Capacity Loss Due to Sub-Optimal Water-Pouring

by

Brett E. Schein

B.S. Electrical Engineering
University of California, Berkeley
May 1994

Submitted to the Department of Electrical Engineering and Computer
Science
in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1996

© Massachusetts Institute of Technology 1996. All rights reserved.

Author ....
Department of Electrical Engineering and Computer Science
May 24, 1996

Certified by...
Mitchell D. Trott
Assistant Professor of Electrical Engineering
Thesis Supervisor

Accepted by ....
F. R. Morgenthaler
Chairman, Departmental Committee on Graduate Students

Capacity Loss Due to Sub-Optimal Water-Pouring

by

Brett E. Schein

Abstract

We consider a particular sub-optimal "water-pouring" scheme applied to an additive, colored Gaussian noise channel. The scheme allocates power uniformly over an optimized band of frequencies. Within the set of all noise power spectral densities yielding a fixed water-pouring capacity, we find the noise spectrum yielding the smallest mutual information between input and output when the transmitter is restricted to use the sub-optimal scheme. The loss at high SNR is a fraction of a bit per Hz; the loss at low SNR is less than 15%.

Thesis Supervisor: Mitchell D. Trott
Title: Assistant Professor of Electrical Engineering

Acknowledgements

I thank my parents for their love, advice, and support in all aspects of my life. I thank my fiance, Joan, for being who she is. Without her, I might not even be here, and I certainly would not be as happy.

I thank my advisor, Prof. Mitch Trott, for his patience and willingness to listen to whatever comes to mind. The luxury to freely compare thoughts was instrumental in our work. I also thank him for dedicating so much time and effort to me.

I thank Prof. Amos Lapidoth for his helpful suggestions and for letting us bother him with our problems.

I thank my officemates, Marc Ibanez, Patrick Maurer, Aradhana Narula, James Sarvis, and Don Wilcoxson, for allowing me to bounce ideas off them, draw pictures on the whiteboard, and whine.

Contents

List of Figures

Chapter 1

Introduction

From an information-theoretic standpoint, the problem of selecting the optimal power spectral density for communication over an additive, colored Gaussian noise channel was solved by Shannon in his original paper [9] and treated rigorously by Pinsker [8]. An essentially equivalent problem, for filtered channels with white Gaussian noise, was solved by Holsinger [5]. Given a fixed power budget, the "water-pouring" solution maximizes the rate of reliable communication. This maximum rate is called the channel capacity [4]. Calculating the capacity, however, requires full characterization of the communication channel. In a practical communication environment, channel characteristics must be estimated, permitting only an approximation to the water-pouring solution. Channel information must also be fed back to the transmitter, possibly requiring a feedback path of significant bandwidth. (Consider describing an arbitrary noise power spectral density $S_{nn}(\omega)$ on $\omega \in [0, \pi]$.) Furthermore, implementing a close approximation to the water-pouring solution may be computationally expensive or excessively complex. The latter concern is not of great significance for a DSP-based multi-carrier system, but may be important for a single-carrier system or when decoding delay is critical.

Sub-optimal power allocation schemes have thus been popularly adopted that do not require exact channel identification or complicated transmission spectrums. Most schemes have been justified by an empirical verification of small capacity loss, assuming a reasonable approximation to the noise power spectrum is available. Existing

analytical work supports this assessment for asymptotically large input power[1] [2]. This phenomenon has inspired a folk theorem that such schemes are relatively efficient in terms of the loss of the maximum achievable rate of reliable communication.

In this research, we find the maximum capacity loss incurred by a particular sub-optimal power allocation scheme. The scheme we investigate restricts the input power spectrum to be uniform over a set of frequency bands. Intuition suggests that an intelligent choice of frequencies should keep the capacity loss modest. We show that this is indeed true by computing the worst-case channel for every SNR. For low SNR, where normalized loss is important, the capacity loss is at most 14.25%. For high SNR, where absolute loss is important, the capacity loss is at most 0.1874 bits per channel use.

While we do not consider any other sub-optimal power allocation scheme, we expect comparable results for any related method. Similarly, we do not solve the more difficult problem of bounding the loss in capacity due to imperfect channel knowledge. To even define this problem requires an assumption on how the noise power spectral density is measured; it is not clear what (if any) model is most natural.

The thesis is organized as follows. In Chapter 2 we describe the water-pouring solution for an additive, colored Gaussian noise channel. In Chapter 3 we define our research problem as a constrained non-linear optimization. In the remainder of the chapter we discuss why standard approaches to its solution are not applicable. In Chapter 4 we characterize the solution to a sub-problem that allows only bounded, piecewise constant noise power spectral densities. In Chapter 5 we use continuity arguments to show that the piecewise constant solution is in fact the solution to the original problem. In Chapter 6 we present the results and discuss their implications.

---

[1]The analysis in [2] uses an error-free zero-forcing decision feedback equalizer, a physically unrealizable device. With an ideal DFE, sub-optimal power allocation has no loss in performance at high SNR. As shown later in the thesis, an exact analysis reveals a high SNR loss of 0.3748 bits/sec/Hz.

Chapter 2

Background

We discuss in Section 2.1 the physical motivation for the colored Gaussian noise channel model. We then describe in Section 2.2 a fundamental result from information theory which asserts that there is a maximum rate of reliable communication over this channel. The discussion naturally leads to the well-known water-pouring solution to maximizing the achievable rate.

2.1   Additive Colored Gaussian Noise Model

We base our research on a discrete-time additive colored Gaussian noise channel. This model arises from the sampled baseband representation of a continuous-time, strictly band-limited, time-invariant linear channel with memory. A typical system is diagrammed in Figure 2-1(a). Digital to analog conversion is modeled as ideal discrete-to-continuous impulse train conversion followed by convolution with a pulse $p(t)$. The channel, indicated by the dashed box in the figure, is modelled as a linear filter $h(t)$ followed by zero-mean additive white Gaussian noise $n(t)$ with power density $\frac{N_0}{2}$ Watts/Hz.

The receiver is a whitened matched filter and sampler (which results in no loss of information [7]). The resulting complex baseband discrete-time equivalent channel is shown in Figure 2-1(b), where $G(z)$ is the composite sampled channel, and where the discrete-time complex circular additive white Gaussian noise $\tilde{n}[k]$ has constant power
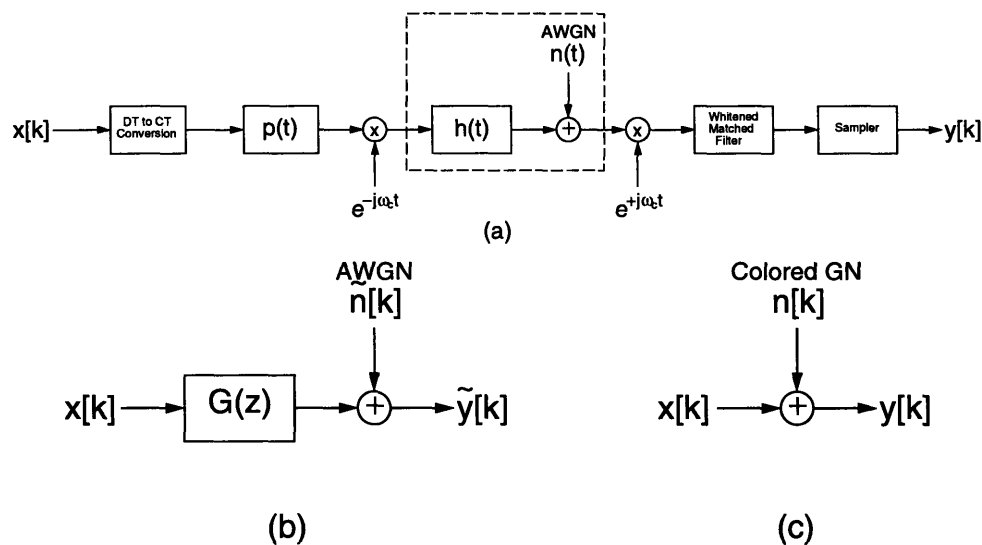
Figure 2-1: Channel Model

spectral density $S_{\tilde{n}\tilde{n}}(\omega) = \mathcal{N}_0 \left[\int \left[(p * h)(t) \right]^2 dt \right]^{\frac{1}{2}}$, $-\pi \leq \omega \leq \pi$.

With a simple change of basis, we can reformulate the filtered channel with white noise as a memoryless channel with colored noise. The resulting channel is pictured in Figure 2-1(c), where the colored Gaussian noise has power spectral density

$$S_{nn}(\omega) = S_{\tilde{n}\tilde{n}}(\omega) \left| G^{-1}(\omega) \right|^2. \tag{2.1}$$

Because the real and imaginary components of the noise are independent, we will henceforth consider only the case of a real signal $x[k]$. That is, one real-valued channel use corresponds to $\frac{1}{2}$ bit per second per Hz.

