

# The Capacity of the Frequency/Time-Selective Fading Channel

Antonia Tulino  
Bell Labs

Giuseppe Caire  
USC

Shlomo Shamai  
Technion

Sergio Verdú  
Princeton

**Abstract**—We find the capacity of discrete-time channels subject to both frequency-selective and time-selective fading, where the channel output is observed in additive Gaussian noise. A coherent model is assumed where the fading coefficients are known at the receiver. Capacity depends on the first-order distributions of the fading processes in frequency and in time, which are assumed to be independent of each other, and a simple formula is given when one of the processes is iid and the other one is sufficiently mixing. When the frequency-selective fading coefficients are known also to the transmitter, we show that the optimum normalized power spectral density is the waterfilling power allocation for a reduced signal-to-noise ratio, where the gap to the actual signal-to-noise ratio depends on the fading distributions.

**Index Terms**—Channel capacity, random matrices, frequency-selective fading, frequency-flat fading, coherent communications, additive Gaussian noise, waterfilling, OFDM.

## I. INTRODUCTION

The simplest discrete-time additive-noise channel subject to fading is the time-selective coherent model:

$$y_i = \sqrt{\gamma} A_i x_i + n_i, \quad i = 1, \dots, n \quad (1)$$

where the complex-valued input codeword  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{C}^n$  is subject to a unit average power constraint,  $\{n_i \in \mathbb{C}\}$  is a unit variance iid complex Gaussian random process, and  $\{A_i \in \mathbb{C}\}$  is a stationary and ergodic fading process known at the receiver. In vector form, (1) becomes

$$\mathbf{y} = \sqrt{\gamma} \mathbf{A} \mathbf{x} + \mathbf{n}, \quad (2)$$

where  $\mathbf{A} = \text{diag}\{A_1, \dots, A_n\}$ . If the decoder (but not the encoder) knows the actual fading realization, the capacity of (1) is equal to (e.g. [1])

$$C(\gamma) = \mathbb{E} [\log (1 + \gamma |A|^2)] \quad (3)$$

where the expectation is with respect to the random variable  $A$  distributed according to the first-order marginal distribution of the fading process  $\{A_i\}$ .

Another important model is the discrete-time frequency-selective fading channel, given by

$$\mathbf{y} = \sqrt{\gamma} \mathbf{F} \mathbf{G} \mathbf{F}^\dagger \mathbf{x} + \mathbf{n}, \quad (4)$$

where  $\mathbf{F}$  is an  $n \times n$  unitary matrix with coefficients

$$\mathbf{F}_{i,k} = \frac{1}{\sqrt{n}} e^{-\frac{j2\pi}{n}(i-1)(k-1)}. \quad (5)$$

The columns of  $\mathbf{F}$  form an  $n$ -dimensional unitary discrete-time Fourier basis, and the fading coefficients affecting the transmitted signal frequency components are denoted by  $\mathbf{G} = \text{diag}\{G_1, \dots, G_n\}$ . Note that the random channel matrix  $\mathbf{F} \mathbf{G} \mathbf{F}^\dagger$  is circulant.

The model in (4) encompasses the random linear time-invariant channel

$$y_i = \sqrt{\gamma} \sum_{\ell=0}^L h_\ell x_{i-\ell} + z_i, \quad i = 1, \dots, n, \quad (6)$$

where  $\{h_\ell\}$  denotes the (random) channel impulse response, under the assumption of cyclic prefix pre-coding and  $L \ll n$  [2].

