Optimal Signature Sets for Transmission of Correlated Data over a Multiple Access Channel

Joydeep Acharya, Ritabrata Roy, Jasvinder Singh, Christopher Rose
Wireless Information Network Laboratory (WINLAB)
73 Brett Rd, Piscataway, NJ 08854
{joy, rito, jasingh, crose}@winlab.rutgers.edu

***Abstract*— For multiple transmitters sending independent data to a single receiver, the problem of optimizing transmitter codewords to maximize capacity has been addressed in [11][12]. This paper considers an analogous scenario when the information sent by the transmitters is correlated. The optimal codeword set and power allocation which minimizes TMSE (total mean square error) at the receiver under a total power constraint have been derived. The equivalence between TMSE and sum capacity is also shown, in the sense that minimizing the former corresponds to maximizing the latter.**

I. Introduction

The information-theoretic capacity of a single cell symbol synchronous white gaussian noise CDMA system was derived in terms of the correlations between user signature waveforms by Verdu [11]. Subsequently, a lot of research work has been done in the area of signature waveform optimization for single cell CDMA type systems [12].

Assuming a finite dimensional signal space, the signature waveforms of users can be described as L-dimensional vectors (codewords) where L is called the spreading gain of the system.

For an average power constraint on symbols of all users, Massey *et al.* [8] showed that the capacity maximizing codewords for the single cell symbol synchronous system are same as the WBE sequences. Viswanath *et al.* [12] generalized the result to the case where the user power constraints are unequal. Further extensions include the colored noise case [13] and joint optimizations of codeword/power levels for fading channels [3].

In a CDMA system, the user's symbols are assumed to be independent of each other and all the above work maintains this assumption. However, there may be scenarios in which transmitters send correlated data to a receiver. For example, in the literature for sensor networks [1], one readily comes across a scenario where sensor nodes (analogous to users in CDMA) measure a common physical phenomenon and send their observations (which are correlated) to a central node.

In many cases, the physical phenomenon under observation can be abstracted as a 2-dimensional information source with spatially varying information density. Since sensor nodes usually have a non–replenishable source of energy, it is highly desirable to keep the transmission powers at their minimum levels. Moreover, since sensors are usually assumed to be deployed in very large numbers, measurements of spatially closer sensors will have a high degree of correlation. It is therefore likely that a fewer number of sensors are sufficient to sense the entire region. Hence, minimizing the total transmission power of a cluster of sensors is more meaningful than optimizing the individual transmit powers.

We assume a sensor network model where nodes use signature waveforms (codewords) to transmit their data to a common receiver and find the optimal codewords which minimize a performance metric (TMSE, defined later) under a total power constraint.

The rest of this paper is arranged as follows. We present the system model in Section II and derive the relevant TMSE expression. In Section III we introduce the notion of majorization and some related results that are required for our analysis. In Section IV we derive the optimal transmitter codewords, power levels and receiver filters by minimizing the TMSE and in Section V we establish an equivalence between between TMSE minimization and sum capacity maximization. Finally, we conclude with a summary and discussion of possible future research in Section VI.

II. Problem Statement

Assuming M users transmitting symbols using unit–norm codewords of length L in an additive white Gaussian channel, the signal at the receiver is given by:

$$\mathbf{r} = \mathbf{S}\mathbf{P}^{\frac{1}{2}}\mathbf{b} + \mathbf{n} \tag{1}$$

where,

$\mathbf{P}_{M\times M}$	: diag $(p_1 p_2 \dots p_M)$
p_i	: transmit power of ith transmitter
$\mathbf{S}_{L\times M}$	: $[\mathbf{s}_1\mathbf{s}_2 \dots \mathbf{s}_M]$
$\mathbf{s}_i$	: unit norm signature codeword of ith transmitter
$\mathbf{b}$	: symbol vector
$\mathbf{n}$	: zero–mean Gaussian noise with variance $\sigma^2\mathbf{I}_L$

$\mathbf{B}_{M\times M} = \mathrm{E}\left[\mathbf{b}\mathbf{b}^\top\right]$ is defined as the symbol correlation matrix.