2.2 Information-Theoretic Results

2.2.1 Mutual Information

Information theory tells us that if the user of the discrete-time channel described in Section 2.1 has an average input power constraint, there is a well-defined maximum rate of reliable communication. Reliable communication is defined as the ability of

the receiver to reproduce one of a finite set of messages chosen by the transmitter with an arbitrarily low probability of error. The average power constraint typically arises either from channel constraints (e.g., FCC or AT&T regulations) or user constraints (e.g., battery life).

Given a wide-sense stationary stochastic input process X_k, the output process Y_k is also stationary. Hence, for a fixed input power spectral density (PSD) $S_{xx}(\omega), \omega \in [0, \pi]$, the mutual information is maximized by a Gaussian input process [6]. The mutual information between input and output may be defined as

$$I(X, Y) = \lim_{n \to \infty} \frac{1}{n} I(X_1, \ldots, X_n; Y_1, \ldots, Y_n).$$ (2.2)

The mutual information is then

$$I(X, Y) = \frac{1}{2\pi} \int_0^\pi \log_2 \left(1 + \frac{S_{xx}(\omega)}{S_{nn}(\omega)} \right) d\omega.$$ (2.3)

The integrand in (2.3) is defined to be zero when both the input spectrum and the noise spectrum are zero at the same frequency.

Using this expression, we define the *capacity* of the channel as

$$C = \max_{S_{xx}(\omega):E[X^2] \le P} I(X; Y),$$ (2.4)

with an input power constraint

$$P = E\left[X^2 \right] = \frac{1}{\pi} \int_0^\pi S_{xx}(\omega) \, d\omega.$$ (2.5)

The units of capacity are bits per channel use.

Information theory tells us two critical things regarding the capacity of the channel. First, if we want to communicate reliably at any rate R (bits/channel use), where $R < C$, then we can use the capacity-achieving input process to generate an appropriate codebook. Second, if we attempt to communicate at any rate $R > C$, then we are guaranteed to have a non-zero lower-bound on the probability of error.

15
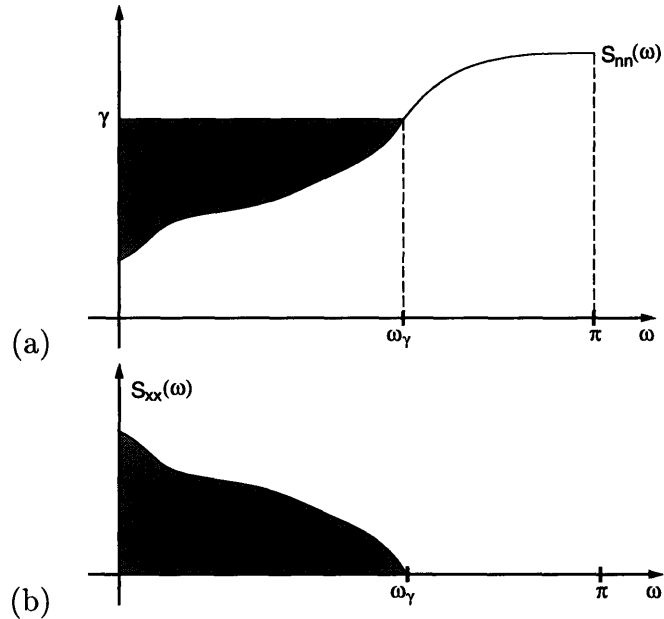
Figure 2-2: Water-Pouring for a Colored Channel.

2.2.2 Water-Pouring Solution

As mentioned in Section 2.2.1, the capacity of this channel, subject to the average input power constraint P, is the maximum mutual information achievable using a stationary Gaussian input process. The well-known maximizing solution is called the *water-pouring solution*. The maximizing equations are (from [3]):

$$C(S_{nn}) = \frac{1}{2\pi} \int_0^\pi \log\left(1 + \frac{(\gamma - S_{nn}(\omega))^+}{S_{nn}(\omega)}\right) d\omega \qquad (2.6)$$

where the *water-pouring level* γ is chosen to satisfy the power constraint

$$P = \frac{1}{\pi} \int_0^\pi (\gamma - S_{nn}(\omega))^+ d\omega. \qquad (2.7)$$

The function $(x)^+$ is defined as $(x)^+ = \max(0, x)$. These equations have been trivially simplified by accounting for the even noise spectrum.

Looking at the solution graphically, as in Figure 2-2, it is clear why the solution is called "water-pouring." If the power constraint P is interpreted as a volume of water,

the input PSD which maximizes the mutual information is determined by the height of the water in its minimum potential energy configuration. A graphical example is shown in Figure 2-2(a), and the resulting input PSD is shown in (b).

Chapter 3

Problem Definition

3.1 Basic Description of the Scheme

A common approach to implementing an approximation to the water-pouring solution for a known channel is to use a uniform power allocation over an intelligent choice of bandwidth. Unlike the water-pouring approach, the input spectrum is not weighted according to the noise.

Consider again the water-pouring solution pictured in Figure 2-2 — the input power density is largest where the noise power is smallest. The assumption, backed by empirical observation, is that a uniform power approach never reduces the mutual information between input and output too far from capacity (as calculated in (2.3) and (2.6)).

We take a game-theoretic approach to calculating the capacity loss due to a uniform power allocation. Nature selects a channel (i.e., a noise power spectral density S_{nn}) from among those that yield a fixed capacity C using a fixed power P, where capacity is calculated via the water-pouring equations (2.6) and (2.7). Given P and S_{nn}, the user then selects a set of frequency bands to maximize the mutual information $I(X;Y)$ when uniform power allocation is enforced. We pit nature against man by finding the worst possible channel to give the user.

We note in passing that the water-pouring frequency bands are usually not the best bands to use with a uniform power allocation. For example, the water-pouring

19
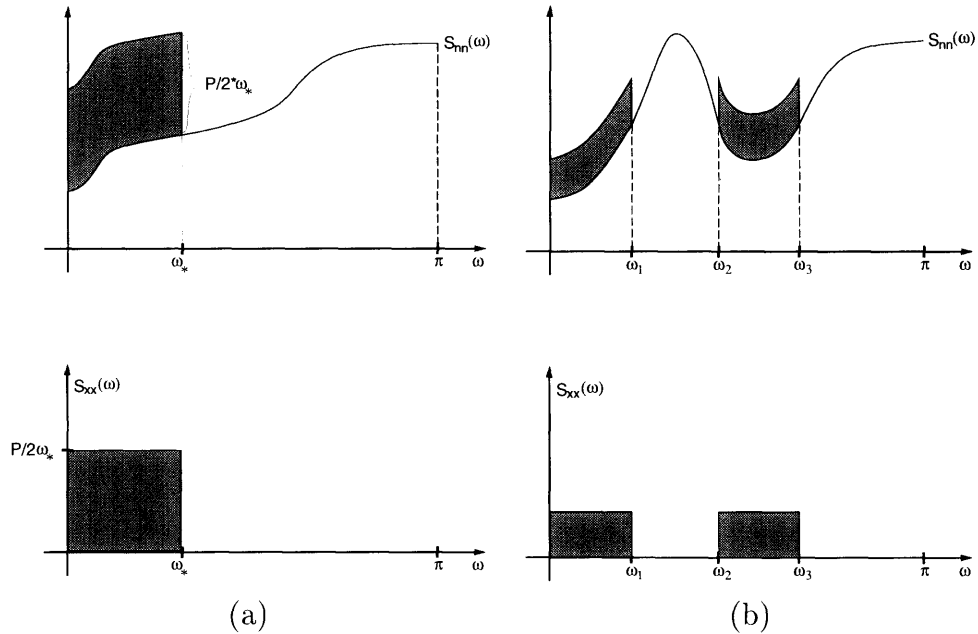
Figure 3-1: Uniform Power Allocation

solution might use a small amount of power over a large bandwidth; a uniform power allocation restricted to use the same bandwidth would be highly sub-optimal. This restriction does not seem to be particularly natural (nor particularly interesting). Hence we consider the more general problem where the user may select bands freely.

Note also that we do not restrict the user to allocate all power over a contiguous band. For example, the allocations sketched in Figure 3-1 are both allowed, though our restriction to a uniform power allocation requires that the height of the input spectrum over both bands in (b) be equal.

We now precisely define the minimax problem.

3.2   Detailed Problem Definition

We parametrize the problem by considering all noise power spectral densities which achieve some fixed capacity C. We assume that all noise PSD's are Lebesgue measurable. Though we would hope for analytical reasons to deal only with continuous functions, it will turn out that the worst-case noise PSD is piecewise continuous. We

do not, however, assume this a priori.

