# WHEN TO ADD ANOTHER DIMENSION WHEN COMMUNICATING OVER MIMO CHANNELS


*S. Goparaju, A. R. Calderbank* — Department of Electrical Engineering, Princeton University

*W. R. Carson, M. R. D. Rodrigues* — Instituto de Telecomunicações, Department of Computer Science, University of Porto

*F. Pérez-Cruz* — Department of Signal Theory and Communications, University Carlos III in Madrid



**ABSTRACT**

This paper introduces a *divide and conquer* approach to the design of transmit and receive filters for communication over a Multiple Input Multiple Output (MIMO) Gaussian channel subject to an average power constraint. It involves conversion to a set of parallel scalar channels, possibly with very different gains, followed by coding per sub-channel (i.e. over time) rather than coding across sub-channels (i.e. over time and space). The loss in performance is negligible at high signal-to-noise ratio (SNR) and not significant at medium SNR. The advantages are reduction in signal processing complexity and greater insight into the SNR thresholds at which a channel is first allocated power. This insight is a consequence of formulating the optimal power allocation in terms of an upper bound on error rate that is determined by parameters of the input lattice such as the minimum distance and kissing number. The resulting thresholds are given explicitly in terms of these lattice parameters. By contrast, when the optimization problem is phrased in terms of maximizing mutual information, the solution is mercury waterfilling, and the thresholds are implicit.



***Index Terms—*** Lattices, MIMO, power allocation, error probability, mutual information


## 1. INTRODUCTION

Finite dimensional lattices play an important role in communication over the Additive White Gaussian Noise (AWGN) channel. Closure under addition means that error probability is largely independent of the transmitted signal, and trellis codes based on lattices and cosets [1] make it possible to trade rate for reliability with very low complexity. When the channel is subject to Inter-Symbol Interference (ISI) the lattice structure makes it possible to separate the problem of error correction from that of inverting the partial response channel (see Laroia et al. [2]).

The design of transmit and receive filters for communication over a MIMO Gaussian channel subject to an average power constraint was considered 20 years ago by Kasturia et al. [3]. Their approach, which they call *vector coding*, is to invert the channel matrix. The disadvantage, which becomes increasingly visible at medium to low SNR, is noise amplification if the gains associated with the eigenvectors of the channel are very different. Much more recently, Bergman and Ottersten [4] have looked to address noise amplification by using lattice invariant transformations to match the shape of the multidimensional input constellation to the different channel gains.

It is natural to approach the problem of filter design by first using the Singular Value Decomposition (SVD) to diagonalize the channel, and then implementing the optimal solution for each of the parallel scalar channels that results. In the context of maximization of mutual information, Telatar [5] proved that this strategy did not result in any performance loss and was optimal for Gaussian inputs. Payaró and Palomar demonstrated that for non-Gaussian inputs the optimal precoder did diagonalize the channel to the extent that the left singular vectors of the optimal precoder can be chosen to coincide with the matrix of right singular vectors of the channel [6]. However, the right singular vectors of the optimal precoder did induce correlations across the channels and the loss in performance due to diagonalizing the channels was described by Pérez-Cruz *et al.* in [7].

We shall avoid the complexity of coding across sub-channels by diagonalizing the channel matrix. This in turn simplifies the objective function used for optimization, which is a sum of upper bounds on the error probabilities of the individual sub-channels. The objective function used by Bergman and Ottersten [4] also takes the form of an upper bound on error probability, but it is complicated by the coupling of component sub-channels.

The objective function is also a lower bound to the mutual information (both bounds become tight at high SNR), but has the advantage that the SNR thresholds are given very explicitly in terms of lattice parameters. This is in contrast with the optimal "mercury waterfilling" (MWF) solutions [8], where thresholds are implicit and specified in terms of the inverse Minimum Mean Squared Error (MMSE) function. The loss in performance is negligible at high SNR and not significant at medium SNR, where it can be further mitigated by non-equiprobable signaling.The slight loss in mutual information can be attributed to allocating power more conservatively than in MWF.

