Intersymbol Interference with Flat Fading: Channel Capacity

Antonia Tulino
Universitá di Napoli, Federico II
Napoli, ITALY 80125
atulino@princeton.edu

Sergio Verdú
Princeton University
Princeton, NJ 08544, USA
verdu@princeton.edu

Giuseppe Caire
University of Southern California
Los Angeles, CA 90089, USA
caire@usc.edu

Shlomo Shamai
Technion
Haifa, ISRAEL 32000
sshlomo@ee.technion.ac.il

***Abstract*—This paper finds the capacity of a linear time-invariant system with a given transfer function, observed in additive Gaussian noise through a memoryless fading channel. A coherent model is assumed where the fading coefficients are known at the receiver (but not the transmitter). We show that the optimum normalized power spectral density is the waterfilling solution for reduced signal-to-noise ratio, where the gap to the actual signal-to-noise ratio depends on both the fading distribution and the channel transfer function.**

I. Channel Model

In this paper we obtain the capacity of a channel with memory where the complex-valued input codeword $\mathbf{x} = (x_1, \ldots x_n) \in \mathbb{C}^n$ is subject to an average power constraint and goes through a deterministic linear time-invariant discrete-time linear system with transfer function $H(f)$. The outputs of the linear system $u_i \in \mathbb{C}$ are multiplied by a memoryless stationary fading process $\mathsf{A}_i \in \mathbb{C}$, known at the decoder but not the encoder. The decoder observes the resulting process contaminated by white Gaussian noise $n_i \in \mathbb{C}$.

$$y_i = \sqrt{\gamma}\,\mathsf{A}_i u_i + n_i, \quad i = 1, \ldots, n \tag{1}$$

$$u_i = \sum_{\ell=0}^{i-1} h[\ell] x_{i-\ell} \tag{2}$$

$$h[i] = \int_{-1/2}^{1/2} H(f) e^{\mathrm{j}2\pi f i}\,\mathrm{d}f \tag{3}$$

where n_i are independent proper complex Gaussian with unit variance; A_i is an i.i.d sequence with uniformly distributed phase in $[0, 2\pi)$ and whose magnitude has unit second moment and finite higher order moments. The codewords are restricted to satisfy

$$\frac{1}{n}\sum_{i=1}^{n} |x_i|^2 \leq 1 \tag{4}$$

This channel model incorporates simultaneously two key features of digital communication systems, namely, intersymbol interference and flat fading. It arises in systems (e.g. [6]) where the memory in the channel is due to a deterministic effect, while the received amplitude is random. For example, in magnetic recording the intersymbol interference coefficients are known beforehand, while the instantaneous amplitude is subject to random fluctuations due to the variations in the distance between the recoded medium and the read/write head (in that case, the real-valued channel model counterpart can be used). Powerline communication also incorporates deterministic intersymbol interference in addition to noise strength subject to rapid fluctuations. Another application is found in networks with backbone-interconnected base stations [10] subject to different levels of interference or different backhaul capacities.

Related, but less general, channels whose capacity has been analyzed before include the randomly spread multicarrier CDMA channel with fading [8] and the Gaussian-erasure channel [7].

II. Summary of Results

Before stating the general result, it is instructive to consider those special cases whose capacity is known:

- *No intersymbol interference* [1]
$$C(\gamma) = \mathbb{E}[\log\left(1 + \gamma|\mathsf{A}|^2\right)] \tag{5}$$
- *No fading* [5]
$$C(\gamma) = \int_{-1/2}^{1/2} [\log\left(\zeta\gamma|H(f)|^2\right)]^+\,\mathrm{d}f \tag{6}$$
where $1 < \zeta < \infty$ is chosen so that the spectral density
$$\bar{S}_x(f) = \left[\zeta - \frac{1}{\gamma|H(f)|^2}\right]^+ \tag{7}$$
has unit area.
- *Gaussian-Erasure channel* [7], corresponding to the case where the distribution of $|\mathsf{A}|$ has two masses, one of which is at 0.

As in the above cases, the capacity of the channel in Section I is achieved by stationary Gaussian inputs [12]. It is advisable to obtain the mutual information rate achieved by an arbitrary stationary Gaussian input, and then find the capacity-achieving input spectral density. In fact, the capacity-achieving input spectral density is one of the main results of the paper as it reveals a remarkable robustness of the classical waterfilling solution for deterministic channels with intersymbol interference.

