Optical distribution networks:
signal-to-noise ratio optimization
and
distributed erbium-doped fiber amplifiers

by

Shayan Mookherjea

B.S. California Institute of Technology (1999)

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

June 2000

Author .....
Department of Electrical Engineering and Computer Science
March 1, 2000

Certified by .....
Vincent W.S. Chan
Joan & Irwin M. Jacobs Professor
Thesis Supervisor

Accepted by ..
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Optical distribution networks: signal-to-noise ratio optimization and distributed erbium-doped fiber amplifiers

by

Shayan Mookherjea

Submitted to the Department of Electrical Engineering and Computer Science
on March 1, 2000, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering and Computer Science

Abstract

Two important issues related to the use of remotely-pumped erbium-doped fiber amplifiers in optical distribution networks are analyzed. We investigate the optimal allocation of gain and loss among the stages of a lumped amplifier chain, using the concepts of a Nash solution and Pareto optimality. The propagation of optical channels along an erbium-doped fiber is derived from basic physical considerations, and compared with the well-known Desurvire model. We demonstrate a simple method of constructing bus distribution networks, but this approach is particularly sensitive to the numerical values of the parameters. A second approach is then discussed, extending the analysis of Sun et al. from ab-initio principles to model the effect of detectors (users) along a distribution network. Theoretical closed-form results indicate that the number of optical receivers that can be supported using this scheme is at least two orders of magnitude higher than without optical amplification. Finally, we analyze the effect of dynamic perturbations in the power that is extracted at the receivers. Though our focus is on bus networks, we also discuss tree distribution networks as extensions of the basic models. Our results have implications on the architecture of optical distribution networks.

Thesis Supervisor: Vincent W.S. Chan
Title: Joan & Irwin M. Jacobs Professor

This work has been funded by DARPA under the ONRAMP Consortium with partners: MIT, MIT/LL, AT&T, Cabletron and JDS/Uniphase.

Optical Distribution Networks:

SNR Optimization and Distributed EDFAs

Shayan Mookherjea

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, June 2000*

It was like that then, the island, thought Cam,
once more drawing her fingers through the waves.
She had never seen it from out at sea before.
It lay like that on the sea, did it, with a dent in the middle
and two sharp crags, and the sea swept in there,
and spread away for miles and miles on either side of the island.
It was very small; shaped something like a leaf stood on end.
So we took a little boat, she thought, beginning to tell herself
a story of adventure about escaping from a sinking ship.
But with the sea streaming through her fingers,
a spray of seaweed vanishing behind them,
she did not want to tell herself seriously a story;
it was the sense of adventure and escape that she wanted,
for she was thinking, as the boat sailed on,
how her father's anger about the point of the compass,
James's obstinacy about the compact, and her own anguish,
all had slipped, all had passed, all had streamed away.
What then came next? Where were they going?

—V. WOOLF, To the Lighthouse (1927)

Acknowledgements

The framework for the topics explored in this thesis arose from discussions with my thesis advisor, Professor Vincent W.S. Chan, Joan and Irwin Jacobs Professor at MIT. I am also thankful to him and the Electrical Engineering & Computer Science department at MIT for financial support as a Research Assistant during the academic year 1999–2000. Dr. Stephen R. Chinn at the MIT Lincoln Laboratory provided numerous helpful suggestions.

Fellow students and the faculty and staff of the Laboratory for Information and Decision Systems (LIDS) contributed to making this experience a memorable one for me. For their friendship and advice, my sincere thanks.

More so than any other person, Professor Amnon Yariv, Summerfield Professor of Applied Physics at Caltech, encouraged and supported my learning and guided my research for the past three years—I am deeply indebted to him for unsparing advice and invaluable friendship. It's a pleasure and an honor to acknowledge those teachers in whose classes I learnt optics, optical electronics and guided waves: Prof. Demetri Psaltis and Prof. William B. Bridges of Caltech.

Professors Hermann Haus and Erich Ippen generously encouraged me to participate in their research group meetings at the MIT Research Laboratory of Electronics enabling me to gain further insight into current research in this exciting field.

This work is dedicated, with love, to my parents, who through their lives have given and taught me far more than I could have ever wished.

Shayan Mookherjea

Cambridge, MA

Contents

Chapter 1

Introduction

A second class of applications, no less important, is that of distribution systems with a very large number of subscribers. The use of optical amplifiers makes it possible to maintain the power arriving at a subscriber's premises at sufficiently high levels so as not to be degraded by the receiver noise. The number of subscribers that can thus be served by a single laser can be increased by anywhere from 1 to 3 orders of magnitude.

