

Design of optimal lattice space-time codes

N. Prasad and I. Berenguer and X. Wang and M. Madihian
 e-mail: {prasad,inaki,wangx,madiahian}@nec-labs.com
 NEC Labs America, Princeton, NJ 08540

Abstract — In this paper we propose a systematic procedure for designing optimal lattice (space-time) codes. By employing stochastic optimization techniques we design lattice codes with minimum error rates when lattice decoders are employed at the receiver. Our design methodology can be tailored to obtain optimal lattice (space-time) codes for any fading statistics and SNR of interest. Further, we obtain fundamental lower bounds on the error probabilities yielded by lattice decoders and characterize their asymptotic behavior.

I. INTRODUCTION

Space-time block code (STBC) design for wireless fading channels has been the focus of intensive research for the past several years. As a result, several powerful STBCs such as orthogonal designs [1] and linear dispersion (LD) codes [2] have been discovered. Algebraic number theoretic tools for code design have also been employed for the i.i.d. Rayleigh fading model with great success [3]. Recently, [4] used the real-baseband model to show that all STBCs proposed in the literature are in-fact lattice codes. This reveals that the traditional STBC design where input information symbols are drawn from QAM constellations (or equivalently PAM in the real representation) result in lattice codes with sub-optimum (in terms of energy efficiency) shaping regions. Thus, there is a possibility to further improve performance by designing lattice codes with optimized shaping regions. On the other hand, a benefit of fixing input information symbols to be QAM symbols is efficient maximum-likelihood (ML) decoding via the sphere-decoder [5]. Unfortunately the complexity of ML decoding can significantly increase for lattice codes with optimized shaping due to the problem of boundary control [4]. One way to balance this trade-off is to employ lattice decoders, which avoid boundary control and hence the increase in complexity, to decode optimized lattice codes. Thus an interesting and important problem which we address here, is the design of optimal (in terms of error-rate) lattice codes for MIMO systems where the receiver employs lattice decoders. We note that no such systematic design procedure has been proposed previously and the examples given in [4] were obtained through random search.

One of the main problems in our quest for optimal lattice codes is that obtaining closed-form objective functions needed for deterministic optimization or other analytical techniques seem intractable, even for the simple albeit impractical i.i.d Rayleigh fading model. For such class of problems a promising approach is to use stochastic optimization based on the well known Robbins-Monro algorithm [6]. We adopt this technique and propose several formulations that can be used to obtain optimal lattice codes for arbitrary fading statistics and any SNR of interest.

II. SYSTEM DESCRIPTIONS

Consider an M -transmit N -receive multiple-input multiple-output (MIMO) channel with no channel state information (CSI) at the transmitter and perfect CSI at the receiver. The wireless channel is assumed to be quasi-static and flat fading and can be represented by an $N \times M$ matrix \mathbf{H}^c , which is assumed to remain fixed for $t = 1, \dots, T$. The complex-baseband model of the received signal can be expressed as

$$\mathbf{y}_t^c = \sqrt{\frac{\rho}{M}} \mathbf{H}^c \mathbf{x}_t^c + \mathbf{w}_t^c, \quad (1)$$

where $\sqrt{\frac{\rho}{M}} \mathbf{x}_t^c \in \mathbb{C}^M$ is the transmitted signal at time t , $t = 1, \dots, T$, $\mathbf{y}_t^c \in \mathbb{C}^N$ is the received signal and $\mathbf{w}_t^c \in \mathbb{C}^N$ denotes the i.i.d. circularly symmetric Gaussian noise, with complex Gaussian elements of zero mean and unit variance, i.e., $\mathbf{w}_t^c \sim \mathcal{CN}(0, \mathbf{I})$. The random variables in \mathbf{H}^c are assumed to be drawn from some continuous joint distribution.

The equivalent real-valued channel model corresponding to (1) can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \quad (2)$$

where $\mathbf{x} = [\mathbf{x}_1^T, \dots, \mathbf{x}_T^T]^T \in \mathbb{R}^{2MT}$ is a codeword belonging to a codebook \mathcal{C} with

$$\mathbf{x}_t = \begin{bmatrix} \Re\{\mathbf{x}_t^c\} \\ \Im\{\mathbf{x}_t^c\} \end{bmatrix}, \quad \mathbf{H} = \mathbf{I}_T \otimes \begin{bmatrix} \Re\{\mathbf{H}^c\} & -\Im\{\mathbf{H}^c\} \\ \Im\{\mathbf{H}^c\} & \Re\{\mathbf{H}^c\} \end{bmatrix}. \quad (3)$$

The goal of this paper is the design of a lattice codebook $\mathcal{C} \subseteq \mathbb{R}^{2MT}$ satisfying either the average energy constraint

$$\frac{1}{|\mathcal{C}|} \sum_{\mathbf{x} \in \mathcal{C}} \|\mathbf{x}\|^2 \leq TM, \quad (4)$$

or the peak energy constraint

$$\|\mathbf{x}\|^2 \leq TM, \quad \forall \mathbf{x} \in \mathcal{C}. \quad (5)$$

Note that the rate of the code is $R = \frac{1}{T} \log_2 |\mathcal{C}|$ bits per channel use and ρ represents the average transmit power.

