Optimal Signal Constellations for Space-Time Rayleigh Channels

ALFRED O. HERO AND THOMAS L. MARZETTA[0]

December 1, 1999

Abstract

We characterize signal constellations which attain single-user computational cut-off rate for the complex Rayleigh flat fading spatio-temporal channel under a peak power constraint. For this channel a constellation is a set of codeword matrices coded over both space (number of antenna elements) and time (temporal block length). The cut-off rate is the maximum of an average decoding error exponent over all possible space-time codeword probability distributions. This error exponent is monotone decreasing in a measure of dissimilarity between all pairs of codeword matrices. For low SNR the dissimilarity function reduces to a trace norm of differences between outerproducts of pairs of codewords. Under the practical constraint of finite dimensional codeword constellations, different characterizations of the optimal constellation are obtained. For arbitrary finite dimension, the optimal constellation must admit an equalizer distribution, i.e., a positive set of codeword probabilities which equalizes the conditional decoding error probabilities. The cut-off rate attaining constellation is determined by maximization of a quadratic form over the feasible set of equalizer constellations. Properties of this optimal constellation are obtained under different operating regimes depending on the number of transmit and receive antennas, the number of codewords in the constellation, and the received SNR.

1 Introduction

Computational cut-off rate analysis has been adopted to establish practical coding limits [7, 2] for many types of communications channels. The cut-off rate specifies the highest information rate beyond which sequential decoding becomes impractical [6, 8], provides a lower bound on channel capcaity, and it is frequently simple to compute when capacity is intractible. In this paper we investigate the cut-off rate for the Rayleigh flat fading spatio-temporal channel model introduced by Marzetta and Hochwald [5] under a maximum peak transmitted power constraint. Codewords $\{S_i\}$ for this channel are complex $T \times M$ matrices whose T rows represent temporal coordinates and whose M columns represent spatial coordinates, indexed over T transmitted time samples and M transmitter antennas, respectively. The set of peak constrained codewords $\mathcal{S}$ are the set of complex $T \times M$ matrices which satisfy the peak power constraint: $\|S_i\|^2 \leq TM$, where $\|S\|^2 = \mathrm{tr}\{SS^H\}$.

The receiver is an N element antenna array which, for L transmitted codeword matrices $\{S_i\}_{i=1}^L$, $S_i \in \mathcal{S}$, produces the sequence of $T \times N$ observation matrices

$$X_i = \sqrt{\eta} S_i H_i + W_i, \quad i = 1, \ldots, L \tag{1}$$

where $\eta = \rho/M$ is the normalized signal-to-noise ratio (SNR) with $\rho > 0$ the SNR per-entry of S_i, H_i is an $M \times N$ matrix of complex channel coefficients, and W_i is a $T \times N$ matrix of complex noises. The piecewise constant Rayleigh flat fading model corresponds to taking the $LN(T + M)$ elements of the matrices $\{H_i\}_{i=1}^L$ and $\{W_i\}_{i=1}^L$ to be i.i.d. zero mean Gaussian random variables with unit variance. The following results are proven in [3].

[0]Alfred Hero (corresponding author) is with the Dept. of EECS, University of Michigan, Ann Arbor MI, 48109-2122 `hero@eecs.umich.edu`. Tom Marzetta is with the Mathematical Sciences Research Center, Bell Laboratories, Lucent Technologies, 600 Mountain Avenue, Murray Hill, NJ 07974 `tlm@research.bell-labs.com`. Much of this work was performed while the first author was at Bell Laboratories in the summer of 1999.