Define the class of admissible noise power spectral densities as

$$\mathcal{A}(C) = \{S_{nn}(\cdot) \mid C(S_{nn}) = C\}. \tag{3.1}$$

The capacity is calculated according to the general water-pouring equations (2.6) and (2.7), assuming a power constraint of P.

The sub-optimal power allocations available to the user have the form

$$S_{xx}(\omega) = \begin{cases} \frac{P}{2\mu(\mathcal{W})} & \text{if } \omega \in \mathcal{W} \\ 0 & \text{if } \omega \notin \mathcal{W}, \end{cases} \tag{3.2}$$

where $\mathcal{W} \subseteq [0, \pi]$ is some (not necessarily contiguous) subset of the frequency band and $\mu(\cdot)$ is Lebesgue measure.

From (2.3) applied to a uniform power allocation, the mutual information is

$$R(S_{nn}, \mathcal{W}) = \frac{1}{2\pi} \int_{\mathcal{W}} \log\left(1 + \frac{P}{2\mu(\mathcal{W})S_{nn}(\omega)}\right) d\omega. \tag{3.3}$$

For every noise PSD, the user optimizes the choice of signalling frequencies $\mathcal{W}$. We denote the maximized mutual information $R^\star(S_{nn})$:

$$R^\star(S_{nn}) = \max_{\mathcal{W} \subseteq [0,\pi]} R(S_{nn}, \mathcal{W}) \tag{3.4}$$

$$= \max_{\mathcal{W} \subseteq [0,\pi]} \frac{1}{2\pi} \int_{\mathcal{W}} \log\left(1 + \frac{P}{2\mu(\mathcal{W})S_{nn}(\omega)}\right) d\omega. \tag{3.5}$$

Note that we can assume any fixed value for P without loss of generality, because the power constraint P has no effect on the set of rates

$$\{R^\star(S_{nn}) \mid S_{nn} \in \mathcal{A}(C)\}. \tag{3.6}$$

To see this, assume that S_{nn} has capacity C and rate $R^\star(S_{nn})$ with power constraint P. If the power constraint is changed to αP, then αS_{nn} achieves the same capacity

C (with water-pouring level $\alpha\gamma$) and the same rate $R^\star(S_{nn})$. For the remainder of the thesis, we assume $P = 2$ (which slightly simplifies the equations).

3.3 Monotone Rearrangement

We can greatly simplify our problem by realizing that we need only consider monotonically non-decreasing noise power spectra rather than arbitrary measurable ones. Looking at the capacity and sub-optimal rate equations ((2.6) and (3.5), respectively), we see that both depend only on the distribution of the noise power spectra.

Given any noise spectrum S_{nn}, we can find an equivalent nondecreasing noise spectrum $S_{n'n'}$ with the same distribution [10, App B.2], hence S_{nn} and $S_{n'n'}$ will have the same capacity and sub-optimal rate [1].

The governing equations simplify for monotonically non-decreasing noise functions:

$$C(S_{nn}) = \frac{1}{2\pi}\int_0^{\omega_\gamma} \log\left(\frac{\gamma}{S_{nn}(\omega)}\right) d\omega, \tag{3.7}$$

$$P = 2 = \frac{1}{\pi}\int_0^{\omega_\gamma} (\gamma - S_{nn}(\omega))d\omega, \tag{3.8}$$

$$\mathcal{A}(C) = \left\{ S_{nn}(\cdot) \mid C(S_{nn}) = C \right\}, \tag{3.9}$$

$$R(S_{nn}, \omega_{\text{bw}}) = \frac{1}{2\pi}\int_0^{\omega_{\text{bw}}} \log\left(1 + \frac{1}{\omega_{\text{bw}} S_{nn}(\omega)}\right) d\omega, \tag{3.10}$$

and finally,

$$R^\star(S_{nn}) = \max_{\omega_{\text{bw}}} R(S_{nn}, \omega_{\text{bw}}). \tag{3.11}$$

The water-pouring bandwidth ω_γ is the smallest frequency for which $\gamma - S_{nn}(\omega) \leq 0$.

In writing the above equation for $R(S_{nn}, \omega_{\text{bw}})$, we have assumed the user allocates his power over the frequency band of measure ω_{bw} with lowest noise power. Also, for notational purposes, define $\omega_\star$ to be the bandwidth ω_{bw} which maximizes (3.11). The

final minimax optimization problem then becomes:

$$L(C) \;=\; C - \min_{\substack{S_{nn}(\cdot) \in \mathcal{A}(C)}} \; \max_{\substack{S_{xx}\ \text{uniform} \\ E[X^2] \le 2}} \; I(X;Y) \tag{3.12}$$

$$\;=\; C - \min_{\substack{S_{nn}(\cdot) \in \mathcal{A}(C)}} \; R^{\star}(S_{nn}). \tag{3.13}$$

3.4   What is Hard?

The constrained optimization problem (3.13) is surprisingly difficult to solve.

First, the calculus of variations cannot easily be applied because the noise spectra we consider may be discontinuous at any number of points. We consider discontinuous functions because, in a sense, they are the completion of the continuous functions. While discontinuous spectra are not physically meaningful, restricting the optimization to continuous spectra is of no help because the optimizing spectrum turns out to be discontinuous.

Second, the constraint set $\mathcal{A}(C)$ of noise power spectral densities yielding capacity equal to C is not convex. Furthermore, loosening the constraint set to $\mathcal{B}(C) = \{S_{nn}(\cdot) \mid C(S_{nn}) \ge C\}$ still results in a non-convex set. Consider Figure 3-2, where both noise spectra $S_{n_1 n_1}$ and $S_{n_2 n_2}$ yield capacity equal to C. Then the convex combination $S_{n_3 n_3} = \frac{1}{2} S_{n_1 n_1} + \frac{1}{2} S_{n_2 n_2}$ necessarily yields a lower capacity. Thus $S_{n_3 n_3}$ is not in $\mathcal{A}$ or $\mathcal{B}$.

Third, the functionals $C(S_{nn})$ and $R^{\star}(S_{nn})$ are integrals with varying (and different) end-points. The capacity functional integrates over the water-pouring bandwidth $[0, \omega_\gamma]$, while the sub-optimal rate functional integrates over an optimized bandwidth $[0, \omega_\star]$. Recall that both ω_γ and $\omega_\star$ vary with S_{nn}.

Fourth, neither $C(S_{nn})$ nor $R^{\star}(S_{nn})$ is continuous in any of the $L_p[0, \pi]$ spaces. We solve this problem in the next chapter by bounding all noise PSD's below (almost everywhere) by an arbitrarily small but fixed lower bound b.
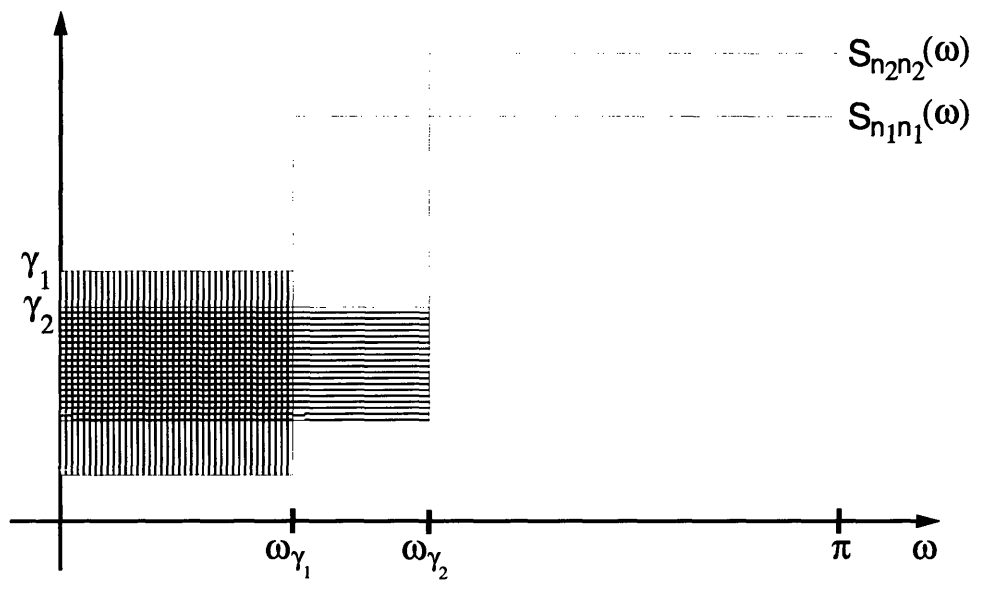
Figure 3-2: Nonconvexity of Admissible Sets

Chapter 4

Necessary and Sufficient Conditions

As mentioned in Section 3.4, we consider a sub-problem of the original. We bound the noise functions above by an arbitrarily large but fixed upper bound B, and we bound them below by an arbitrarily small but fixed lower bound b (for continuity).