A. Lattice Space-Time (LAST) Codes

An n -dimensional lattice Λ is defined by a set of n basis (column) vectors $\mathbf{g}_1, \dots, \mathbf{g}_n$ in \mathbb{R}^n . The lattice is composed of all integral combinations of the basis vectors, i.e., $\Lambda = \{\mathbf{x} = \mathbf{G}\mathbf{z} : \mathbf{z} \in \mathbb{Z}^n\}$, where $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$, and $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n]$. In the Euclidean space, the closest lattice point quantizer $\mathcal{Q}(\cdot)$ associated with Λ is defined by

$$\mathcal{Q}(\mathbf{r}) = \mathbf{x} \in \Lambda, \quad \text{if} \quad \|\mathbf{r} - \mathbf{x}\| \leq \|\mathbf{r} - \mathbf{x}'\|, \quad \forall \mathbf{x}' \in \Lambda, \quad (6)$$

where the ties are broken arbitrarily. The Voronoi cell of Λ is the set of points in \mathbb{R}^n closest to the zero codeword, i.e., $\mathcal{V}_0(\mathbf{G}) = \{\mathbf{r} \in \mathbb{R}^n : \mathcal{Q}(\mathbf{r}) = \mathbf{0}\}$. The Voronoi cell associated with each $\mathbf{x} = \mathbf{G}\mathbf{z} \in \Lambda$ is a shift of $\mathcal{V}_0(\mathbf{G})$ by \mathbf{x} and is denoted by $\mathcal{V}_{\mathbf{z}}(\mathbf{G})$. The (n -dimensional) volume of the Voronoi cell is given by $\text{Vol}(\mathcal{V}_0(\mathbf{G})) = \sqrt{\det(\mathbf{G}^T \mathbf{G})}$ [7].

Let $n = 2MT$. A finite set of points in the n -dimensional translated lattice ($\Lambda + \mathbf{u}$, $\mathbf{u} \in \mathbb{R}^n$) can be used as codewords of a codebook \mathcal{C} and for a rate R bits per channel use, the codebook will contain $|\mathcal{C}| = 2^{T \cdot R}$ such points. Here, we specify a lattice code using the three-tuple $\{\mathbf{G}, \{\mathbf{z}_i\}_{i=1}^{|\mathcal{C}|}, \mathbf{u}\}$, with $\mathbf{G} \in \mathbb{R}^{n \times n}$, $\mathbf{z}_i \in \mathbb{Z}^n$, $i = 1, \dots, |\mathcal{C}|$, and $\mathbf{u} \in \mathbb{R}^n$. \mathbf{u} is referred to as the translation vector and $\{\mathbf{z}_i\}$ are called the coordinate vectors. For a given \mathbf{G} and \mathbf{u} , the code will be referred to as a *spherical lattice code* if the coordinate vectors $\{\mathbf{z}_i\}$ are a solution to

$$\min_{\{\mathbf{z}_i \in \mathbb{Z}^n\}_{i=1}^{|\mathcal{C}|}} \sum_{i=1}^{|\mathcal{C}|} \|\mathbf{G}\mathbf{z}_i + \mathbf{u}\|^2. \quad (7)$$

For ease in exposition, unless otherwise stated we assume $N = M$. Moreover, we will assume that $\mathbf{H}\mathbf{G}$ has full column rank with probability one.

B. Lattice Decoders

In lattice decoding, the receiver assumes that any point in the infinite lattice could have been transmitted. For a given lattice, the *naive lattice decoder* determines

$$\hat{\mathbf{z}} = \arg \min_{\mathbf{z} \in \mathbb{Z}^n} \|\mathbf{y} - \mathbf{H}\mathbf{u} - \mathbf{H}\mathbf{G}\mathbf{z}\|. \quad (8)$$

Note that this decoder should be distinguished from the nearest-codeword decoder, i.e., the maximum-likelihood (ML) decoder. The absence of boundary-control results in substantial savings in complexity. An interesting property of lattice codes with naive lattice decoder is that owing to the lattice symmetry (geometric uniformity), the error probability is invariant to conditioning on a particular transmitted lattice codeword and only depends on the lattice generator. Thus from the energy efficiency point of view, selecting the codewords with minimum norm minimizes the average transmit power and hence *spherical lattice codes are optimal for the naive lattice decoder*.

It has been shown in [4] that an MMSE-GDFE front-end can dramatically improve the performance of the lattice decoding algorithms in MIMO systems. This MMSE lattice decoder determines an upper triangular matrix \mathbf{B} from the Cholesky decomposition of the matrix $\mathbf{I}_n + \mathbf{H}^T \mathbf{H}$ and a matrix $\mathbf{F} = (\mathbf{H}\mathbf{B}^{-1})^T$ and returns

$$\hat{\mathbf{z}} = \arg \min_{\mathbf{z} \in \mathbb{Z}^n} \|\mathbf{F}\mathbf{y} - \mathbf{B}\mathbf{u} - \mathbf{B}\mathbf{G}\mathbf{z}\|. \quad (9)$$

Consider the statistics of the equivalent system model for the MMSE lattice decoder. Defining $\tilde{\mathbf{y}} = \mathbf{F}\mathbf{y} - \mathbf{B}\mathbf{u}$, we have that

$$\tilde{\mathbf{y}} = \mathbf{B}\mathbf{G}\mathbf{z} + \underbrace{(\mathbf{F}\mathbf{H} - \mathbf{B})(\mathbf{G}\mathbf{z} + \mathbf{u}) + \mathbf{F}\mathbf{w}}_{\mathbf{v}} \quad (10)$$

Assuming $\mathbf{x} = \mathbf{G}\mathbf{z} + \mathbf{u}$ to be zero-mean with $\mathbb{E}[\mathbf{x}\mathbf{x}^T] = 1/2\mathbf{I}$ and using the fact that $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, 1/2\mathbf{I})$ is independent of \mathbf{x} , it can be shown that [4] $\mathbb{E}[\mathbf{v}\mathbf{v}^T] = 1/2\mathbf{I}$. Although \mathbf{v} contains a signal (\mathbf{z}) dependent term, assuming $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, 1/2\mathbf{I})$ is very effective. Henceforth, we will make this assumption and then the error probability yielded by the MMSE lattice decoder is identical to that of a naive lattice decoder operating on (10) but where \mathbf{v} is independent AWGN so that spherical lattice codes are optimal for the MMSE lattice decoder as well.