Denote the input power spectral density by $S_x(f)$ and the output power spectral density by

$$S(f) = S_x(f)|H(f)|^2 \tag{8}$$

978-1-4244-2571-6/08/$25.00 ©2008 IEEE

For any scaled power spectral density $\gamma S(f)$, define the positive function (note it does not depend on the fading distribution) $\daleth_0(y,\gamma)$, $0 \le y \le 1$, that solves the following fixed-point equation:

$$\frac{1}{1+\daleth_0(y,\gamma)} = \int_{-1/2}^{1/2} \frac{1}{1+\,y\,\gamma S(f) + (1-\,y)\daleth_0(y,\gamma)}\,\mathrm{d}f \quad (9)$$

The input-output mutual information is

$$I(\gamma) \;=\; \int_0^1 \log\left(1+\daleth_0(y,\bar\gamma)\right) dy \quad (10)$$

where $\bar\gamma$ (which depends on both γ and y) is the solution to

$$\frac{1}{1+\daleth_0(y,\bar\gamma)} = \mathbb{E}\left[\frac{\bar\gamma}{\bar\gamma + \daleth_0(y,\bar\gamma)(\bar\gamma(1-y)+\gamma y\,|\mathsf{A}|^2)}\right] \quad (11)$$

It can be shown that we can recover the abovementioned cases of intersymbol interference, fading, and erasures by particularizing (9), (10), (11) to $\mathsf{A} = 1$, $S(f) = 1$, and $P[\mathsf{A}=0] = 1 - P[\mathsf{A} = (1-\mathsf{e})^{-1/2}]$, respectively.

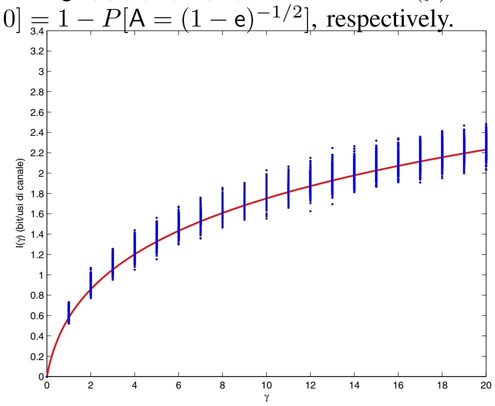

Fig. 1. Mutual information rate for $h[i] = e^{-0.2i^2}$ and Rayleigh fading, with white Gaussian input. Also shown are realizations of the normalized mutual information conditioned on the fading coefficients with $n = 200$.

Trying to optimize (9 -11) with respect to the input power spectral density is a challenging problem. Fortunately, the desired optimum input power spectral density admits a very compact characterization.

The effect of flat fading on the capacity-achieving input power spectral density is tantamount to a power penalty: For all $\gamma > 0$,

$$S_x^*(f,\nu\,\gamma) = \bar S_x(f,\gamma) \quad (12)$$

where $\bar S_x$ is the waterfilling solution in (7) and $\nu \ge 1$ satisfies:

$$\frac{1}{\zeta_\gamma} = \mathbb{E}\left[\frac{|\mathsf{A}|^2\,(\zeta_\gamma - 1)}{\nu + |\mathsf{A}|^2\,(\zeta_\gamma-1)}\right] \quad (13)$$

and ζ_γ is the fading-free water level for γ.

Denoting the S-transform of the distribution of $|\mathsf{A}|^2$ by [8]

$$\mathcal{G}(x) = -\frac{x+1}{x}\eta^{-1}_{|\mathsf{A}|^2}(1+x) \quad (14)$$

where the η-transform of the distribution of $|\mathsf{A}|^2$ is [8]

$$\eta_{|\mathsf{A}|^2}(t) = \mathbb{E}\left[\frac{1}{1+t|\mathsf{A}|^2}\right] \quad (15)$$

It can be shown that (13) it is equivalent to:

$$\nu = \mathcal{G}\left(\frac{-1}{\zeta_\gamma}\right) \quad (16)$$

The minimum energy per bit and wideband slope $\mathcal{S}_0$ [9] of the spectral efficiency of the Gaussian ISI channel with flat fading are given by

$$\left(\frac{E_b}{N_0}\right)_{\min} \;=\; \frac{\ln 2}{G_{\max}} \quad (17)$$

$$\mathcal{S}_0 \;=\; \frac{2}{\mathbb{E}[|\mathsf{A}|^4] + \frac{1-B_{\max}}{B_{\max}}}$$

where $G_{\max}$ is the maximum channel gain i.e.

$$G_{\max} = \max_f |H(f)|^2. \quad (18)$$

and $B_{\max} = \mu(\{f : |H(f)|^2 = G_{\max}\})$.