2 An Integral Cutoff Rate Representation

We first obtain an integral representation for the cut-off rate R_o which depends on a pairwise dissimilarity measure $D(S_i\|S_j)$ over the set of codewords $\mathcal{S}$:

$$R_o = \max_{P\in\mathcal{P}} -\ln \int_{S_1\in\mathcal{S}} dP(S_1) \int_{S_2\in\mathcal{S}} dP(S_2)\, e^{-ND(S_1\|S_2)} \quad \text{nats/symbol.} \tag{2}$$

where

$$D(S_1\|S_2) \quad \stackrel{\text{def}}{=} \tfrac{1}{2}\ln \frac{\left|I_T+\frac{\eta}{2}(S_1S_1^H+S_2S_2^H)\right|^2}{\left|I_T+\eta S_1S_1^H\right|\left|I_T+\eta S_2S_2^H\right|}. \tag{3}$$

In (2) the maximization is performed over a suitably constrained set $\mathcal{P}$ of probability distributions P defined over the set of peak constrained codewords $\mathcal{S}$.

The dissimilarity measure $D(S_i\|S_j)$ is a pseudo-metric, i.e. it is lower bounded by a true metric, and is a decreasing function of the spatial correlation between prewhitened versions of pairs of codeword matrices. For low SNR the dissimilarity measure reduces to a distance metric equal to the trace norm of pairwise differences between outerproducts of the codeword matrices.

3 Finite (K) Dimensional Cut-off Rate

For practical coding schemes it is of interest to restrict attention to finite sets of codewords. Let K be a prespecified finite positive integer. The cut-off rate (2) restricted to discrete distributions concentrated on at most K codewords will be called the K dimensional cut-off rate and takes the form

$$\tilde{R}_o(K) \quad = \quad \max_{\{P_i,S_i\}_{i=1}^K\in\mathcal{G}^K} -\ln \tilde{Q}\left(\{P_i\}_{i=1}^K,\{S_i\}_{i=1}^K\right) = -\ln \min_{\{P_i,S_i\}_{i=1}^K\in\mathcal{G}^K} \tilde{Q}\left(\{P_i\}_{i=1}^K,\{S_i\}_{i=1}^K\right) \tag{4}$$

where $\tilde{Q}$ is the quadratic form

$$\tilde{Q}\left(\{P_i\}_{i=1}^K,\{S_i\}_{i=1}^K\right) = \sum_{i=1}^K P_i \sum_{j=1}^K P_j\, e^{-ND(S_i\|S_j)}, \tag{5}$$

and where $\mathcal{G}^K$ is a suitably constrained set of codewords $\{S_i\}_{i=1}^K$ and codeword probabilities $\{P_i\}_{i=1}^K$.

When $P_i=0$ the codeword S_i is never transmitted which motivates a natural definition of a codeword constellation.

Definition 1 *A set of matrices $\{S_i\}_{i=1}^K$ in $\mathbb{C}^{T\times M}$ is a codeword constellation if all assigned codeword probabilities P_i are strictly positive, $i=1,\ldots,K$.*

The K dimensional cut-off rate, defined as the cut-off rate for constellations whose dimension does not exceed K, is the appropriate limiting factor for finite precision coding schemes.

Define the feasibility set $\tilde{\mathcal{S}}^K_{\text{peak}}$ of peak constrained K-dimensional constellations

$$\tilde{\mathcal{S}}^K_{\text{peak}} = \left\{\{S_i\}_{i=1}^K : S_i\in\mathbb{C}^{TM},\ \|S_i\|^2\le TM, E_K^{-1}\underline{1}_K\in\mathbb{R}_+^K,\ S_iS_i^H\neq S_jS_j^H, i\neq j\right\}. \tag{6}$$

where $\underline{1}_k=[1,\ldots,1]^T\in\mathbb{R}^K$ and $E_K=\left(\left(e^{-ND(S_i\|S_j)}\right)\right)_{i,j=1}^K$ is the $K\times K$ *dissimilarity matrix* of the constellation. E_k is positive definite as long as the outerproduct matrices $\{S_iS_i^H\}_{i=1}^k$ are distinct. It can be shown that $\tilde{\mathcal{S}}^K_{\text{peak}}$ is the set of "equalizer" constellations for which there exists an "equalizer probability vector" $\underline{P}_K$ satisfying

$$E_K\underline{P}_K = c\underline{1}_k$$

for some $c>0$.

