CUT-OFF RATE AND SIGNAL DESIGN FOR THE RAYLEIGH FADING SPACE-TIME CHANNEL

Alfred O. Hero *

Dept. EECS
University of Michigan
1301 Beal Avenue
Ann Arbor, MI 48109-2122
hero@eecs.umich.edu

Thomas L. Marzetta

Mathematical Sciences Research Center
Bell Laboratories
Lucent Technologies
600 Mountain Avenue
Murray Hill, NJ 07974
tlm@research.bell-labs.com

ABSTRACT

We consider the single-user computational cut-off rate for the complex Rayleigh flat fading spatio-temporal channel under a peak power constraint. Determination of the cut-off rate requires maximization of an average error exponent over all possible space-time codeword probability distributions. This error exponent is monotone decreasing in a measure of dissimilarity between pairs of codeword matrices. For low SNR the dissimilarity function reduces to a trace norm of differences between outerproducts of pairs of codewords. We characterize the cut-off rate and the rate achieving constellation 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

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 $\|S\|^2 = \mathrm{tr}\{SS^H\}$.

Cut-off rate analysis has frequently been adopted to establish practical coding limits [7, 2] as the cut-off rate specifies the highest information rate beyond which sequential decoding becomes impractical [6, 8] and as it is frequently simpler to calculate than channel capacity.

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

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

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 presented. Proofs are given in [3].

For the flat Rayleigh fading model under a peak transmitted power constraint there is no advantage to using more transmitting antennas than time samples (Proposition 1). Furthermore, there is no advantage to transmitting signals that are not spatially orthogonal, i.e. one might as well transmit mutually orthogonal temporal waveforms at each antenna element. These parallel results of Marzetta and Hochwald for average power constraints [5].

An integral representation for the cut-off rate is obtained which depends on a pairwise dissimilarity measure over the set of signal matrices. This dissimilarity measure is a decreasing function of the spatial correlation between pairs of signal matrices. For low SNR the dissimilarity measure reduces to a distance metric equal to the trace norm of pairwise differences between outerproducts of the signal matrices.

A lower bound is given on the largest possible minimum distance for arbitrary sets of signal matrices of fixed finite

*This work was performed, in part, while the first author was visiting the Mathematical Sciences Research Center.

dimension. This is also a lower bound on the maximum distance of signals in the optimal cut-off rate attaining signal set.

A necessary and sufficient condition for a signal probability distribution to attain cut-off rate is that it equalize the decoder error rate over all possible signal matrices. We call this the equalization condition and it plays a central role in this work.

The determination of the K dimensional cut-off rate reduces to maximization of a quadratic form over the set of feasible constellations, defined as those constellations which satisfy both the peak power constraint and the finite dimensional equalization condition. This quadratic form is similar to that arising in the Capon/MVDR method for adaptive beamforming arrays. If the feasible set of K dimensional constellations is empty then the optimal constellation is necessarily of dimension less than K.

For low symbol-rate the optimal constellation is a set of scaled mutually orthogonal unitary matrices in $\mathbb{C}^{T\times M}$. This constellation also maximizes minimum distance over all constellations of the same dimension. When SNR is low the rank of the signal matrices in the constellation is one and cut-off rate is achieved by applying all power to a single antenna element at a time. As the SNR increases the rank of the signal matrices increases and more and more antenna elements are utilized. Interestingly, the number of receive antennas N plays no role whatsoever in determining how many transmit antennas should be used.

2. CUTOFF RATE REPRESENTATIONS

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 $\mathcal{S}$:

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

where

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

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 decreasing function of the spatial correlation between prewhitened versions of pairs of codeword matrices. It can be shown that $D(S_1||S_2) = \eta^2/8||S_1S_1^H - S_2S_2^H||^2 + o(\eta^2)$, so that 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.

We end this subsection with a result that parallels Theorems 1 and 2 of Marzetta and Hochwald [5], but covers the case of peak power constrained signal sets.

Proposition 1 *Assume that the transmitted signal S is constrained to satisfy the peak power constraint $||S||^2 \leq MT$. The peak power constrained cut-off rate attained with $M > T$ transmit antennas is the same as that attained with $M = T$ antennas. Therefore, there is no advantage to using more than T transmit antennas. Furthermore, for $M \leq T$ the signal matrix which achieves peak power constrained cut-off rate can be expressed as $S = V\Lambda$ where V is a $T \times T$ unitary matrix, $\Lambda = [\Lambda_M, O]^T$ is a $T \times M$ matrix, and Λ_M is a diagonal $M \times M$ matrix.*

Readers familiar with Theorems 1 and 2 of [5] might suspect that characterization of the statistical distribution P of the optimal cut-off achieving signal matrix S can be obtained. Indeed, paralleling the arguments of [5], it can be shown that, as $tr\{SS^H\} \leq TM$ is invariant to unitary premultiplication of S and as the maximization in the definition of R_o is over a concave function of P, the peak-power constrained cut-off rate is attained by random matrices of the form $S = V\Lambda$ where V is a $T \times T$ isotropically distributed matrix, $\Lambda = [\Lambda_M, 0]^T$ is a random $T \times M$ diagonal matrix, and V and Λ are statistically independent.