For large SNR, capacity behaves like

$$C(\gamma) = \mathcal{S}_\infty\left(\log_2\gamma - \mathcal{L}_\infty\right) + o(1) \quad (19)$$

(expressed in bits per complex dimension), where $\mathcal{S}_\infty$ and $\mathcal{L}_\infty$ are known as the high-SNR slope and the high-SNR dB offset respectively [4]. We have shown the following expressions:

$$\mathcal{S}_\infty \;=\; \min\{P\{|\mathsf{A}|^2 \ne 0\}, B\} \quad (20)$$

$$\mathcal{L}_\infty = -\int_0^1 \log_2 \digamma(y)\;\mathrm{d}y \quad (21)$$

where $\mathcal{I} = \{f : |H(f)|^2 > 0\}$, $B = \mu(\mathcal{I})$ denotes the generalized bandwidth of the linear system, and $\digamma(y)$ is the solution of the fixed point equation

$$B - y\mathcal{S}_\infty = \int_{\mathcal{I}}\left(1+\frac{y|H(f)|^2}{(1-y)\mathcal{G}(-\mathcal{S}_\infty y)\digamma(y)B}\right)^{-1}\mathrm{d}f$$

III. Main Result

Theorem 1 *The input-output mutual information rate achieved by Gaussian u_i with power spectral density $S(f)$ is*

$$I(\gamma) \;=\; \int_0^1 \log\left(1+\daleth_0(y,\gamma\,\alpha(y,\gamma))\right)\mathrm{d}y \quad (22)$$

where $\daleth_0$ is defined in (9) and $\alpha(y,\gamma)$ is the solution to

$$\alpha\,\mathcal{G}\left(\frac{-y\,\daleth_0(y,\gamma\,\alpha)}{1+\daleth_0(y,\gamma\,\alpha)}\right) = 1 \quad (23)$$

Proof: We will make use of the following auxiliary result.[1]

[1] We refer the reader to [8] for standard terminology used in random matrix theory.

Lemma 1 *[7, Theorem 1] Let $\mathbf{B}$ be an $n \times n$ nonnegative definite random matrix. Let $\rho = \lim_{n\to\infty} rank(\mathbf{B})/n$. The Shannon transform and η transforms are related through*

$$\mathcal{V}_{\mathbf{B}}(\gamma) \quad = \quad \rho \int_0^1 \log\left(1 + \beth(y,\gamma)\right)\, dy \tag{24}$$

where $\beth$ is defined by the fixed-point equation

$$\rho y \frac{\beth(y,\gamma)}{1+\beth(y,\gamma)} = 1 - \eta_{\mathbf{B}}\left(\frac{\gamma y}{1+(1-y)\beth(y,\gamma)}\right) \tag{25}$$

In addition, the proof of Theorem 1 makes use of the following new key result whose proof is sketched in the Appendix.

Theorem 2 *Denote the diagonal matrix of fading coefficients by*

$$\mathbf{A} = diag\{\mathsf{A}_1, \ldots, \mathsf{A}_n\}. \tag{26}$$

Let $\mathbf{\Psi}$ be circulant non-negative definite with an asymptotic spectral distribution. The η-transform of $\mathbf{A\Psi A}^\dagger$ is given by

$$\eta_{\mathbf{A\Psi A}^\dagger}(\gamma) = \eta \tag{27}$$

where (η, α) is the solution of the following coupled fixed-point equations:

$$\eta \quad = \quad \eta_{\mathbf{\Psi}}\left(\gamma\alpha\right) \tag{28}$$

$$\eta \quad = \quad \eta_{\mathbf{AA}^\dagger}\left(\frac{1-\eta}{\alpha\eta}\right) \tag{29}$$

At this point, we can proceed with the proof of Theorem 1. Consider the product of $n \times n$ Toeplitz matrices $\mathbf{\Sigma} = \mathbf{H\Sigma}_x\mathbf{H}^\dagger$ where $\mathbf{\Sigma}_x$ is the covariance matrix of the stationary input process, and the (i,j)-th element of $\mathbf{H}$ is equal to $h[i-j]$ defined in (3). Using Theorem 2 and [7, Lemma 1], we can readily show that η-transform of $\mathbf{A\Sigma A}^\dagger$ is the same as the η-transform of $\mathbf{A\Psi A}^\dagger$, where the circulant marix $\mathbf{\Psi}$ has the asymptotic eigenvalue distribution of $\mathbf{\Sigma}$.