—A. YARIV, Optical Electronics in Modern Communications (1997)

The impact of erbium-doped fiber amplifiers (EDFAs) on lightwave communications has been perhaps the most unprecedented, unexpected and significant technological achievement of this decade. While single-mode silica glass fiber is far less lossy compared to most other optically transparent media, the signal absorption coefficient is still significant enough to prevent signal transmission for communications purposes over more than a hundred or so kilometers: hardly enough to traverse the width of the country. As noted by Desurvire [6], the performance of communications systems (e.g. in Gbits/s-km) has been growing by a factor of ten every four years or so over the last two decades—which is no small achievement considering the degree of maturity in electromagnetics (e.g. radar) and communication theory already achieved.

The new technology has quickly found applications in a variety of formats. For example, the degree of flexibility offered by wavelength-division multiplexed (WDM) systems in terms of bandwidth is truly unprecedented, and solitons have demonstrated error-free and penalty-free propagation for over a million kilometers. Com-

munications engineers often take the properties of erbium-doped fiber for granted in designing networks. With maturing technology and increased applications, however, we come across problems where a textbook approach isn't quite feasible, or when a "traditional" layered approach to network design is far less robust. One such problem, we feel, is that faced by designers of distribution networks.

The majority of the backbone long-haul networks in the US and in several other countries use optical fiber as the transmission medium, and EDFAs play a major role in extending the span of such channels. By aggregating traffic from a number of sources, it's commercially feasible for network managers to provide careful control of these few strands of fiber that traverse long distances, often along carefully demarcated paths, and sheltered from undesirable conditions (e.g. temperature fluctuations that degrade the EDFA pump sources). The problem with distribution networks is that, by definition, they serve a large number of users in a relatively short geographic span. There are many more access points to the fiber, and far higher need for robust equipment that is simple to maintain. Furthermore, a distribution network must use its resources as efficiently as possible—one cannot aggregate the data traffic from a household, for instance, to fill a fiber to capacity; nevertheless, the household cannot be served with *less* than one fiber leading to its doorstep.

What *can* be shared in such a situation is the gain that is added to the optical signals for long-distance propagation: a number of households (more generally, let's call them "users") can be served on different WDM channels which are amplified simultaneously at one EDFA. The cost of the EDFA maintenance is then shared among these users.

In typical distribution networks, and more so in commercial neighborhoods where real estate is at a premium, it's still not economically viable to set up little huts at intervals along the fiber route to house the EDFAs and equip these huts with temperature control and monitoring equipment. The component of EDFAs that requires most attention is typically the (semiconductor laser diode) pump. The gain processes

take place by the interaction of the pump and the signal beams in a specially doped section of fiber. In the absence of a pump beam (which is at a specific wavelength), the signal beams view this doped section as just another length of fiber. It's only when a pump beam is present that the induced gain overcomes the natural loss due to absorption.

The common EDFA package integrates the pump diode and the specially doped section (erbium is the doping element), but there is no physical barrier to isolating these two elements and using a section of undoped fiber to connect them. In fact, the same section of signal-carrying single mode fiber (SMF) that couples to the erbium-doped fiber (EDF) can be used to feed in the pump beam as well. This is called a remotely-pumped amplifier, and vastly simplifies network design. The pump sources can be conveniently located at the start of the distribution network, and sections of EDF inserted along the length of the EDF where appropriate. Turning on the pump diodes (e.g. at a network operations building) will result in gain at all the EDF sections along the fiber, just as if an individual EDFA were installed at each EDF location. Adjusting and monitoring the pump sources, not only from a maintenance perspective but also to effectively manage the traffic flow in the network, is now less of a problem.

It's not obvious that such a scheme will, in fact, work. The pump will itself be absorbed by the SMF before it can reach the EDF sections to provide gain to the signal. We need to investigate what fraction of the pump is unused after the first stage: is it reasonable to expect more than one stage of remote amplification? And how many users can be supported this way? The answers to these questions form this report.

At the end of this chapter, we show that a passive bus distribution network cannot support more than a handful of users. The rest of our discussion will analyze a simple way to enhance this number by more than two orders of magnitude.

