathbf{I}+\frac{1}{2}\mathbf{X}_i^*\mathbf{X}_i]^{-1/2}$ and $\tilde{\mathbf{X}}_j = \mathbf{X}_j[\mathbf{I}+\frac{1}{2}\mathbf{X}_j^*\mathbf{X}_j]^{-1/2}$. Using Proposition 1 in (23), we get

$$\mathrm{CR} = -\log\left\{\sum_i\sum_j P_iP_j \times \frac{|\mathbf{I}+\mathbf{V}_i^2|^{N_r/2}\,|\mathbf{I}+\mathbf{V}_j^2|^{N_r/2}}{|\mathbf{I}+\frac{1}{2}\mathbf{V}_i^2|^{N_r}\,|\mathbf{I}+\frac{1}{2}\mathbf{V}_j^2|^{N_r}\,|\mathbf{I}-\boldsymbol{\mathcal{F}}_{ij}^*\boldsymbol{\mathcal{F}}_{ij}|^{N_r}}\right\} \quad (24)$$

It can be seen that CR depends on $\{\mathbf{\Phi}_i\}_{i=1}^L$ only through terms of the form $|\mathbf{I}-\boldsymbol{\mathcal{F}}_{ij}^*\boldsymbol{\mathcal{F}}_{ij}|$. By the property of the positive semidefinite ordering [24] $|\mathbf{I}-\boldsymbol{\mathcal{F}}_{ij}^*\boldsymbol{\mathcal{F}}_{ij}| \le |\mathbf{I}| = 1$, and it can be seen that $\boldsymbol{\mathcal{F}}_{ij}^*\boldsymbol{\mathcal{F}}_{ij} = \mathbf{0}$ when $\{\mathbf{\Phi}_i\}_{i=1}^L$ are a set of orthogonal matrices, i.e., $\mathbf{\Phi}_i^*\mathbf{\Phi}_i = \mathbf{I}$ and $\mathbf{\Phi}_i^*\mathbf{\Phi}_j = \mathbf{0}\ \forall\, i,j$. Therefore CR is maximized for any $\{P_i\}_{i=1}^L$ and $\{\mathbf{V}_i\}_{i=1}^L$ when $\{\mathbf{\Phi}_i\}_{i=1}^L$ are a set of orthogonal matrices. This set can be obtained since $L \le \frac{T}{N_t}$. Thereafter, the optimization over $\{P_i\}_{i=1}^L$ and $\{\mathbf{V}_i\}_{i=1}^L$ may be carried out through dc programming as in Section III-B using a set of orthogonal matrices as the base constellation to obtain the globally optimal solution and the cutoff rate R_0 as defined in (3).

D. Minimum $\frac{E_b}{N_0}$ Designs

In this section, following common notation, we use $\mathrm{SNR} = \frac{E_s}{N_0}$ in place of P with E_s denoting received energy per symbol and let E_b denote energy per bit. The bit-energy normalized by the noise power is denoted by $\frac{E_b}{N_0}$ and the spectral-efficiency as $\mathrm{C}(\frac{E_b}{N_0})$ (b/s/Hz). The quantity $\frac{E_b}{N_0}$ as a function of spectral-efficiency is obtained from the Shannon capacity $\mathrm{C}(\frac{E_b}{N_0}) = C(\mathrm{SNR})$, where $\frac{E_b}{N_0} = \frac{\mathrm{SNR}}{C(\mathrm{SNR})}$. As noted in [8], the bit-energy/spectral-efficiency tradeoff is the key to understanding the tradeoff between bandwidth and power at low SNR. A key insight provided by [8] is that for the average power limited noncoherent channel, achieving the $(\frac{E_b}{N_0})_{\min}$ is very demanding in terms of bandwidth and the peak-to-average ratio of the transmitted signal. As a result, the authors in [25] consider the SISO noncoherent Rician fast fading channel under both an average and peak-to-average power constraint and characterize the bit-energy/spectral-efficiency curve for different Rician factors. In order to obtain these curves, the capacity $C(\mathrm{SNR})$ is obtained using numerical methods as there are no closed form expressions. For MIMO channels, it is much harder to obtain the bit-energy/spectral-efficiency curve as the mutual information is not known in closed form, the capacity at general SNR is not known, and numerical methods to evaluate it are also intractable. Our approach will be to obtain constellations that maximize the cutoff rate expression at a general SNR as in the previous section. From Monte-Carlo simulations, we obtain the mutual informations of the resulting constellations $I^*(\mathrm{SNR})$. Using these in $\frac{E_b}{N_0} = \frac{\mathrm{SNR}}{I^*(\mathrm{SNR})}$ and denoting the spectral-efficiency by $\mathrm{I}^*(\frac{E_b}{N_0}) = I^*(\mathrm{SNR})$ b/s/Hz, we can obtain a bit-energy/spectral-efficiency tradeoff curve using optimal constellations indicated by this design technique. The minimum normalized energy per bit obtained through this constellation design technique would be an upper bound on the $(\frac{E_b}{N_0})_{\min}$ achievable over all possible schemes. In Section V, we show through numerical simulations that interesting insights on modulation and coding can be obtained through this method.

E. Comparison With Constellations of Borran et al. [12]

In this section, we compare our design technique with another prominent work, viz. Borran *et al.* [12], which also proposes a numerical technique to design constellations at a general SNR for the noncoherent Rayleigh-fading MIMO channel. One similarity between [12] and this work is that both leverage existing unitary constellations, that are well suited for the high SNR regime, to design constellations for the low/medium SNR regime. The design criterion in [12] is different from that used here. Moreover, the symbol error probability of uncoded constellations was used as the performance measure throughout that work. In this correspondence, we use a more fundamental performance measure, viz. mutual information, and we will show that the cutoff-rate optimized constellations improve significantly upon the Borran *et al.* constellations, especially in the low-medium SNR regime.

Through many examples, it was shown in [12] that the KL-distance between constellation points is a more suitable distance measure than the Euclidean distance measure for the noncoherent MIMO Rayleigh-fading channel. Borran *et al.* [12] proposed adopting existing unitary designs in a multi-level manner, and demonstrated gains in the low/medium SNR range. For their single antenna designs, the equiprobable constellation points were constrained to lie on concentric spheres and after making a sub-optimal approximation on the allowed arrangement of points, the KL-distance between the closest points was maximized. Specifically, it is assumed that between any two concentric shells on which the constellation points are constrained to lie, there is at least one pair that is collinear with the origin. This assumption is a major drawback, since it renders the constellations strictly suboptimal and not even locally optimal in most cases, as a mild perturbation of the points along a shell typically improves the minimum distance. Even though the designs are for any specified SNR, they are different from the base unitary designs primarily in the low/medium SNR regime, where the constellation points occupy multiple levels with a point typically at the origin. The technique was extended to multiple-antenna constellations as well, while still inheriting the sub-optimal approximation.

Another major drawback of that work is that it restricts the constellation points to be equi-probable. In the crucial low SNR regime, it is well known that signals with relatively large PAPRs are required to approach capacity for the noncoherent Rayleigh fading channel. Therefore, unequal probabilities for constellation points are essential, and this is incorporated in our work resulting in significant improvement.