From the definition of the S-transform in terms of the η-transform (14), we notice that (29) is equivalent to:

$$\alpha\, \mathcal{G}\left(\eta - 1\right) = 1. \tag{30}$$

Using (28) and (30) we obtain

$$\eta = \eta_{\mathbf{\Sigma}}\left(\frac{1}{\mathcal{G}\left(\eta-1\right)}\gamma\right). \tag{31}$$

Furthermore using (27), (28) and the η-transform of Topelitz matrices, we can write

$$\eta_{\mathbf{A\Sigma A}^\dagger}(\gamma) = \int_{-1/2}^{1/2}\left(1 + \frac{\gamma S(f)}{\mathcal{G}\left(\eta_{|\mathsf{A}|^2}(\gamma)-1\right)}\right)^{-1} \mathrm{d}f \tag{32}$$

Because of space limitations, we give the proof sketch (full details are in [12]) in the case where the fading distribution puts no mass at 0, and $S(f) > 0$ for $-1/2 < f < 1/2$ and, thus, the normalized rank of $\mathbf{A\Sigma A}^\dagger$ equals 1.

Applying Lemma 1 to $\mathbf{B} = \mathbf{A\Sigma A}^\dagger$ we obtain the Shannon transform as:

$$\mathcal{V}_{\mathbf{A\Sigma A}^\dagger}(\gamma) = \int_0^1 \log\left(1+\beth(y,\gamma)\right) dy \tag{33}$$

with $\beth$ satisfying:

$$y\frac{\beth(y,\gamma)}{1+\beth(y,\gamma)} = 1 - \eta_{\mathbf{A\Sigma A}^\dagger}\left(\frac{\gamma y}{1+(1-y)\beth(y,\gamma)}\right) \tag{34}$$

and with $\eta_{\mathbf{A\Sigma A}^\dagger}$ satisfying the equation:

$$\eta_{\mathbf{A\Sigma A}^\dagger}(t) = \eta_{\mathbf{\Sigma}}\left(\frac{t}{\mathcal{G}(\eta_{\mathbf{A\Sigma A}^\dagger}(t)-1)}\right) \tag{35}$$

Choosing

$$t = \frac{\gamma y}{1+(1-y)\beth(y,\gamma)} \tag{36}$$

we get

$$\alpha(y,\gamma)\Sigma_{\mathbf{AA}^\dagger}\left(\eta_{\mathbf{A\Sigma A}^\dagger}\left(\frac{\gamma y}{1+(1-y)\beth(y,\gamma)}\right)-1\right) = 1 \tag{37}$$

and

$$y\frac{\beth(y,\gamma)}{1+\beth(y,\gamma)} \quad = \quad 1 - \eta_{\mathbf{\Sigma}}\left(\frac{\alpha(y,\gamma)\, y\gamma}{1+(1-\,y)\beth(y,\gamma)}\right). \tag{38}$$

Writing explicitly the η-transform of $\mathbf{\Sigma}$, we have

$$\frac{\beth(y,\gamma)}{1+\beth(y,\gamma)} = \int_{-1/2}^{1/2}\frac{\alpha(y,\gamma)\gamma S(f)df}{1+(1-y)\beth(y,\gamma)+\alpha(y,\gamma)\gamma y S(f)} \tag{39}$$

Comparing (39) with the definition of $\beth_0$ in (9) we conclude that $\beth(y,\gamma) = \beth_0(y, \gamma\alpha(\gamma,y))$

Furthermore, from (34), (36) and (37), we obtain that $\alpha(\gamma, y)$ satisfies (23). ■

IV. Optimality of Waterfilling with Power Penalty

Theorem 3 *The capacity-achieving input power spectral density is given by*

$$S_x^*(f,\gamma) = \frac{1}{\theta(\zeta)}\left[\zeta - \frac{1}{\gamma|H(f)|^2}\right]^+ \tag{40}$$

where

$$\theta(\zeta) = \zeta\,\mathbb{E}\left[\frac{|\mathsf{A}|^2(\zeta-\theta(\zeta))}{1+|\mathsf{A}|^2(\zeta-\theta(\zeta))}\right] \tag{41}$$

where $0 \le \theta(\zeta) \le E[|\mathsf{A}|^2]$, and ζ is chosen such that the integral of (40) equals 1.