## 2. SYSTEM MODEL

Since we are dealing with lattices, we consider directly the real-valued MIMO Gaussian channel model equivalent to the underlying complex-valued model [4]:

$$\mathbf{y} = \sqrt{\mathsf{g}} \cdot \mathbf{H_R}\, \mathbf{H}\, \mathbf{H_T}\, \mathbf{x} + \mathbf{H_R} \mathbf{n}, \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^N$ is the $N$-dimensional vector of receive symbols, $\mathbf{x} \in \mathbb{R}^N$ is the vector of transmit symbols, $\mathbf{n} \in \mathbb{R}^N$ is a realization of a multivariate Gaussian random variable with mean $\mathbf{0}$ and covariance matrix $\mathbf{I}$, and $\mathbf{H} \in \mathbb{R}^{N\times N}$ is the channel matrix which is assumed to be known perfectly to the transmitter and the receiver. The matrices $\mathbf{H_T} \in \mathbb{R}^{N\times N}$ and $\mathbf{H_R} \in \mathbb{R}^{N\times N}$ represent linear transmit and receive filters. We use throughout the SVD of the channel matrix $\mathbf{H} = \mathbf{U_H}\boldsymbol{\Lambda}_\mathbf{H}\mathbf{V}_\mathbf{H}^\mathsf{T}$, where $\mathbf{U_H}$ and $\mathbf{V_H}$ are orthonormal matrices, and $\boldsymbol{\Lambda}_\mathbf{H} = \mathsf{diag}\left\{\lambda_{\mathbf{H}_n}\right\}$ is a $N \times N$ diagonal matrix containing the singular values of $\mathbf{H}$. The scaling factor $\mathsf{g}$ denotes the SNR at which we transmit $\mathbf{x}$.

We sample the channel input $\mathbf{x}$ from an $N$-dimensional constellation $\mathbf{X}$ with cardinality $M$ that has zero mean and unit power per dimension, i.e., $\mathbb{E}\left\{\mathbf{X}\right\} = \mathbf{0}$ and $\mathbb{E}\left\{\mathbf{X}\mathbf{X}^\mathsf{T}\right\} = \mathbf{I}$. Furthermore, we assume that an $M_n$-PAM signal is transmitted over the $n^{th}$ channel, such that the input vectors are equivalently drawn from an integer lattice. The lattice is rescaled so that it has unit-power per channel and we can express the minimum distance:

$$\mathsf{d}_{\mathsf{min},n} = \sqrt{\frac{12}{M_n^2 - 1}}, \tag{2}$$

and the kissing number $K_n$ on the $n^{th}$ channel is double the number of pairs of points at minimum distance $\mathsf{d}_{\mathsf{min},n}$, since the distance is counted once for each point of a pair:

$$K_n = 2(M_n - 1). \tag{3}$$

### 2.1. Design constraint: Channel diagonalization

The objective is to design the transmit and receive filters such that:

$$\sqrt{\mathsf{g}}\, \mathbf{H_R}\, \mathbf{H}\, \mathbf{H_T} = \mathbf{D}, \tag{4}$$

where $\mathbf{D}$ is some diagonal matrix. Although this diagonalization constraint admits the possibility of some loss in performance compared to an unconstrained design, as shown in [7] at high SNR it considerably simplifies the optimization procedure. However, it is sufficiently robust so as to circumvent the noise amplification problems away from high SNR associated with the special case of channel inversion, i.e., when $\mathbf{D} = \mathbf{I}$.

The general form of the transmit and receive filters follows immediately from (4). We choose a receive filter, which results in no noise coloring, as follows:

$$\mathbf{H_R} = \mathbf{U}_\mathbf{H}^\mathsf{T}. \tag{5}$$

In turn, we choose a transmit filter to satisfy the diagonal constraint as follows:

$$\mathbf{H_T} = \mathbf{V_H}\boldsymbol{\Lambda}_{\mathbf{H_T}}, \tag{6}$$

where $\boldsymbol{\Lambda}_{\mathbf{H_T}} = \mathsf{diag}\left(\lambda_{\mathbf{H_T}n}\right)$ is some $N \times N$ diagonal matrix with elements $\lambda_{\mathbf{H_T}n}$. It is clear that the transmit and receive filters perform two standard operations: *i*) conversion of the MIMO channel into a bank of parallel independent channels, whose gains $\left(|h_n|^2\right)$ depend on the original channel singular values; *ii*) power allocation over the parallel independent channels. Consequently, we introduce the following definitions to give helpful intuition $\lambda^2_{\mathbf{H_T}n} = p_n, n = 1, \dots, N$, and $\lambda^2_{\mathbf{H}_n} = |h_n|^2, n = 1, \dots, N$.