The performance measure used throughout in [12] is the average symbol error probability (when each block of symbols is decoded separately). However, a more fundamental performance measure is the mutual information, since it accounts for both modulation as well as channel coding across independent coherence blocks. Moreover, in the low/medium SNR regime and with relatively smaller coherence lengths, coding across blocks may be considered necessary to improve the reliablity. As a result, we use the mutual information of the constellations, which are obtained through Monte-Carlo simulations. We show via simulations that the cutoff-rate optimized constellations outperform those in [12] over the entire range of SNRs. Fig. 6 is a representative plot comparing the mutual informations. It can be seen that the gains by using cutoff-rate optimized constellations are in the crucial low/medium SNR regime, with significant gains at low SNR.

IV. Extension to Spatially Correlated Channels

In this section, we show how the main results obtained in this correspondence may be extended to more general spatially correlated channels. For simplicity, we state our results in terms of the so-called separable transmit and receive correlation model. We point out however, that the main results also hold similarly for a more general model proposed in [26], which subsumes the separable model as well as the virtual channel representation model of [27].

The received signal is modeled as

$$\mathbf{R} = \mathbf{S}\mathbf{H} + \mathbf{W} \tag{25}$$

where $\mathbf{S}$ is the transmitted symbol matrix and $\mathbf{W}$ is the noise matrix. Here, the symbols $\{\mathbf{S}\}$ are normalized so that the average transmit power constraint $\frac{1}{T}\mathrm{E}[\mathrm{tr}(\mathbf{S}\mathbf{S}^*)] \leq P$ holds. It is assumed that $\mathbf{W}$ has i.i.d. $\mathcal{CN}(0,1)$ entries. As per the separable channel model, $\mathbf{H}$ has correlated, circularly symmetric, complex Gaussian entries and is represented by

$$\mathbf{H} = \mathbf{\Sigma}_t^{1/2}\mathbf{H}_w\mathbf{\Sigma}_r^{1/2} \tag{26}$$

where $\mathbf{\Sigma}_t$ and $\mathbf{\Sigma}_r$ are the transmit and receive array correlation matrices respectively. $\mathbf{H}_w$ has i.i.d. $\mathcal{CN}(0,1)$ entries. Using the spectral decompositions $\mathbf{\Sigma}_t = \mathbf{U}_t\mathbf{\Lambda}_t\mathbf{U}_t^*$ and $\mathbf{\Sigma}_r = \mathbf{U}_r\mathbf{\Lambda}_r\mathbf{U}_r^*$, where the eigenvalues are arranged in the descending order along the diagonal, we have $\mathbf{H} = \mathbf{U}_t\mathbf{\Lambda}_t^{1/2}\mathbf{U}_t^*\mathbf{H}_w\mathbf{U}_r\mathbf{\Lambda}_r^{1/2}\mathbf{U}_r^* = \mathbf{U}_t\tilde{\mathbf{H}}\mathbf{U}_r^*$, where $\tilde{\mathbf{H}} = \mathbf{\Lambda}_t^{1/2}\mathbf{U}_t^*\mathbf{H}_w\mathbf{U}_r\mathbf{\Lambda}_r^{1/2}$. The normalizations in (25) are such that $\sum_{i=1}^{N_t}\lambda_i^t = N_t$ and $\sum_{i=1}^{N_r}\lambda_i^r = N_r$. The output of the channel in (25) can be written as

$$\mathbf{R} = \mathbf{S}\mathbf{U}_t\tilde{\mathbf{H}}\mathbf{U}_r^* + \mathbf{W} \tag{27}$$

After post-multiplying (27) by $\mathbf{U}_r$, denoting $\mathbf{S}\mathbf{U}_t$ by $\mathbf{X}$, and denoting $\mathbf{R}\mathbf{U}_r$ by $\mathbf{Y}$, we get

$$\mathbf{Y} = \mathbf{X}\tilde{\mathbf{H}} + \mathbf{N} \tag{28}$$

which represents the sufficient statistics of the received signal. Clearly, $\mathrm{E}[\mathrm{tr}(\mathbf{S}\mathbf{S}^*)] = \mathrm{E}[\mathrm{tr}(\mathbf{X}\mathbf{X}^*)]$, and, hence, the precoded code $\{\mathbf{X}\}$ satisfies the same average transmit power constraint as the original code $\{\mathbf{S}\}$. We may hence consider (28) to be our effective channel model, and use the notation $\mathbf{X}$ to denote a codeword precoded by the eigenmatrix of $\mathbf{\Sigma}_t$.

The cutoff rate expression for this channel model can be easily shown to be (cf. [28])

$$\mathrm{CR} = -\log\Bigg\{\sum_i\sum_j P_iP_j \times \prod_{n=1}^{N_r}\frac{\left|\mathbf{I}+\mathbf{X}_i\mathbf{\Lambda}_t\mathbf{X}_i^*\lambda_n^r\right|^{1/2}\left|\mathbf{I}+\mathbf{X}_j\mathbf{\Lambda}_t\mathbf{X}_j^*\lambda_n^r\right|^{1/2}}{\left|\mathbf{I}+\frac{1}{2}\left(\mathbf{X}_j\mathbf{\Lambda}_t\mathbf{X}_j^*+\mathbf{X}_i\mathbf{\Lambda}_t\mathbf{X}_i^*\right)\lambda_n^r\right|}\Bigg\}. \tag{29}$$

From (29) and (4), it is clear that the cutoff rate expressions for the i.i.d. and correlated channels have similar forms. Due to this similarity, it can be verified that results analogous to Proposition 1 and Theorems 1 and 2 are true for the correlated channel as well. For the sake of completeness but to avoid repetition, we just state the results for the spatially correlated fading channel without proofs.

Proposition 3: Let the rank of $\mathbf{\Sigma}^t$ be $M \leq N_t$ and $M^* = \min(T, M)$. Then, any constellation $\mathcal{C} = \{\mathbf{X}_i\}_{i=1}^L$ with respective priors $\{P_i\}_{i=1}^L$ may be expressed equivalently as a constellation $\mathcal{C}'$ of the form $\{\mathbf{\Phi}_i\mathbf{V}_i\}_{i=1}^L$ with the same priors such that it satisfies the same average and peak power constraints and has the same cutoff rate expression value. Here, $\mathbf{\Phi}_i \in \mathbb{C}^{T\times N_t}$ is a matrix with orthonormal columns, ie. $\mathbf{\Phi}_i^*\mathbf{\Phi}_i = \mathbf{I}\ \forall\, i = 1,\ldots,L$ and $\mathbf{V}_i \in \Re^{N_t\times N_t}$ is a nonnegative diagonal matrix with at most M^* nonzero diagonal entries. □

Consider a constellation $\mathcal{C} = \{\mathbf{\Phi}_i\mathbf{V}_i\}_{i=1}^L$, which is in the form indicated in Proposition 3. If we define the positive semidefinite diagonal matrix $\mathbf{D}_i = \mathbf{V}_i\mathbf{\Lambda}_t\mathbf{V}_i^*$, the average power constraint over $\mathcal{C}$ can be expressed as $\frac{1}{T}\sum_i P_i \mathrm{tr}(\mathbf{D}_i\mathbf{\Lambda}_t^\dagger) \leq P$, and the peak power constraint is $\frac{1}{N_tT}\mathrm{tr}(\mathbf{D}_i\mathbf{\Lambda}_t^\dagger) \leq \mathcal{K}, \forall i$. We define the constraint set $\mathcal{G}$ over which we find the optimal constellations as

$$\Bigg\{\{P_i,\mathbf{V}_i\}_{i=1}^L : P_i \geq \epsilon \sum_i P_i = 1,$$
$$\mathbf{D}_i = \mathbf{V}_i\mathbf{\Lambda}_t\mathbf{V}_i^* \geq \mathbf{0}\, \frac{1}{TN_t}\mathrm{tr}\left(\mathbf{D}_i\mathbf{\Lambda}_t^\dagger\right) \leq \mathcal{K}\, \frac{1}{T}\sum_i P_i\ \mathrm{tr}\left(\mathbf{D}_i\mathbf{\Lambda}_t^\dagger\right) \leq P\Bigg\}.$$

Again, $\epsilon > 0$ is chosen to be small, so that the optimal constellations over $\mathcal{G}$ are as close as we please to the optimal constellations without the constraints $P_i \geq \epsilon\ \forall i$.