Define a constellation as follows

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, $i = 1, \ldots, K$.*

In the following sections we specialize to the case of discrete signal constellations for which equalizer distributions are always optimal.

3. DISCRETE CONSTELLATIONS

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 practical coding schemes.

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

$$\tilde{\mathcal{S}}_{\text{peak}}^K = \{\{S_i\}_{i=1}^K : S_i \in \mathbb{C}^{TM},\ ||S_i||^2 \leq TM,\ E_K^{-1}\underline{1}_K \in \mathbf{R}_+^K,\ S_iS_i^H \neq S_jS_j^H, i \neq j\}. \quad (4)$$

where $\underline{1}_K = [1, \ldots, 1]^T \in \mathbf{R}^K$ and $E_K = \left((e^{-ND(S_i \| S_j)}) \right)_{i,j=1}^K$ is the $K \times K$ *dissimilarity matrix* of the constellation. E_K is positive definite as long as the outerproduct matrices $\{S_i S_i^H\}_{i=1}^K$ are distinct. It can be shown that $\tilde{S}_{\text{peak}}^K$ 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 2 *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 $\tilde{S}_{\text{peak}}^k$, $k = 1, \ldots, K$, and*

$$\tilde{R}_o(K) \;=\; \ln \left\{ \max_{0<k\leq K} \max_{\{S_i\}_{i=1}^k \in \tilde{S}_{\text{peak}}^k} \underline{1}_k^T E_k^{-1} \underline{1}_k \right\}. \quad (5)$$

Observe that by taking $K = \infty$ in Proposition 2, 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 (5) 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].

4. SPECIAL LIMITING CASES

Here we specialize the cutoff rate to several limiting regimes.

4.1. Large Dimension K

Recall the definition $D_{\min} = \min_{i \neq j} D(S_i \| S_j)$. When $K \leq \lfloor T/M \rfloor$ we will see (Proposition 5) that a set of signal matrices $\{S_i\}_{i=1}^K$ exists for which $D(S_i \| S_j) = D_{\min}$ for all $i \neq j$, and which simultaneously attains the cut-off rate $\tilde{R}_o(K)$ and attains the maximum possible value $D_{\min}$. The following result establishes a lower bound on the largest possible value $D_{\min}^{**} = \max_{\{S_i\}_{i=1}^K : S_i \in S_{\text{peak}}^K} \min_{i \neq j} D(S_i \| S_j)$. A constellation $\{S_i\}_{i=1}^K$ achieving $D_{\min} = D_{\min}^{**}$ is called a maximin constellation of dimension K.

For positive integer p let $\mathcal{L}_{p,T}$ be the uniform lattice of 2^{pT} points covering the T-dimensional unit cube: $\mathcal{L}_{p,T} =$

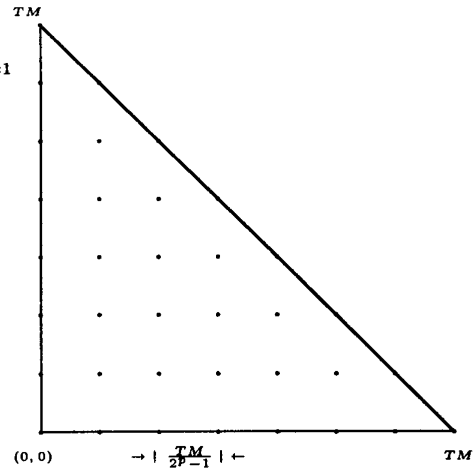

Figure 1: For $g_T(2^{p-1} - 1) < K \leq g_T(2^p - 1)$ and p a positive integer the vectors of singular values of the constellation are optimal (maximum $D_{\min}$) when they form a uniform lattice and there are exactly $K = g_T(2^p - 1)$ signals in the constellation. In the figure $T = 2$ and $p = 3$. 3.