Assuming a linear receiver filter, $\mathbf{c}_i$, corresponding to the i^{th} transmitter, the filter output is given by:

$$y_i = \mathbf{r}^\top\mathbf{c}_i \tag{2}$$

The mean square error (MSE) corresponding to the i^{th} transmitter is given by,

$$\text{MSE}_i = \text{E}\left[\left(\mathbf{r}^\top \mathbf{c}_i - b_i\right)^2\right] \tag{3}$$

which allows us to define total MSE as

$$\begin{aligned}\text{TMSE} &= \sum_{i=1}^{M} \text{MSE}_i \\ &= \sum_{i=1}^{M} \mathbf{c}_i^\top \left(\mathbf{SP}^{\frac{1}{2}}\mathbf{BP}^{\frac{1}{2}}\mathbf{S}^\top + \sigma^2 \mathbf{I}_L\right) \mathbf{c}_i + M \\ &\quad - 2\sum_{i=1}^{M} \mathbf{c}_i^\top \mathbf{SP}^{\frac{1}{2}} \text{E}\left[\mathbf{b} b_i\right] \\ &= \text{tr}\Big(\mathbf{C}^\top \mathbf{SP}^{\frac{1}{2}}\mathbf{BP}^{\frac{1}{2}}\mathbf{S}^\top \mathbf{C} + \sigma^2 \mathbf{C}^\top \mathbf{C} \\ &\quad -2\mathbf{C}^\top \mathbf{SP}^{\frac{1}{2}}\mathbf{B} + \mathbf{I}_M\Big)\end{aligned} \tag{4}$$

The optimization problem can then be stated as follows:

$$\min_{\mathbf{S},\mathbf{P},\mathbf{C}} \text{TMSE} \quad \text{subject to } \text{tr}\left(\mathbf{P}\right) = P_{\text{tot}} \tag{5}$$

III. Majorization: Definitions and Some Key Results

This section outlines certain mathematical relationships that are needed in obtaining the results in this paper. A detailed survey of these inequalities and their properties may be found in [4].

Definition 1: Let $\mathbf{x} = (x_{[1]}, x_{[2]}, \ldots, x_{[n]})$ and $\mathbf{y} = (y_{[1]}, y_{[2]}, \ldots, y_{[n]})$ be decreasing sequences of real numbers. Then, $\mathbf{x}$ is *majorized* by $\mathbf{y}$ (denoted by $\mathbf{x} \prec \mathbf{y}$) if

$$\sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]}, \quad k = 1, 2, \ldots, n-1$$

$$\text{and,} \quad \sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}$$

Thus, majorization of $\mathbf{x}$ by $\mathbf{y}$ suggests that the components of $\mathbf{x}$ are "less spread out" or "more nearly equal" than the components of $\mathbf{y}$.

An important example of majorization between two vectors is the following:

Example 1: For every $\mathbf{a} \in \Re^n$ such that $\sum_{i=1}^{n} a_i = 1$,

$$(a_1, a_2, \ldots, a_n) \succ \left(\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}\right)$$

Definition 2: A real–valued function $\phi : \Re^n \to \Re$, defined on a set $\mathcal{A} \subset \Re^n$, is Schur–convex on $\mathcal{A}$ if

$$\mathbf{x} \prec \mathbf{y} \text{ on } \mathcal{A} \Rightarrow \phi(\mathbf{x}) \leq \phi(\mathbf{y})$$

The function ϕ is strictly Schur–convex if $\mathbf{x} \prec \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$ implies that $\phi(\mathbf{x}) < \phi(\mathbf{y})$. Also, the function ϕ is Schur–concave if $-\phi$ is Schur–convex.

An important class of Schur–convex functions is the following:

Example 2: If $g : \Re \to \Re$ is convex and increasing, then $\phi(\mathbf{x}) = \sum_{i=1}^{n} g(x_i)$ is increasing and Schur–convex.

IV. Optimal Transmitter Codewords, Power Levels and Receiver Structure

It is well–known [10] that the structure of the optimum linear receiver that minimizes the MSE is the MMSE receiver. For this problem, the expression for the optimum receiver was obtained as:

$$\mathbf{C}^\star = \left(\mathbf{SP}^{\frac{1}{2}}\mathbf{BP}^{\frac{1}{2}}\mathbf{S}^\top + \sigma^2 \mathbf{I}_L\right)^{-1} \left(\mathbf{SP}^{\frac{1}{2}}\mathbf{B}\right) \tag{6}$$