Theorem 3: Consider a constellation $\mathcal{C} = \{\mathbf{\Phi}_i\mathbf{V}_i\}_{i=1}^L$ with corresponding transmission probabilities $\{P_i\}_{i=1}^L$. Given $\{\mathbf{\Phi}_i\}_{i=1}^L$, the optimal set of $\{P_i\}_{i=1}^L$ and $\{\mathbf{V}_i\}_{i=1}^L$ in $\mathcal{G}$ may be obtained through a dc program. □

In the fixed probability case, let the constraint set $\mathcal{H}$, over which we find the optimal constellations be

$$\Bigg\{\{\mathbf{V}_i\}_{i=1}^L : \mathbf{D}_i = \mathbf{V}_i\mathbf{\Lambda}_t\mathbf{V}_i^* \geq \mathbf{0},\, \frac{1}{N_tT}\mathrm{tr}\left(\mathbf{D}_i\mathbf{\Lambda}_t^\dagger\right) \leq \mathcal{K}\ \forall i,$$
$$\frac{1}{T}\sum_i P_i\ \mathrm{tr}\left(\mathbf{D}_i\mathbf{\Lambda}_t^\dagger\right) \leq P\Bigg\}.$$

Theorem 4: Consider a constellation with the structure $\{\mathbf{\Phi}_i\mathbf{V}_i\}_{i=1}^L$ as given in Proposition 1. Let the matrices be transmitted with probabilities $\{P_i\}_{i=1}^L$. Then, given $\{\mathbf{\Phi}_i\}_{i=1}^L$ and $\{P_i\}_{i=1}^L$, the optimal set of $\{\mathbf{V}_i\}_{i=1}^L$ in $\mathcal{H}$ may be obtained through either

i) a dc program over a polytope, or
ii) a separable concave minimization program, or
iii) a convex minimization program with an additional separable reverse convex constraint.

□

V. Numerical Examples

For the optimizations, we use TOMLABs *glcCluster* module to solve general constrained mixed-integer global optimization problems. The module uses a clustering algorithm to generate a good and large set of initial points and then conducts local searches (using module *SNOPT*) from each of them before choosing the best point. The solution is not guaranteed to be optimal, but is usually optimal or very close and is also reached more quickly compared to many other optimization programs. To obtain the mutual informations of the constellations we use Monte-Carlo simulations. The mutual information is

$$\sum_i P_i \int_{\mathbf{y}} p(\mathbf{y}/i)\log\left(\frac{p(\mathbf{y}/i)}{p(\mathbf{y})}\right) = \sum_i P_i \mathrm{E}_{\mathbf{y}/i}\left[\log\left(\frac{p(\mathbf{y}/i)}{\sum_j P_j\ p(\mathbf{y}/j)}\right)\right]. \quad (30)$$

The expectations in (30) are calculated using Monte Carlo integrations.

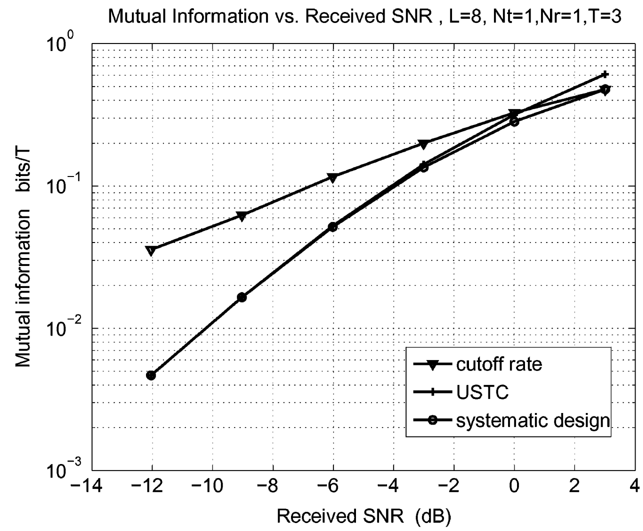

Fig. 2. Comparison of Mutual Informations: $\mathrm{L} = 8, \mathrm{Nt} = 1, \mathrm{Nr} = 1, \mathrm{T} = 3$.

In all simulations with $T > 2$, we use the corresponding systematic unitary design [9] as our base constellation. For the case when $T = 2$, we use the real two dimensional constellations in [29] for ease of representation. The reason for including many examples involving two-dimensional real constellations is for graphically representing them to illustrate some key features of the optimal constellations better. We emphasize here that any suitable unitary design (even with dimension $T > 2$, may be used and other options include [30], the recently proposed designs in [10], and the complex Givens codes in [11] for $T = 2$ which has the same low-complexity decoder as in [29] but with the advantage of using the complex dimensions as well. The advantage of using the designs in [10] is that the constellations are algebraic by construction, and QPSK modulation can be used to transmit the entries of the matrices. This makes it more convenient to design coding schemes using iterative decoding techniques. Designing suitable coding schemes for the unequiprobable constellations obtained here is a separate problem requiring further consideration.

A. General Probabilities and Amplitudes

In this subsection, we present examples of constellations obtained by optimizing jointly over probabilities $\{P_i\}_{i=1}^L$ and amplitudes $\{\mathbf{V}_i\}_{i=1}^L$.

In Fig. 2, the mutual informations of the optimal 8-point constellations for the $N_t = 1, N_r = 1, T = 3$ system are plotted at different SNRs. The USTC is also plotted alongside as a reference. It can be seen that the mutual informations of the new constellations are significantly higher than the base systematic constellations and the USTC at low SNRs. For SNRs over 3 dB, the curves corresponding to the optimized constellations and the systematic designs coincide indicating that equal probability and amplitude constellations suffice in this regime. In Fig. 3, a similar optimization is carried out on a 16-point constellation for a $N_t = 2, N_r = 2, T = 4$ system. Again, significant gains are observed at low SNRs over both the base unitary constellations and the USTC. To give examples of optimal constellations and for easy representation, we plot optimal constellations for $T = 2$ using the constellations in [29] as the base unitary constellations, over a wide range of SNRs in Figs. 7 and 8. For the low SNR examples, we introduced the restriction $\mathcal{K} = 50 \times P$. It should be noted that only two quadrants are used in plotting the constellation points since $\mathbf{X}$ and $-\mathbf{X}$ are indistinguishable using the cutoff rate expression. These examples show the smooth transition from a constellation with large peak-to-average power ratio at low SNR and having a point at zero transmitted with high probability, to a constellation resembling the base unitary

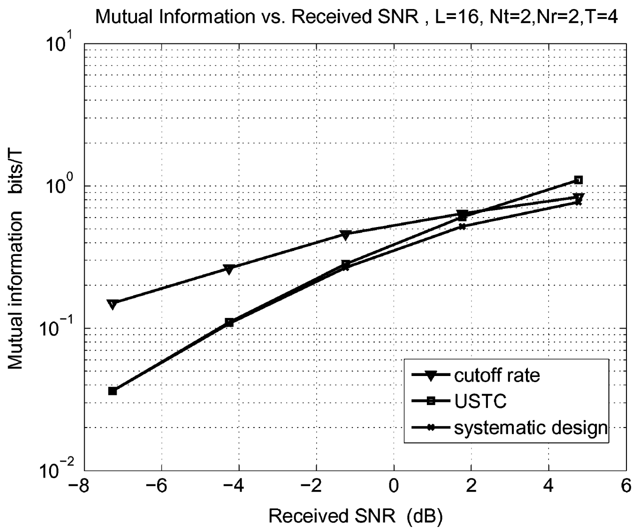

Fig. 3. Comparison of Mutual Informations: $\mathrm{L}=16, \mathrm{Nt}=2, \mathrm{Nr}=2, \mathrm{T}=4$.

