The Gaussian Erasure Channel

4 authors, including:

Sergio Verdú
Princeton University
118 PUBLICATIONS   **8,955** CITATIONS

SEE PROFILE

Giuseppe Caire
Technische Universität Berlin
575 PUBLICATIONS   **24,651** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Rate Splitting for MIMO Wireless Networks: A Promising PHY-Layer Strategy for 5G and beyond View project

Project   Software defined Radio Platform View project

The Gaussian Erasure Channel

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 linear time-invariant systems observed in additive Gaussian noise through a memoryless erasure channel. This problem requires obtaining the asymptotic spectral distribution of a submatrix of a nonnegative definite Toeplitz matrix obtained by retaining each column/row independently and with identical probability. 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 erasure probability and the channel transfer function. We find asymptotic expressions for the capacity in the sporadic erasure and sporadic non-erasure regimes as well as the low and high signal-to-noise regimes.

I. Introduction

The erasure channel plays an important role in information theory and coding theory. It is a very useful idealization of situations where the symbols observed by the receiver have either very high or very low reliability. Applications of erasure channels range from communication subject to jamming to packet-switched store-and-forward networks, from magnetic recording to wireless communications subject to fading, from powerline communications subject to impulsive noise to frequency-hopped multiaccess channels. For many applications, discrete erasure channels where symbols are either received without error or erased are rather coarse idealizations. Within the paradigm of discrete memoryless noisy channels, it is straightforward to find the capacity of channels that incorporate errors as well as erasures. Even in the presence of memory in the erasures it has been shown recently [1] that the capacity of the concatenation of a discrete memoryless channel with capacity C and an erasure channel (possibly with memory) with erasure rate e is equal to $(1-e)C$. Also straightforward is to deal with power-constrained memoryless Gaussian channels observed through memoryless erasure channels: the capacity is also equal to the capacity of the memoryless Gaussian channel times the proportion of non-erased symbols [2]. A much more difficult case is the concatenation of a *channel with memory* the output of which is observed through a *memoryless* erasure channel. This paper deals with a particularly important instance of this class, where the channel with memory is a standard input-power constrained linear Gaussian channel with white noise and given transfer function. This model is relevant for example in the case of a frequency selective Gaussian channel with impulsive noise where the power of the impulses is much larger than the average received power. The Gaussian erasure model can also be used to assess the throughput in uplink cellular systems subject to topological randomness: the uplink cellular system is modelled by the classical Wyner model [3] and cell sites are either "on" or "off" in an independent and identically distributed manner.

This paper finds the capacity of the Gaussian erasure channel introduced above. This problem requires obtaining the asymptotic spectral distribution of a submatrix of a nonnegative definite Toeplitz matrix obtained by randomly deleting rows and columns. We also find explicitly the optimal (capacity-achieving) input power spectral density, easily computable upper and lower bounds to the mutual information and asymptotic expressions for the capacity in various regimes.

Due to space limitation, proofs are omitted and are given in full detail in [8].

II. Problem Setup

In this paper we analyze the channel with memory where the input codeword $\mathbf{x} = (x_1, \ldots x_n) \in \mathbb{R}^n$ is subject to an average power constraint and goes through a linear time-invariant discrete-time linear system with transfer function $H(f)$ the output of the linear system $(u_1, \ldots u_n) \in \mathbb{R}^n$ is contaminated by independent identically distributed Gaussian noise $(n_1, \ldots n_n) \in \mathbb{R}^n$; finally, a process of erasures $\mathbf{e} = (e_1, \ldots e_n) \in \{0, 1\}^n$, known to the receiver, controls which noisy outputs are available to the receiver. Because the noise is memoryless, this setup is is equivalent to
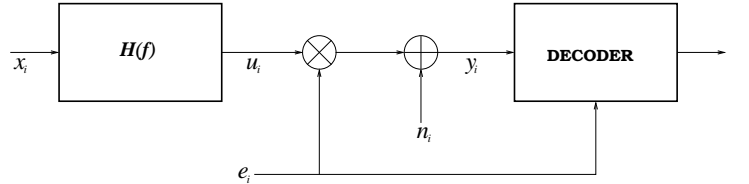