The following representation theorem asserts that $\tilde{R}_o(K)$ is attained by an equalizer constellation which maximizes a simple quadratic form.

Proposition 1 *Let K be a positive integer. The peak power constrained K dimensional cut-off rate $\tilde{R}_o(K)$ is attained by a constellation in one of the feasible sets $\tilde{\mathcal{S}}^k_{\text{peak}}$, $k=1,\ldots,K$, and*

$$\tilde{R}_o(K) \quad = \quad \ln\left\{\max_{0<k\leq K}\max_{\{S_i\}_{i=1}^k\in\tilde{\mathcal{S}}^k_{\text{peak}}} \underline{1}_k^T E_k^{-1}\underline{1}_k\right\}. \tag{7}$$

Observe that by taking $K=\infty$ in Proposition 1, we obtain the cut-off rate of constellations of countable, but possibly infinite, dimension. The objective function $\underline{1}_k^T E_k^{-1}\underline{1}_k$ maximized in (7) is similar to the criterion used in Capon's method, also known as minimum variance distortionless response (MVDR), for adapting the weights of a beamforming array of antenna elements and for high resolution spectral estimation [4]. The proof of the proposition uses the Kuhn Tucker conditions for constrained optimization.

4 Special Limiting Cases

4.1 Large Dimension K

Let $D^{**}_{\min} = \max_{\{S_i\}_{i=1}^K : S_i\in\mathcal{S}} \min_{i\neq j} D(S_i\|S_j)$ be the maximin distance of any K dimensional constellation of $T\times M$ complex codeword matrices subject to the peak power constraint. The following lower bound on the average distance of the optimal K dimensional cut-off rate attaining codeword set is obtained via a sphere packing argument.

Proposition 2 *Let $\{S_i^*\}_{i=1}^K$ be a constellation that attains the K dimensional cut-off rate $\tilde{R}_o(K)$. Then, if $K\gg\lfloor T/M\rfloor$*

$$\overline{D(S_i^*\|S_j^*)} \geq D^{**}_{\min} \geq \frac{\eta^2\Delta_T^{1/T}TM}{8}\,K^{-1/T},$$

where $\overline{D(S_i^*\|S_j^*)}$ is the mean value

$$\overline{D(S_i^*\|S_j^*)} = \frac{\sum_{i\neq j}P_i^*P_j^*D(S_i^*\|S_j^*)}{\sum_{i\neq j}P_i^*P_j^*},$$

and Δ_T is the Minkowski lower bound on sphere packing density in $\mathbb{R}^T$ [1].

The previous two propositions can be combined to yield the following useful lower bound on $\tilde{R}_o(K)$ for $K\gg\lfloor T/M\rfloor$

$$\tilde{R}_o(K) \geq \max_{1\leq k\leq K}\ln\left(\frac{k}{1+(k-1)e^{-\gamma k^{-1/T}}}\right).$$

where $\gamma = N\eta^2\Delta_T^{1/T}TM/8$. Numerical evaluation indicates that the bound on $\tilde{R}_o(K)$ and the maximizer $k=\text{argmax}_k(\cdot)$ of the bound, are both slowly increasing in η, N and T.

4.2 Large Number of Receivers N

Proposition 3 *Let $\{S_i\}_{i=1}^K$ be a set of peak constrained codewords with distinct outerproducts $\{S_iS_i^H\}_{i=1}^K$. Define the finite positive integer N_o*

$$N_o = \lfloor\frac{\ln(K-1)}{D_{\min}}\rfloor + 1, \tag{8}$$

where $D_{\min} = \min_{i\neq j} D(S_i\|S_j)$. Then, for $N\geq N_o$

$$\tilde{R}_o(K) = \max_{\{S_i\}\in\mathcal{S}^K_{\text{peak}}} -\ln\frac{1}{K^2}\sum_{i,j=1}^K e^{-ND(S_i\|S_j)} + O(\delta^{N-N_o+1}),$$

where $\delta = \exp(-D_{\min})$.