Substituting (6) in (4), the TMSE expression reduces to:

$$\begin{aligned}\text{TMSE} &= M - \text{tr}\left[\mathbf{BP}^{\frac{1}{2}}\mathbf{S}^\top \left(\sigma^2 \mathbf{I}_L + \mathbf{SP}^{\frac{1}{2}}\mathbf{BP}^{\frac{1}{2}}\mathbf{S}^\top\right)^{-1} \mathbf{SP}^{\frac{1}{2}}\mathbf{B}\right] \\ &= M - \text{tr}\left[\frac{\mathbf{BP}^{\frac{1}{2}}\mathbf{S}^\top}{\sigma^2} \left\{\mathbf{I}_L - \frac{\mathbf{SP}^{\frac{1}{2}}\mathbf{BP}^{\frac{1}{2}}\mathbf{S}^\top}{\sigma^2} \right.\right. \\ &\quad \left.\left. + \left(\frac{\mathbf{SP}^{\frac{1}{2}}\mathbf{BP}^{\frac{1}{2}}\mathbf{S}^\top}{\sigma^2}\right)^2 - \cdots\right\} \mathbf{SP}^{\frac{1}{2}}\mathbf{B}\right] \\ &= M - \text{tr}(B) + \sigma^2 \text{tr}\left[\left(\sigma^2 \mathbf{B}^{-1} + \mathbf{P}^{\frac{1}{2}}\mathbf{S}^\top \mathbf{SP}^{\frac{1}{2}}\right)^{-1}\right]\end{aligned} \tag{7}$$

Note that $\mathbf{SP}^{\frac{1}{2}}\mathbf{BP}^{\frac{1}{2}}\mathbf{S}^\top$ is positive definite, which implies that $\left(\mathbf{SP}^{\frac{1}{2}}\mathbf{BP}^{\frac{1}{2}}\mathbf{S}^\top + \sigma^2 \mathbf{I}_L\right)$ is invertible. Also, it has been assumed in the above analysis that $\mathbf{B}^{-1}$ exists. However, it will be argued at the end of this section that invertibility of $\mathbf{B}$ is not necessary since it does not affect the structure of the optimum codewords.

Let $\mathbf{B} = \mathbf{U}_1 \mathbf{\Sigma}_1 \mathbf{U}_1^\top$ and $\mathbf{A} = \mathbf{SP}^{\frac{1}{2}} = \mathbf{U}_2 \mathbf{\Sigma}_2 \mathbf{V}_2^\top$
where $\mathbf{\Sigma}_1 = \text{diag}\left(\lambda_1, \lambda_2, \ldots, \lambda_M\right)$
such that, $\lambda_1 > \lambda_2 > \ldots > \lambda_M$
and $\mathbf{\Sigma}_2 = \left[\text{diag}(\mu_1, \mu_2, \ldots, \mu_L), \mathbf{0}_{L\times(M-L)}\right]$

Note that $\mathbf{S}$ and $\mathbf{P}^{\frac{1}{2}}$ can be obtained from $\mathbf{A}$ as the normalized columns and norms of columns of $\mathbf{A}$ respectively.

Then, the optimization problem can be rewritten as:

$$\min_{\mathbf{A}\in\mathcal{A}} \text{tr}\left[\left(\sigma^2 \mathbf{B}^{-1} + \mathbf{A}^\top \mathbf{A}\right)^{-1}\right] \tag{8}$$

where, $\mathcal{A}$ is the set of all $L \times M$ matrices such that

$$\text{tr}(\mathbf{A}^\top \mathbf{A}) = \sum_{j=1}^{L} \mu_j^2 = P_{\text{tot}}$$

Lemma 1: $\forall \mathbf{A} \in \mathcal{A}$, $\exists \widetilde{\mathbf{A}} \in \mathcal{A}$ such that $\text{TMSE}(\widetilde{\mathbf{A}}) \leq \text{TMSE}(\mathbf{A})$ and $\widetilde{\mathbf{A}}^\top \widetilde{\mathbf{A}}$ commutes with $\mathbf{B}$.

Proof: Marshall and Olkin [4, Lemma 9.G.4] state the following:

$$\det\left(\mathbf{G} + \mathbf{H}\right) \leq \prod_{i=1}^{n} \left(\lambda_{[i]}(\mathbf{G}) + \lambda_{[n+1-i]}(\mathbf{H})\right) \tag{9}$$

Define a function $\theta(\mathbf{A}) = \det\left(\sigma^2\mathbf{B}^{-1} + \mathbf{A}^\top\mathbf{A}\right)$.