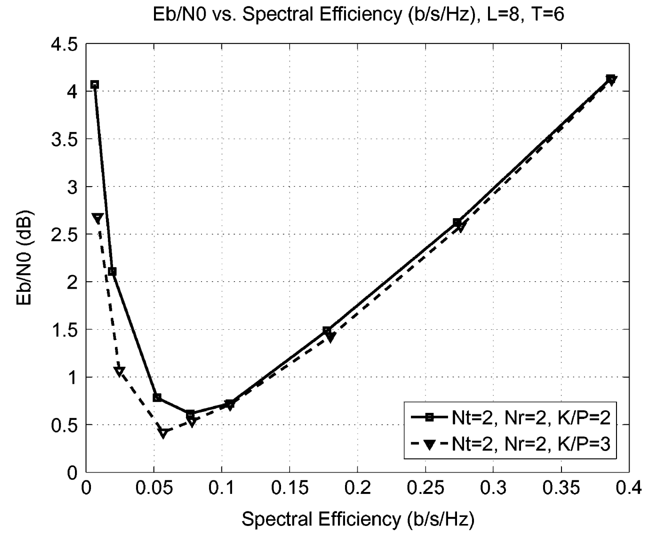

Fig. 5. $\mathrm{Eb}/\mathrm{N0}$ versus spectral efficiency.

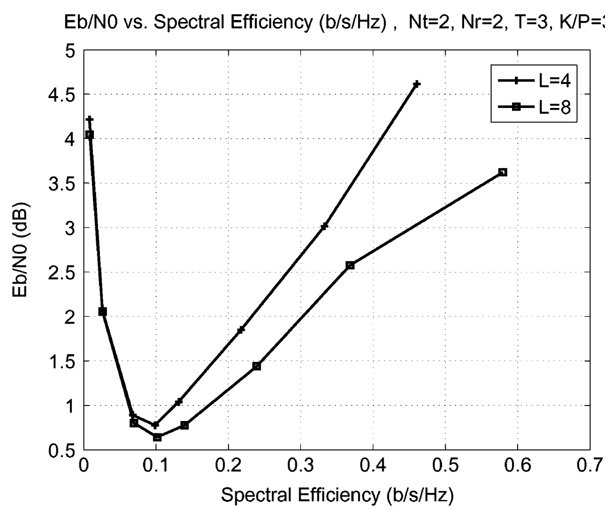

Fig. 4. $\mathrm{Eb}/\mathrm{N0}$ versus spectral efficiency.

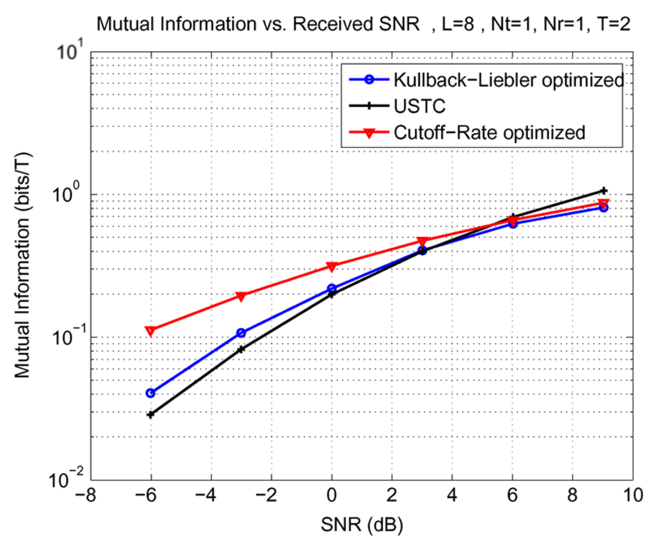

Fig. 6. Mutual Information versus SNR.

constellation with equiprobable points at relatively high SNR. In some of the plots, not all points are explicitly plotted since they are either assigned a zero probability of transmission, or a zero amplitude which results in multiple points coinciding at the origin. The optimization selects the best subset of $\{\mathbf{\Phi}_i\}_{i=1}^{L}$ to transmit along, while nulling out the rest (by assigning zero amplitudes).

For average power limited channels with a fixed peak-power constraint, the bit energy required for reliable communications decreases monotonically with decreasing spectral efficiency. Hence, the minimum bit energy is achieved at zero spectral efficiency. However, under an average power constraint P and when the peak power $\mathcal{K}$ is a constant times P, it can be verified that certain regularity conditions given in [4] are satisfied. The main result in [4] states that the capacity in such a case would grow as $O(\mathrm{SNR}^2)$. As a result, $(\frac{E_b}{N_0})_{\min}$ would occur at a nonzero spectral efficiency I_0^*. This choice of the peak-power constraint has the effect of limiting the peak-to-average power ratio of the optimal constellations. Fig. 4 gives some typical bit-energy/spectral-efficiency plots wherein the minimum energy per bit occurs at a nonzero spectral efficiency. It can be seen that the bit-energy increases with decreasing spectral-efficiency for all spectral efficiencies less than I_0^*. Hence in this regime and for a fixed rate, increasing the bandwidth increases the minimum bit energy. This is undesirable and hence operating in this region should be avoided. Indeed, every $\frac{E_b}{N_0} > (\frac{E_b}{N_0})_{\min}$ would correspond to two constellations which have spectral efficiencies I_1^* and I_2^* with $\mathrm{I}_1^* > \mathrm{I}_0^*$ and $\mathrm{I}_2^* < \mathrm{I}_0^*$. Clearly, the constellation corresponding to I_1^* should be used. Similar insights are obtained in [25] by analyzing the bit-energy/spectral-efficiency tradeoff for the SISO Rician fast-fading channel.

In Fig. 4, the optimized constellations for $N_t = N_r = 2$ and $T = 3$ are obtained subject to the average and peak power constraints as in Proposition 1 with $\mathcal{K} = 3 \times P$. It can be seen that for $L = 4$, the curve has a minimum at $\mathrm{I}_0^* \approx 0.098$, with $(\frac{E_b}{N_0})_{\min} \approx 0.778$ dB and the SNR ≈ -9.29 dB. The minimum bit-energy is lower for the eight-point constellation and occurs at $(\frac{E_b}{N_0})_{\min} \approx 0.642$ dB, indicating that it is better to use more constellation points for power-efficiency.