III. DESIGN ALGORITHM

We propose to use a stochastic gradient descent algorithm to optimize the probability of error (or its bounds) over a feasible set of generator matrices. We provide a brief description of the algorithm in its general form and elaborate on particular designs in the next section. Let ψ denote a random vector defined over some sample space. Also let ϕ denote the vector of parameters lying in a feasible set Ω . Our objective is to minimize $f(\phi) \triangleq \mathbb{E}[g(\phi, \psi)]$ over Ω using “noisy” but unbiased estimates of $f'(\phi) \triangleq \nabla_{\phi} f(\phi) = \mathbb{E}[\nabla_{\phi} g(\phi, \psi)]$.¹ Then the stochastic gradient descent algorithm works as follows. Let ϕ_k denote the vector of parameters at the k^{th} step. Then the $(k+1)^{\text{th}}$ iteration proceeds as follows.

1. Draw L samples ψ_1, \dots, ψ_L .
2. Obtain unbiased gradient estimate:
 $\hat{f}'(\phi_k) = \frac{1}{L} \sum_{i=1}^L \nabla_{\phi} g(\phi, \psi_i)|_{\phi=\phi_k}$.
3. Update: $\phi_{k+1} = \Pi_{\Omega}(\phi_k - a_k \hat{f}'(\phi_k))$.

The step-size sequence $\{a_k\}$ is (usually) chosen as the harmonic series $a_k = c/k$, where c is a positive scalar. $\Pi_{\Omega}(\cdot)$ resembles a projection operator² in that it finds a point in the feasible set close to the input argument when the latter falls outside the feasible set. For the problem at hand several objective functions which are defined in the sequel can be used since their gradients are derived in the required form, i.e., their unbiased estimates can be obtained via simulations. In the subsequent subsections we will consider various choices for the feasible set Ω and the associated $\Pi_{\Omega}(\cdot)$.

A. Average Energy Constraint

In order to satisfy the power constraints, the feasible set of generator matrices is

$$\Theta_{\text{avg}} = \left\{ \mathbf{G} : \min_{\{\mathbf{z}_i \in \mathbb{Z}^n\}, \mathbf{u} \in \mathbb{R}^n} \left\{ \frac{\sum_{i=1}^{2^{RT}} \|\mathbf{G}\mathbf{z}_i + \mathbf{u}\|^2}{2^{RT}} \right\} \leq MT \right\} \quad (11)$$

It seems intractable to parameterize the set Θ_{avg} . In fact for a given \mathbf{G} and efficient way of obtaining an optimal spherical code, i.e. an optimal set of codewords (which minimize the average energy) is also not known. As a consequence, we adopt the following sub-optimum approach. We set $\phi = \mathbf{G}$ and $\Omega = \Theta_{\text{avg}}$ and define $\Pi_{\Omega}(\cdot)$ as follows. For a given input $\mathbf{G} \in \mathbb{R}^{n \times n}$, we determine a “good” spherical code (having low average energy) using an iterative technique suggested in [12], which converges to a fixed point very fast in about 2-3 iterations.

¹The exchange of derivative and the expectation is required for this method.

²Usually this method is employed to solve unconstrained problems. In the constrained version there is no universal rule or method to enforce the constraints.

With the codewords $\{\mathbf{G}\mathbf{z}_i + \mathbf{u}\}$ so determined we compute a scaling factor $\beta = \sqrt{\frac{MT2^{RT}}{\sum_{i=1}^{2^{RT}} \|\mathbf{G}\mathbf{z}_i + \mathbf{u}\|^2}}$. If $\beta \geq 1$ then $\mathbf{G} \in \Theta_{\text{avg}}$ and we let $\Pi_{\Omega}(\mathbf{G}) = \mathbf{G}$; otherwise we set $\Pi_{\Omega}(\mathbf{G}) = \beta\mathbf{G}$. As will be revealed in the simulation results, this method works well in practice.

B. Peak Energy Constraint

We now consider the design under a peak energy constraint, where all codewords must satisfy $\|\mathbf{G}\mathbf{z}_i + \mathbf{u}\|^2 \leq MT$, $\forall i$. Letting $\mathcal{S}_n(\sqrt{MT})$ denote a n -dimensional sphere centered at the origin and of radius \sqrt{MT} , we can leverage a result stated in [4], which says that if $\text{Vol}(\mathcal{V}_0(\mathbf{G})) \leq \frac{\text{Vol}(\mathcal{S}_n(\sqrt{MT}))}{2^{RT}}$, then there exists a translation vector \mathbf{u} and coordinate vectors $\{\mathbf{z}_i\}$ such that $\|\mathbf{G}\mathbf{z}_i + \mathbf{u}\|^2 \leq MT$, $1 \leq i \leq 2^{RT}$. Moreover, since the error probability and its bounds for a given generator $\alpha\mathbf{G}$ monotonically decrease in α , we consider the set of lattice generators

$$\Theta_{\text{peak}} = \left\{ \mathbf{G} : |\mathbf{G}^T \mathbf{G}|^{1/2} = \frac{\text{Vol}(\mathcal{S}_n(\sqrt{MT}))}{2^{RT}} \right\}. \quad (12)$$