Choose $\mathbf{G} = \sigma^2\mathbf{B}^{-1}$ and $\mathbf{H} = \mathbf{A}^\top\mathbf{A}$ following a similar argument as in [13]. Define $\widetilde{\mathbf{A}} = \mathbf{A}\mathbf{Q}$, where $\mathbf{Q}$ is an orthogonal matrix chosen so that $\sigma^2\mathbf{B}^{-1}$ and $\widetilde{\mathbf{A}}^\top\widetilde{\mathbf{A}}$ commute and the eigenvector corresponding to the ith largest eigenvalue of $\sigma^2\mathbf{B}^{-1}$ is the same as that corresponding to the $(n+1-i)$th largest eigenvalue of $\widetilde{\mathbf{A}}^\top\widetilde{\mathbf{A}}$.

Note that $\widetilde{\mathbf{A}} \in \mathcal{A}$ since $\mathrm{tr}(\widetilde{\mathbf{A}}^\top\widetilde{\mathbf{A}}) = \mathrm{tr}\left(\mathbf{Q}^\top\mathbf{A}^\top\mathbf{A}\mathbf{Q}\right) = P_{\text{tot}}$.

Using (9), $\theta(\widetilde{\mathbf{A}}) \geq \theta(\mathbf{A})$. Since $\theta(\mathbf{A})$ is Schur–concave and TMSE is Schur–convex in the eigenvalues of $\left(\sigma^2\mathbf{B}^{-1} + \mathbf{A}^\top\mathbf{A}\right)$, it follows that $\text{TMSE}(\widetilde{\mathbf{A}}) \leq \text{TMSE}(\mathbf{A})$. ■

Lemma 1, combined with the fact that two matrices commute if and only if they share the same eigenvectors [9], restricts the optimization space to that subset of $\mathcal{A}$ for which the condition $\mathbf{V}_2 = \mathbf{U}_1$ holds. Note that this condition is sufficient but not necessary.

Substituting $\mathbf{V}_2 = \mathbf{U}_1$ in (7), the following two cases arise.

1) $M \geq L$:

$$\begin{aligned}\text{TMSE} &= M - \mathrm{tr}\left(\mathbf{U}_1\boldsymbol{\Sigma}_1\mathbf{U}_1^\top +\right)\\ &\quad \sigma^2\mathrm{tr}\left[\left(\sigma^2\mathbf{U}_1\boldsymbol{\Sigma}_1^{-1}\mathbf{U}_1^\top + \mathbf{U}_1\boldsymbol{\Sigma}_2^\top\boldsymbol{\Sigma}_2\mathbf{U}_1^\top\right)^{-1}\right]\\ &= M - \sum_{i=1}^{M}\lambda_i + \sigma^2\sum_{i=1}^{L}\frac{1}{\frac{\sigma^2}{\lambda_i} + \mu_i^2} + \sum_{i=L+1}^{M}\frac{\lambda_i}{\sigma^2}\end{aligned} \tag{10}$$

The Lagrangian corresponding to the optimization problem at hand can be written as follows:

$$\mathcal{L}\left(\mu_1^2, \ldots, \mu_L^2, \beta\right) = \text{TMSE} + \beta\left(\sum_{i=1}^{L}\mu_i^2 - P_{\text{tot}}\right)$$

It is required that $\dfrac{\partial\mathcal{L}}{\partial\mu_i} = 0$ and $\dfrac{\partial\mathcal{L}}{\partial\beta} = 0$.

Using Kuhn-Tucker conditions [2], this leads to the following optimal solution:

$$\mu_i = \sqrt{\max\left(0, \frac{P_{\text{tot}}}{L} + \frac{\sigma^2}{L}\sum_{i=1}^{L}\frac{1}{\lambda_i} - \frac{\sigma^2}{\lambda_i}\right)} \tag{11}$$

Note that the optimal solution depends only on the first L eigenvalues of B, *i.e.*, $\{\lambda_i\}_{i=1}^{L}$. Also, the optimal solution has the property that if $\lambda_i \geq \lambda_j$, then $\mu_i \leq \mu_j$ as described in the proof for Lemma 1. It will now be shown that the ordering $O_1 : \lambda_1 > \lambda_2 > \ldots > \lambda_M$ achieves the optimal solution.