$\{(i_1, \ldots, i_T) : \, i_l \in \{0, 1/(2^p - 1), 2/(2^p - 1), \ldots, 1\}\}$. Define the integer valued function

$$g_T(q) = \sum_{i_T=0}^{q} \sum_{i_{T-1}=0}^{q-i_T} \cdots \sum_{i_1=0}^{q-i_T-\cdots-i_2} .$$

For example, $g_2(q) = (q+1)(q+2)/2$ and $g_3(q) = (q+1)(q+2)(q+3)/6$. $g_T(2^p - 1)$ is the number of lattice points of $\mathcal{L}_{p,T}$ which are inside of the T-dimensional unit simplex $\left\{ (u_1, \ldots, u_T) : \, \sum_{l=1}^T u_l \leq 1, \; u_l \in [0, 1] \right\}$ (See Fig. 1).

Proposition 3 *For given $K > 1$ let p be the unique integer for which $g_T(2^{p-1} - 1) < K \leq g_T(2^p - 1)$. Then*

$$D_{\min}^{**} \geq \frac{\eta^2}{8} \frac{(TM)^2}{(2^p - 1)^2} + o(\eta^2) > \frac{\eta^2 (TM)^2}{128} K^{-2/T} + o(\eta^2),.$$

The first inequality is tight when $K = g_T(2^p - 1)$ and the anti-diagonal matrices $Z_i = S_i S_i^H - \operatorname{diag}(S_i S_i^H)$ satisfy $\|Z_i - Z_j\| = 0$, $i \neq j$.

The proposition says that, for sufficiently small SNR η^2, D_{min}^{**} cannot asymptotically decrease to zero faster than rate $K^{-2/T}$. This asymptotic rate can be achieved, for example, in the case that $M = T$, the $\{S_i S_i^H\}_{i=1}^K$'s are diagonal and form the constellation of a linearly constrained lattice code [1] in $\mathbf{R}^T$. Specifically, each S_i is a square matrix with orthogonal rows and the vectors of diagonals of the

$S_iS_i^H$ matrices are locations of the tie points of the K point lattice shown in Fig. 1.

The above result leads to a lower bound on the average distance $\overline{D(S_i^*\|S_j^*)}$ of the optimal K-dimensional cut-off rate achieving constellation, $\{S_i^*\}_{i=1}^K$, where

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

We can show [3] that $\overline{D(S_i^*\|S_j^*)} \geq D_{\min}^{**}$ so that

$$\overline{D(S_i^*\|S_j^*)} > \frac{\eta^2(TM/2)^2}{8} K^{-2/T} + o(\eta^2).$$

4.2. Low Dimensional Constellations

When the number of signal matrices K to be considered is sufficiently small significant simplification of the cut-off rate computation is possible. In particular, one obtains optimality of a set of scaled mutually orthogonal unitary signal matrices and a simple form form for $\tilde{R}_o(K)$.

The first result specifies the solution to optimization of the dissimilarity measure $D(S_i\|S_j)$ defined in (3).

For given η, T and M define the integer M_o

$$M_o = \text{argmax}_{m\in\{1,\ldots,M\}} \left\{ m \ln \frac{(1+\eta TM/(2m))^2}{1+\eta TM/m} \right\}. \quad (7)$$

We will see below that under some conditions M_o is the rank of the signal matrices S_i in the optimal K-dimensional constellation.

The proof of the following proposition is based on the alternative but equivalent representation for $D(S_1\|S_2)$

$$D(S_1\|S_2) = \frac{1}{2}\ln \frac{\left|I_M+\frac{\eta}{2}S_1^HS_1\right|^2 \left|I_M+\frac{\eta}{2}S_2^HS_2\right|^2}{\left|I_M+\eta S_1^HS_1\right|\left|I_M+\eta S_2^HS_2\right|} \left|I_M - \kappa^H\kappa\right|^2,$$

where κ is a $M\times M$ multiple signal correlation matrix

$$\kappa = \tilde{S}_2^H \tilde{S}_1$$

Proposition 4 *Let* $2M \leq T$. *Then*

$$D_{\max} \stackrel{\text{def}}{=} \max_{S_1,S_2\in S_{\text{peak}}^K} D(S_1\|S_2) = M_o \ln \frac{(1+\eta TM/(2M_o))^2}{1+\eta TM/M_o}.$$

Furthermore, the optimal signal matrices which attain $D_{\max}$ can be taken as scaled rank M_o mutually orthogonal unitary $T\times M$ matrices of the form

$$S_1 = \sqrt{TM/M_o}\,\Phi_1, \qquad S_2 = \sqrt{TM/M_o}\,\Phi_2$$

where, for $j = 1, 2$,

$$\Phi_j^H\Phi_j = I_{M_o}, \quad \text{and} \quad \Phi_i^H\Phi_j = 0,\ i\neq j.$$

The assumption $2M \leq T$ is critical and ensures that the singular vectors of S_1 and S_2 can be chosen as mutually orthogonal for any set of singular values.