In most physically meaningful channel models (see [1] and references therein) the diagonal coefficients of  $\mathbf{G}$  are identically distributed. If, moreover, they are cyclically stationary (the joint distribution is invariant to cyclic shifts), then the impulse response coefficients are uncorrelated, which is a common assumption. Using the fact that  $\mathbf{F}$  is unitary and under ergodicity and stationarity assumptions on the fading coefficients, the capacity of (4) is given by (again, assuming knowledge of  $\mathbf{G}$  at the decoder but not at the encoder)

$$C(\gamma) = \mathbb{E} [\log (1 + \gamma |\mathbf{G}|^2)]. \quad (7)$$

Both (3) and (7) are achieved by Gaussian iid input vectors  $\mathbf{x}$ . When the encoder knows  $\mathbf{G}$ , then it allocates power according to the waterfilling formula. In fact, in the familiar case of a deterministic linear time-invariant system with transfer function  $H(f)$ ,  $-\frac{1}{2} \leq f \leq \frac{1}{2}$ , the mutual information achieved by a stationary Gaussian input process with power spectral density  $S_x(f)$  is equal to the right side of (7) with  $G = S_x(U) |H(U)|^2$  and  $U$  uniformly distributed on  $[-\frac{1}{2}, \frac{1}{2}]$ .

A general discrete-time coherent fading model is given by the noisy version of the output of a linear

time-varying system with random impulse response  $\{h_{i,\ell}\}$  known at the receiver:

$$y_i = \sqrt{\gamma} \sum_{\ell=0}^L h_{i,\ell} x_{i-\ell} + z_i, \quad i = 1, \dots, n \quad (8)$$

or, equivalently, in vector form

$$\mathbf{y} = \sqrt{\gamma} \mathbf{H} \mathbf{x} + \mathbf{n}, \quad (9)$$

where  $\mathbf{H}$  is the matrix representation of the convolution operator in (8). Subject to suitable stationarity and ergodicity assumptions on  $\{h_{i,\ell}\}$  the capacity is given by

$$C(\gamma) = \lim_{n \rightarrow \infty} \sup_{\text{tr}(\Sigma_x) \leq n} \frac{1}{n} \mathbb{E} [\log \det (\mathbf{I} + \gamma \mathbf{H} \Sigma_x \mathbf{H}^\dagger)] \quad (10)$$

A general closed-form formula for (10) in terms of the statistics of  $\{h_{i,\ell}\}$  has not been found yet either with or without knowledge of  $\mathbf{H}$  at the transmitter.

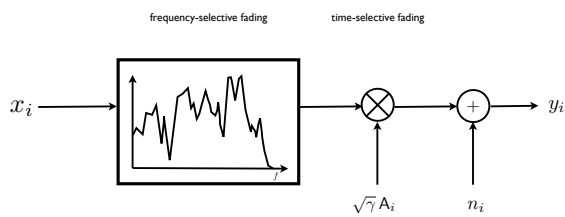


Fig. 1. Frequency-selective time-selective fading channel.

Since most mobile wireless systems are subject to both frequency-selective fading (e.g., due to multipath) and to time-selective fading (e.g., due to shadowing), it is of interest to consider a channel model that incorporates both effects. In this paper we consider the following model (Figure 1)

$$\mathbf{y} = \sqrt{\gamma} \mathbf{A} \mathbf{F} \mathbf{G} \mathbf{F}^\dagger \mathbf{x} + \mathbf{n}, \quad (11)$$

obtained by concatenating a *random circulant* matrix  $\mathbf{F} \mathbf{G} \mathbf{F}^\dagger$ , with a time-domain diagonal fading matrix, where, as defined before,  $\mathbf{A}$  and  $\mathbf{G}$  are random diagonal matrices modeling the time-selective and frequency-selective fading coefficients, respectively. Note that (11) is a special case of (8), which captures some interesting features of time and frequency selectivity. For example, we may consider a case where signaling takes place over a set of orthogonal carriers (as in OFDM), each attenuated by a random coefficient, with the whole signal then subject to a form of time-selective fading. Examples of time-selective (frequency-flat) fading include shadowing, impulsive noise/jamming that saturates the receiver input thereby

erasing some of the received values [9], and satellite communication with the presence of a line-of-sight path modeled as a Markov chain [5].