The nice feature of the set Θ_{peak} is that any $\mathbf{G} \in \Theta_{\text{peak}}$ can be expressed as a differentiable function of a parameter vector ϕ . To see this, let $\mathbf{G} = \mathbf{U}\mathbf{R}$ be the QR decomposition of $\mathbf{G} \in \Theta_{\text{peak}}$, where \mathbf{U} is unitary and \mathbf{R} is lower triangular with positive diagonal elements. Then since the orthogonal group is a differentiable manifold, we can express \mathbf{U} as a differentiable function $n(n-1)/2$ parameters. Setting $\{R_{k,k} = \exp(t_k)\}_{k=1}^{n-1}$ and $R_{n,n} = \frac{\text{Vol}(\mathcal{S}_n(\sqrt{MT}))}{2^{RT} \prod_{k=1}^{n-1} \exp(t_k)}$ we see that \mathbf{R} is a differentiable function of $n(n+1)/2 - 1$ parameters ($\{t_k\}$ and all its strictly lower triangular elements). Collecting all the $n(n+1)/2 - 1 + n(n-1)/2 = n^2 - 1$ parameters into a vector ϕ we have that $\mathbf{G} = \mathbf{G}(\phi)$ is a differentiable function of ϕ . Thus we can implement an unconstrained stochastic gradient-descent search.

C. Average Energy Constraint: Continuous Approximation

We invoke a popular approximation, see for instance [8], which is accurate for high rates. The main idea is that at high rates a codeword $\mathbf{x} = \mathbf{G}\mathbf{z} + \mathbf{u}$ of a spherical lattice code can be (approximately) considered as a realization of a spherically uniform random vector. Using this approximation, we can take the feasible set to be

$$\Theta_{\text{avg-cont}} = \left\{ \mathbf{G} : |\mathbf{G}^T \mathbf{G}|^{1/2} = \frac{\text{Vol}(\mathcal{S}_n(\sqrt{MT+1}))}{2^{RT}} \right\} \quad (13)$$

The advantage is that we again have an unconstrained gradient descent algorithm. Also a ‘‘good’’ spherical code is determined only for the final (optimized) generator \mathbf{G} and then scaled to satisfy the average energy constraint.

IV. LATTICE DESIGN: OBJECTIVE FUNCTIONS AND GRADIENTS

We first derive three objective functions along with their gradients for the naive decoder, which can be employed in the

gradient-descent based design algorithm. The first one is the exact error probability whereas the second and third ones are an upper and a lower bound, respectively. Let us start with the exact error probability. As mentioned earlier, for the naive lattice decoder without loss of generality we can assume that $\mathbf{z} = 0$ is the transmitted coordinate vector. Then letting $\text{Pe}(\mathbf{G})$ denote the error probability (averaged over the channel realizations) of a lattice code with generator \mathbf{G} , from the decision rule in (8) we see that

$$\begin{aligned} \text{Pe}(\mathbf{G}) &= E_{\mathbf{H}}[Pr(\mathbf{w} \notin \mathcal{V}_0(\mathbf{H}\mathbf{G})|\mathbf{H})] \\ &= 1 - \frac{1}{\pi^{n/2}} \int_{\mathcal{V}_0(\mathbf{H}\mathbf{G})} \exp(-\|\mathbf{w}\|^2) d\mathbf{w} \end{aligned} \quad (14)$$

Unfortunately, as noted in [8] the integral in (14) is in general impossible to obtain in a closed form. However its derivative can be estimated. To show this, first we invoke a result from [9] also used in [10], which says that for a random vector \mathbf{q} with n i.i.d components having uniform $U[0, 1]$ distribution, the vector $\mathbf{H}\mathbf{G}\mathbf{q} - \mathcal{Q}(\mathbf{H}\mathbf{G}\mathbf{q})$ is uniformly distributed in $\mathcal{V}_0(\mathbf{H}\mathbf{G})$. As a result we can express $\text{Pe}(\mathbf{G})$ as

$$1 - E_{\mathbf{H}} \left[\frac{|\mathbf{G}^T \mathbf{H}^T \mathbf{H} \mathbf{G}|^{1/2}}{\pi^{n/2}} E_{\mathbf{q}}[\exp(\min_{\mathbf{z} \in \mathcal{Z}^n} \|\mathbf{H}\mathbf{G}\mathbf{q} - \mathbf{H}\mathbf{G}\mathbf{z}\|^2)] \right] \quad (15)$$

We offer the following proposition. A proof is omitted here due to space constraints.

Lemma 1: Suppose $\mathbf{G}(\phi)$ is a differentiable function of ϕ and $\phi \in \Phi$. Then $\frac{d}{d\phi} E_{\mathbf{q}}[\exp(\min_{\mathbf{z} \in \mathcal{Z}^n} \|\mathbf{H}\mathbf{G}\mathbf{q} - \mathbf{H}\mathbf{G}\mathbf{z}\|^2)]$ equals

$$\sum_{\mathbf{z} \in \mathcal{Z}} E_{\mathbf{q}} \left[\mathbf{1}(\mathbf{H}\mathbf{G}\mathbf{q} \in \mathcal{V}_{\mathbf{z}}(\mathbf{H}\mathbf{G})) \frac{d}{d\phi} \exp(\|\mathbf{H}\mathbf{G}(\phi)\mathbf{q} - \mathbf{H}\mathbf{G}(\phi)\mathbf{z}\|^2) \right] \quad (16)$$

where \mathcal{Z} is any finite set of coordinate vectors such that $\cup_{\mathbf{z} \in \mathcal{Z}} \mathcal{V}_{\mathbf{z}}(\mathbf{H}\mathbf{G})$ covers the (bounded) fundamental parallelotope $\{\mathbf{H}\mathbf{G}\mathbf{q}, \mathbf{q} \in [0, 1]^n\}$.