We solve this sub-problem in the following chapters. We first solve a simpler, related sub-problem that is restricted to piecewise constant noise spectra. We then use continuity to prove the result is the same for the general bounded sub-problem.

The bounded and piecewise constant sub-problems are defined in Sections 4.1 and 4.2. In Section 4.3, we argue that the infimum over piecewise constant noise spectra is achievable. In Section 4.4, we derive necessary conditions on the solution; with achievability, this implies that the solution has one of a small number of forms.

4.1   Bounded Power Spectra

We define a sub-problem of the original constrained optimization problem stated in Equation (3.13). We will call this sub-problem the *bounded problem*. The set of allowable noise power spectral densities is

$$\mathcal{A}(C, b, B) = \{S_{nn} \mid C(S_{nn}) = C, \ b \leq S_{nn}(\omega) \leq B\}, \tag{4.1}$$

where $0 < b < B$ are arbitrary (but fixed) lower and upper bounds. The bounded problem is:

$$L_{b,B}(C) = \max_{S_{nn} \in \mathcal{A}(C,b,B)} C - R^\star(S_{nn}). \tag{4.2}$$

We find that the maximizing power spectrum S_{nn} does not track either bound as they are loosened appropriately ($b \to 0$ and $B \to \infty$).

The upper bound allows the noise spectrum to be uniformly approximated with a series of piecewise constant functions (see Section 5.1) and guarantees the existence of a solution to the simplified problem defined below (see Section 4.2). The lower bound is needed for the continuity of capacity in the L^∞ norm (see Section 5.2). The upper bound does not appear to be an essential restriction, since the water-pouring and uniform allocation bandwidths ω_γ and $\omega_\star$ automatically avoid regions with sufficiently large noise power. The lower bound eliminates the possibility of a noise spectrum for which all capacity is achieved in the neighborhood of a spectral null. We consider this to be a physically unreasonable case and thus ignore it. Unlike the upper bound, however, our arguments cannot be easily extended to eliminate the lower bound.

4.2   Piecewise Constant Power Spectra

Rather than solving (4.2) directly for the maximizing noise power spectrum, we search for a solution within the class of piecewise constant power spectra with a finite number of steps. While this restricted class is small, we will see in Section 5.4 that it is rich enough to solve the bounded problem (4.2).

Let $T(M)$ be the set of bounded, non-decreasing step functions with at most M steps, and let

$$\mathcal{K}(C, M) \;=\; \mathcal{A}(C, b, B) \bigcap T(M) \tag{4.3}$$

be the set of bounded M-step power spectra that have capacity C. The bounds
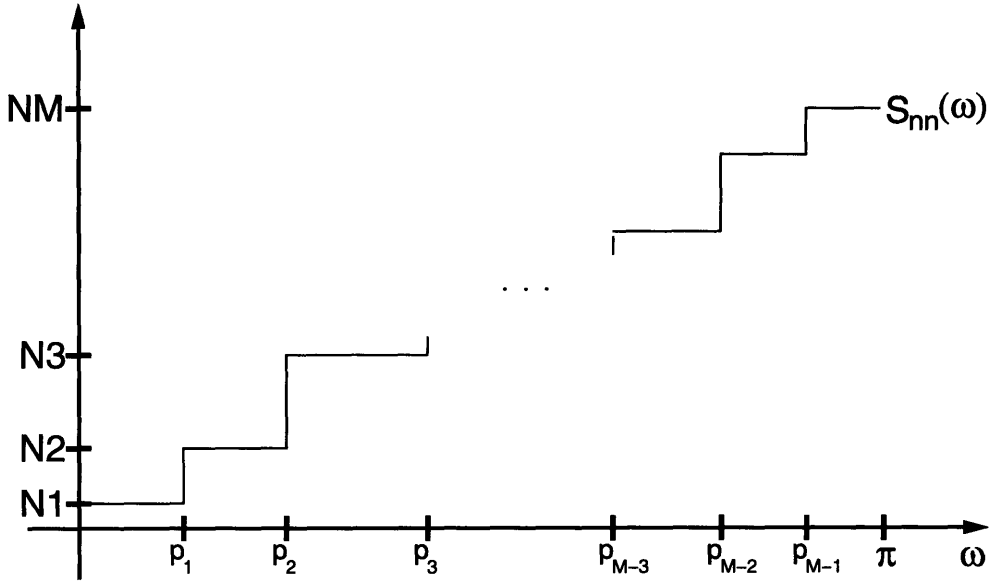
Figure 4-1: Representative Member of $T(M)$

(b, B) are omitted from the notation for brevity. To avoid confusion with general $S_{nn} \in \mathcal{A}(C, b, B)$, we denote members of $\mathcal{K}(C, M)$ by $S_{\tilde{n}\tilde{n}}$.

We reduce (4.2) to the following simplified problem:

$$W_M(C) = \min_{S_{\tilde{n}\tilde{n}} \in \mathcal{K}(C,M)} R^\star(S_{\tilde{n}\tilde{n}}). \qquad (4.4)$$

Since the noise spectra in $\mathcal{K}(C, M)$ are non-decreasing step functions, this is a $2M$-dimensional constrained optimization problem. A representative step function is sketched in Figure 4-1. We parametrize $S_{\tilde{n}\tilde{n}}$ by the step heights $\mathbf{N} = [N_1, N_2, \ldots, N_M]$ and step positions $\mathbf{p} = [p_1, p_2, \ldots, p_M]$, and we refer interchangeably to $(\mathbf{N}, \mathbf{p})$ and the corresponding $S_{\tilde{n}\tilde{n}}$. To solve (4.4), we take the following approach. First, we show that an achievable solution exists (i.e., the minimum is not just an infimum). We do this by showing that $C(\mathbf{N}, \mathbf{p})$ and $R^\star(\mathbf{N}, \mathbf{p})$ are continuous with respect to the Euclidean norm on the (bounded) $(\mathbf{N}, \mathbf{p})$ space. Second, there are only two possibilities for the solution to (4.4): either $\omega_\star$ is on the edge of a step, or $\omega_\star$ is on the interior of a step. For each case, we derive necessary conditions for local extrema and conclude that only a small number of configurations are possible.

4.3 Achievability

We now establish the continuity of $C(\mathbf{N}, \mathbf{p})$ and $R^\star(\mathbf{N}, \mathbf{p})$. Given any $S_{\tilde{n}\tilde{n}} \in \mathcal{K}$, we know from the water-pouring solution that ω_γ must lie at the end of some step. Assume $\omega_\gamma = p_r$, the end of the rth step. Then the equation for capacity simplifies to

$$
\begin{aligned}
C(\mathbf{N}, \mathbf{p}) &= C(S_{\tilde{n}\tilde{n}}) & (4.5)\\
&= \sum_{i=1}^{r} (p_i - p_{i-1}) \ln\left(\frac{\gamma}{N_i}\right) & (4.6)\\
&= \sum_{i=1}^{r} (p_i - p_{i-1}) \ln\left(\frac{1}{\omega_\gamma}\left(1 + \sum_{j=1}^{r}(p_j - p_{j-1})N_j\right)\right) - \ln(N_i). & (4.7)
\end{aligned}
$$

Some implicit constraints are $\gamma \le N_r$, $p_0 = 0$, $p_M = \pi$, $0 \le p_1 \le p_2 \le \cdots \le p_M$, and $b \le N_1 \le N_2 \le \cdots \le N_M \le B$.

From (4.7), it is clear that $C(\mathbf{N}, \mathbf{p})$ is a continuous function with respect to the Euclidean norm on $\Re^{2M}$. Furthermore, the bounded $2M$-dimensional space is closed and thus compact. Hence, the subset of $\Re^{2M}$ corresponding to $\mathcal{K}(C, M)$ is compact.