Fig. 5 indicates that as the ratio $\mathcal{K}/P$, which is a measure of the peak-to-average power ratio increases, the minimum bit-energy and I_0^* also decreases. This indicates that a peakier signal is more power efficient when operating at appropriately low spectral efficiencies. Since the curves merge above a certain spectral-efficiency, this indicates that in that regime, the least peaky signal may be used without any loss in

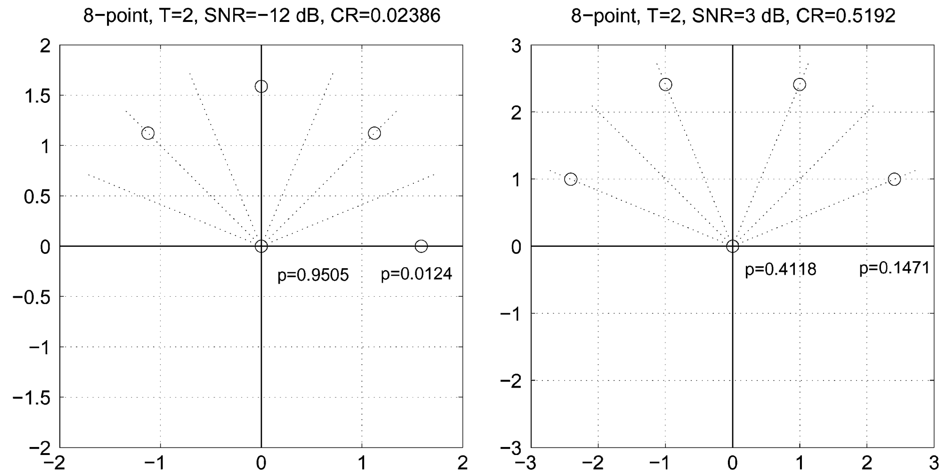

Fig. 7. Constellations obtained through the unequiprobable formulation for eight points at SNRs − 12 and 3 dB.

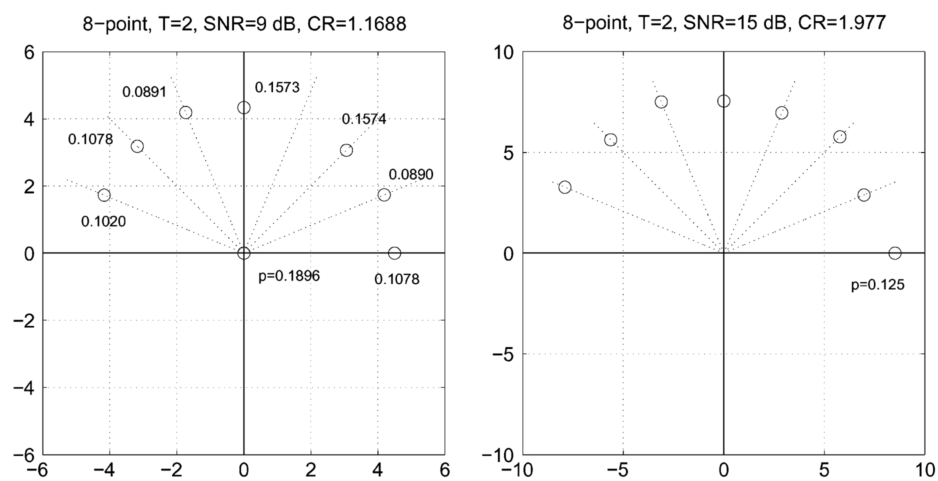

Fig. 8. Constellations obtained through the unequiprobable formulation for eight points at SNRs 9 and 15 dB.

power-efficiency. Fig. 6 is a representative plot comparing the mutual informations of the cutoff-rate optimized constellations against those in Borran *et al.*. [12]. It can be seen that the gains by using cutoff-rate optimized constellations are in the crucial low/medium SNR regime, with significant gains at low SNR.

B. Equiprobable Constellations

For the examples in this section, we obtain optimal constellations assuming that the transmission probability of each point is $1/L$, $\mathcal{K} = 50 \times P$, and optimizing over the amplitudes. Even in this case, it can be seen from Figs. 9 and 10 that in the low SNR regime, many of the amplitudes are assigned the value 0 and all these points coincide at the origin. Such a solution provides the constellation as high a peak-to-average power as possible and is hence intuitively expected for the low-medium SNR range. As in the unequal probability formulation, one can see the smooth transition in the optimal constellations from the low SNR regime to the high SNR regime in Figs. 9 and 10. In some of the plots, not all 16 points are explicitly plotted since multiple points are assigned to a zero amplitude and these points coincide at the origin and yield one point with correspondingly higher probability.

With a suitable constellation expansion to compensate for the resulting loss in source entropy, we find through numerical simulations that constellations designed this way can be nearly as good as constellations designed by optimizing jointly over probabilities and amplitudes. The advantage of using the equiprobable formulation is that at certain SNRs more regular constellations can be obtained as compared to those obtained by optimizing both $\{P_i\}_{i=1}^{L}$ and $\{\mathbf{V}_i\}_{i=1}^{L}$ jointly with nearly the same cutoff rate expression values. In fact, the optimizations indicate that only the point at zero would have a higher probability than $1/L$ since many points become coincident there, while every other point would have a probability of $1/L$. The resulting constellation hence has at most L points with the same probability, while usually just the zero point with a different probability. This simpler asymmetry in the constellation may help to significantly simplify design of coding schemes over those constellations with general pmfs.

VI. Conclusion

We considered a discrete input and continuous output MIMO channel and proposed a constellation design technique at general SNR. This technique leverages existing dense unitary designs to generate constellations for general SNRs using the cutoff rate expression as the design criterion. The mutual informations of constellations thus designed exhibit significant to moderate improvements in the low-medium SNR regime over the base unitary designs as well as the USTC. The mutual informations generated by the optimal constellations thus provide a good benchmark and the tightest known lower bound on the capacity of noncoherent MIMO channels in the low-medium SNR range. Using the mutual informations of these

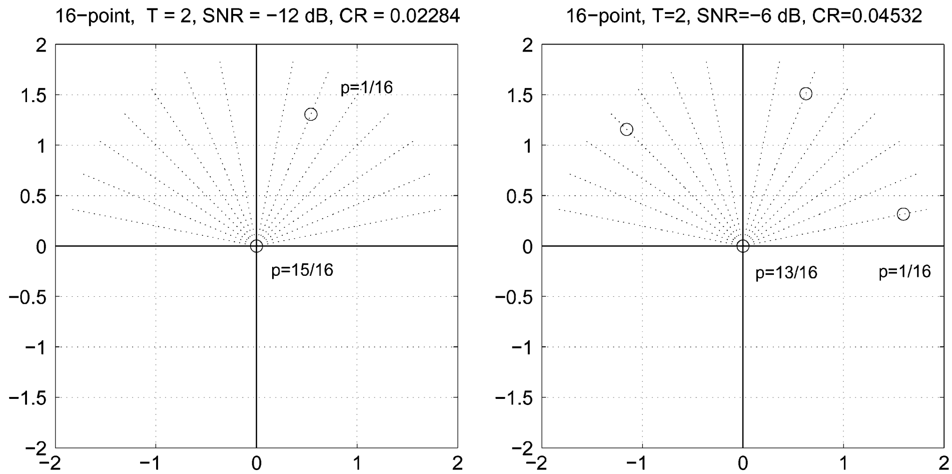

Fig. 9. Constellations obtained through the equiprobable formulation for 16 points at SNRs -12 and -6 dB.