Throughout this paper, we assume that the discrete-time fading random processes  $\{A_i : i \in \mathbb{Z}\}$  and  $\{G_i : i \in \mathbb{Z}\}$  are mutually independent, stationary and ergodic. Furthermore, either the time-domain fading or the frequency-domain fading is assumed to be iid, while the other is *strong mixing*. We denote by  $A$  and  $G$  two independent random variables with the same first-order marginal distributions of  $\{A_i\}$  and  $\{G_i\}$ , respectively. Notice that  $A$  and  $G$  are independent but are allowed to have different distributions.

The main technical advance required to solve the capacity of the channel model (11) is the asymptotic spectral distribution of the matrix  $\mathbf{A} \Sigma \mathbf{A}^\dagger$ , when  $\Sigma$  is a random symmetric non-negative definite circulant matrix independent of  $\mathbf{A}$ . When the fading is known to the receiver only, the capacity is given by Theorem 1, which represents the main result of this paper. Also, in Theorem 2 we show that when the frequency-domain fading is known also to the transmitter, the capacity achieving power allocation on the channel frequency components takes on the form of the well-known “waterfilling” solution for a scaled channel signal-to-noise ratio (SNR), where the scaling coefficient can be characterized as the solution of a fixed-point equation. We also provide a number of easily computable upper and lower bounds to capacity, and simple formulas for the asymptotic behavior of capacity in the limit of small SNR are also presented. Because of space limitations, all proofs, as well as our results on the high SNR regime, are omitted.

## II. CHANNEL CAPACITY RESULTS

### A. Main Results

**Theorem 1** *The capacity of the channel model (11) with fading unknown to the transmitter is given by*

$$C(\gamma) = \mathbb{E} [\log (1 + \alpha \gamma |G|^2)] + \mathbb{E} [\log (1 + \nu \gamma |A|^2)] - \log(1 + \alpha \nu \gamma) \quad (12)$$

where

$$0 \leq \alpha \leq \mathbb{E} [|A|^2] \quad (13)$$

$$0 \leq \nu \leq \mathbb{E} [|G|^2] \quad (14)$$

are coefficients that depend on  $\gamma$  and on the fading distributions, and are defined by the solution to

$$\mathbb{E} \left[ \frac{1}{1 + \alpha \gamma |G|^2} \right] = \frac{1}{1 + \alpha \nu \gamma} = \mathbb{E} \left[ \frac{1}{1 + \nu \gamma |A|^2} \right] \quad (15)$$

Notice the interesting duality between the frequency and time domains: the distributions of the random

variables  $|A|^2$  and  $|G|^2$  play exactly the same role in the evaluation of capacity. In that respect, note that if  $\Sigma_x$  in (10) is a multiple of the identity, the determinant is the same whether  $\mathbf{H} = \mathbf{A}\mathbf{F}\mathbf{G}\mathbf{F}^\dagger$  or  $\mathbf{H} = \mathbf{G}^\dagger\mathbf{F}^\dagger\mathbf{A}^\dagger\mathbf{F}$ . In the absence of time-domain fading ( $|A|^2$  deterministic) the solution to (15) satisfies  $\alpha = |A|^2$ . In the absence of frequency-domain fading ( $|G|^2$  deterministic) the solution to (15) satisfies  $\nu = |G|^2$ . Intuitively,  $\alpha$  captures the effect of time-domain fading variations, and  $\nu$  captures the effect of frequency-domain fading variations.

Consider the following special cases of the setup in Section I:

- **Frequency-selective fading.** In the absence of time-domain fading ( $|A|^2 = 1$ ), the solution to (15) satisfies  $\alpha = 1$  and the second and third term in (12) cancel, recovering (7).
- **Time-selective fading.** For a deterministic frequency-flat channel,  $|G|^2 = 1$ . Thus, (15) is solved by  $\nu = |G|^2 = 1$ , in which case the first and third terms in (12) cancel and we obtain (3).
- **Frequency-selective fading with on-off time-selective fading.** In the special case where  $|A|^2$  takes on the values 0 or 1 with probability  $e$  and  $1 - e$ , we obtain

$$C_e(\gamma) = \mathbb{E} [\log (1 + (1 - \hat{e})\gamma|G|^2)] + d(e||\hat{e}) \quad (16)$$

where the binary divergence is defined as

$$d(a||b) = a \log \frac{a}{b} + (1 - a) \log \frac{1 - a}{1 - b} \quad (17)$$

and  $\hat{e} \geq e$  is the ( $\gamma$ -dependent) solution to

$$\frac{e}{\hat{e}} = \mathbb{E} \left[ \frac{1}{1 + (1 - \hat{e})\gamma|G|^2} \right] \quad (18)$$