The derivative of (15) can now be computed by first exchanging it with expectation over \mathbf{H} (which is justified by the bounded convergence theorem) and then using (16).

Next we consider the union upper bound on the conditional error probability, $\text{Pe}(\mathbf{G}, \mathbf{H})$, given by [7]

$$\text{Pe}^{\text{ub}}(\mathbf{G}, \mathbf{H}) \triangleq \sum_{\mathbf{z} \in \mathcal{R}} Q \left(\sqrt{\frac{\mathbf{z}^T \mathbf{G}^T \mathbf{H}^T \mathbf{H} \mathbf{G} \mathbf{z}}{2}} \right), \quad (17)$$

where \mathcal{R} is the set of all *relevant* coordinate vectors for a given $\mathbf{H}\mathbf{G}$ such that $\{\mathbf{H}\mathbf{G}\mathbf{z}\}_{\mathbf{z} \in \mathcal{R}}$ determine all the facets of $\mathcal{V}_0(\mathbf{H}\mathbf{G})$. An algorithm to determine all such coordinate vectors is given in [11]. Now, for a lattice generator in n dimensions, it is known that the maximum number of relevant vectors is $2^{n+1} - 2$, [11]. For our design algorithm we would like to be able to estimate the gradient of the upper bound. Our next result allows us to do just that.

Lemma 2: Let \mathbf{z} be a relevant coordinate vector for the lattice generated by $\mathbf{H}\mathbf{G}(\phi)$. Then $\exists \Delta > 0$ small enough such that $\forall \delta \in [-\Delta, \Delta]$, \mathbf{z} remains a relevant coordinate vector for the lattice generated by $\mathbf{H}\mathbf{G}(\phi + \delta)$.

Proof: Using Proposition 2 from [11] we note that $\mathbf{z}_i \in \mathcal{R}$ if and only if

$$\|\mathbf{H}\mathbf{G}(\phi)\mathbf{z}_i/2\| < \|\mathbf{H}\mathbf{G}(\phi)\mathbf{z}_i/2 - \mathbf{H}\mathbf{G}(\phi)\mathbf{z}_j\|, \forall \mathbf{z}_j \in \mathcal{Z}^n \setminus \{\mathbf{z}_i\} \quad (18)$$

Further since the lattice is a sphere packing we must have that for each $\mathbf{z}_i \in \mathcal{R}$

$$\|\mathbf{HG}(\phi)\mathbf{z}_i/2\| = \min_{\mathbf{z}_j \in \mathbb{Z}^n \setminus \{0, \mathbf{z}_i\}} \{ \|\mathbf{HG}(\phi)\mathbf{z}_i/2 - \mathbf{HG}(\phi)\mathbf{z}_j\| \} - \delta_i \quad (19)$$

for some $\delta_i > 0$. Next, to check if \mathbf{z}_i is relevant for $\mathbf{HG}(\phi + \delta)$, using (18) we see that it is sufficient to check if no coordinate vector lies in the set

$$\mathcal{Z}_{i,\delta} = \{ \mathbf{z} \in \mathbb{Z}^n \setminus \{0, \mathbf{z}_i\} : \|\mathbf{HG}(\phi + \delta)\mathbf{z}_i/2 - \mathbf{HG}(\phi + \delta)\mathbf{z}\| \leq \|\mathbf{HG}(\phi + \delta)\mathbf{z}_i/2\| \}$$

Letting $s_{\max}(\cdot)$, $s_{\min}(\cdot)$ denote the maximum and minimum singular values of the matrix argument, it can be seen that

$$\mathcal{Z}_{i,\delta} \subseteq \left\{ \mathbf{z} \in \mathbb{Z}^n \setminus \{0, \mathbf{z}_i\} : \|\mathbf{z}_i/2 - \mathbf{z}\| \leq \|\mathbf{z}_i/2\| \frac{s_{\max}(\mathbf{HG}(\phi + \delta))}{s_{\min}(\mathbf{HG}(\phi + \delta))} \right\}$$

Using the continuity of $\frac{s_{\max}(\mathbf{HG}(\phi + \delta))}{s_{\min}(\mathbf{HG}(\phi + \delta))}$ in δ we can define a small interval $\delta \in [-\tilde{\Delta}, \tilde{\Delta}]$, $\tilde{\Delta} > 0$, where this ratio is bounded above by a positive constant $c < \infty$. Letting $\mathcal{Z}'_i \triangleq \{ \mathbf{z} \in \mathbb{Z}^n \setminus \{0, \mathbf{z}_i\} : \|\mathbf{z}_i/2 - \mathbf{z}\| \leq \|\mathbf{z}_i/2\|c \}$, we set $\mathcal{Z}' = \cup_{i \in \mathcal{R}} \mathcal{Z}'_i$. Thus, we have obtained a large enough but *finite* set \mathcal{Z}' of coordinate vectors and $\tilde{\Delta} > 0$ small enough such that $\forall \delta \in [-\tilde{\Delta}, \tilde{\Delta}]$ & $\mathbf{z}_i \in \mathcal{R}$, if $\min_{\mathbf{z} \in \mathcal{Z}' \setminus \{0, \mathbf{z}_i\}} \{ \|\mathbf{HG}(\phi + \delta)\mathbf{z}_i/2 - \mathbf{HG}(\phi + \delta)\mathbf{z}\| \}$ is greater than $\|\mathbf{HG}(\phi + \delta)\mathbf{z}_i/2\|$, then \mathbf{z}_i remains relevant for $\mathbf{HG}(\phi + \delta)$. Next, using the continuity of \mathbf{G} in ϕ we can show that for any $\mathbf{z}_j \in \mathcal{Z}' \setminus \{0, \mathbf{z}_i\}$