### 2.2. Performance Metric: Error Probability

The performance metric used is an upper bound to the error probability. First we upper bound the true error probability $\Pr(e)$ using the union bound which results in a sum of subchannel-wise error probabilities, $\Pr(e_n), n = 1, \dots, N$:

$$\Pr(e) = \Pr\left(e_1 \vee e_2 \vee \dots \vee e_N\right) \leq \sum_{n=1}^{N} \Pr\left(e_n\right). \tag{7}$$

We can then upper bound the subchannel-wise error probabilities (assuming equiprobable input symbols, i.e., $\Pr\left(x_n = x_n^i\right) = 1/M_n$) as follows:

$$\begin{aligned} \Pr\left(e_n\right) &= \sum_{i=1}^{M_n} \Pr\left(e_n | x_n = x_n^i\right) \cdot \Pr\left(x_n = x_n^i\right) \\ &= \frac{K_n}{M_n} Q\left(\sqrt{\mathsf{g}\frac{p_n|h_n|^2\mathsf{d}^2_{\mathsf{min},n}}{4}}\right) \\ &\leq \frac{1}{2}\frac{K_n}{M_n}\exp\left(-\mathsf{g}\frac{p_n|h_n|^2\mathsf{d}^2_{\mathsf{min},n}}{8}\right), \end{aligned} \tag{8}$$

where $\Pr\left(e_n | x_n = x_n^i\right)$ is the probability of error when $x_n^i$ is transmitted and $Q(\cdot)$ is bounded for $x \geq 0$ by [9]:

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right) dy \leq \frac{1}{2}\exp\left(-\frac{x^2}{2}\right). \tag{9}$$

Consequently, an upper bound to the error probability is:

$$P_{e,UB} = \frac{1}{2}\sum_{n=1}^{N}\frac{K_n}{M_n}\exp\left(-\mathsf{g}\frac{p_n|h_n|^2\mathsf{d}^2_{\mathsf{min},n}}{8}\right). \tag{10}$$

## 3. OPTIMAL POWER ALLOCATION POLICY

We define a channel metric $a_n$ that measures the strength of the $n^{th}$ sub-channel and governs the SNR where it turns on:

$$a_n = \frac{K_n}{M_n} \cdot \frac{|h_n|^2 \mathsf{d}^2_{\mathsf{min},n}}{8}. \tag{11}$$

Note that $a_n$ is a function of the kissing number and the minimum distance of $M_n$-PAM. The design of transmit and receive filters that optimize $P_{e,UB}$ subject to a transmit power constraint can thus be expressed as the following convex optimization problem in terms of $p_n$:

$$\min_{p_1,\ldots,p_N} \sum_{n=1}^{N} \frac{K_n}{M_n} \cdot \exp\left(-\mathsf{g} p_n \frac{|h_n|^2 \mathsf{d}^2_{\mathsf{min},n}}{8}\right), \tag{12}$$

subject to:

$$\mathbb{E}\left\{\mathsf{tr}\left(\mathbf{H_T}\mathbf{x}\mathbf{x}^{\mathsf{T}}\mathbf{H_T^{\mathsf{T}}}\right)\right\} = \sum_{n=1}^{N} p_n \leq 1 \tag{13}$$

$$p_n \geq 0, \quad n = 1, \ldots, N. \tag{14}$$

The KKT conditions imply that a channel is allocated power once its channel metric exceeds the threshold $\lambda/\mathsf{g}$:

$$a_n \cdot \mathsf{g} < \lambda, \qquad p_n^* = 0 \tag{15}$$

$$a_n \cdot \mathsf{g} \geq \lambda, \qquad p_n^* = \frac{8}{\mathsf{g}|h_n|^2 \mathsf{d}^2_{\mathsf{min},n}} \cdot \ln\left(\frac{1}{\lambda} a_n \mathsf{g}\right), \tag{16}$$

where $\lambda$ is such that (13) is satisfied with equality.