- **Independent Rayleigh fading and Markov-correlated shadowing.** Here we have iid Rayleigh fading in the frequency domain and a two-state Markov shadowing process in the time domain. In particular,  $|G_i|^2$  is a sequence of independent exponential random variables with mean 1, and  $|A_i|^2$  is a Markov chain with two states, taking values  $a_0 = 0.1$  and  $a_1 = 1$ , with transition probabilities  $\mathbb{P}[a_0 \rightarrow a_1] = 0.3$  and  $\mathbb{P}[a_1 \rightarrow a_0] = 0.7$ . In order to solve (15) for  $\alpha$  and  $\nu$  as a function of  $\gamma$ , we proceed as follows. For any given  $\eta \in [0, 1]$ , let  $\alpha$  be the solution of the equation

$$\mathbb{E} \left[ \frac{1}{1 + \gamma\alpha|G|^2} \right] = \frac{e^{\frac{1}{\gamma\alpha}}}{\gamma\alpha} \int_{1/\alpha\gamma}^{\infty} \frac{e^{-t}}{t} dt = \eta \quad (19)$$

and let  $\nu$  be the solution of the equation

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{1 + \gamma\nu|A|^2} \right] &= \frac{\pi_0}{1 + \gamma\nu a_0} + \frac{\pi_1}{1 + \gamma\nu a_1} \\ &= \eta \end{aligned} \quad (20)$$

where  $(\pi_0, \pi_1)$  is the stationary distribution of the two-state Markov chain. Then, using the second equality in (15), we find (e.g., using the bisection method) the value of  $\eta$  that satisfies

$$\alpha\nu\gamma\eta = 1 - \eta \quad (21)$$

Finally, using the values of  $\alpha$  and  $\nu$  and  $\eta$  so obtained, calculate  $C(\gamma)$  using (12). Figure 2 shows the comparison between  $C(\gamma)$  and Monte Carlo simulation of the finite-dimensional mutual information formula

$$\frac{1}{n} \mathbb{E} [\log \det (\mathbf{I} + \gamma \mathbf{A}\mathbf{F}\mathbf{G}\mathbf{G}^\dagger\mathbf{F}^\dagger\mathbf{A}^\dagger)] \quad (22)$$

for  $n = 64$ . We also show the realization of the normalized log-det without the expectation, in order to give an idea of the spread of the finite-dimensional mutual information for given (random) realization of the fading processes. We notice that the agreement between simulation and the result of Theorem 1 is remarkable for even a relatively small value of  $n$ .

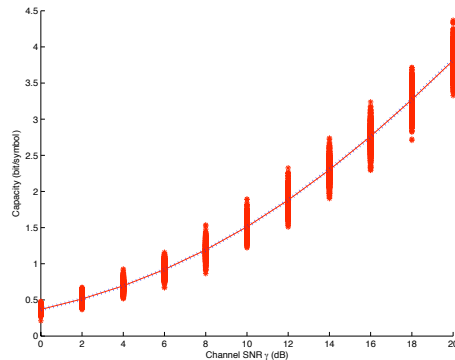


Fig. 2. Rayleigh frequency-selective fading and two-state Markov shadowing. Solid line: solution to Theorem 1; Dotted line: Monte Carlo evaluation of (22) for  $n = 64$ ; the clouds of points correspond to realizations of the random variable inside the expectation in (22). (1000 points per cluster).