$$\begin{aligned} \|\mathbf{HG}(\phi + \delta)\mathbf{z}_i/2 - \mathbf{HG}(\phi + \delta)\mathbf{z}_j\| &\geq \|\mathbf{HG}(\phi)\mathbf{z}_i/2\| + \delta_i - o_\delta^1(i, j) \\ &\geq \|\mathbf{HG}(\phi + \delta)\mathbf{z}_i/2\| + \delta_i - o_\delta^1(i, j) - o_\delta^2(i), \end{aligned}$$

where $o_\delta^1(i, j)$, $o_\delta^2(i)$ are positive terms that goes to zero as $\delta \rightarrow 0$. Since the sets \mathcal{R} and \mathcal{Z}' are finite, we can now conclude that $\forall \delta \in [-\Delta, \Delta]$, where $0 < \Delta < \tilde{\Delta}$ is small enough, all the coordinate vectors in \mathcal{R} remain relevant for $\mathbf{HG}(\phi + \delta)$ as well. ■

Now, suppose for a given $\mathbf{HG}(\phi)$ the number of relevant coordinate vectors in \mathcal{R} is equal to the upper bound $2^{n+1} - 2$. Then as a consequence of the Lemma 2, we can conclude that $\forall \delta \in [-\Delta, \Delta]$, the set \mathcal{R} contains *all the relevant coordinate vectors of $\mathbf{HG}(\phi + \delta)$* . This observation allows us to take the derivative of the upper bound (conditioned on \mathbf{H}) assuming \mathcal{R} to be a fixed set and we have that $\frac{d}{d\phi} \text{Pe}^{\text{ub}}(\mathbf{G}, \mathbf{H})$ equals

$$\sum_{\mathbf{z} \in \mathcal{R}} \left(\frac{-\exp(-\mathbf{z}^T \mathbf{G}^T \mathbf{H}^T \mathbf{HG} \mathbf{z} / 4)}{\sqrt{\pi \mathbf{z}^T \mathbf{G}^T \mathbf{H}^T \mathbf{HG} \mathbf{z}}} \mathbf{z}^T \mathbf{G}^T \mathbf{H}^T \mathbf{H} \left(\frac{d}{d\phi} \mathbf{G}(\phi) \right) \mathbf{z} \right) \quad (20)$$

Finally, we derive a lower bound and obtain its derivative in the desired form. Suppose the kissing number of lattice generated by $\mathbf{HG}(\phi)$ is 2 i.e. there are exactly two shortest (non-zero) vectors in the lattice. Letting $\mathbf{HG}(\phi)\mathbf{z}_1$ and $\mathbf{HG}(\phi)\mathbf{z}_2$ denote these vectors, we must have that $\mathbf{z}_2 = -\mathbf{z}_1$ and that $\mathbf{z}_1, \mathbf{z}_2$ are relevant³. Then since the half-spaces $\{ \mathbf{y} : \|\mathbf{y} - \mathbf{HG}(\phi)\mathbf{z}_1\|^2 \leq \|\mathbf{y}\|^2 \}$ and $\{ \mathbf{y} : \|\mathbf{y} + \mathbf{HG}(\phi)\mathbf{z}_1\|^2 \leq \|\mathbf{y}\|^2 \}$ do not overlap we can obtain a conditional lower bound, $\text{Pe}^{\text{lb}}(\mathbf{G}, \mathbf{H})$, given by

$$Q \left(\sqrt{\frac{\mathbf{z}_1^T \mathbf{G}^T \mathbf{H}^T \mathbf{HG} \mathbf{z}_1}{2}} \right) + Q \left(\sqrt{\frac{\mathbf{z}_2^T \mathbf{G}^T \mathbf{H}^T \mathbf{HG} \mathbf{z}_2}{2}} \right). \quad (21)$$

³The kissing number can be no less than 2. Further, the fact that both $\mathbf{z}_1, \mathbf{z}_2 = -\mathbf{z}_1$ are relevant follows from the necessary and sufficient condition given in (18).

Also, using arguments similar to those used before we can show that the derivative of (21) equals

$$\sum_{i=1}^2 \frac{-\exp(-\mathbf{z}_i^T \mathbf{G}^T \mathbf{H}^T \mathbf{HG} \mathbf{z}_i / 4)}{\sqrt{\pi \mathbf{z}_i^T \mathbf{G}^T \mathbf{H}^T \mathbf{HG} \mathbf{z}_i}} \mathbf{z}_i^T \mathbf{G}^T \mathbf{H}^T \mathbf{H} \left(\frac{d}{d\phi} \mathbf{G}(\phi) \right) \mathbf{z}_i \quad (22)$$

In our case, the fading matrix is drawn from a continuous distribution and the generator \mathbf{G} has no structure so that \mathbf{HG} always has the maximum number of relevant vectors and the minimum kissing number. Thus, the unconditional upper and lower bounds and their derivatives can be obtained after averaging (17), (21) and (20), (22) over \mathbf{H} , respectively⁴.

A better upper bound, referred to as the improved upper bound, can also be obtained as

$$\Pr(\mathcal{O}) + E_{\mathbf{H}}[\mathbf{1}(\mathbf{H}^c \in \mathcal{O}^c) \text{Pe}^{\text{ub}}(\mathbf{G}, \mathbf{H})], \quad (23)$$

where $\mathcal{O} \triangleq \{ \mathbf{H}^c : \log \left| \frac{\rho}{M} (\mathbf{H}^c)^\dagger \mathbf{H}^c \right| < R \}$ is called the outage set. (23) is obtained by taking the upper bound to be one when $\mathbf{H}^c \in \mathcal{O}$. The improved bound is dramatically tighter for the naive decoder since the conditional upper bound in (17) almost always exceeds one when $\mathbf{H}^c \in \mathcal{O}$. Further, since the set \mathcal{O} is independent of \mathbf{G} the derivative of (23) is readily obtained using (20).