We begin our analysis by identifying the main constraint on the span of an am-

plified bus distribution network: spontaneous emission noise that is inherent in the process of amplification itself. The basic principles are well-known, and the major portion of our discussion of noise in optical amplifier chains focuses on solving an optimization problem from a mathematical perspective: how do we allocate the gains among the various stages of an amplifier chain to maximize the signal-to-noise ratio of the final stage? This is equivalent to extending the span of the network as much as possible by reallocating gains and losses among the stages. The problem has been solved under various conditions in the literature, and most textbooks [30, 6] mention it. But in order to keep the discussions as general as possible (and fun!), and yet lose no rigor in our analysis, we use game theory—specifically the remarkable theorems of Nash with regard to cooperative games—to present a new approach to solving this basic problem.

Nevertheless, that's not our goal—we'll show in the remainder of the report that the optimum SNR solution is but one step in maximizing the number of users. We demonstrate a method for constructing such bus networks, but the numerical examples in that section will clearly show that it's difficult to apply optimization arguments to this method. In fact, some of its basic assumptions might be questioned, and so we present a second approach, based on a first-principles derivation of the (rate) equations that govern the entire system.

The reader who wishes to preview the results may read through the Conclusion before the intervening chapters. We believe that the game-theory approach to optimizing SNR, the use of a spectral density function to model the receivers along the network, and the perturbation theory results of the tap fraction are all original contributions to the field. Needless to say, we're already considering extensions to the models described here.

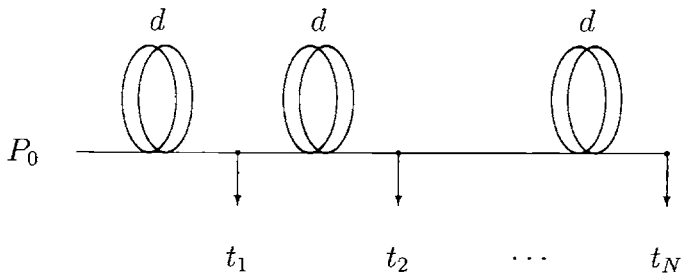

Figure 1-1: A passive distribution bus consisting of M repeated sections. Section $i = 1, 2, \ldots, M$ is characterized by a tap t_i and preceded by a section of single mode fiber, SMF of length d.

1.1 Passive distribution bus

Since our goal is to highlight and quantify the advantages that amplification can bring to optical distribution bus networks, we'll first discuss the passive distribution bus. We assume that the input signal power is P_0, and the individual users are situated along this passive fiber at intervals of length d.

As the signal propagates down the fiber, it is attenuated (absorption coefficient α_s'). At each receiver k, a fraction t_k of the signal power is coupled out of the fiber for detection. We call this a 'tap' and t_k denotes the tap fraction of the k^{th} receiver.

When we are free to set the tap fractions as we wish, each receiver taps no more power than is necessary to perform (relatively) error-free demodulation of the optical carrier. We describe this in more detail in the next chapter: for our purposes here, we assume that this translates into a minimum detectable power P_{min}, which is constant for all receivers.

From Figure 1-1, the first tap fraction is

$$t_1 = \frac{P_{min}}{P_0\, e^{-\alpha_s' d}} \tag{1.1}$$

The remaining signal power $(1-t_1)e^{-\alpha'_s d}P_0$ propagates a further distance d before encountering the second tap, which is set according to

$$t_2 = \frac{P_{min}}{(1-t_1)\,P_0\,e^{-2\alpha'_s d}} \tag{1.2}$$

It's easy to see that there's a simple relationship between any two successive tap fractions

$$t_{k+1} = \frac{t_k}{(1-t_k)e^{-\alpha'_s d}} \tag{1.3}$$

Since these are, after all, tap *fractions*, they cannot exceed 1, and we reach the limit on M, the number of receivers that can be supported, when

$$t_M = \frac{P_{min}/P_0}{\left(\prod_{k=1}^{M-1}(1-t_k)\right)\,e^{-M\alpha'_s d}} \geq 1 \tag{1.4}$$

and all previous $t_k < 1$.

As typical numerical values, consider $\alpha'_s = 0.5$ dB/km for a signal channel near 1550 nm in typical single-mode fiber (SMF). Let the input signal power P_0 be -5 dBm and the minimum detection threshold (receiver sensitivity) P_{min} be -35 dBm. We must also account for the finite losses at each tap beyond the tap fraction (e.g. coupling between sections of fiber). For example, if we assume a further loss of $T = -0.5$ dB at each tap, which is about what can we expect with present day technology, then [1.4] becomes

$$t_M = \frac{P_{min}/P_0}{\left(\prod_{k=1}^{M-1}(1-t_k)\right)\,T^{M-1}\,e^{-M\alpha'_s d}} \geq 1 \tag{1.5}$$

In this case, for $d = 50$ m, we have $M = 30$ whereas for $d = 10$ m, $M = 38$.