For $\omega_{\text{bw}} \in [p_{s-1}, p_s]$, the equation for the sub-optimal rate is

$$
\begin{aligned}
R(S_{\tilde{n}\tilde{n}}, \omega_{\text{bw}}) = R(\mathbf{N}, \mathbf{p}, \omega_{\text{bw}}) &= \sum_{i=1}^{s-1}(p_i - p_{i-1}) \ln\left(1 + \frac{1}{\omega_{\text{bw}} N_i}\right) \\
&\quad + (\omega_{\text{bw}} - p_{s-1}) \ln\left(1 + \frac{1}{\omega_{\text{bw}} N_s}\right) & (4.8)\\
R^\star(S_{\tilde{n}\tilde{n}}) = R^\star(\mathbf{N}, \mathbf{p}) &= \max_{\omega_{\text{bw}}} R(S_{\tilde{n}\tilde{n}}, \omega_{\text{bw}}). & (4.9)
\end{aligned}
$$

The optimizing bandwidth $\omega_\star$ lies on some step s, so that $\omega_\star \in [p_{s-1}, p_s]$. Unlike ω_γ, $\omega_\star$ need not be an endpoint of a step. For $\omega_{\text{bw}} \in (p_{s-1}, p_s)$,

$$
\begin{aligned}
\left.\frac{\partial R}{\partial \omega_{\text{bw}}}\right|_{\omega_\star} &= \ln\left(1 + \frac{1}{\omega_\star N_s}\right) - \sum_{i=1}^{s-1} \frac{(p_i - p_{i-1})}{\omega_\star(\omega_\star N_i + 1)} \\
&\quad - \frac{(\omega_\star - p_{s-1})}{\omega_\star(\omega_\star N_s + 1)}. & (4.10)
\end{aligned}
$$

At the left and right end-points $\omega_{\text{bw}} = p_{s-1}$ and $\omega_{\text{bw}} = p_s$, (4.10) remains valid for right and left partial derivatives, respectively. Thus, if $\omega_\star \in (p_{s-1}, p_s)$, the continuity
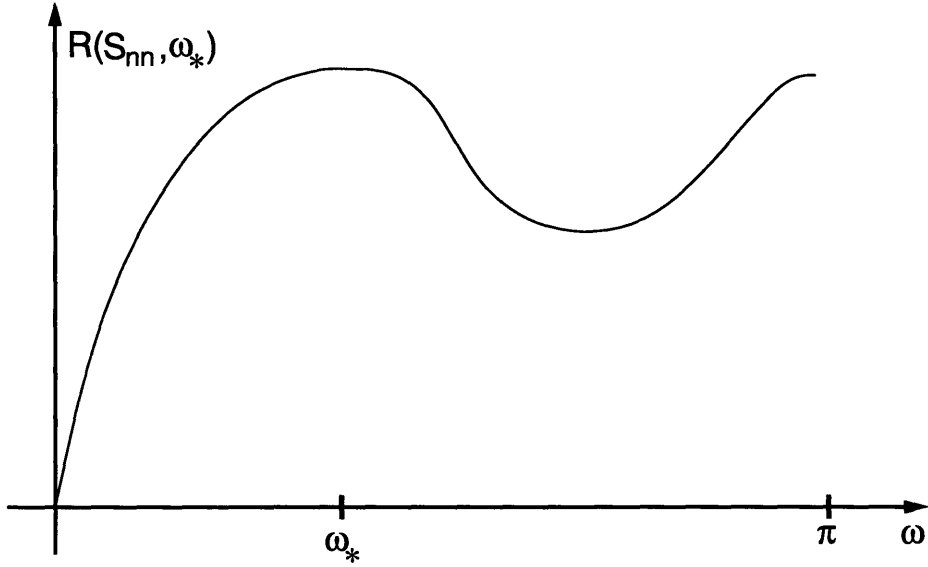
Figure 4-2: Multiple Peaks in $R(S_{\tilde{n}\tilde{n}}, \omega_\star)$

of (4.10) in ω_{bw} implies that $\left.\frac{\partial R}{\partial \omega_{\text{bw}}}\right|_{\omega_\star} = 0$. If $\omega_\star = p_s$, however, we know only that $\left.\frac{\partial R}{\partial \omega_{\text{bw}}}\right|_{\omega_\star -} \geq 0$ and $\left.\frac{\partial R}{\partial \omega_{\text{bw}}}\right|_{\omega_\star +} \leq 0$.

To establish the continuity of $R^\star(\mathbf{N}, \mathbf{p})$, we rewrite it as

$$R^\star(\mathbf{N}, \mathbf{p}) = R(\mathbf{N}, \mathbf{p}, \omega_\star(\mathbf{N}, \mathbf{p})), \tag{4.11}$$

where

$$\omega_\star(\mathbf{N}, \mathbf{p}) = \arg\max_{\omega_{\text{bw}}} R(\mathbf{N}, \mathbf{p}, \omega_{\text{bw}}). \tag{4.12}$$

It suffices to show that both $R(\mathbf{N}, \mathbf{p}, \omega_{\text{bw}})$ and $\omega_\star(\mathbf{N}, \mathbf{p})$ are continuous. The former is clear. The latter is intuitively reasonable, but we have not established this rigorously. One way in which $\omega_\star(\mathbf{N}, \mathbf{p})$ can fail to be continuous is if $R(\mathbf{N}, \mathbf{p}, \omega_{\text{bw}})$ has multiple peaks in ω_{bw}, as sketched in Figure 4-2. However, we believe that $\left.\frac{\partial^2 R}{\partial \omega_{\text{bw}}^2}\right|_{\omega_\star} < 0$, which would imply that $R(\mathbf{N}, \mathbf{p}, \omega_{\text{bw}})$ is unimodal in ω_{bw}. (Even if $\omega_\star(\mathbf{N}, \mathbf{p})$ is not continuous, still $R^\star(\mathbf{N}, \mathbf{p})$ is almost certainly so.) Granted continuity of $R^\star$, a global minimum noise spectrum is achievable within the M-step problem (4.4).

4.4   Necessary Conditions

4.4.1   Sub-Case: Necessary Conditions When $\omega_\star \in (p_{s-1}, p_s)$

We now derive necessary conditions that must be satisfied by any local extremum $S_{\tilde{n}\tilde{n}} \in \mathcal{K}(C, M)$ for which $\omega_\star$ is on the interior of a step, $\omega_\star \in (p_{s-1}, p_s)$. In such cases, we have already established that $C(\mathbf{N}, \mathbf{p})$, $R(\mathbf{N}, \mathbf{p}, \omega_{\text{bw}})$, and $\frac{\partial R}{\partial \omega_{\text{bw}}}(\mathbf{N}, \mathbf{p}, \omega_{\text{bw}})$ are continuous. Hence, Lagrange multipliers may be used.

Form the functional

$$J = R(\mathbf{N}, \mathbf{p}, \omega_{\text{bw}}) + \lambda C(\mathbf{N}, \mathbf{p}) + \lambda' \frac{\partial R}{\partial \omega_{\text{bw}}}(\mathbf{N}, \mathbf{p}, \omega_{\text{bw}}). \tag{4.13}$$

Provided that $(\mathbf{N}, \mathbf{p})$ is a regular point (an issue we do not address), we are guaranteed that for any local extremum $(\mathbf{N}, \mathbf{p})$, a pair of constants (λ, λ') exist such that J is stationary. Since in this sub-case $\omega_\star$ is at a point of zero derivative,

$$0 = \frac{\partial J}{\partial \omega_{\text{bw}}} \tag{4.14}$$

$$= \frac{\partial R}{\partial \omega_{\text{bw}}} + \lambda \frac{\partial C}{\partial \omega_{\text{bw}}} + \lambda' \frac{\partial^2 R}{\partial \omega_{\text{bw}}^2} \tag{4.15}$$

$$= \lambda' \frac{\partial^2 R}{\partial \omega_{\text{bw}}^2}. \tag{4.16}$$

The last equation follows because $\frac{\partial R}{\partial \omega_{\text{bw}}}$ is zero, while the second term is zero because capacity is independent of ω_{bw}. Under the assumption (still to be verified) that the second derivative is not zero, from (4.16) we conclude that $\lambda' = 0$. The functional in (4.13) simplifies to

$$J = R(\mathbf{N}, \mathbf{p}, \omega_{\text{bw}}) + \lambda C(\mathbf{N}, \mathbf{p}). \tag{4.17}$$

We proceed by deriving necessary conditions on the local extrema in this subcase. Recall that $\omega_\gamma = p_r$ for some step r.

Result 1: If $\omega_\gamma < \omega_\star$, which implies $\gamma \leq N_{r+1}$, then

$$\frac{\partial J}{\partial N_{r+1}} = \frac{\partial R}{\partial N_{r+1}} + \lambda \frac{\partial C}{\partial N_{r+1}} \tag{4.18}$$

$$= \frac{\partial R}{\partial N_{r+1}} \tag{4.19}$$

$$= \min\left((p_{r+1} - p_r), (\omega_\star - p_r)\right) \frac{1}{N_{r+1} + \frac{1}{\omega_\star}} \tag{4.20}$$

$$\neq 0. \tag{4.21}$$

The minimum term above takes care of both $s = r+1$ and $s > r+1$. Since $\frac{\partial J}{\partial N_{r+1}} = 0$ for local extrema, local extrema must satisfy $\omega_\gamma \geq \omega_\star$. The water-pouring bandwidth ω_γ is always at least as large as the uniform allocation bandwidth $\omega_\star$.