For ordering O_1, the eigenvalues $\{\gamma_i\}_{i=1}^{M}$ of $\sigma^2\mathbf{B}^{-1} + \mathbf{A}^\top\mathbf{A}$ are given by:

$$\gamma_i\left(O_1\right) = \begin{cases} \frac{\sigma^2}{L}\sum_{j=1}^{L}\frac{1}{\lambda_j} + \frac{P_{\text{tot}}}{L}, & i \leq L_1 \leq L\\ \frac{\sigma^2}{\lambda_i}, & i > L_1\end{cases}$$

It can be verified that for any other ordering O_2,

$$\{\gamma_i\}_{i=1}^{M}\left(O_2\right) \succ \{\gamma_i\}_{i=1}^{M}\left(O_1\right) \tag{12}$$

Now consider the function $f(x) = \frac{a}{x}$. It can be shown that $f(x)$ is convex if $a, x \in \Re^+$. Using Example 2, it follows that TMSE is a Schur–convex function in the eigenvalues $\{\gamma_i\}_{i=1}^{M}$ of $\sigma^2\mathbf{B}^{-1} + \mathbf{A}^\top\mathbf{A}$, which in conjunction with (12) implies that O_1 achieves the optimal solution.

2) $M < L$:

It can be verified that only the first M μ_is need to be optimized, and the remaining $(L - M)$ eigenvalues may be set to zero for obtaining the optimal solution.

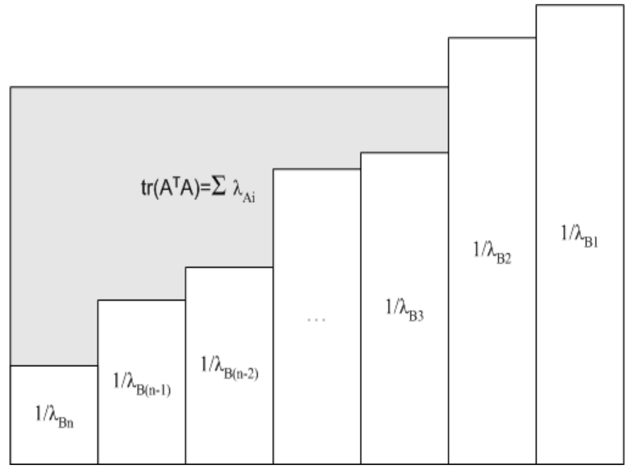

Fig. 1. Waterfilling is achieved by distributing the sum of the eigenvalues of $\mathbf{A}^\top\mathbf{A}$ over the eigenvalues of $\mathbf{B}^{-1}$.

In other words, for any M and L, the optimal solution corresponds to waterfilling (Fig. 1) the smallest $K(= \min(L, M))$ eigenvalues of $\mathbf{B}^{-1}$ with those of $\mathbf{A}^\top\mathbf{A}$, and aligning the eigenvectors of $\mathbf{A}^\top\mathbf{A}$ and $\mathbf{B}$ as described in the proof of Lemma 1.

The above analysis assumed that $\mathbf{B}$ is invertible. However, the result holds even for a non–invertible $\mathbf{B}$ since it can be made invertible by adding an infinitesimally small perturbation matrix (while ensuring that $\mathbf{B}$ is still a correlation matrix). As a result, previously non–zero eigenvalues of $\mathbf{B}^{-1}$ will suffer very little change, while the other eigenvalues (previously zero) will now attain large finite values, but the corresponding dimensions will be avoided by the waterfilling solution [5].

V. Relationship Between TMSE and Sum Capacity

Verdu [11] derives the information theoretic capacity region for a white Gaussian synchronous CDMA system. Proceeding in a similar manner, the sum capacity for the system under consideration can be expressed as:

$$C_{\text{sum}} = \frac{1}{2}\log\left[\det\left(\sigma^2\mathbf{I}_L + \mathbf{S}\mathbf{P}^{\frac{1}{2}}\mathbf{B}\mathbf{P}^{\frac{1}{2}}\mathbf{S}^\top\right)\right] - \frac{L}{2}\log\sigma^2 \tag{13}$$

where it is assumed that the symbols b_i are jointly Gaussian with known covariance $\mathbf{B}$.