Proof: Using [7, Theorem 12], we can write the objective function as

$$C(\gamma) \;=\; \lim_{n\to\infty}\frac{1}{n}\max_{\Lambda_x}\mathbb{E}\left[\log\det\left(\mathbf{I}+\gamma\mathbf{AF\Lambda}_H\mathbf{\Lambda}_x\mathbf{\Lambda}_H\mathbf{F}^\dagger\mathbf{A}^\dagger\right)\right] \tag{42}$$

where the maximization is over the set of nonnegative diagonal matrices with trace equal to n; $\mathbf{\Lambda}_H$ is the diagonal matrix of the singular values of $\mathbf{H}$; and

$$\mathbf{F} = \frac{1}{\sqrt{n}}\left[e^{-\mathrm{j}\frac{2\pi}{n}(i-1)(p-1)}\,\middle|\, \begin{array}{l} i=1,\ldots,n \\ p=1,\ldots,n\end{array}\right] \tag{43}$$

To solve (42) we make use of the key non-asymptotic optimization result:

Theorem 4 *[7, Theorem 4]: Let $\mathbf{\Phi}$ be an $m \times n$ complex valued random matrix whose ith column is denoted by ϕ_i. Consider the optimization problem*

$$\max_{\mathbf{D}} \mathbb{E}\left[\log\det\left(\mathbf{I} + \gamma \mathbf{\Phi}\mathbf{D}\mathbf{\Phi}^\dagger\right)\right] \tag{44}$$

where the maximum is over all diagonal matrices whose trace is equal to a constant ξ. Then, for $i = 1, \ldots, n$, d_i^*, the ith diagonal element of the diagonal matrix $\mathbf{D}^*$ that achieves the maximum in (44) is the positive solution to

$$\mathbb{E}\left[\frac{Z_i}{1+\gamma d_i^* Z_i}\right] = \frac{1}{\nu\gamma} \tag{45}$$

$$Z_i = \phi_i^\dagger \left(\mathbf{I} + \gamma \sum_{j\neq i} d_j^* \phi_j \phi_j^\dagger\right)^{-1} \phi_i \tag{46}$$

if it exists (i.e. if $\nu\gamma\mathbb{E}[Z_i] > 1$); otherwise, $d_i^* = 0$. The parameter ν is chosen so that $\sum_{i=1}^n d_i^* = \xi$.

Letting

$$\mathbf{Q} = \mathbf{A}\mathbf{F} \tag{47}$$

$$\mathbf{\Phi} = \mathbf{A}\mathbf{F}\mathbf{\Lambda}_H \tag{48}$$

$$\mathbf{D} = \mathbf{\Lambda}_x \tag{49}$$

and denoting the columns of $\mathbf{Q}$ by $\mathbf{q}_i$, (46) takes the form

$$Z_i = H_i^2 \mathbf{q}_i^\dagger \left(\mathbf{I} + \gamma \sum_{j\neq i} H_j^2 d_j^* \mathbf{q}_j \mathbf{q}_j^\dagger\right)^{-1} \mathbf{q}_i \tag{50}$$

Using [7, Lemma 1] and Lemma 2 in the Appendix we can show almost sure convergence of (50) to

$$\lim_{n\to\infty} \mathbf{q}_i^\dagger \left(\mathbf{I} + \gamma \sum_{j\neq i} H_j^2 d_j^* \mathbf{q}_j \mathbf{q}_j^\dagger\right)^{-1} \mathbf{q}_i = \alpha \tag{51}$$

Thus,

$$\mathbb{E}\left[\frac{Z_i}{1+\gamma d_i^* Z_i}\right] \to \frac{H_i^2\alpha}{1+\gamma d_i^* H_i^2 \alpha} \tag{52}$$