### 3.1. SNR Thresholds

The SNR value at which the $i^{th}$ channel is switched on is denoted by $g_i$. We relabel the channels such that $a_1 \geq a_2 \geq \cdots \geq a_N$, because the power allocation at low SNR and the SNR thresholds $g_i$ become more intuitive relative to this order. Note that when the $i^{th}$ sub-channel just switches on,

$$\lambda = g_i \cdot a_i. \tag{17}$$

Combining (17) with (16), and using the power constraint (13) with equality, we obtain the SNR value at which the $i^{th}$ 'strongest' sub-channel turns on:

$$g_i = g_i \cdot 1 = g_i \sum_{n=1}^{N} p_n = \sum_{n=1}^{i-1} \left(\frac{8}{|h_n|^2 \mathsf{d}^2_{\mathsf{min},n}} \ln \frac{a_n}{a_i}\right). \tag{18}$$

### 3.2. Low-SNR Regime

If we specialize the optimal power allocation policy to the regime of low SNR ($\mathsf{g} \to 0$) where we can use the approximation $e^x \approx 1 + x$ for small $x$ so that our objective function becomes:

$$P_{e,UB} \approx \frac{1}{2} \sum_{n=1}^{N} \left(\frac{K_n}{M_n} - \mathsf{g} p_n a_n\right), \tag{19}$$

this expression is minimized, when the available power is allocated to the strongest channel, i.e., the channel with the largest $a_n$ value.

### 3.3. High-SNR Regime

In general, when any two channels $i$ and $j$ are active, we can use (16) and equate $\lambda$ to get:

$$p_i^* \frac{|h_i|^2 \mathsf{d}^2_{\mathsf{min},i}}{8} = \frac{1}{\mathsf{g}} \ln\left(\frac{a_i}{a_j}\right) + p_j^* \frac{|h_j|^2 \mathsf{d}^2_{\mathsf{min},j}}{8}. \tag{20}$$

At high SNR, as $\mathsf{g} \to \infty$ the first term on the right-hand side in (20) tends to zero and we get a simple expression that describes channel equalization:

$$p_i^* = \frac{|h_j|^2 \mathsf{d}^2_{\mathsf{min},j}}{|h_i|^2 \mathsf{d}^2_{\mathsf{min},i}} p_j^*. \tag{21}$$

This policy aims to reverse the stretching of the input lattice due to the channel ($|h_n|^2$) and unit-power normalization ($\mathsf{d}^2_{\mathsf{min},n}$) in (2) so that the action of the channel is to "pass through" the integer lattice so that the received lattice is a translate of a scaled integer lattice.

## 4. RESULTS

We compare the results obtained using our objective function $P_{e,UB}$ with those using the true error probability $P_e$ and mutual information (MWF). The channel is a randomly generated $4 \times 4$ real diagonal matrix. The input is a $\mathbb{Z}^4$ lattice, with each sub-channel being used to transmit a 4-PAM signal. Fig. 1 shows that the loss in performance is extremely low at all SNR. This is significant because $P_{e,UB}$ splits into a sum of functions of individual sub-channels, unlike $P_e$, a fact which expresses itself in Fig. 2. For $P_{e,UB}$ and mutual information optimization, the channels turn on one after another, whereas for $P_e$ optimization, the power is shared between all the channels for all SNR. The plots show a rapid convergence between various techniques once all the channels are turned on, i.e., we recover the same solution at high SNR.

A notable phenomenon is seen in Fig. 3 where we compare the power allocations of MWF and $P_{e,UB}$. Since optimizing $P_{e,UB}$ also maximizes a lower bound to mutual information, the curves display similar characteristics. However, the channels in our solution turn on at higher SNR than in MWF. An intuitive explanation for this can be that our objective function is more conservative when turning on channels. We refrain from using a channel until the SNR is good enough, to keep the $P_{e,UB}$ low, whereas MWF turns a channel on as soon as some code exists for reliable communication.

## 5. CONCLUSIONS

We have expressed the problem of linear filter design for MIMO Gaussian channels using lattice parameters. Our