### B. Optimality of waterfilling with power penalty

If the transmitter knows the frequency-domain fading coefficients, then it can choose the input covariance matrix  $\Sigma_x$  as a function of  $\mathbf{G}$  in order to maximize the mutual information. It is sufficient to consider a circulant input covariance in the form  $\Sigma_x = \mathbf{F}\mathbf{Y}\mathbf{F}^\dagger$ . In the absence of time-domain fading, maximizing the mutual information of the frequency-selective fading channel given in (6) and (4) with respect to the input power spectral density yields (e.g. [3]) the well-known waterfilling formula

$$C(\gamma) = \mathbb{E} [\log (1 + \gamma \bar{S}_x(\gamma, |G|^2) |G|^2)] \quad (23)$$

$$= \mathbb{E} [\log (\zeta_\gamma \gamma |G|^2)]^+ \quad (24)$$

where  $\bar{S}_x(\gamma, \cdot)$  is the waterfilling power allocation function:

$$\bar{S}_x(\gamma, z) = \left[ \zeta_\gamma - \frac{1}{\gamma z} \right]^+, \quad z \in \mathbb{R}_+ \quad (25)$$

and the water level  $1 < \zeta_\gamma < \infty$  is chosen in order to satisfy the transmit power constraint, i.e., such that

$$\mathbb{E} \left[ \left[ \zeta_\gamma - \frac{1}{\gamma |G|^2} \right]^+ \right] = 1 \quad (26)$$

The input power spectral density is implicitly given by the function  $\bar{S}_x(\gamma, \cdot)$  and by the realization of the frequency selective fading ( $|G_1|^2, \dots, |G_n|^2$ ). In particular, for any given blocklength  $n$  and Discrete Fourier Transform (DFT) frequencies  $\{f_i = (i-1)/n : i = 1, \dots, n\}$ , the input energy associated with the  $i$ -th frequency component is given by  $S_x(f_i) = \bar{S}_x(\gamma, |G_i|^2)$  and satisfies

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}[S_x(f_i)] = 1 \quad (27)$$

Letting  $n \rightarrow \infty$  yields the power spectral density defined on the discrete-time frequency domain, that without loss of generality can be taken to be the interval  $[0, 1]$ . In the presence of time-domain fading, the capacity-achieving input power spectral density is given by the following result.

**Theorem 2** For all  $\gamma > 0$  and  $z \geq 0$ , the capacity-achieving input power spectral density is given by

$$S_x^*(\gamma, z) = \bar{S}_x(\gamma', z) \quad (28)$$

where  $\bar{S}_x(\cdot, \cdot)$  is the waterfilling power allocation in (25) and  $\gamma = \mu(\gamma')\gamma'$  with  $\mu(\gamma') \geq 1$  being the solution of the equation:

$$\mathbb{E} \left[ \frac{\zeta_{\gamma'}}{\zeta_{\gamma'} - 1 + \mu |A|^2} \right] = 1 \quad (29)$$

where  $\zeta_{\gamma'}$  is the fading-free water level in (25) for the reduced SNR  $\gamma'$ .

We notice that the power allocation function  $S_x^*(\cdot, \cdot)$  coincides with the waterfilling power allocation function  $\bar{S}_x(\cdot, \cdot)$  for the case without time-domain fading, calculated for a lower value of the SNR parameter: namely,  $\gamma'$  instead of  $\gamma$ . In order to evaluate the capacity  $C(\gamma)$  when the frequency-domain fading is known to the transmitter, we search for the value  $\gamma' \in (0, \gamma]$  such that  $\gamma = \mu(\gamma')\gamma'$ . Then,  $C(\gamma)$  is equal to (12) with the modified fading random variable given by  $|G|^2 \left[ \zeta_{\gamma'} - \frac{1}{\gamma' |G|^2} \right]^+$ .

### C. Bounds

**Theorem 3** The capacity (12) is lower bounded by

$$C(\gamma) \geq \mathbb{E} [\log(1 + \gamma |A|^2 |G|^2)] \quad (30)$$

**Theorem 4** The capacity in (12) is lower bounded by

$$C(\gamma) \geq \mathbb{E} [\log(1 + \alpha \gamma |G|^2)] \quad (31)$$

$$C(\gamma) \geq \mathbb{E} [\log(1 + \nu \gamma |A|^2)] \quad (32)$$

$$C(\gamma) \geq \log(1 + \nu \alpha \gamma) \quad (33)$$

where  $(\alpha, \nu)$  are given in Theorem 1.