Fig. 1. Linear Gaussian Erasure Channel.

$$y_i = \sqrt{\gamma}\, e_i u_i + n_i, \quad i = 1, \ldots, n \qquad (1)$$

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

$$h[i] = \int_{-1/2}^{1/2} H(f) e^{j2\pi f i}\, \mathrm{d}f \qquad (3)$$

where n_i are independent Gaussian with unit variance; e_i is independent sequence with $P[e_i = 0] = e$, and the codewords are restricted to satisfy

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 \le 1 \qquad (4)$$

Since the receiver knows the location of the erasures, the capacity is given by

$$C(\gamma) = \lim_{n\to\infty}\frac{1}{2n}\max_{\boldsymbol{\Sigma}}\mathbb{E}\left[\log\det\left(\mathbf{I}+\gamma\mathbf{E}\boldsymbol{\Sigma}\mathbf{E}\right)\right] \qquad (5)$$

where we have denoted the random matrix

$$\mathbf{E} = \text{diag}\{e_1,\ldots e_n\} \qquad (6)$$

and (5) follows from the optimality of Gaussian inputs since conditioned on $\mathbf{E}$, the channel is Gaussian; the maximum in (5) is over all matrices that can be written as

$$\boldsymbol{\Sigma} = \mathbf{H}\boldsymbol{\Sigma}_x\mathbf{H}^T \qquad (7)$$

with $\mathbf{H}$ denoting the Toeplitz channel matrix whose (i,j) entry is $h[i-j]$ and $\text{tr}\{\boldsymbol{\Sigma}_x\} = n$. It can be shown (see [8]) that there is no loss of optimality in restricting the input to be stationary (with power spectral density $S_x(f)$), and we can restrict attention to Toeplitz matrices $\boldsymbol{\Sigma}_x$. The power spectral density of the signal at the output of the linear system is denoted by

$$S(f) = S_x(f)|H(f)|^2. \qquad (8)$$

For brevity we will denote the mutual information achieved with output power spectral density $S(f)$ by $I_e(\gamma)$. Note that in the absence of erasures ($e = 0$),

$$I_0(\gamma) = \frac{1}{2}\int_{-1/2}^{1/2}\log(1+\gamma S(f))\,df \qquad (9)$$

which when maximized with respect to the input power spectral density yields capacity (e.g. [4])

$$C(\gamma) = \frac{1}{2}\int_{-1/2}^{1/2}\log\left(1+\gamma S_x^*(f)|H(f)|^2\right)\,df \qquad (10)$$

where $S_x^*(f)$ is the waterfilling input power spectral density given by

$$S_x^*(f) = \left[\zeta - \frac{1}{\gamma|H(f)|^2}\right]^+ \qquad (11)$$

and the water level $1 < \zeta < \infty$ is chosen so that $S_x^*(f)$ has unit area.

The solution (10), known since [5] can be justified by means of the Grenander-Szëgo theorem on the distribution of the eigenvalues of large Toeplitz matrices (e.g.[6]). In view of (5), obtaining the capacity of the linear Gaussian erasure channel involves analyzing the asymptotic distribution of the eigenvalues of the random matrix $\mathbf{E}\boldsymbol{\Sigma}\mathbf{E}$ where $\boldsymbol{\Sigma}$ is a product of deterministic Toeplitz matrices, and $\mathbf{E}$ is a random 0-1 diagonal matrix. In other words, the central problem is to obtain the asymptotic spectral distribution of a submatrix of

a nonnegative definite Toeplitz matrix obtained by retaining each column/row independently and with identical probability. Such a result, at the intersection of the asymptotic eigenvalue distribution of Toeplitz matrices (e.g. [6]) and of random matrices (e.g. [7]), is the main contribution of this paper.

III. RANDOM MATRIX THEORY: η AND SHANNON TRANSFORMS

The η-transform and the Shannon transform were motivated by the application of random matrix theory to various problems in the information theory of noisy communication channels [7]. These transforms, intimately related with each other and with the Stieltjes transform traditionally used in random matrix theory [7], characterize the spectrum of a random matrix while carrying certain engineering intuition.