Plugging (52) in Theorem 4, the sought-after power spectral density satisfies

$$|H(f)|^2\alpha\nu\gamma = 1 + \gamma\alpha S_x^*(f)|H(f)|^2 \tag{53}$$

if $\alpha\nu\gamma|H(f)|^2 > 1$, and $S_x^*(f) = 0$ otherwise. Using (51), we get

$$S_x^*(f,\gamma) = \left[\nu - \frac{1}{\gamma|H(f)|^2\left(\frac{1}{\mathcal{G}(\eta_{\mathbf{A}\mathbf{\Sigma}\mathbf{A}^\dagger}(\gamma)-1)}\right)}\right]^+ \tag{54}$$

Choosing the water level so that the integral of (54) is equal to 1, leads to

$$\nu = \frac{1}{1-\eta_{\mathbf{\Sigma}}(\alpha\gamma)} \tag{55}$$

$$= \frac{1}{1-\eta_{\mathbf{A}\mathbf{\Sigma}\mathbf{A}^\dagger}(\gamma)} \tag{56}$$

where we have used Theorem 2.

To obtain the final result (40), we change variables and let $\zeta = \nu\alpha$, thereby expressing (54) as

$$S_x^*(f,\gamma) = \frac{1}{\alpha}\left[\zeta - \frac{1}{\gamma|H(f)|^2}\right]^+ \tag{57}$$

where in view of (51) and (56), α satisfies the equation

$$\alpha = -\mathcal{G}^{-1}\left(\frac{1}{\alpha}\right)\zeta \tag{58}$$

whose solution in the interval $[0, P(|\mathsf{A}| \neq 0)]$ is denoted by $\theta(\zeta)$ and is given in (41). ■

Finally, we show in this section that the form of the result given in (12) and (13) follows from Theorem 3.

Let

$$\rho\bar{\zeta} = \zeta_\gamma \tag{59}$$

$$= \frac{1}{-\mathcal{G}^{-1}(\rho)} \tag{60}$$

where (60) follows by solving for ζ_γ in (13). Comparing (60) to (58) we conclude that $1 = \rho\,\theta(\bar{\zeta})$, and consequently:

$$\frac{\bar{\zeta}}{\theta(\bar{\zeta})} = \zeta_\gamma \tag{61}$$

$$\gamma\theta(\bar{\zeta}) = \frac{\gamma}{\rho} \tag{62}$$

Thus, (12) follows in view of (40) and its particularization to the conventional case without fading (7).

V. Appendix: Proof of Theorem 2

To prove Theorem 2 we need the following lemma which is the most technically challenging result in this work. The proof is omitted because of space limitations.

Lemma 2 *Suppose that the sequence of non-negative scalars $\mathbf{\Lambda} = diag(\lambda_1, \ldots, \lambda_n)$ has a limiting empirical distribution. Then, the columns of $\mathbf{Q}$ defined in (47) satisfy*

$$\mathbf{q}_i^\dagger \left(\mathbf{I} + \gamma \sum_{j\neq i} \lambda_j \mathbf{q}_j \mathbf{q}_j^\dagger\right)^{-1} \mathbf{q}_i \overset{a.s.}{\to} \alpha \tag{63}$$

for all i, where α depends on the fading distribution and on the asymptotic distribution of $\mathbf{\Lambda}$ but it does not depend on i.

Using the symmetric nature of $\mathbf{F}$, it is possible to interchange the roles of the matrices $\mathbf{A}$ and $\mathbf{\Lambda}^{1/2}$ in Lemma 2 to show that the columns of $\bar{\mathbf{Q}} = \mathbf{\Lambda}^{1/2}\mathbf{F}$ satisfy

$$\bar{\mathbf{q}}_i^\dagger \left(\mathbf{I} + \gamma \sum_{j\neq i} |\mathsf{A}_j|^2 \bar{\mathbf{q}}_j \bar{\mathbf{q}}_j^\dagger\right)^{-1} \bar{\mathbf{q}}_i \overset{a.s.}{\to} \nu \tag{64}$$

where ν depends on the fading distribution and on the asymptotic distribution of $\mathbf{\Lambda}$ but it does not depend on i. Using Lemma 2 we obtain

$$\frac{1}{n}\mathrm{tr}\left\{\left(\mathbf{I}+\gamma\mathbf{A}\boldsymbol{\Psi}\mathbf{A}^\dagger\right)^{-1}\right\} = \frac{1}{n}\mathrm{tr}\{(\mathbf{I}+\gamma\mathbf{Q}\boldsymbol{\Lambda}^\dagger\mathbf{Q}^\dagger)^{-1}\} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+\gamma\lambda_i\mathbf{q}_i^\dagger\mathbf{B}_i^{-1}\mathbf{q}_i}$$

where

$$\mathbf{B}_i = \mathbf{I}+\gamma\sum_{j\neq i}\lambda_j\mathbf{q}_j\mathbf{q}_j^\dagger \quad (65)$$

$$= \mathbf{I}+\gamma\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\dagger-\gamma\lambda_i\mathbf{q}_i\mathbf{q}_i^\dagger \quad (66)$$