The significance of the above proposition is that if K is no greater than $\exp(ND_{\min})+1$ the cutoff rate acheiving constellation has equi-probable codewords ($P_i = 1/K$) and it minimizes the arithmetic mean of the exponentiated pairwise distances $\frac{1}{K^2}\sum_{i,j=1}^{K} e^{-ND(S_i\|S_j)}$. Furthermore, for $K \geq \lfloor T/M \rfloor$, by Proposition 2 we have the bound $N_o \leq \lfloor 8K^{1/T}\ln(K-1)/(\eta^2 TM)\rfloor$.

4.3 High SNR η

Define

$$D_{\max} = M \ln \frac{1+\eta/2}{\sqrt{1+\eta}}, \tag{9}$$

Proposition 4 *Assume that $M \leq T/2$ and $K \leq \lfloor T/M \rfloor$. If $\eta > \eta_T$ the peak constrained K dimensional cut-off rate is*

$$\tilde{R}_o(K) = \ln\left(\frac{K}{1+(K-1)e^{-ND_{\max}}}\right)$$

and $D_{\max}$ is given in (9). Furthermore, the optimal constellation attaining $\tilde{R}_o(K)$ is the set $\{S_i^*\}_{i=1}^K$ with $S_i^* = \sqrt{T}\Phi_i$ where $\{\Phi_i\}_{i=1}^K$ are a set of $T\times M$ full rank mutually orthogonal unitary matrices: $\Phi_i^H\Phi_i = I_M$, and $\Phi_i^H\Phi_j = 0$, $i \neq j$, and the optimal probability assignment to $\{S_i^*\}_{i=1}^K$ is uniform: $P_i^* = 1/K$, $i = 1,\ldots,K$.

Thus, assuming that the number of transmit antenna elements M is less than half of the period T over which fade coefficients are assumed constant, that the received SNR exceeds the specified threshold, and that the constellation has at most T/M codewords, the optimal constellation is a set of scaled mutually orthogonal unitary matrices in $\mathbb{C}^{T\times M}$. It can be shown that this constellation also maximizes minimum distance over all constellations of the same dimension. The high SNR regime of operation occurs under either of the following conditions: high transmitted power, low receiver noise, low rms fade-induced signal attenuation, small number of transmit antennas, or large constant fade period T. Interestingly, the number of receive antennas N plays no role whatsoever in determining the SNR threshold.

References

[1] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*, Springer-Verlag, New York, NY, 1988.

[2] J. Hagenauer, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. on Inform. Theory*, vol. IT-42, no. 2, pp. 429–445, Mar. 1996.

[3] A. O. Hero and T. L. Marzetta, "On computational cut-off rate for space time coding," Technical report, Bell Laboratories, Lucent Technologies, Murray Hill, NJ, 1999.

[4] S. M. Kay, *Modern Spectral Estimation: Theory and Application*, Prentice-Hall, Englewood-Cliffs N.J., 1988.

[5] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh fading," *IEEE Trans. on Inform. Theory*, vol. IT-45, pp. 139–158, Jan. 1999.

[6] J. E. Savage, "Sequential decoding – the computation problem," *Bell Syst. Tech. Journ.*, vol. 45, pp. 149–175, Jan. 1966.

[7] F.-Q. Wang and D. J. Costello, "Probabilistic construction of large constraint length trellis codes for sequential decoding," *IEEE Trans. on Communications*, vol. COM-43, no. 9, pp. 2439–2448, Sept. 1995.

[8] J. M. Wozencraft and R. S. Kennedy, "Modulation and demodulation for probabilistic coding," *IEEE Trans. on Inform. Theory*, vol. IT-12, pp. 291–297, July 1966.