The following result yields upper bounds to capacity and shows that in the presence of one type of fading, the fading in the other domain is deleterious.

**Theorem 5** The capacity in (12) is upper bounded by:

$$C(\gamma) \leq \mathbb{E} [\log(1 + \gamma \mathbb{E}[|G|^2] |A|^2)] \quad (34)$$

$$C(\gamma) \leq \mathbb{E} [\log(1 + \gamma |G|^2 \mathbb{E}[|A|^2])] \quad (35)$$

### D. Asymptotics

1) *Low SNR Asymptotics*: In this subsection we characterize the behavior of capacity for vanishing  $\gamma$ . We define the *kurtosis* of a real random variable  $Z$  as

$$\kappa(Z) = \frac{\mathbb{E}[Z^4]}{\mathbb{E}^2[Z^2]} \quad (36)$$

**Theorem 6** In the absence of channel state information at the transmitter, the minimum energy per bit and the wideband slope  $\mathcal{S}_0$  [10] of the spectral efficiency of channel (11) are:

$$\left( \frac{E_b}{N_0} \right)_{\min} = \frac{\ln 2}{\mathbb{E}[|A|^2] \mathbb{E}[|G|^2]} \quad (37)$$

$$\mathcal{S}_0 = \frac{2}{\kappa(|G|) + \kappa(|A|) - 1} \quad (38)$$

When the transmitter knows the frequency-domain fading coefficients, then

$$\left( \frac{E_b}{N_0} \right)_{\min} = \frac{\ln 2}{\mathbb{E}[|A|^2] G_{\max}} \quad (39)$$

$$\mathcal{S}_0 = \frac{2}{\kappa(|A|) + \frac{1}{B_{\max}} - 1} \quad (40)$$

where  $G_{\max}$  is the essential supremum of the frequency-selective fading, defined as

$$G_{\max} = \sup\{z : \mathbb{P}(|G|^2 \leq z) < 1\} \quad (41)$$

and  $B_{\max} = \mathbb{P}(|G|^2 = G_{\max})$  is the probability mass at  $G_{\max}$ .

We notice that  $G_{\max}$  takes on the meaning of the ‘‘peak’’ of the frequency-domain channel fading transfer function and  $B_{\max}$  corresponds to the ‘‘bandwidth’’

(i.e., the Lebesgue measure of the set of frequencies) over which the fading takes on its maximum value. When the transmitter has knowledge of the frequency-domain fading channel, the optimal power allocation of Theorem 2 puts constant power  $1/B_{\max}$  over the frequency components  $f_i = (i-1)/n$  for which  $|G_i|^2 = G_{\max}$  and zero power elsewhere. This explains the quite different behavior of  $(E_b/N_0)_{\min}$  and  $\mathcal{S}_0$  in the cases of unknown or known frequency-domain fading at the transmitter.

### III. CONCLUDING REMARKS

We have obtained the channel capacity of a channel model that captures the effect of fading in both the time-domain and frequency-domain. The central technical result of this paper is the asymptotic freeness of the random diagonal matrix  $\mathbf{A}\mathbf{A}^\dagger$  and the random circulant matrix  $\mathbf{F}\mathbf{G}\mathbf{G}^\dagger\mathbf{F}^\dagger$ , when the coefficients in  $\mathbf{G}$  are iid independent of those of  $\mathbf{A}$  which satisfy relatively mild assumptions (or viceversa). This allows us to obtain the asymptotic eigenvalue distribution of  $\mathbf{A}\mathbf{F}\mathbf{G}\mathbf{G}^\dagger\mathbf{F}^\dagger\mathbf{A}^\dagger$  in terms of its  $\eta$ -transform, which yields the channel capacity in the case where the transmitter has no information about the realization of the fading, but only knows its statistics and the channel SNR.