A distribution network that can support only 40 users has limited applications.

But we will see that it's possible to substantially increase the number of receivers that can be supported when we use amplification to counteract the propagation loss and tap fractions. Our discussion will show that this amplification can be achieved in a way that makes it technologically feasible to implement amplified distribution networks: by physically separating the elements that consume power and are susceptible to degradation (the pump diodes) from the location along the network where the process of amplification takes place (a section of erbium-doped fiber).

Chapter 2

Detection of optical signals

Occasional use has been made of certain gambling terms,
particularly with reference to poker.
While it seems safe to assume a passing familiarity with games such as chess,
I am assured that it is unreasonable to expect universal acquaintance
with the rules and terminology of poker.
Admittedly, this shook my preconceptions to the core,
but I will endeavour to sketch the general idea!
—A.J. JONES, Game Theory: Mathematical Models of Conflict (1980)

One of the most fundamental limitations on optical communication is imposed by the accumulation of noise from the active elements (lasers, amplifiers, modulators) along the light path. We consider a typical communication model where the channel is described by an optical amplifier chain—alternating sections of optical amplifiers and passive optical fiber, with a signal source (laser) at the transmitter end and a detector (e.g. avalanche photo-diode) at the receiver. Electromagnetic waves are attenuated in propagation through real media (such as optical fiber), and to compensate for this loss, we must periodically amplify the signal along the path.

As will be discussed in the next chapter, this process of amplification adds noise—spontaneous emission noise—at each amplifier stage. Along an amplifier chain, the noise added by a particular amplifier is itself amplified by the remaining amplifiers further along the chain. This phenomena of amplified spontaneous emission (ASE) results in a decrease of the signal-to-noise ratio (SNR) with each increasing stage. At a typical detector, a bit-error-rate (BER) threshold translates into a lower bound on

the SNR, and therefore, the length of such a optical amplifier chain is bounded, even in the absence of nonlinearities, by ASE accumulation.

In the first section, we describe the process of detection of an optical signal using a semiclassical model. The question of optimizing the gain of each amplifier along such a chain so that we maximize the SNR is addressed in some detail in Section 2.2. We show that under ideal conditions, the best performance is achieved for a single gain stage, placed as far as possible from the detector. Of course, in any real scenario, the gain of a single stage will be limited, and a sub-optimal implementation may be necessary.

Using erbium-doped fibers both as an amplifying and a transmission medium, it is possible to characterize the entire channel as a single gain element. Section 2.3 shows that the noise performance of such a setup is worse than that of the one described earlier. Nevertheless, this model will be of considerable importance to us in later chapters.

2.1 Fundamentals of detection

The noise processes in optical communications have been studied extensively for decades. For those who have some familiarity with this topic, we'd like to point out at the start what we will *not* be concerned with in this report.

We won't refer to nonclassical (e.g. squeezed) light, though the noise properties of such states are often very attractive [24]. Further, we'll consider only the simplest form of optical amplifier—the traveling-wave (TW) optical amplifier model. Parametric and wave-mixing amplifiers are not discussed [12].

While we do mention, and account for, the noise induced by non-ideal detection, the bulk of our argument will focus on the most fundamental (classical) sources of noise in optical communication systems: shot noise and amplified spontaneous emission. What we're after is not a complete description of the noise generating processes themselves, but rather an understanding of the way these processes interfere with

communications. The starting point for the major portion of our discussion will the concept of a signal-to-noise ratio (SNR).

In discussing the basic physics of noise in optical amplification, there are several levels of detail: we'll restrict ourselves to a semiclassical model and a simple derivation of the signal-to-noise ratio in an optical amplifier chain. Our goal is to justify, as simply as we can, the definition [2.7] as a physically-intuitive figure of merit for our optimization problem. For a detailed quantum-mechanical treatment, see [15] or [3].

Our discussion in this section is limited to introducing the terminology and physical concepts that we'll use in this chapter. For a more complete discussion, the reader is referred to the many excellent books that cover this subject in detail [29, 30, 6]. The reader who is willing to take [2.7] on faith may skip the following paragraphs.

The concept of a signal to noise ratio arises only in the context of a receiver i.e. when the electromagnetic waves are converted to electrical impulses for signal processing. Therefore, let's begin our survey by describing a standard detection mechanism for optical signals.