Finally as a consequence of our assumption, *the corresponding bounds and gradients for the MMSE lattice decoder are obtained after replacing \mathbf{H} by \mathbf{B}* . Thus for the improved upper bound we take $\mathcal{O} \triangleq \{ \mathbf{H}^c : \log \left| \mathbf{I} + \frac{\rho}{M} (\mathbf{H}^c)^\dagger \mathbf{H}^c \right| < R \}$. The intuition behind defining the outage sets in this manner rose from the asymptotic lower bound analysis which is presented below.

V. FUNDAMENTAL LOWER BOUNDS

In this section we allow $N \geq M$ and first consider the naive decoder and obtain firm lower bounds for the sets Θ_{peak} and $\Theta_{\text{avg-cont}}$. For a given n and rate R , let \mathbf{G} be any matrix in Θ_{peak} . Next, for a given channel realization \mathbf{H} , we note that

$$\text{Pe}(\mathbf{G}, \mathbf{H}) = \Pr(\mathbf{w}' \notin \mathcal{V}_0(\Sigma^{1/2}) | \mathbf{HG}) \quad (24)$$

where $\Sigma = (\mathbf{HG})^T \mathbf{HG} > 0$. (24) follows after using a change of variables $\mathbf{w}' = \Sigma^{-1/2} (\mathbf{HG})^T \mathbf{w}$. We next consider the RHS of (24). Note that $\Sigma^{1/2}$ generates a n dimensional lattice with $\text{Vol}(\mathcal{V}_0(\Sigma^{1/2})) = |\Sigma^{1/2}| = |\mathbf{G}^T \mathbf{G}|^{1/2} |\mathbf{H}^T \mathbf{H}|^{1/2}$. To obtain the lower bound, we note that $\mathbf{w}' \sim \mathcal{N}(0, 1/2 \mathbf{I}_n)$ and use Shannon's classical idea developed for lattices in [8], which yields

$$\begin{aligned} \Pr(\mathbf{w}' \notin \mathcal{V}_0(\Sigma^{1/2}) | \mathbf{HG}) &\geq \Pr(\mathbf{w}' \notin \mathcal{S}_{\text{eff}}(r(\mathbf{HG})) | \mathbf{HG}) \\ &= \exp(-r^2(\mathbf{HG})) \sum_{k=0}^{n/2-1} \frac{(r^2(\mathbf{HG}))^k}{k!}, \end{aligned} \quad (25)$$

where $\mathcal{S}_{\text{eff}}(r(\mathbf{HG}))$ is a sphere in n dimensions, centered at the origin and having the same volume as $\mathcal{V}_0(\Sigma^{1/2})$ so that

$$r(\mathbf{HG}) = \left(\frac{|\mathbf{G}^T \mathbf{G}|^{1/2} |\mathbf{H}^T \mathbf{H}|^{1/2}}{V_n} \right)^{1/n}. \quad (26)$$

⁴Again the exchange of derivative and the expectation over \mathbf{H} can be rigorously justified.

Clearly the lower bound is decreasing in $r(\mathbf{HG})$. Using the fact $|\mathbf{G}^T \mathbf{G}|^{1/2} \leq \text{Vol}(\mathcal{S}_n(\sqrt{MT})) / 2^{RT}$ we see that

$$r(\mathbf{HG}) \leq r(\mathbf{H}^c) \triangleq \frac{(MT)^{1/2} |\frac{\rho}{M} (\mathbf{H}^c)^\dagger \mathbf{H}^c|^{1/(2M)}}{2^{R/(2M)}} \quad (27)$$

Thus a lower bound valid $\forall \mathbf{G} \in \Theta_{\text{peak}}$ is given by

$$E_{\mathbf{H}} \left[\exp(-r^2(\mathbf{H}^c)) \sum_{k=0}^{n/2-1} \frac{(r^2(\mathbf{H}^c))^k}{k!} \right]. \quad (28)$$

The corresponding bound for $\forall \mathbf{G} \in \Theta_{\text{avg-cont}}$ follows after replacing MT in the numerator of (27) by $MT + 1$.

The following theorem determines the *best achievable* performance of the naive decoder over the set $\Theta_{\text{avg-cont}}$ in the limit of large code-lengths, i.e., $T \rightarrow \infty$ or equivalently $n \rightarrow \infty$.

Proposition 1: The error probabilities of any sequence of lattice codes of unbounded length with generators $\{\mathbf{G}_n\}$, $\mathbf{G}_n \in \Theta_{\text{avg-cont},n}$, $\forall n$ obtained via the naive lattice decoder, denoted by $\{\text{Pe}(\mathbf{G}_n)\}$, must satisfy $\liminf_{n \rightarrow \infty} \text{Pe}(\mathbf{G}_n) \geq \Pr(\log |\frac{\rho}{M} (\mathbf{H}^c)^\dagger \mathbf{H}^c| < R)$. Further, there exists a sequence for which $\lim_{n \rightarrow \infty} \text{Pe}(\mathbf{G}_n) = \Pr(\log |\frac{\rho}{M} (\mathbf{H}^c)^\dagger \mathbf{H}^c| < R)$.