Result 2: If $\omega_\gamma > p_s$, that is, if the water-pouring bandwidth is past the step containing $\omega_\star$, then

$$\frac{\partial J}{\partial N_{s+1}} = \frac{\partial R}{\partial N_{s+1}} + \lambda \frac{\partial C}{\partial N_{s+1}} \tag{4.22}$$

$$= \lambda \frac{\partial C}{\partial N_{s+1}} \tag{4.23}$$

$$\neq 0. \tag{4.24}$$

This follows because λ cannot be 0 (or the capacity constraint would be inactive), $\frac{\partial R}{\partial N_{s+1}} = 0$ since $\omega_\star < p_s$, and $\frac{\partial C}{\partial N_{s+1}} < 0$ since $\omega_\gamma \geq p_{s+1}$ by assumption. Thus we see that local extremum must satisfy $r = s$. In other words, $\omega_\star \in (p_{s-1}, p_s)$ and $\omega_\gamma = p_s$.

Result 3: For any $i < s$,

$$0 = \frac{\partial J}{\partial N_i} \tag{4.25}$$

$$= \frac{\partial R}{\partial N_i} + \lambda \frac{\partial C}{\partial N_i} \tag{4.26}$$

$$= \left(\frac{1}{N_i + \frac{1}{\omega_\star}} - \frac{1}{N_i}\right) + \lambda \left(\frac{1}{\gamma} - \frac{1}{N_i}\right). \tag{4.27}$$

This equation must be satisfied by every $i < s$ for the same λ. Fixing $i = 1$, we can solve for λ as a function of N_1. Solving this equation for any $1 < i < s$ with the given

31

λ results in a quadratic equation whose two roots are $N_i = N_1$ and $N_i = \gamma - \frac{1}{\omega_\star} - N_1$. This means that for all local extrema, there are at most two levels less than N_s, namely N_1 and $\gamma - \frac{1}{\omega_\star} - N_1$. In other words, $s \leq 3$ for every local extremum.

Result 4: If $p_s < \pi$, then

$$0 = \frac{\partial J}{\partial p_s} \tag{4.28}$$

$$= \frac{\partial R}{\partial p_s} + \lambda \frac{\partial C}{\partial p_s} \tag{4.29}$$

$$= \lambda \frac{\partial C}{\partial p_s} \tag{4.30}$$

$$\neq 0. \tag{4.31}$$

This follows because $\omega_\star < p_s$ implies that $\frac{\partial R}{\partial p_s} = 0$. Furthermore, since $\omega_\gamma = p_s$ implies that $\gamma > N_s$, clearly $\frac{\partial C}{\partial p_s} > 0$. Thus every local extremum must satisfy $\omega_\gamma = p_s = p_{s+1} = \cdots = p_M = \pi$. In other words, there are at most three distinct steps (of non-zero width).

Finally, note that we can immediately eliminate the case where $\omega_\star$ is on the first step. If this were the case, then $s = 1$ and $S_{\tilde{n}\tilde{n}}$ is constant. The obvious water-pouring and uniform allocation solutions are identical: $\omega_\gamma = \omega_\star = \pi$, which contradicts the assumption that $\omega_\star$ is on the interior of a step.

Combining all points above, we have a relatively small set of functions satisfying the necessary conditions for local extremum when $\frac{\partial R}{\partial \omega_{bw}}\big|_{\omega_\star} = 0$ and $\omega_\star \in (p_{s-1}, p_s)$. We summarize the left-over possibilities in Figure 4-3. Note that the figure is not complete in that we have not considered the bounds (b, B). When we take these bounds into account, the resulting possibilities allow two extra steps with noise power densities equal to the bounds (b, B).
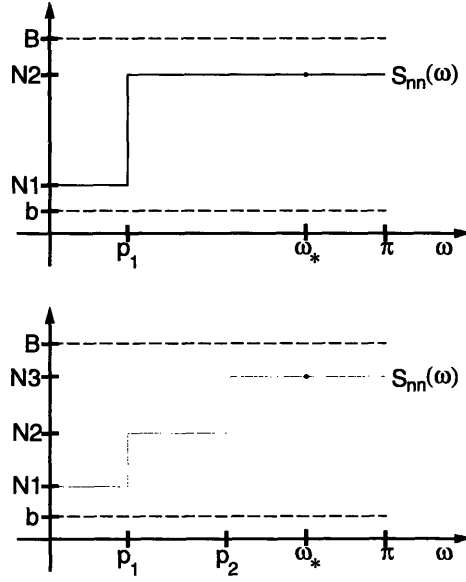
Figure 4-3: Local Extrema Possibilities for First Sub-Case

4.4.2 Sub-Case: Necessary Conditions When $\omega_\star = p_s$

We now consider the remaining possibility, where $\omega_\star$ is on the boundary of a step, $\omega_\star = p_s$. There are two possibilities for $\frac{\partial R}{\partial \omega_{bw}}$ at $\omega_\star$: either

$$\left.\frac{\partial R}{\partial \omega_{bw}}\right|_{\omega_\star -} = 0 \quad \text{and} \quad \left.\frac{\partial R}{\partial \omega_{bw}}\right|_{\omega_\star +} < 0, \qquad (4.32)$$

or

$$\left.\frac{\partial R}{\partial \omega_{bw}}\right|_{\omega_\star -} > 0 \quad \text{and} \quad \left.\frac{\partial R}{\partial \omega_{bw}}\right|_{\omega_\star +} \leq 0. \qquad (4.33)$$

Without loss of generality, we interpret $\omega_\star = p_s$ as meaning $\omega_\star$ is located at the right end of step s rather than the left end of step $s + 1$.

In both cases, Results 1–3 above from the interior point sub-case still apply. To summarize, then, the necessary conditions are $\omega_\gamma = p_s$ and $s \leq 3$. Since $\omega_\star$ is assumed to equal p_s in this sub-case, these necessary conditions imply $\omega_\gamma = \omega_\star$.

Additionally, note that if $\left.\frac{dR}{d\omega_{bw}}\right|_{\omega_\star -} = 0$ and $\left.\frac{dR}{d\omega_{bw}}\right|_{\omega_\star +} < 0$, then Result 4 also applies. Specifically, $\frac{\partial J}{\partial p_s} \neq 0$ unless $p_s = \pi$, and this implies that $\omega_\gamma = \pi$.
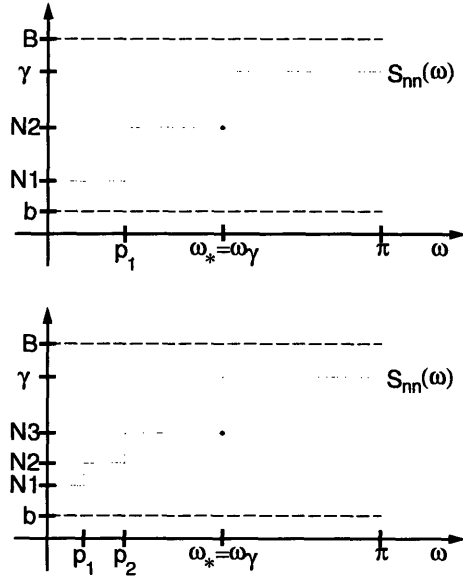
33

Figure 4-4: Local Extrema Possibilities for Second Sub-Case

Unfortunately, in the second case, the analysis leading to Result 4 does not apply. This is clear because increasing p_s necessarily would increase the rate R^* — after all, $\left(\frac{dR}{d\omega_{bw}}\right)_{\omega_*-} > 0$. Some simplification is still possible, however. Any number of steps to the right of ω_γ is equivalent to a single step since they affect neither R^* nor capacity. Furthermore, $\omega_\gamma = p_s$ implies that $N_{s+1} \geq \gamma$. Without any loss of generality, we need only consider the case $N_{s+1} = N_{s+2} = \cdots = \gamma$.

In parallel with the discussion at the end of Section 4.4.1, we can immediately eliminate the case $\omega_* = p_1$. This is in fact a local extremum, but it is a maximum of $R(\mathbf{N}, \mathbf{p}, \omega_*)$ not a minimum: $R(\mathbf{N}, \mathbf{p}, \omega_*) = C(\mathbf{N}, \mathbf{p})$.

Thus, the picture does not change much from the case where $\omega_* \in (p_{s-1}, p_s)$. All additional extremal possibilities are summarized in Figure 4-4. Again, if the bounds (b, B) are active, we must insert additional steps at b and B.

Remarkably, regardless of the value of M (and taking the bounds into account), no more than six steps are required to minimize the uniform allocation rate for a fixed capacity. This will prove essential in the analysis that follows. The results of a partial numerical search over the possibilities is given in Chapter 6.