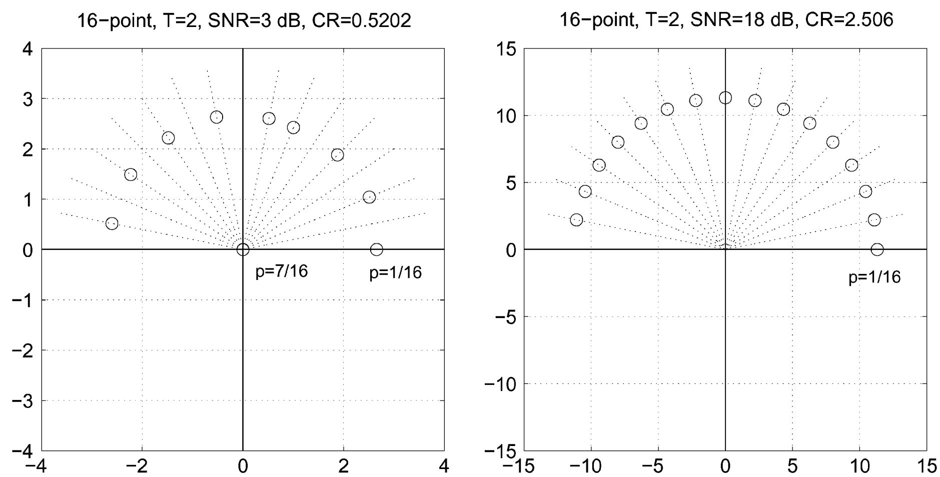

Fig. 10. Constellations obtained through the equiprobable formulation for 16 points at SNRs 3 and 18 dB.

optimal constellations, we also obtain bit-energy/spectral-efficiency tradeoff curves from which insights on the tradeoff between bandwidth, rate and bit-energy can be obtained. From such tradeoff curves, we can also obtain a firm upper bound on the minimum bit-energy required for reliable communication, and good estimates of the ideal spectral efficiency region to operate in.

APPENDIX

A. Properties of the Cutoff Rate

In this Appendix, we prove that the cutoff rate is an increasing function of SNR and N_r. In order to do so, we need the following lemma which is an interesting result on its own.

Lemma 3: $f(\mathbf{A}) = \mathrm{tr}((\mathbf{I}+\mathbf{A})^{-1}\mathbf{A})$ is concave over positive semidefinite $\mathbf{A}$.

Proof: $f(\mathbf{A})$ is concave iff $g(t) = f(\mathbf{A}+t\mathbf{B})$ is concave over $t \geq 0$ [20], with $\mathbf{B}$ being positive semidefinite. We have

$$\begin{aligned} g(t) &= \mathrm{tr}((\mathbf{I}+\mathbf{A}+t\mathbf{B})^{-1}(\mathbf{A}+t\mathbf{B})) \quad (31)\\ &= \mathrm{tr}((\mathbf{I}+\mathbf{A})^{-1/2}(\mathbf{I}+t(\mathbf{I}+\mathbf{A})^{-1/2}\\ &\quad \times \mathbf{B}(\mathbf{I}+\mathbf{A})^{-1/2})^{-1}(\mathbf{I}+\mathbf{A})^{-1/2}\mathbf{A})\\ &\quad + \mathrm{tr}((\mathbf{I}+\mathbf{A})^{-1/2}(\mathbf{I}+t(\mathbf{I}+\mathbf{A})^{-1/2}\\ &\quad \times \mathbf{B}(\mathbf{I}+\mathbf{A})^{-1/2})^{-1}(\mathbf{I}+\mathbf{A})^{-1/2}t\mathbf{B}) \quad (32)\\ &= \mathrm{tr}((\mathbf{I}+\mathbf{A})^{-1/2}\mathbf{A}(\mathbf{I}+\mathbf{A})^{-1/2}\\ &\quad \times (\mathbf{I}+t(\mathbf{I}+\mathbf{A})^{-1/2}\mathbf{B}(\mathbf{I}+\mathbf{A})^{-1/2})^{-1})\\ &\quad + \mathrm{tr}(t(\mathbf{I}+\mathbf{A})^{-1/2}\mathbf{B}(\mathbf{I}+\mathbf{A})^{-1/2}\\ &\quad \times (\mathbf{I}+t(\mathbf{I}+\mathbf{A})^{-1/2}\mathbf{B}(\mathbf{I}+\mathbf{A})^{-1/2})^{-1}). \quad (33) \end{aligned}$$

Let $(\mathbf{I}+\mathbf{A})^{-1/2}\mathbf{B}(\mathbf{I}+\mathbf{A})^{-1/2} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*$ and $(\mathbf{I}+\mathbf{A})^{-1/2}\mathbf{A}(\mathbf{I}+\mathbf{A})^{-1/2} = \mathbf{C}$. Substituting these in (33) and simplifying, we get

$$\begin{aligned} g(t) &= \mathrm{tr}(\mathbf{C}\mathbf{U}(\mathbf{I}+t\mathbf{\Lambda})^{-1}\mathbf{U}^*)\\ &\quad + \mathrm{tr}(t\mathbf{U}\mathbf{\Lambda}\mathbf{U}^*\mathbf{U}(\mathbf{I}+t\mathbf{\Lambda})^{-1}\mathbf{U}^*) \quad (34)\\ &= \mathrm{tr}(\mathbf{U}^*\mathbf{C}\mathbf{U}(\mathbf{I}+t\mathbf{\Lambda})^{-1})\\ &\quad + \mathrm{tr}(t\mathbf{\Lambda}(\mathbf{I}+t\mathbf{\Lambda})^{-1}) \quad (35)\\ &= \sum_i \frac{[\mathbf{U}^*\mathbf{C}\mathbf{U}]_{ii}}{1+t\lambda_i} + t\sum_i \frac{\lambda_i}{1+t\lambda_i}. \quad (36) \end{aligned}$$

It can be seen that $\frac{d^2 g}{dt^2} = \frac{-2\lambda_i^2(1-[\mathbf{U}^*\mathbf{C}\mathbf{U}]_{ii})}{(1+t\lambda_i)^3}$. If the eigen value decomposition of $\mathbf{A} = \mathbf{W}\mathbf{\Theta}\mathbf{W}^*$, then

$$\begin{aligned} [\mathbf{U}^*\mathbf{C}\mathbf{U}]_{ii} &= [\mathbf{U}^*\mathbf{W}\mathbf{\Theta}(\mathbf{I}+\mathbf{\Theta})^{-1}\mathbf{W}^*\mathbf{U}]_{ii} \quad (37)\\ &= \sum_k \frac{\theta_k}{1+\theta_k}|[\mathbf{U}^*\mathbf{W}]_{ik}|^2 \quad (38) \end{aligned}$$

$$\leq \sum_k |[\mathbf{U}^* \mathbf{W}]_{ik}|^2 \tag{39}$$
$$= 1. \tag{40}$$

Therefore, $\frac{d^2 g}{dt^2} \leq 0$ and hence, g is concave. Therefore, f is concave.□

Lemma 4: Let $f(\mathbf{A},\mathbf{B}) = \frac{|\mathbf{I}+\gamma\mathbf{A}|^{1/2}|\mathbf{I}+\gamma\mathbf{B}|^{1/2}}{|\mathbf{I}+\frac{\gamma}{2}(\mathbf{A}+\mathbf{B})|}$ where $\mathbf{A}$ and $\mathbf{B}$ are positive semidefinite. Then,

(i) $f(\mathbf{A},\mathbf{B}) \leq 1$ with equality iff $\mathbf{A} = \mathbf{B}$.
(ii) $f(\mathbf{A},\mathbf{B})$ is a decreasing function of γ.