The rank M_o of the optimal matrices S_1 and S_2 increases from 1 to M as the SNR parameter ηTM increases from 0 to ∞ (see Fig. 2). Numerical evaluation has shown that the functional relationship between M_o and SNR is well approximated by the relation

$$M_o \approx \max\left(1, \lfloor a\eta TM + b + 0.48\rfloor\right)$$

where a, b are the slope and intercept of the least squares linear fit to the function $y(x) = \text{argmax}_{m=1,2,\ldots} m \ln[(1+x/(2m))^2/(1+x/m)]$. The approximation is a lower bound and underestimates the exact value of M_o, given by (7), by at most 1 over less than 0.5% of the SNR range shown in Fig. 2 ($0 < \eta TM \leq 120$). If the SNR is sufficiently large, e.g. (from Fig. 2) $\eta TM \geq 17$ for $M = 6$ and $T \geq 12$, $M_o = M$ and the optimal signal matrices utilize all M transmit antennas. On the other hand for small SNR, i.e. (from Fig. 2) $\eta TM < 4$, $M_o = 1$ and the optimal signal matrices apply all available transmit power to a single antenna element over the coherent fade interval T.

The final result of this section is an expression for the cut-off rate.

Proposition 5 *Let $2M \leq T$ and let M_o be as defined in (7). Suppose that $M_o \leq \min\{M, T/K\}$. Then 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 by (8). Furthermore, the optimal constellation attaining $\tilde{R}_o(K)$ is the set of K rank M_o mutually orthogonal unitary matrices and the optimal probability assignment is uniform: $P_i^* = 1/K$, $i = 1,\ldots,K$.

Any unitary transformation on the columns (spatial coordinates) of a set of signal matrices produces a set of signal matrices with identical $D_{\min}$. In particular, any set of K mutually orthogonal $T\times M_o$ permutation matrices has optimal distance properties. This simple set of signal matrices corresponds to transmitting energy on a single antenna element at a time, among a total of $M_o \leq M$ elements, in each of the available T time slots. Since $\tilde{R}_o(K)$ is increasing in K the maximum cut-off rate achievable using these mutually orthogonal unitary matrices is obtained by using the maximum possible number of them: $K = \lfloor T/M_o \rfloor$. Observe that the resulting optimal constellation may correspond to a code of quite low symbol rate, e.g. for $M_o = M = T/2$ the symbol rate is only 1 bit-per-symbol.

It is noteworthy that the optimal peak constrained signal constellation specified by Proposition 5 does not include the zero valued signal matrix $S_i = O$. Including zero in the signal constellation would allow signalling using on-off keying. On-off keying is often proposed for average power constrained signalling over low SNR channels since it permits energy discrimination at the receiver. As contrasted with on-off keying all signals in the optimal peak constrained signal set have equal power. We conjecture that the zero signal would result from replacing the peak power constraint with an average power constraint in Proposition 5.

4.3. Conclusions

We have derived representations for the computational cut-off rate for space time coding under the Rayleigh flat fading channel model under a peak transmitted power constraint. For finite dimensional constellations the cut-off rate and the optimal signal distribution were specified as a solution to a quadratic optimization problem and it was shown that optimal constellations have codeword distributions which satisfy an equalization condition. This characterization of optimality motivated us to study properties of the set of feasible constellations which satsify the equalization property. Easily verifiable necessary and sufficient conditions were given for validating that a given signal constellation lies in the feasible set. Based on one of these conditions in [3] a greedy procedure was proposed for recursively constructing or refining a good feasible constellation.

5. 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, 2000.

[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.

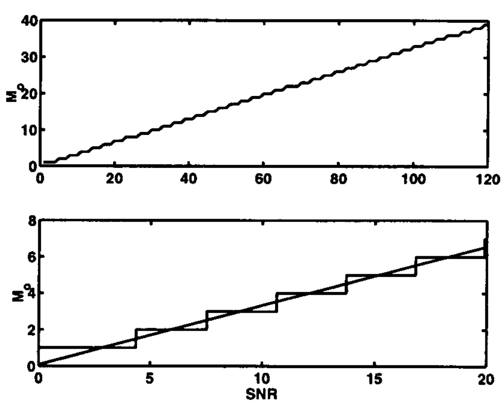

Figure 2: Top panel shows M_o given by (7) as a function of the SNR parameter ηTM. Bottom panel is blow up of first panel over a reduced range of SNR.