It will now be shown that TMSE minimization and sum capacity maximization are equivalent problems. Using the notation defined previously,

$$C_{\text{sum}} = \frac{1}{2}\log\left[\det\left(\sigma^2\mathbf{I}_L + \mathbf{A}\mathbf{B}\mathbf{A}^\top\right)\right] - \frac{L}{2}\log\sigma^2 \qquad (14)$$

Lemma 2: $\forall\mathbf{A} \in \mathcal{A}$, $\exists\widetilde{\mathbf{A}} \in \mathcal{A}$ such that $C_{\text{sum}}(\widetilde{\mathbf{A}}) \geq C_{\text{sum}}(\mathbf{A})$ and $\widetilde{\mathbf{A}}^\top\widetilde{\mathbf{A}}$ commutes with $\mathbf{B}$.

Proof: Similar to Lemma 1. ■

As in the case of TMSE, Lemma 2 when combined with the fact that two matrices commute if and only if they share the same eigenvectors [9], restricts the optimization space to that subset of $\mathcal{A}$ for which the condition $\mathbf{V}_2 = \mathbf{U}_1$ holds. Again, this condition is sufficient but not necessary.

A similar analysis reveals that sum capacity is Schur–concave under the total power constraint, and hence minimizing TMSE is equivalent to maximizing C_{sum}.

VI. Conclusion And Future Work

This paper considered a sensor network model where sensors transmit correlated information to a receiver using a set of signature waveforms. The optimal signature waveforms and transmit power levels for minimizing the TMSE at the receiver under a total power constraint were derived. Furthermore, the equivalence between TMSE and sum capacity for our system was shown, in the sense that minimizing the former corresponds to maximizing the latter under a total power constraint.

An important area of future work is to carefully compare the efficiency of correlated data transmission using the scheme presented in this paper to that using distributed source coding [6] and define suitable metrics for comparing and contrasting the two. Also, throughout the paper we have assume that different transmitters operate under a total power constraint. Search of optimal codewords under individual power constraints is another important problem and one where we expect ideas from [12], [7] to prove especially useful.

References

[1] I.F. Akyildiz, Y. Sankarasubramaniam, and E. Cayirci. A Survey on Sensor Networks. *IEEE Communications Magazine*, 40(8):102–114, August 2002.

[2] R. Fletcher. *Practical Methods of Optimization*. John Wiley, New York, NY, 1988.

[3] O. Kaya and S. Ulukus. Jointly Optimal Power and Signature Sequence Allocation for Fading CDMA. December 2003.

[4] A.W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York, NY, 1979.

[5] D. C. Popescu and C. Rose. Codeword Optimization for Uplink CDMA Dispersive Channels. *IEEE Transactions on Wireless Systems*, to appear.

[6] S.S. Pradhan and K. Ramchandran. Distributed Source Coding Using Syndromes (DISCUS). *IEEE Transactions on Information Theory*, 49(3):626–643, March 2003.

[7] C. Rose. CDMA Codeword Optimization: Interference Avoidance and Convergence via Class Warfare. *IEEE Transactions on Information Theory*, 47(7):2368–2382, September 2001.

[8] M. Rupf and J.L. Massey. Optimum Sequence Multisets for Synchronous Code-Division Multiple-Access Channels. *IEEE Transactions on Information Theory*, 40(4):1226–1266, July 1994.

[9] Gilbert Strang. *Linear Algebra and its Applications*. International Thompson Publishing, Singapore, 1988.

[10] Sennur Ulukus and Aylin Yener. Iterative Transmitter and Receiver Optimization for CDMA Networks. *IEEE Transactions on Wireless Communications*, to appear.

[11] Sergio Verdu. Capacity Region of Gaussian CDMA Channels: The Symbol–Synchronous Case. *Proceedings of the 24th Allerton Conference on Communication, Control and Computing*, 1989.

[12] Pramod Viswanath and Venkat Anantharam. Optimal Sequences and Sum Capacity of Synchronous CDMA Systems. *IEEE Transactions on Information Theory*, 45(6):1984–1991, September 1999.

[13] Pramod Viswanath and Venkat Anantharam. Optimal Sequences for CDMA Under Colored Noise: A Schur–Saddle Function Property. *IEEE Transactions on Information Theory*, 48(6):1295–1318, June 2002.