Proof:

(i) Follows by the strict concavity of $\log|\mathbf{I} + \mathbf{S}|$ over positive semidefinite matrices $\mathbf{S}$.
(ii) $\log(f(\mathbf{A},\mathbf{B})) = \frac{1}{2}\log|\mathbf{I}+\gamma\mathbf{A}| + \frac{1}{2}\log|\mathbf{I}+\gamma\mathbf{B}| - \log|\mathbf{I} + \frac{\gamma}{2}(\mathbf{A}+\mathbf{B})|$. Using the standard formula $\frac{d\log|\mathbf{X}+t\mathbf{Y}|}{dt} = \mathrm{tr}((\mathbf{X}+t\mathbf{Y})^{-1}\mathbf{Y})$ to differentiate $\log(f(\mathbf{A},\mathbf{B}))$ with respect to γ, we get $\gamma\frac{d\log(f(\mathbf{A},\mathbf{B}))}{d\gamma}$

$$\begin{aligned} &= \frac{1}{2}\mathrm{tr}\left((\mathbf{I}+\gamma\mathbf{A})^{-1}\gamma\mathbf{A}\right) + \frac{1}{2}\mathrm{tr}((\mathbf{I}+\gamma\mathbf{B})^{-1}\gamma\mathbf{B}) \\ &\quad - \mathrm{tr}\left(\left(\mathbf{I}+\frac{\gamma}{2}(\mathbf{A}+\mathbf{B})\right)^{-1}\frac{\gamma}{2}(\mathbf{A}+\mathbf{B})\right) \end{aligned} \tag{41}$$
$$\leq 0 \tag{42}$$

by the concavity of $\mathrm{tr}((\mathbf{I}+\mathbf{A})^{-1}\mathbf{A})$ over positive semidefinite matrices $\mathbf{A}$ from Lemma 3. Hence, by the monotonicity of $\log(\cdot)$, f is a decreasing function of γ. □

In (3), we may normalize all constellation matrices by $\sqrt{P}$, so that the argument of $\log(\cdot)$ consists of terms of the from in Lemma 4 with $\gamma = P$. Hence by Lemma 4, we can see that the cutoff rate expression is an increasing function of P. Moreover, increasing N_r only decreases each term in the argument of $-\log(\cdot)$ in (3) by Lemma 4 (i). Therefore, the cutoff rate expression is an increasing function of N_r.

B. DC Decomposition of Power Constraint

Consider the function $g = \sum_i P_i \sum_k D_{ik} + b(\sum_i P_i^2 + \sum_{ik} D_{ik}^2)$ and the vector $\mathbf{h} = [P_1\ P_2\ \ldots\ P_L\ \mathbf{d}_1^T\ \ldots\ \mathbf{d}_L^T]$, where $\mathbf{d}_i = [D_{i1}\ D_{i2}\ \ldots\ D_{iN_t}]^T$. The function $b(\sum_i P_i^2 + \sum_{ik} D_{ik}^2)$ is a convex function of the vector $\mathbf{h}$ for any positive constant b. Define $\mathbf{a} = [1\ 1\ \ldots\ 1]^T$. The Hessian of this function g with respect to the vector $\mathbf{h}$ is given by

$$\begin{bmatrix} 2b\,\mathbf{I}_{L\times L} & \mathbf{I}_{L\times L}\otimes\mathbf{a}^T \\ \mathbf{I}_{L\times L}\otimes\mathbf{a} & 2b\,\mathbf{I}_{LN_t\times LN_t} \end{bmatrix}.$$

Since for any $b > 0$, $2b\,\mathbf{I}_{L\times L}$ is positive definite, the Hessian is positive semidefinite iff the Schur-complement of the Hessian with respect to $2b\,\mathbf{I}_{L\times L}$ is positive semidefinite. The Schur-complement is

$$2b\,\mathbf{I}_{LN_t\times LN_t} - (\mathbf{I}_{L\times L}\otimes\mathbf{a})\frac{1}{2b}\mathbf{I}_{L\times L}(\mathbf{I}_{L\times L}\otimes\mathbf{a}^T)$$
$$= 2b\,\mathbf{I}_{L\times L}\otimes\mathbf{I}_{N_t\times N_t} - \frac{1}{2b}(\mathbf{I}_{L\times L}\otimes\mathbf{a}\mathbf{a}^T) \tag{43}$$
$$= \mathbf{I}_{L\times L}\otimes\left\{2b\mathbf{I}_{N_t\times N_t} - \frac{1}{2b}\mathbf{a}\mathbf{a}^T\right\}. \tag{44}$$

Since the Kronecker product of two positive semidefinite matrices is positive semidefinite, it is sufficient if $2b\mathbf{I}_{N_t\times N_t} - \frac{1}{2b}\mathbf{a}\mathbf{a}^T$ is positive semidefinite. The only nonzero eigenvalue of $\mathbf{a}\mathbf{a}^T$ is $\mathbf{a}^T\mathbf{a}$ and therefore the Schur complement is Hermitian and has all eigenvalues positive if $4b^2 > \mathbf{a}^T\mathbf{a}$, or $b > \frac{1}{2}\sqrt{N_t}$, which makes the Schur-complement positive definite. Therefore, $\{\sum_i P_i \sum_k D_{ik} + b(\sum_i P_i^2 + \sum_{ik} D_{ik}^2)\} - \{b(\sum_i P_i^2 + \sum_{ik} D_{ik}^2)\}$ is the difference of convex functions for $b > \frac{1}{2}\sqrt{N_t}$. In the proof of Theorem 1 we take b to be $\sqrt{N_t}$.

REFERENCES

[1] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.
[2] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.
[3] X. Wu and R. Srikant, "MIMO Channels in the Low SNR Regime: Communication Rate, Error Exponent and Signal Peakiness," in *Proc. IEEE Int. Inf. Theory Workshop*, San Antonio, TX, Oct. 2004.
[4] C. Rao and B. Hassibi, "Analysis of multiple-antenna wireless links at Low SNR," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 2123–2130, Sep. 2004.
[5] L. Zheng, D. N. C. Tse, and M. Medard, "Channel coherence in the low SNR regime," *IEEE Trans. Inf. Theory*, 2005, submitted for publication.
[6] S. G. Srinivasan and M. K. Varanasi, "STORM: Optimal constellations for noncoherent MIMO communications at low SNR under PAPR constraints," in *Proc. 44th Annu. Allerton Conf. Communications, Control and Computing*, Monticello, IL, Sep. 2006.
[7] B. Hassibi and T. L. Marzetta, "Multiple-antennas and isotropically-random unitary inputs: The received signal density in closed-form," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1473–1484, June 2002, Special Issue on Shannon Theory: Perspective, Trends, and Applications.
[8] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002, Special Issue on Shannon Theory: Perspective, Trends, and Applications.
[9] 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, Sep. 2000.
[10] A. Ashikhmin and A. R. Calderbank, "Space-Time Reed-Muller Codes for Noncoherent MIMO Transmission," in *Proc. IEEE Int. Symp. Inf. Theory*, Adelaide, Australia, Sep. 2005.
[11] P. Dayal, M. Brehler, and M. K. Varanasi, "Leveraging coherent space-time codes for noncoherent communication via training," *IEEE Trans. Inf. Theory*, pp. 2058–2080, Sept. 2004.
[12] M. J. Borran, A. Sabharwal, and B. Aazhang, "On design criteria and construction of noncoherent space-time constellations," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2332–2351, Oct. 2003.
[13] A. O. Hero, III and T. L. Marzetta, "Cut-off rate and signal design for the Rayleigh fading space—time channel," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2400–2416, Sep. 2001.
[14] I. C. Abou-Faycal, M. D. Trott, and S. Shamai (Shitz), "The capacity of discrete-time memoryless Rayleigh fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.
[15] M. C. Gursoy, V. Poor, and S. Verdu, "The noncoherent Rician fading channel—Part I: Structure of the capacity-achieving input," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2193–2206, Sep. 2005.
[16] T. H. Chan, S. Hranilovic, and F. R. Kschischang, "Capacity-achieving probability measure for conditionally gaussian channels with bounded inputs," *IEEE Trans. Inf. Theory*, vol. 51, pp. 2073–2088, Jun. 2005.
[17] J. Huang and S. P. Meyn, "Characterization and computation of optimal distributions for channel coding," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2336–2351, Jul. 2005.
[18] R. Horst, P. M. Pardalos, and N. V. Thoai, *Introduction to Global Optimization*. Boston, MA: Kluwer, 2000.
[19] C. A. Floudas, *Deterministic Global Optimization*. Boston, MA: Kluwer, 2000.
[20] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[21] R. Horst and H. Tuy, *Global Optimization*. New York: Springer, 1996.
[22] H. P. Benson, "Separable concave minimization via partial outer approximation and branch and bound," *Oper. Res. Lett.*, vol. 9, no. 6, pp. 389–394, 1990.
[23] H. Tuy, "Convex programs with an additional reverse convex constraint," *J. Optimiz. Theory Applicat.*, vol. 52, no. 3, pp. 463–486, Mar. 1987.
[24] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1993.
[25] M. C. Gursoy, H. V. Poor, and S. Verdú, "Noncoherent Rician fading channel—Part II: Spectral efficiency in the low-power regime," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2207–2221, Sep. 2005.