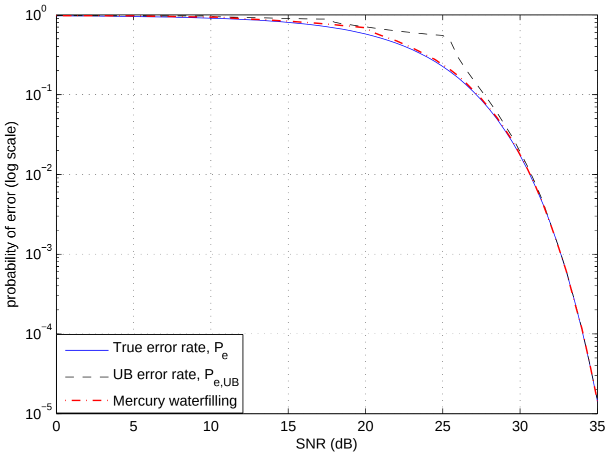


**Fig. 1**. BER performance comparison for the filters which optimize $P_e$, $P_{e,UB}$ and mutual information.

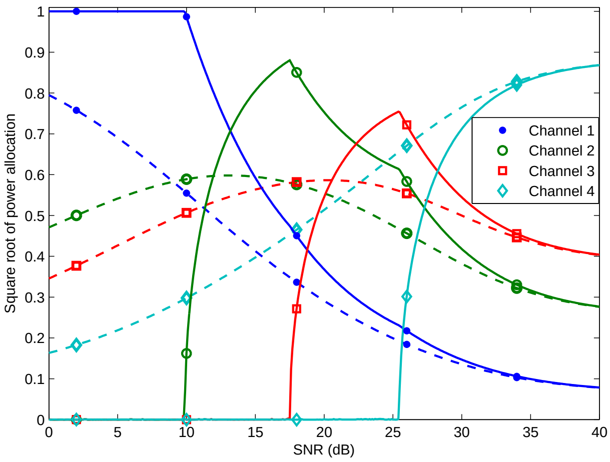


**Fig. 2**. Optimal power allocation comparison for optimal filters which optimize $P_e$ (dashed lines) and $P_{e,UB}$ (solid lines).

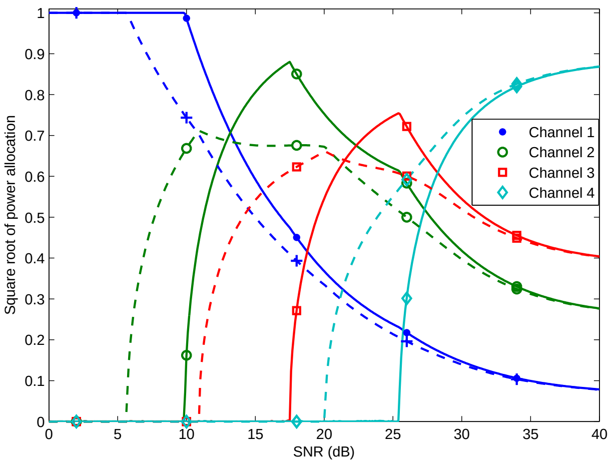


**Fig. 3**. Optimal power allocation comparison for the filters which optimize mutual information (dashed lines) and $P_{e,UB}$ (solid lines).

framework decouples the problem and attacks each sub-channel independently. This not only simplifies the relevant calculations, but also endows the set of sub-channels with a total order leading to their turning on sequentially as SNR increases. The optimization of an upper bound to the detection error probability instead of the true error probability eases the process further without significant losses and provides us with intuitive expressions for the power allocation policy. The current framework paves path to either introducing non-equiprobable signaling or using 'better' lattices like the 8-dimensional Gosset lattice, a large number of which are expressible as a union of integer lattice translates.

## 6. REFERENCES

[1] A. R. Calderbank and N. J. A. Sloane, "New trellis codes based on lattices and cosets," in *IEEE Trans. Inf. Theory*, Mar. 1987, pp. 177–195.

[2] R. Laroia, S. Tretter, and N. Farvardin, "A simple and effective precoding scheme for noise whitening on intersymbol interference channels," in *IEEE Trans. Commun.*, Oct. 1993, pp. 1460–1463.

[3] S. Kasturia, J. Aslanis, and J. Cioffi, "Vector coding for the partial response channel," in *IEEE Trans. Inf. Theory*, July 1990, pp. 741–762.

[4] S. Bergman and B. Ottersten, "Lattice-based linear precoding for MIMO channels with transmitter CSI," in *IEEE Trans. Signal Process.*, July 2008, pp. 2902–2914.

[5] E. Teletar, "Capacity of multi-antenna Gaussian channels," in *AT&T-Bell Labs, Tech. Rep.*, 1995.

[6] M. Payaró and D. P. Palomar, "On optimal precoding in linear vector Gaussian channels with arbitrary input distribution," in *IEEE ISIT '09*, July 2009, pp. 1085–1089.

[7] F. Pérez-Cruz, M. R. Rodrigues, and S. Verdú, "MIMO Gaussian channels with arbitrary inputs: Optimal precoding and power allocation," in *IEEE Trans. Inf. Theory*, Mar. 2010, pp. 1070–1084.

[8] A. Lozano, A. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," in *IEEE Trans. Inf. Theory*, July 2006, pp. 3033–3051.

[9] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*, Wiley, 1965.