For the case when the frequency-domain fading is known to the transmitter, we found the optimal frequency-domain power allocation function that takes on the form of a modified waterfilling power allocation for an SNR value lower than the actual channel SNR. Intuitively, this means that in the presence of time-domain fading it is more convenient to spread the signal energy on a smaller bandwidth than in the case of no time-domain fading, in order to “concentrate” the transmitter energy only on those frequency components with very high gain and at the same time have a signal that is more correlated in the time domain, in order to cope with the time-selectivity of the channel.

The capacity formula of Theorem 1 is given in terms of the solution of coupled fixed-point equations. Although the numerical computation of such formulas is quite straightforward, we have also provided simple upper and lower bounds that can be computed from their closed-form expressions. Finally, we have provided simple and closed-form expressions for the low-SNR approximation in terms of the fundamental asymptotic parameters  $(E_b/N_0)_{\min}$  and  $\mathcal{S}_0$ .

As illustrated numerically, and typical in random matrix theory, the convergence of

$$\frac{1}{n} \mathbb{E} [\log \det (\mathbf{I} + \gamma \mathbf{A}\mathbf{F}\mathbf{G}\mathbf{G}^\dagger\mathbf{F}^\dagger)]$$

is very fast. Just like with multiantenna systems where large-size asymptotic formulas are useful proxies for

even small arrays, in the present case, the main result gives an accurate approximation to the capacity of standardized OFDM (number of carriers ranging from 52 in IEEE802.11a to 6817 in DVB).

### ACKNOWLEDGMENTS

This work was partially supported by VIDES Project, Misura 3.17 of Regione Campania, Italy, the U. S. National Science Foundation Theoretical Foundations Research Grant CCF-0728445 and the U.S.-Israel Binational Science Foundation under Grant 2008269.

### REFERENCES

- [1] E. Biglieri, J. Proakis, and S. Shamai, “Fading channels: Information-theoretic and communications aspects,” *IEEE Trans. on Information Theory*, vol. 44, no. 6, pp. 2619-2692
- [2] W. Hirt and J. L. Massey, “Capacity of the discrete-time Gaussian channel with intersymbol interference,” *IEEE Trans. on Information Theory* vol. 34, N. 3, pp. 38 - 38, May 1988
- [3] T. Cover and J. Thomas, *Elements of Information Theory*. 2nd Ed., New York: Wiley, 2006
- [4] R. M. Gray, “Toeplitz and Circulant Matrices: A Review,” *Foundations and Trends in Communications and Information Theory*, vol. 2, issue 3, pp 155-239, 2006
- [5] E. Lutz, D. Cygan, M. Dippold, F. Dolainsky, W. Papke, “The land mobile satellite communication channel-recording, statistics, and channel model,” *IEEE Transactions on Vehicular Technology*, Volume 40, no. 2, pp. 375 - 386, May 1991
- [6] S. Shamai and S. Verdú, “The Effect of Frequency-flat Fading on the Spectral Efficiency of CDMA,” *IEEE Trans. Information Theory*, vol. 47, no. 4, pp. 1302-1327, May 2001.
- [7] A. M. Tulino and S. Verdú, “Random Matrix Theory and Wireless Communications,” *Foundations and Trends In Communications and Information Theory*, vol. 1, no. 1, pp. 1-184, 2004.
- [8] A. M. Tulino and A. Lozano and S. Verdú, “Capacity-Achieving Input Covariance for Single-user Multi-Antenna Channels,” *IEEE Trans. on Wireless Communications*, vol. 5, no. 3, pp. 662-671, Mar. 2006.
- [9] A. M. Tulino, S. Verdú, G. Caire and S. Shamai, “Capacity of the Gaussian Erasure Channel,” *2007 IEEE International Symposium on Information Theory*, Nice, France, June 2007.
- [10] S. Verdú, “Spectral Efficiency in the Wideband Regime” *IEEE Trans. Information Theory*, vol. 48, no. 6, pp. 1319-1343, June 2002.