Definition 1 *Given a nonnegative definite random matrix $\mathbf{A}$, its η-transform is*

$$\eta_{\mathbf{A}}(\gamma) = \mathbb{E}\left[\frac{1}{1+\gamma X}\right] \qquad (12)$$

where X is a nonnegative random variable whose distribution is the asymptotic Empirical Spectral Distribution (ESD) of $\mathbf{A}$ while γ is a nonnegative real number.

Let $\boldsymbol{\Sigma}$ be the $n \times n$ nonnegative definite Toeplitz matrix:

$$\Sigma_{i,j} = \sigma_{|i-j|} \qquad (13)$$

for an absolutely summable sequence $\sigma_0, \sigma_1, \ldots$. The η-transform of $\boldsymbol{\Sigma}$ is given by:

$$\eta_{\boldsymbol{\Sigma}}(\gamma) = \int_{-1/2}^{1/2}\frac{1}{1+\gamma S(f)}\,df \qquad (14)$$

Definition 2 *Given a nonnegative definite random matrix $\mathbf{A}$, its Shannon transform is defined as*

$$\mathcal{V}_{\mathbf{A}}(\gamma) = \mathbb{E}[\log(1+\gamma X)] \qquad (15)$$

where X is a nonnegative random variable whose distribution is the asymptotic ESD of $\mathbf{A}$ while γ is a nonnegative real number.

Note that

$$I_e(\gamma) = \frac{1}{2}\mathcal{V}_{\mathbf{E}\boldsymbol{\Sigma}\mathbf{E}}(\gamma) \qquad (16)$$

The η and Shannon transforms are related through

$$\frac{d}{d\gamma}\mathcal{V}_{\mathbf{A}}(\gamma) = \frac{1-\eta_{\mathbf{A}}(\gamma)}{\gamma}\log e \qquad (17)$$

Another relation between the η-transform and the Shannon transform, which we introduce here, is the following general result which in addition to playing a key role in our analysis is of independent interest.

Theorem 1 *Let* $\mathbf{A}$ *be a nonnegative definite random matrix. Let* $\rho = \lim_{n\to\infty} rank(\mathbf{A})/n$. *The Shannon transform and* η *transforms are related through*

$$\mathcal{V}_{\mathbf{A}}(\gamma) = \rho \int_0^1 \log\left(1 + \beth(y,\gamma)\right) dy \qquad (18)$$

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

$$\rho y \frac{\beth(y,\gamma)}{1 + \beth(y,\gamma)} = 1 - \eta_{\mathbf{A}}\left(\frac{\gamma y}{1 + (1-y)\beth(y,\gamma)}\right) \qquad (19)$$

Some properties of the solution to (19) are:
1) $\beth(y,0) = 0$, $\beth(y,\gamma)$ is monotonically increasing with γ and monotonically decreasing with y.
2) $\lim_{\gamma\to\infty} \frac{\beth(y,\gamma)}{\gamma} = F(y)$, where $F(y)$ is the solution of

$$\eta_{\mathbf{A}}\left(\frac{y}{(1-y)F}\right) = 1 - \rho y \qquad (20)$$

IV. ASYMPTOTIC EIGENVALUE DISTRIBUTION OF $\mathbf{E\Sigma E}$

The first step is to show that, as in the conventional case without erasures, the asymptotic eigenvalue distribution is the same as if $\mathbf{\Sigma}$ were replaced by a circulant matrix. The sufficient condition in the following lemma is satisfied for (7) because of the conventional asymptotic equivalence of products of Toeplitz matrices to circulant matrices [6, Thm. 5.3].