Chapter 5

Continuity Solves the Bounded Problem

We now show how to use the general M-step problem to solve the bounded problem (4.2). We show that any noise spectrum $S_{nn} \in \mathcal{A}(C, b, B)$ considered in the bounded problem may be approximated by a piecewise constant function $S_{\tilde{n}\tilde{n}}$ with a finite number of steps M. We then show that the capacity achieved with $S_{\tilde{n}\tilde{n}}$ is close to $C(S_{nn})$. In the previous chapter we found that the minimum achievable rate for a step function of M steps achieving a fixed capacity does not depend on M (for $M > 5$) — $W_M(C) = W_6(C)$ for all $M > 5$. We use this result to prove that the noise spectrum which solves the bounded problem is simply that which solves $W_6(C)$.

5.1 Continuity of M-step Approximations

Given a non-decreasing noise function S_{nn} bounded above by B, we can find a series of piecewise constant, non-decreasing functions S_i, each with a finite number of steps, such that $B \geq S_1(\omega) \geq S_2(\omega) \geq \cdots \geq S_{nn}(\omega) \geq b \ \forall \omega \in [0, \pi]$ and such that $S_i(\omega) \to S_{nn}(\omega)$ uniformly [1].

We rewrite this in a useful form for our continuity argument. Given any S_{nn} and $\epsilon > 0$, we can find a piecewise constant function S_i with $M(\epsilon)$ or fewer steps, such that $B \geq S_i(\omega) \geq S_{nn}(\omega)$ and $\|S_i - S_{nn}\|_\infty < \epsilon$.

5.2 Continuity of Capacity

Given S_{nn} and any other noise power spectral density $S_i \geq S_{nn}$ with $\|S_i - S_{nn}\|_\infty < \epsilon$, we can bound $|C(S_{nn}) - C(S_i)|$. Using (3.7) for capacity,

$$C(S_i) = \int_0^{\omega_{\gamma_i}} \ln\left(\frac{\gamma_i}{S_i(\omega)}\right) d\omega. \tag{5.1}$$

Since $S_i \geq S_{nn}$, $\gamma_i \geq \gamma$ and $\omega_{\gamma_i} \geq \omega_\gamma$. Thus

$$C(S_i) \geq \int_0^{\omega_\gamma} \ln\left(\frac{\gamma}{S_i(\omega)}\right) d\omega. \tag{5.2}$$

Now, $S_i \leq S_{nn} + \epsilon$ implies

$$C(S_i) \geq \int_0^{\omega_\gamma} \ln\left(\frac{\gamma}{S_{nn}(\omega) + \epsilon}\right) d\omega. \tag{5.3}$$

Since $\ln\left(\frac{\alpha}{\beta + x}\right)$ is a convex function of x for positive α and β,

$$C(S_i) \geq \int_0^{\omega_\gamma} \ln\left(\frac{\gamma}{S_{nn}(\omega)}\right) - \frac{1}{\gamma S_{nn}(\omega)} \epsilon \, d\omega. \tag{5.4}$$

Finally, since $S_{nn} \geq b$, and thus $\gamma \geq b$ as well,

$$C(S_i) \geq \int_0^{\omega_\gamma} \ln\left(\frac{\gamma}{S_{nn}(\omega)}\right) - \frac{1}{b^2} \epsilon \, d\omega \tag{5.5}$$

$$\geq C(S_{nn}) - \frac{\pi}{b^2} \epsilon. \tag{5.6}$$

We thus have the desired continuity of capacity in the L^∞ sense.

5.3 Continuity of M-Step Sub-Problem

Recall that we defined the following problem in Chapter 4:

$$W_M(C) = \min_{S_{\tilde{n}\tilde{n}} \in \mathcal{K}(C,M)} R^\star(S_{nn}). \tag{5.7}$$

36

In that chapter we found that the minimizing noise spectrum has no more than six steps, which implies that $W_M(C)$ is constant in M for $M \geq 6$. In other words, we can define

$$W_F(C) \quad = \quad \min_{S_{\tilde{n}\tilde{n}} \in \cup_{M=1}^{\infty} \mathcal{K}(C,M)} R^{\star}(S_{\tilde{n}\tilde{n}}). \tag{5.8}$$

$W_F(C)$ is the minimum $R^{\star}(S_{\tilde{n}\tilde{n}})$ over all noise power spectral densities with capacity C and a finite number of steps. Then $W_F(C) = W_6(C) = W_M(C)$ for all $M \geq 6$.

Furthermore, we assert that $W_F(C)$ is a continuous, strictly increasing function of capacity. Both results are critical to proving that $W_F(C)$ solves the bounded problem.

5.4 Simplifying the Sub-Problem via Continuity

Consider any noise spectrum S_{nn} which achieves a water-pouring capacity $C(S_{nn})$ and a uniform allocation rate $R^{\star}(S_{nn})$. We will prove, using the continuity results above, that we can find an upper bound on $C(S_{nn}) - R^{\star}(S_{nn})$ that depends only on $C(S_{nn})$. Furthermore, we achieve this upper bound using the spectrum that solves $W_F(C(S_{nn}))$.

Fix $\chi > 0$. By the continuity of $W_F(C)$, we can find an $\epsilon > 0$ such that

$$W_F(C(S_{nn}) + \frac{\pi}{b^2}\epsilon) - W_F(C(S_{nn})) < \chi. \tag{5.9}$$

Using the continuity of step-function approximations to S_{nn}, we can find an M-step noise spectrum $S_{\tilde{n}\tilde{n}} \geq S_{nn}$ such that

$$\|S_{\tilde{n}\tilde{n}} - S_{nn}\|_{\infty} < \epsilon. \tag{5.10}$$

By continuity of the capacity function established in (5.6), this implies

$$C(S_{\tilde{n}\tilde{n}}) > C(S_{nn}) - \frac{\pi}{b^2}\epsilon. \tag{5.11}$$

Returning to the original bounded problem (4.2),

$$C(S_{nn}) - R^\star(S_{nn}) \;=\; (C(S_{nn}) - R^\star(S_{\tilde{n}\tilde{n}})) + (R^\star(S_{\tilde{n}\tilde{n}}) - R^\star(S_{nn})). \quad (5.12)$$

Since $S_{\tilde{n}\tilde{n}} \geq S_{nn}$, the second term in (5.12) is negative, and thus

$$C(S_{nn}) - R^\star(S_{nn}) \;\leq\; C(S_{nn}) - R^\star(S_{\tilde{n}\tilde{n}}). \quad (5.13)$$

By definition of the function W_F,

$$C(S_{nn}) - R^\star(S_{nn}) \;\leq\; C(S_{nn}) - W_F(C(S_{\tilde{n}\tilde{n}})). \quad (5.14)$$

Combining the monotonicity of $W_F(C)$ in capacity with (5.11), we find

$$C(S_{nn}) - R^\star(S_{nn}) \;\leq\; C(S_{nn}) - W_F(C(S_{nn}) - \frac{\pi}{b^2}\epsilon). \quad (5.15)$$

Our choice of ϵ in (5.9) ensures that

$$C(S_{nn}) - R^\star(S_{nn}) \;\leq\; C(S_{nn}) - W_M(C(S_{nn})) + \chi. \quad (5.16)$$

But $\chi > 0$ was arbitrary and $W_M = W_6$. Thus, for any S_{nn},

$$C(S_{nn}) - R^\star(S_{nn}) \;\leq\; C(S_{nn}) - W_6(C(S_{nn})). \quad (5.17)$$

Since S_{nn} is an arbitrary member of $\mathcal{A}(C(S_{nn}), b, B)$, we can solve the original bounded problem for any fixed capacity simply by determining $W_6(C)$ for all C. We again emphasize that this result follows only because $W_M = W_6$ for all $M \geq 6..$

It is interesting to note that we do not need to prove the achievability of a solution in the general bounded optimization problem. Indeed, (5.17) means

$$W_F(C) = \sup_{S_{nn} \in \mathcal{A}(C,b,B)} C(S_{nn}) - R^\star(S_{nn}). \quad (5.18)$$

However, since $W_F(C)$ is achievable with a 6-step solution S_6, and since $S_6 \in \mathcal{A}(C, b, B)$, the supremum is achieved. The supremum in (5.18) thus becomes a maximum. This last fact is not of great importance, however, because the point of this research is simply to find a tight lower bound on the rate loss.

Chapter 6

Numerical Wrap-Up

6.1   Solution Method

In Chapters 4 and 5, we reduced the solution of the bounded problem to the search for the worst-case M-step noise function satisfying the necessary conditions sketched in Figures 4-3 and 4-4 (recall that we must also include the possiblity that the noise functions have two extra steps equal to the bounds (b, B)).