Therefore,

$$\eta_{\mathbf{A}\boldsymbol{\Psi}\mathbf{A}^\dagger}(\gamma) = \lim_{n\to\infty}\frac{1}{n}\mathrm{tr}(\mathbf{I}+\gamma\mathbf{A}\boldsymbol{\Psi}\mathbf{A}^\dagger)^{-1} \quad (67)$$

$$= \lim_{n\to\infty}\frac{1}{n}\mathrm{tr}(\mathbf{I}+\gamma\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\dagger)^{-1} \quad (68)$$

$$= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+\gamma\lambda_i\mathbf{q}_i^\dagger\mathbf{B}_i^{-1}\mathbf{q}_i} \quad (69)$$

$$= \eta_{\boldsymbol{\Psi}}\left(\alpha\gamma\right) \quad (70)$$

where (70) follows from Lemma 2.

Following analogous steps and (64), it can be shown that

$$\eta_{\boldsymbol{\Psi}^{1/2\dagger}\mathbf{A}^\dagger\mathbf{A}\boldsymbol{\Psi}^{1/2}}(\gamma) = \lim_{n\to\infty}\frac{1}{n}\mathrm{tr}(\mathbf{I}+\gamma\boldsymbol{\Psi}^{1/2\dagger}\mathbf{A}^\dagger\mathbf{A}\boldsymbol{\Psi}^{1/2})^{-1}$$

$$= \lim_{n\to\infty}\frac{1}{n}\mathrm{tr}\{(\mathbf{I}+\gamma\bar{\mathbf{Q}}\mathbf{A}\mathbf{A}^\dagger\bar{\mathbf{Q}}^\dagger)^{-1}\}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+\gamma|\mathsf{A}_i|^2\bar{\mathbf{q}}_i^\dagger\mathbf{C}_i^{-1}\bar{\mathbf{q}}_i} \quad (71)$$

where

$$\mathbf{C}_i = \mathbf{I}+\gamma\sum_{j\neq i}|\mathsf{A}_j|^2\bar{\mathbf{q}}_j\bar{\mathbf{q}}_j^\dagger \quad (72)$$

$$= \mathbf{I}+\gamma\bar{\mathbf{Q}}\mathbf{A}\mathbf{A}^\dagger\bar{\mathbf{Q}}^\dagger-\gamma|\mathsf{A}_i|^2\bar{\mathbf{q}}_i\bar{\mathbf{q}}_i^\dagger \quad (73)$$

From (71) and (64) it follows that:

$$\eta_{\boldsymbol{\Psi}^{1/2\dagger}\mathbf{A}^\dagger\mathbf{A}\boldsymbol{\Psi}^{1/2}}(\gamma) = \eta_{\mathbf{A}\boldsymbol{\Psi}\mathbf{A}^\dagger}(\gamma) \quad (74)$$

$$= \eta_{\mathbf{A}\mathbf{A}^\dagger}\left(\nu\gamma\right) \quad (75)$$

where (75) follows from (64). Now notice that

$$\alpha\,\eta_{\mathbf{A}\boldsymbol{\Psi}\mathbf{A}^\dagger}(\gamma) = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbf{q}_i^\dagger\mathbf{B}_i^{-1}\mathbf{q}_i}{1+\lambda_i\,\gamma\mathbf{q}_i^\dagger\mathbf{B}_i^{-1}\mathbf{q}_i^\dagger} \quad (76)$$

$$= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathbf{q}_i^\dagger(\gamma\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\dagger+\mathbf{I})^{-1}\mathbf{q}_i$$

$$= \lim_{n\to\infty}\frac{1}{n}\mathrm{tr}\left(\mathbf{Q}^\dagger(\gamma\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\dagger+\mathbf{I})^{-1}\mathbf{Q}\right)$$