Lemma 1 *Let* $\mathbf{\Sigma}$ *have eigenvalues*

$$\mathbf{\Lambda} = diag\{\lambda_1, \ldots \lambda_n\}. \qquad (21)$$

Further, denote the unitary DFT matrix

$$\mathbf{F} = \frac{1}{\sqrt{n}}\left[e^{-j\frac{2\pi}{n}(\ell-1)(m-1)} \,\middle|\, \begin{matrix} \ell = 1,\ldots,n \\ m = 1,\ldots,n \end{matrix} \right] \qquad (22)$$

and the circulant matrix

$$\mathbf{\Psi} = \mathbf{F\Lambda F}^\dagger \qquad (23)$$

Let $\mathbf{E}$ *be the erasure matrix. If* $\mathbf{\Sigma}$ *and* $\mathbf{\Psi}$ *are asymptotically equivalent, then for all* $\gamma > 0$

$$\eta_{\mathbf{E\Sigma E}}(\gamma) = \eta_{\mathbf{E\Psi E}}(\gamma) \qquad (24)$$

The next theorem yields the desired characterization of the asymptotic eigenvalue distribution of $\mathbf{E\Sigma E}$:

Theorem 2 *Let* $\mathbf{E}$ *and* $\mathbf{\Psi}$ *be matrix sequences defined as in Lemma 1. Then the* η-*transform of* $\mathbf{E\Psi E}$ *is given by*

$$\eta_{\mathbf{E\Psi E}}(\gamma) = \eta$$

where η *is the solution of the fixed-point equation:*

$$\eta = \eta_{\mathbf{\Psi}}\left(\gamma - \gamma\frac{e}{\eta}\right) \qquad (25)$$

Using Lemma 1, Theorem 2 and (14), we can write (25) as

$$\eta_{\mathbf{E\Sigma E}}(\gamma) = \int_{-1/2}^{1/2} \frac{1}{1 + \gamma S(f)(1 - e/\eta_{\mathbf{E\Sigma E}}(\gamma))} \, df \qquad (26)$$

Using (17) with $\mathbf{A} = \mathbf{E\Sigma E}$ the mutual information rate as a function of the signal-to-noise ratio γ can be characterized in terms of the η-transform of $\mathbf{E\Sigma E}$:

$$I_e(\gamma) = \frac{1}{2}\int_0^\gamma \frac{1 - \eta_{\mathbf{E\Sigma E}}(x)}{x} \, dx \qquad (27)$$

An alternative characterization of the capacity is given by our main result:

Theorem 3 *The mutual information rate achieved with output power spectral density* $S(f)$ *and erasure rate* e *is equal to*

$$I_e(\gamma) = \frac{1}{2}\int_0^{1-e} \log\left(1 + \beth_0(y,\gamma)\right) dy \qquad (28)$$

where $\beth_0$ *is the solution to*

$$\frac{\beth_0(y,\gamma)}{1 + \beth_0(y,\gamma)} = \int_{-1/2}^{1/2} \frac{\gamma S(f)}{1 + y\gamma S(f) + (1-y)\beth_0(y,\gamma)} \, df \qquad (29)$$

An important feature Theorem 3 is that the mutual information rate can be computed via a fast algorithm that solves the fixed-point equation iteratively. From (28) and the properties of $\beth_0(y,\gamma)$ we can also show that $I_e(\gamma)$ is a concave decreasing function of $e \in [0,1]$.

In special cases such as the following it is possible to find a closed-form for the capacity.

Example 1 *Suppose the Gaussian stationary random process* $\{u_i\}$ *in (1) has an ideal low-pass power spectral density*

$$S^u(f) = \begin{cases} \frac{1}{B} & |f| \le B/2 \\ 0 & B/2 < |f| \le 1/2 \end{cases} \qquad (30)$$

for some $B \in (0,1]$. *The solution of (29) is*

$$\beth_0^u(y,\gamma) = \frac{1}{2(1-y)}\left[\gamma - 1 - \gamma y/B + \sqrt{(\gamma - 1 - \gamma y/B)^2 + 4\gamma(1-y)}\right] \qquad (31)$$