Based on additional perturbation arguments not developed here, we believe that, for sufficiently small b and large B, the worst-case noise spectrum has only two steps for all C. We searched numerically over this class; as expected, $W_2(C)$ appears to be a continuous, strictly increasing function of capacity.

6.2   Results

We find that, to the limits of numerical precision, the optimizing noise spectrum does not track the bounds as b and B are loosened toward 0 and ∞. We presume, then, that we have actually solved the original (unbounded) optimization problem.

We present our results in the form of two graphs. Refer to Figure 6-1. The top graph plots the fractional rate loss due to uniform power allocation. The bottom graph plots the absolute rate loss. Both are shown with capacity on a logarithmic axis to better illustrate asymptotic behavior.
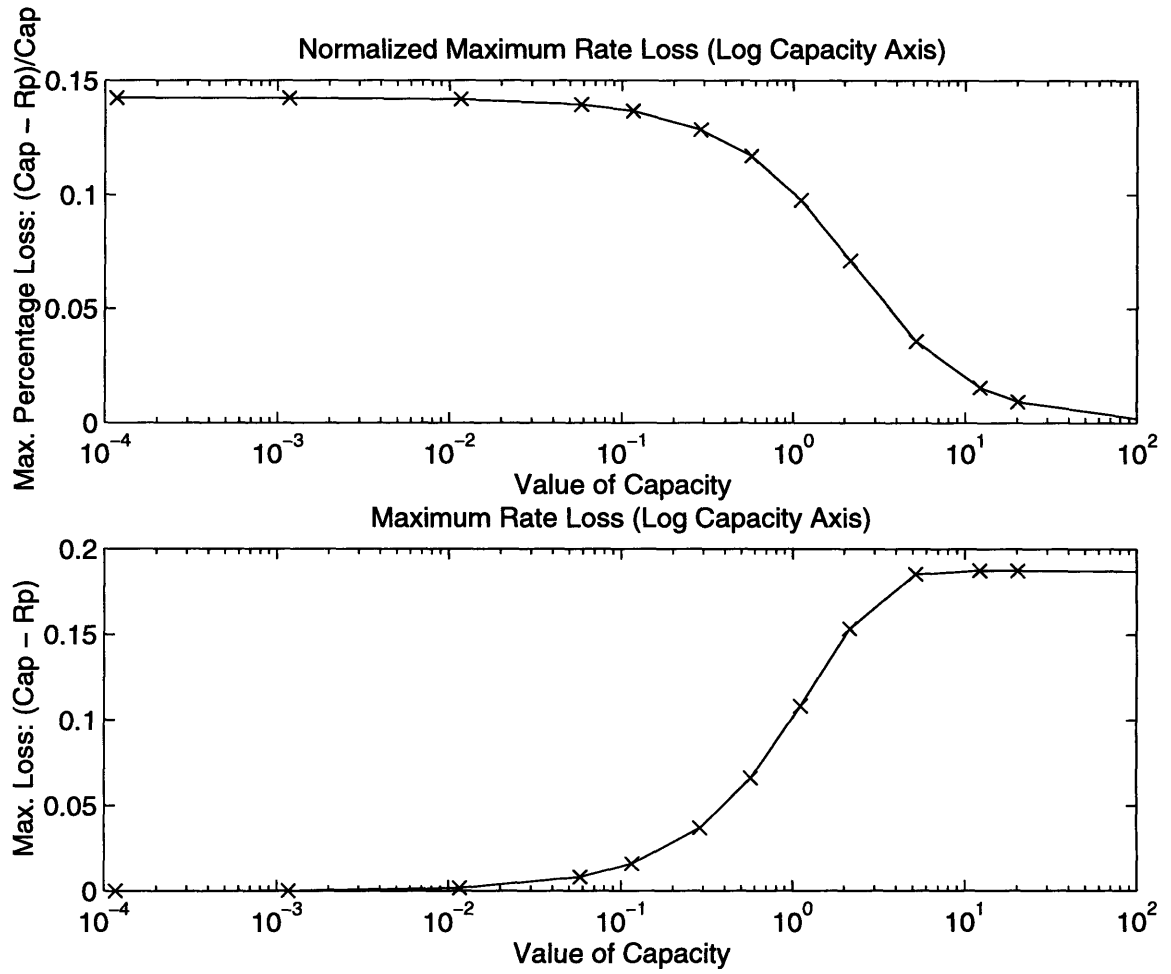
Figure 6-1: Results

6.3 Discussion

Recall that the results presented are calculated for a discrete-time, real, additive Gaussian noise channel. A loss of L bits per channel use corresponds to a loss of $\frac{1}{2}L$ bits/sec/Hz in the passband channel.

We see that the bound on the normalized rate loss reaches a maximum asymptotic value around 14 percent for low-capacity channels. The fractional loss decreases monotonically as capacity increases.

On the other hand, the absolute rate loss reaches a maximum asymptotic value around 0.1874 bits per channel use for very high-capacity channels. The absolute loss increases monotonically with capacity.

Since the absolute loss is bounded by a small number, we see that the issue of precise power allocation is of more importance in moderate- to low-capacity cases. The fractional loss is in excess of 7 percent for all rates below 4 bits/sec/Hz.

Note that we not only bound the sub-optimal capacity loss, but we also characterize particularly bad (worst-case) channels. One might reasonably wonder if the worst-case noise spectrum can arise in practice. Based on limited numerical results, the ratio of highest to lowest noise power density for the optimized spectrum is modest.

6.4 Future Work

Some steps in the derivation of the worst-case capacity loss have not been rigorously addressed. Specifically,

1. $\left.\frac{\partial^2 R}{\partial \omega_{\mathrm{bw}}^2}\right|_{\omega_\star} = 0$, which simplifies the Lagrange multiplier approach.

2. The monotonicity of $R(\mathbf{N}, \mathbf{p}, \omega_{\mathrm{bw}})$ with respect to ω_{bw}.

3. The continuity of $R^\star(\mathbf{N}, \mathbf{p})$ over the bounded space of $(\mathbf{N}, \mathbf{p})$.

4. The regular point issue in the Lagrange multiplier approach.

5. The continuity of $W_F(C) = W_6(C)$ in capacity.

6. The rigorous justification that $W_6(C)$ is just $W_2(C)$.

Regarding Item 1, if we can show that $\frac{\partial^2 R}{\partial \omega_{bw}{}^2}\Big|_{\omega_*} < 0$ for all noise spectra, then we believe that Items 2 and 3 will easily follow. We do not believe that any of 1–6 pose a threat to the conclusion, save perhaps 6 (which could result in a small change in the numerical results for $W_F(C)$). It may be possible to derive a close-form solution to $W_2(C)$ by taking partial derivatives with respect to the step locations p_i, but the expressions are discouragingly complex.

The results of this thesis may be applied to a particular time-varying memoryless channel for which the transmitter is permitted to either use the channel at full power or remain silent at each step. Regardless of the marginal fading statistics, if the transmitter has appropriate information about the channel, the maximum rate loss is again 14 percent. Interestingly, the worst-case channel has only two states; in the less favorable state, the transmitter should adopt a probabilistic strategy which uses the channel a fraction of the time (corresponding to $\frac{\omega_* - p_1}{p_2 - p_1}$).

A more general framework for this research would address the loss in performance due to partial or incorrect knowledge of the channel at the transmitter. This might arise from imperfect channel measurement, limited feedback from the receiver, or time-variation in the channel.

Bibliography

[1] Malcolm Adams and Victor Guillemin. *Measure Theory and Probability*. Birkhäuser, Boston, second edition, 1996.

[2] J. Barry, E. A. Lee, and D. Messerschmitt. Capacity penalty due to ideal zero-forcing decision-feedback equalization. In *Proceedings of ICC '93 — IEEE International Conference on Communications*, pages 422–427. IEEE Communications Society, May 1993.

[3] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, New York, 1991.

[4] Robert G. Gallager. *Information Theory and Reliable Communication*. Wiley, New York, 1968.

[5] J. L. Holsinger. Digital communication over fixed time-continuous channels with memory, with special application to telephone channels. Technical Report 430, M.I.T. Research Lab. of Electronics, 1964.

[6] Shunsuke Ihara. *Information Theory for Continuous Systems*. World Scientific, Singapore, 1993.

[7] Edward A. Lee and David G. Messerschmitt. *Digital Communications*. Kluwer, Massachusetts, second edition, 1994.

[8] M. S. Pinsker. *Calculation and Estimation of the Quantity of Information, The Capacity of a Channel, and the Rate of Production of Information, in terms of the Second Moments of the Distributions*. PhD thesis, University of Moscow, 1957.

[9] Claude E. Shannon. A mathematical theory of communication. *Bell System Tech. Journal*, 27, 1948.

[10] Elias M. Stein. *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press, New Jersey, first edition, 1970.