Proof sketch: [4] proved the existence of a sequence of lattice codes with generators $\{\mathbf{G}_n\}$, $\mathbf{G}_n \in \Theta_{\text{peak},n}$, $\forall n$ such that $\limsup_{n \rightarrow \infty} \text{Pe}(\mathbf{G}_n) \leq \Pr(\log |\frac{\rho}{M} (\mathbf{H}^c)^\dagger \mathbf{H}^c| < R)$. Thus to prove the theorem we prove the converse for the larger set $\Theta_{\text{avg-cont}}$. First note that

$$\Pr \left(\sum_{j=1}^n Z_j^2 > r^2(\mathbf{H}^c) |\mathbf{H}^c| \right) = \exp(-r^2(\mathbf{H}^c)) \sum_{k=0}^{n/2-1} \frac{(r^2(\mathbf{H}^c))^k}{k!} \quad (29)$$

where $\{Z_j^2\}$ are i.i.d $\mathcal{N}(0, 1/2)$ (independent of \mathbf{H}). Then using the weak law of large numbers in (29) along with (27), it is seen that

$$\lim_{n \rightarrow \infty} \Pr \left(\sum_{j=1}^n Z_j^2 > r^2(\mathbf{H}^c) |\mathbf{H}^c| \right) = \begin{cases} 1 & R > \log |\frac{\rho}{M} (\mathbf{H}^c)^\dagger \mathbf{H}^c| \\ 0 & R < \log |\frac{\rho}{M} (\mathbf{H}^c)^\dagger \mathbf{H}^c| \end{cases} \quad (30)$$

Using (30) with the dominated convergence theorem, we can infer that the lower bound in the limit $n \rightarrow \infty$ becomes $\Pr(\log |\frac{\rho}{M} (\mathbf{H}^c)^\dagger \mathbf{H}^c| < R)$.

On the other hand the lower bounds for the sets Θ_{peak} , $\Theta_{\text{avg-cont}}$ and the (assumed equivalent) model in (10) can be derived by simply replacing \mathbf{H} with \mathbf{B} in the corresponding bounds derived above. The limiting value of both the lower bounds can be readily verified to be $\Pr(\log |\mathbf{I} + \frac{\rho}{M} (\mathbf{H}^c)^\dagger \mathbf{H}^c| < R)$.

VI. NUMERICAL RESULTS

Due to space constraints we only provide one illustrative example. In Figure 1 we consider a MIMO link with 4 transmit and 2 receive antennas with $T = 2$ and $R = 4$ bits per channel use, experiencing correlated Rayleigh fading. The correlation parameters were set to model the scenario C (urban). We plot the block error probabilities (BLERs) obtained with the MMSE Lattice decoder for three randomly generated spherical lattice codes along with that of the optimized spherical lattice code. The code was obtained by optimizing the improved upper bound using the method described in Section III-A over the assumed equivalent

model (10). Our optimized code yields a gain of about 5 dB over randomly chosen spherical codes, which highlights the importance of tailoring the code to the channel statistics and shows that optimizing codes for the naive decoder over the model in (10) provides excellent designs for the MMSE lattice decoder.

VII. CONCLUSIONS

We have proposed a systematic method for designing optimal spherical lattice codes for lattice decoders. The design method is universal in the sense that it can be applied to optimize the lattice codes for arbitrary channel statistics and SNR. Simulation results have shown that our optimization method yields low error rate lattice codes that outperform other lattice codes proposed in the literature.

REFERENCES

- [1] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.
- [2] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1804–1824, July 2002.
- [3] P. Dayal and M. K. Varanasi, "An optimal two transmit antenna space-time code and its stacked extensions," in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, Monterey, CA, Nov. 2003.
- [4] H. E. Gamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 6, pp. 968–985, June 2004.
- [5] B. Hassibi and H. Vikalo, "On sphere decoding algorithm: Parts I and II," to appear *IEEE Trans. Signal Processing*.
- [6] J. C. Spall, *Introduction to Stochastic Search and Optimization*, John Wiley, New York, 2003.
- [7] J. Conway and N. J. Sloane, *Sphere Packings, Lattices and Groups*, Springer-Verlag, New York, NY, 1999.
- [8] V. Tarokh, A. Vardy, and K. Zeger, "Universal bound on the performance of lattice codes," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 670–681, Mar. 1999.
- [9] J. Conway and N. Sloane, "On the Voronoi regions of certain lattices," *SIAM J. Algebraic and Discrete Methods*, vol. 5, no. 3, pp. 294–305, Sept. 1984.
- [10] E. Agrell and T. Eriksson, "Optimization of lattices for quantization," *IEEE Trans. Inform. Theory*, vol. 44, no. 5, pp. 1814–1828, Sept. 1998.
- [11] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inform. Theory*, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.
- [12] J. H. Conway and N. J. A. Sloane, "A fast encoding method for lattice codes and quantizers," *IEEE Trans. Inform. Theory*, vol. 29, no. 6, pp. 820–824, Nov. 1983.
- [13] E. A. Lee and D. G. Messerschmitt, *Digital Communication*, Kluwer Academic Publishers, Boston, MA, 2nd edition, 1994.

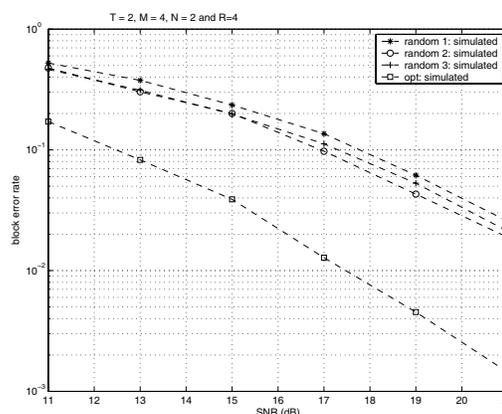


Fig. 1. BLERs of lattice codes with MMSE lattice decoder.