$$= \lim_{n\to\infty}\frac{1}{n}\mathrm{tr}\left(\mathbf{A}(\mathbf{I}+\gamma\mathbf{A}\mathbf{F}\boldsymbol{\Lambda}\mathbf{F}^\dagger\mathbf{A}^\dagger)^{-1}\mathbf{A}^\dagger\right)$$

$$= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\frac{|\mathsf{A}_i|^2}{1+\gamma|\mathsf{A}_i|^2\bar{\mathbf{q}}_i^\dagger\mathbf{C}_i^{-1}\bar{\mathbf{q}}_i}$$

$$= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\bar{\mathbf{q}}_i^\dagger\mathbf{C}_i^{-1}\bar{\mathbf{q}}_i}\,\frac{|\mathsf{A}_i|^2\bar{\mathbf{q}}_i^\dagger\mathbf{C}_i^{-1}\bar{\mathbf{q}}_i}{1+\gamma|\mathsf{A}_i|^2\bar{\mathbf{q}}_i^\dagger\mathbf{C}_i^{-1}\bar{\mathbf{q}}_i}$$

$$= \frac{1}{\nu\gamma}\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\frac{\gamma|\mathsf{A}_i|^2\bar{\mathbf{q}}_i^\dagger\mathbf{C}_i^{-1}\bar{\mathbf{q}}_i}{1+\gamma|\mathsf{A}_i|^2\bar{\mathbf{q}}_i^\dagger\mathbf{C}_i^{-1}\bar{\mathbf{q}}_i} \quad (77)$$

where we have used (64) and

$$\frac{1}{1+\gamma|\mathsf{A}_i|^2\bar{\mathbf{q}}_i^\dagger\mathbf{C}_i^{-1}\bar{\mathbf{q}}_i} = \left(\left[\mathbf{I}+\gamma\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^\dagger\right]^{-1}\right)_{ii} \quad (78)$$

which follows from the Sherman-Morrison formula [11]

$$\left[\mathbf{M}+\mathbf{x}\mathbf{y}^\dagger\right]^{-1} = \mathbf{M}^{-1}-\frac{(\mathbf{M}^{-1}\mathbf{x})(\mathbf{y}^\dagger\mathbf{M}^{-1})}{1+\mathbf{y}^\dagger\mathbf{M}^{-1}\mathbf{x}} \quad (79)$$

According to (77) we obtain

$$\gamma\,\nu\,\alpha\,\eta_{\mathbf{A}\boldsymbol{\Psi}\mathbf{A}^\dagger}(\gamma) = 1-\eta_{\mathbf{A}\boldsymbol{\Psi}\mathbf{A}^\dagger}(\gamma), \quad (80)$$

and solving for $\nu\gamma$ we find

$$\nu\gamma = \frac{1-\eta_{\mathbf{A}\boldsymbol{\Psi}\mathbf{A}^\dagger}(\gamma)}{\alpha\eta_{\mathbf{A}\boldsymbol{\Psi}\mathbf{A}^\dagger}(\gamma)}$$

By using this into (75), together with (70) yields the system of fixed point equations of Theorem 2, and the proof is complete.

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, Oct 1998.

[2] T. Cover and J. Thomas, *Elements of Information Theory*. 2nd Ed., New York: Wiley, 2006

[3] 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

[4] 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.

[5] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–21, Jan. 1949.

[6] W. S. Smith, P. H. Wittke, and L. L. Campbell, "Error probabilities on fading channels with intersymbol interference and noise," *IEEE Trans. Information Theory*, vol. 39, pp. 1598-1607, Sept. 1993.

[7] A. Tulino, S. Verdú, G. Caire, and S. Shamai, "Capacity of the Gaussian-Erasure Channel," submitted for publication, 2007.

[8] A. 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.

[9] S. Verdú, "Spectral Efficiency in the Wideband Regime" *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319-1343, June 2002.

[10] A. D. Wyner, "Shannon-Theoretic Approach to a Gaussian Cellular Multiple Access Channel", *IEEE Trans. Inform. Theory,* Vol. 40, No. 6, pp. 1713-1727, November 1994.

[11] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*. New York: Academic Press, 1985.

[12] A. Tulino, S. Verdú, G. Caire, and S. Shamai, "The combined effect of flat fading and intersymbol interference on channel capacity," manuscript, 2008.

ACKNOWLEDGEMENT

This paper was funded by the National Science Foundation under Grant CCF-0728445, and by the U.S.-Israel Binational Science Foundation under Grant 2004140.