Using (28) and (31), we get

$$\begin{aligned}
I_e^u(\gamma, B) &= \frac{1 - 2B}{4}\log\frac{2(1-B)B}{B - 2B^2 - B\gamma + (1-e)\gamma + \Delta} \\
&+ \frac{1-e}{2}\log\frac{B(2e + \gamma - 1) - (1-e)\gamma + \Delta}{2Be} \\
&+ \frac{1}{4}\log\frac{2(1-B)B\gamma^2 e^2}{\beta} \qquad (32)
\end{aligned}$$

where

$$\Delta = \sqrt{B^2((1+\gamma)^2 - 4(1-e)\gamma) + 2B(1-e)(1-\gamma)\gamma + (1-e)^2\gamma^2} \qquad (33)$$

$$\beta = B(\gamma-1)\gamma(2-e) - B^2(1 - 2(1-e)\gamma + \gamma^2) - (1-e)\gamma^2 + (B+\gamma - B\gamma)\Delta \qquad (34)$$

V. INPUT SPECTRUM OPTIMIZATION

The goal of this section is to find the optimum power spectral density as a function of the signal-to-noise ratio γ and the erasure rate e. Trying to optimize the expression in Theorem 3 with respect to the input power spectral density appears to be a challenging problem. Instead, we give the following general finite-dimensional result which is of independent interest and

is related to a result in [10] obtained through an MMSE representation.

Theorem 4 *Let $\boldsymbol{\Phi}$ be an $m \times n$ complex valued random matrix whose ith column is denoted by $\boldsymbol{\phi}_i$. Consider the optimization problem*

$$\max_{\mathbf{D}} \mathbb{E}\left[\log \det\left(\mathbf{I} + \gamma \boldsymbol{\Phi}\mathbf{D}\boldsymbol{\Phi}^{\dagger}\right)\right] \quad (35)$$

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 (35) is the positive solution to

$$\mathbb{E}\left[\frac{Z_i}{1 + \gamma d_i^* Z_i}\right] = \frac{1}{\nu\gamma} \quad (36)$$

$$Z_i = \boldsymbol{\phi}_i^{\dagger}\left(\mathbf{I} + \gamma \sum_{j \neq i} d_j^* \boldsymbol{\phi}_j \boldsymbol{\phi}_j^{\dagger}\right)^{-1} \boldsymbol{\phi}_i \quad (37)$$

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

Applying Theorem 4 to our problem we obtain the following compact characterization of the optimum input power spectral density.

Theorem 5 *The capacity-achieving input power spectral density, $S_x^*(f, \gamma, \mathrm{e})$ is given by*

$$S_x^*(f, \gamma, \mathrm{e}) = \frac{1}{\theta(\mathrm{e}, \zeta)}\left[\zeta - \frac{1}{\gamma|H(f)|^2}\right]^{+} \quad (38)$$

where

$$\theta(\mathrm{e}, \zeta) = \frac{1}{2}\left[\zeta + 1 - \sqrt{(\zeta - 1)^2 + 4\zeta\mathrm{e}}\right] \quad (39)$$

and ζ is chosen so that the integral of (38) is equal to 1.

As the following corollary to Theorem 5 shows, the effect of erasures on the capacity-achieving input power spectral density is tantamount to a reduction in the signal to noise ratio.

Corollary 1 *For all $0 \leq \mathrm{e} \leq 1$, $\gamma > 0$,*

$$S_x^*(f, \gamma/\kappa, \mathrm{e}) = S_x^*(f, \gamma, 0) \quad (40)$$

where

$$\kappa = 1 - \frac{\mathrm{e}\,\zeta_\gamma}{\zeta_\gamma - 1} \quad (41)$$

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

VI. Bounds

This section presents some easily computable upper and lower bounds on the mutual information of the Gaussian erasure channel, that need no fixed-point equation solution.