[26] J. H. Kotecha and A. M. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 52, no. 2, pp. 546–557, Feb. 2004.
[27] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.
[28] S. G. Srinivasan and M. K. Varanasi, "Optimal correlations for the noncoherent MIMO Rayleigh fading channel," *IEEE Trans. Wireless Commun.*, submitted for publication.
[29] V. Tarokh and I.-M. Kim, "Existence and construction of noncoherent unitary space–time codes," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3112–3117, Dec. 2002.
[30] M. L. McCloud, M. Brehler, and M. K. Varanasi, "Signal design and convolutional coding for noncoherent space—time communication on the block-Rayleigh-fading channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1186–1194, May 2002.

Single- and Multiple-Antenna Constellations for Communication Over Unknown Frequency-Selective Fading Channels

Jochen Giese and Mikael Skoglund, *Senior Member, IEEE*

***Abstract*—Data transmission through frequency-selective block fading channels is considered in the case where neither the transmitter nor the receiver has any knowledge of the channel coefficients. Standard code design approaches for this scenario take channel uncertainty at the receiver into account by splitting the available channel coherence time into a part dedicated to training symbols utilized for channel estimation and a second part using an error-control coding scheme that is designed without channel uncertainty in mind. In contrast, in this correspondence joint codes are designed that are optimized for communication over the unknown channel and operate over the full coherence time. Using an approximation of the union bound on codeword error probability as design criterion, codes based on general complex-valued symbols are obtained with a gradient search optimization technique. Numerical examples for both single antenna as well as multiple-antenna systems illustrate that significant improvement over training-based schemes can be obtained.**

***Index Terms*—Code design, diversity, fading channels, frequency-selective channels, joint channel estimation and data detection, multiple antennas, wireless communication.**

I. Introduction

One goal of code design for communication over noisy channels is to minimize the impact of additive noise which is unknown to both transmitter and receiver. Very powerful error-control coding schemes [1], [2] have been developed to reduce the probability of error at the receiver in the presence of noise. If apart from noise the channel exhibits fading, the system and code design become significantly more complicated, especially if the fading parameters are unknown to transmitter and receiver. Several approaches have been proposed in the literature in order to take into account the uncertainty of channel parameters. A traditional method is to try to estimate the channel parameters using a so-called pilot or training sequence of symbols which are known to both transmitter and receiver [3]. The output of the channel is then used in a known-input known-output algorithm for channel identification and subsequent decoding stages utilize the now (approximately) known channel parameters. Alternative approaches use *blind* algorithms [4] for unknown-input–known-output system identification or *semiblind* methods [5] which use a combination of pilot symbols and data in order to produce channel estimates. Yet another approach is differential modulation where the information is coded in the *difference* between two subsequent codewords or symbols.

The approaches mentioned above basically try to reduce the problem of data transmission through an unknown channel to the problem of data communication through a channel where the channel coefficients are (at least approximately) known to the receiver, resulting in a situation where signal design issues are, so far, much better understood. Some information-theoretic justification for the approach based on training sequences has been provided recently by [3] with similar results in [6]–[8] where it is stated that the usage of optimized pilot sequences implies no or insignificant loss in capacity, at least in the high signal-to-noise ratio (SNR) domain. We emphasize that an argumentation based on capacity results essentially requires that no delay constraints are imposed on the system, because infinite block-lengths are needed to achieve capacity. Thus, in real-time applications with strict delay constraints, the use of separate coding and training has no solid theoretical backup to claim general optimality. In particular for short blocks, training might use up too much of the available coherence time such that the error-correcting code for the data is necessarily short and thus, potentially, weak compared to a code that is designed to exploit the full coherence time while taking channel uncertainty into account.

The above argumentation provides motivation for code design explicitly optimized for the communication over unknown channels. In particular for systems with multiple antennas at the transmitter and at the receiver, the overhead for separated training can become substantial because the channel between each pair of transmitter and receiver antennas has to be estimated. Therefore, designs avoiding separated training are particularly attractive for multiple-antenna systems. The considerations and designs for this scenario in, e.g., [9]–[13] are targeted toward frequency-flat fading channels. When the channel is frequency-selective, the number of unknown channel parameters grows further and therefore the advantages of designing a transmission scheme that avoids separated training is even more apparent. The present correspondence contains a framework for the code design in the general case of multiple antenna systems and unknown frequency-selective channels. We note here that the transmission of data over unknown frequency-selective channels was also discussed in [14] as well as [15], [16] in the context of orthogonal frequency division multiplexing (OFDM) transmission.

In earlier work [17], [18], we designed codes for the transmission over unknown frequency-selective channels based on a binary phase-shift keying (BPSK) modulation alphabet using combined channel decoding and channel estimation at the receiver. The design criterion was based on the union bound on exact error probability. Good codes were found using the method of simulated annealing. In the present correspondence, we drop the restriction of BPSK modulation and allow arbitrary complex-valued symbols in the codewords. The resulting single and multiple antenna signal constellations can thus be interpreted as a *joint coding and modulation* design for a given *joint data detector and channel estimator* at the receiver. Optimization is carried out using a

Manuscript received June 4, 2003; revised December 23, 2005. The material in this correspondence was presented in part at ICC 2003, Anchorage, AS, May 2003 and ISIT 2003, Yokohama, Japan, May 2003.

J. Giese was with the Royal Institute of Technology, Stockholm, Sweden. He is now with Qualcomm CDMA Technologies GmbH, 90408 Nuremberg, Germany (e-mail: giese@ee.kth.se).
M. Skoglund is with the School of Electrical Engineering, Royal Institute of Technology, SE-1000 44 Stockholm, Sweden (e-mail: skoglund@ee.kth.se).

Communicated by R. Urbanke, Associate Editor for Coding Techniques.
Digital Object Identifier 10.1109/TIT.2007.892779