Theorem 6 *Denote*

$$G = \int_{-1/2}^{1/2} S(f)\,\mathrm{d}f. \quad (42)$$

and let B the "generalized bandwidth" of $S(f)$ (i.e. Lebesgue measure of its support $\mathcal{I} \in [-1/2, 1/2]$),. The mutual information rate is lower bounded by

$$I_{\mathrm{e}}(\gamma) \geq I_{\mathrm{e}}^{\mathrm{u}}(G\gamma, B) + \frac{1}{2}\int_{\mathcal{I}} \log\left(\frac{1 + \gamma S(f)}{1 + \gamma G/B}\right)\mathrm{d}f \quad (43)$$

$$I_{\mathrm{e}}(\gamma) \geq (1 - \mathrm{e})I_0(\gamma) \quad (44)$$

$$I_{\mathrm{e}}(\gamma) \geq I_0(\gamma) - \frac{\mathrm{e}}{2}\log(1 + \beth_0(1 - \mathrm{e}, \gamma)) \quad (45)$$

where an explicit expression for $I_{\mathrm{e}}^{\mathrm{u}}(\gamma, B)$ is given in (32).

Theorem 7 *With the same notation as in Theorem 6, the mutual information $I_{\mathrm{e}}(\gamma)$ is upper bounded by*

$$I_{\mathrm{e}}(\gamma) \leq I_{\mathrm{e}}^{\mathrm{u}}(G\gamma, B) \quad (46)$$

$$I_{\mathrm{e}}(\gamma) \leq I_0((1 - \mathrm{e})\gamma) \quad (47)$$

$$I_{\mathrm{e}}(\gamma) \leq \frac{1 - \mathrm{e}}{2}\log(1 + G\gamma) \quad (48)$$

$$I_{\mathrm{e}}(\gamma) \leq I_0(\gamma) + \frac{\mathrm{e}}{2}\log\left(\int_{-1/2}^{1/2} \frac{1}{1 + \gamma S(f)}\,\mathrm{d}f\right) \quad (49)$$

VII. Asymptotics

A. Sporadic Erasures

The following result shows that the upper bound (49) is asymptotically tight.

Theorem 8 *For any output power spectral density $S(f)$,*

$$I_{\mathrm{e}}(\gamma) = I_0(\gamma) - \frac{\mathrm{e}}{2}\log\frac{1}{\eta_{\boldsymbol{\Sigma}}(\gamma)} + o(\mathrm{e}) \quad (50)$$

Using the capacity-achieving power spectral density we obtain

Theorem 9 *In the regime of sporadic erasures the capacity of the Gaussian erasure channel satisfies:*

$$C_{\mathrm{e}}(\gamma) = C_0(\gamma) - \frac{\mathrm{e}}{2}\log\frac{\zeta}{\zeta - 1} + o(\mathrm{e}) \quad (51)$$

where ζ is the water level of the power spectral density that achieves $C_0(\gamma)$.

B. Sporadic non-erasures

As $\mathrm{e} \to 1$, the solution in Theorem 3 takes the limiting expression:

$$\lim_{\mathrm{e}\to 1} \frac{I_{\mathrm{e}}(\gamma)}{1 - \mathrm{e}} = \frac{1}{2}\log\left(1 + \gamma \int_{-1/2}^{1/2} S(f)\,\mathrm{d}f\right) \quad (52)$$

Optimizing (52) over unit input power spectrum $S_x(f)$ with $S(f) = |H(f)|^2 S_x(f)$ results in $S_x(f)$ that places all its power at the most favorable frequency (or frequencies), yielding

$$\lim_{\mathrm{e}\to 1} \frac{C_{\mathrm{e}}(\gamma)}{1 - \mathrm{e}} = \frac{1}{2}\log\left(1 + \gamma G_{\max}\right) \quad (53)$$

where the maximum channel gain is denoted by

$$G_{\max} = \max_f |H(f)|^2. \quad (54)$$

Example 2 *Figure 2 shows the capacity $C_e(\gamma)$ forfixed $\gamma = 10$ dB, as a function ofthe erasure probability e, for a channel with transfer function*

$$|H(f)|^2 = \begin{cases} 10 & |f| < 0.05; \\ 7 & 0.05 \le |f| \le 0.1; \\ 5 & 0.1 < |f| < 0.2; \\ 2 & 0.2 \le |f| \le 0.3; \\ 1 & 0.3 < |f| < 0.4; \\ 0 & 0.4 \le |f| \le 0.5 \end{cases} \qquad (55)$$

Figure 2 shows also the the affine approximationsof capacity for sporadic erasures (51) andsporadic non-erasures (53).
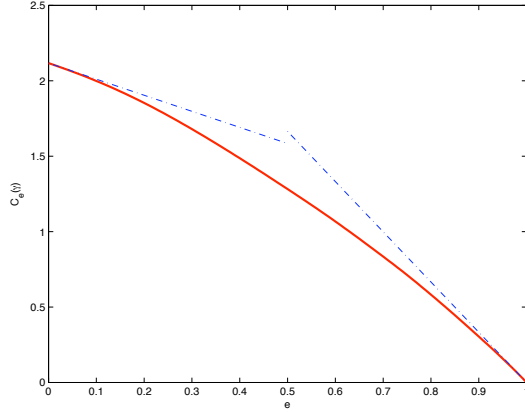

Fig. 2. Capacity versus erasure rate at $\gamma = 10dB$ for the channel given in (55), with its affine asymptotic approximations for small and large erasure probability.

C. Low-SNR

We characterize the behavior of capacity for fixed e and vanishing γ. In order to be consistent with the notation and definitions introduced in [11] to characterize the low-SNR (or *wideband*) regime, we shall consider the complex circularly symmetric version of our channel model, where $\gamma = E_s/N_0$ is the *transmitter* SNR, and N_0 denotes the complex noise variance per component. The system spectral efficiency C (measured in bit/s/Hz) as a function of E_b/N_0, where E_b denotes the transmitted energy per information bit. The wideband slope $\mathcal{S}_0$ is the value of bits per second per Hz per 3dB of energy per bit evaluated at the minimum E_b/N_0.

Theorem 10 *The minimum energy per bit and wideband slope $\mathcal{S}_0$ of the spectral efficiency of the Gaussian erasure channel are given by*

$$\left(\frac{E_b}{N_0}\right)_{\min} = \frac{1}{(1-e)G_{\max}\log_2 e} \qquad (56)$$

$$\mathcal{S}_0 = \frac{2(1-e)B_{\max}}{eB_{\max}+1-e} \qquad (57)$$

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

D. High-SNR

For large SNR, capacity behaves like

$$C_e(\gamma) = \mathcal{S}_\infty \left(\log_2 \gamma - \mathcal{L}_\infty\right) + o(1) \qquad (58)$$

(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 [12].

Theorem 11 *Consider a Gaussian erasure channel with given channel transfer function $H(f)$ and erasure probability e. Let $\mathcal{I} = \{f : |H(f)|^2 > 0\}$ and $B = \mu(\mathcal{I})$ denote its generalized bandwidth. The high-SNR slope is given by*

$$\mathcal{S}_\infty = \min\{1 - e, B\} \qquad (59)$$

and the high-SNR dB offset is given by

$$\mathcal{L}_\infty = -\int_0^1 \log_2 F(y) \ dy \qquad (60)$$

where $F(y)$ is the solution of the fixed point equation

$$B - y\mathcal{S}_\infty = \int_{\mathcal{I}} \frac{1}{1 + \frac{y(1-e-y\mathcal{S}_\infty)}{(1-y)(1-y\mathcal{S}_\infty)F} \frac{|H(f)|^2}{B}} \ df \qquad (61)$$

REFERENCES

[1] S. Verdú and T. Weissman, "The Information Lost in Erasures," *preprint*, 2007.
[2] D. Julian, "Erasure networks," *2002 IEEE International Symp. Information Theory*, p. 138, Lausanne, Switzerland, June 30-July 5, 2002.
[3] 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.
[4] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[5] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–21, Jan. 1949.
[6] R. M. Gray, "Toeplitz and Circulant Matrices: A Review," *Foundations and Trends in Communications and Information Theory*, vol. 3, 2006.
[7] 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.
[8] A. Tulino, S. Verdú, G. Caire, and S. Shamai, "The Capacity of the Gaussian Erasure Channel," *preprint*, 2007
[9] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*. New York: Academic Press, 1985.
[10] 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.
[11] S. Verdú, "Spectral Efficiency in the Wideband Regime" *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319-1343 , June 2002.
[12] 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.