2.1.1 Avalanche photodiode

An optical fiber link will typically terminate in a detector of some sort, and the prototypical model is that of an avalanche photodiode (APD). An APD is a diode sensitive to electromagnetic radiation in the optical frequency range that is operated under reverse bias near its avalanche breakdown point. An optical field incident on an APD will generate pairs of holes and electrons in the semiconductor material, which drift towards opposite terminals of the device. In the region of the p-n junction, these carriers generate additional carriers by impact ionization. This chain reaction leads to a greater current flow than in a normal photodiode, and the multiplicative factor g is called the avalanche gain of the photodiode.

Since impact ionization is a random process, we usually deal with the mean value of the avalanche gain $E\{g\} = \bar{g}$, and its second moment $E\{g^2\}$. A common approxi-

mation, and one we'll use in our discussion, is to write

$$E\{g^2\} = E\{g\}^{2+x} \tag{2.1}$$

where x is called the excess noise factor exponent of the APD and is typically a small positive number (e.g. 10^{-1}). We'll use this approximation in our expressions for a particular contribution to the overall noise—that due to shot noise.

From an electronic circuits perspective, an APD is typically specified as an ideal current source together with a bias resistor and a lumped capacitor to represent the junction capacitance of the device. The current source in this case depends on the incident optical power, and is non-zero even when there is no light incident on the device. This residual contribution is called the 'dark current' of the APD.

As discussed in [30], if an optical signal of power P, is incident on a photodiode, the mean-square value of the resulting current at the output of the photodetector is

$$< {i_s}^2 >= \left(\frac{\bar{g} P e \eta}{h\nu} \right)^2 \tag{2.2}$$

2.1.2 Detection noise sources

For electrical engineers, the concept of shot noise is a familiar one: it arises from the random nature of the processes of generation and flow of quanta of charge (electrons) in a semiconductor. A similar concept holds for quanta of light (photons) as well. The rate of incidence of photons on the surface of a detector approximately follows Poisson statistics. An APD can be thought of as counting these incident photons, and it's easy to see that the sample paths of such a counting process are characterized by discontinuous jumps at the (random) arrival times of the photons. The resultant current at the output of the APD has a power-spectral density associated with this phenomena, and its integral over a relevant bandwidth (e.g. B_e, the electronic bandwidth of the detector) yields a term that has the dimensions of power. This term,

called shot noise, is modeled in circuit terms by an equivalent noise generator, or a source of alternating currents at a particular frequency.

Johnson (Nyquist, thermal) noise is used to describe voltage fluctuations across a dissipative circuit element, e.g. the overall impedance of the detector, which has a non-zero real part. The random motion of charge carriers sets up local charge gradients and an AC voltage. The power dissipation of this voltage is modeled by a voltage source in series with or a current source parallel to a resistor.

2.1.3 Spontaneous emission

This is an important source of noise in laser oscillators and amplifiers. This is mainly an effect of the random de-excitation of the laser ions, which is a random process, and is most consistently understood from quantum mechanics [29, 9].

Excited ions in the upper level of a laser medium are characterized by a finite excited state lifetime (e.g. $\tau = 10$ ms for erbium-doped glass [6], and $\tau = 3$ ms for a ruby laser [29]). When these ions spontaneously transition to the ground state, they emit a photon. Unlike the case of stimulated transitions, these randomly emitted photons have no coherence relationships with the incoming signal light. The collection of these incoherent photons, themselves amplified by the remaining section of amplifier and by successive amplifiers, forms an interfering optical signal which we call amplified spontaneous emission (ASE) noise.

There are at least two prominent effects of spontaneous emission: the spectral broadening of the laser output (the Schawlow-Townes linewidth [29]) and a fundamental limit on the signal-to-noise ratio achievable at the output of optical amplifiers. We'll ignore the former, since in high bit-rate systems, the bandwidth of the optical pulses typically dominate the source laser linewidth [18].

At the output of an optical (traveling-wave type) laser amplifier of gain G, the

ASE power in a single mode of spectral bandwidth B_o is [30]

$$\mathcal{N} = \mu\, h\nu\, B_o(G-1) \tag{2.3}$$

where μ is the atomic inversion factor.

We note that the expression for μ can be a complicated one: indeed, the system of [5.34], [5.36] and [5.39] is in integral form, but we shall make the assumption of constant μ for high input pump power (which is of practical interest). Furthermore, for the purposes of this chapter, the important property is an obvious one: that μ is strictly greater than 1.

Consider a chain of N optical amplifiers, each characterized by gain G_i, where the input signal power to the chain P_0. The first amplifier increases the signal power level to G_1P_0 and also adds $\mathcal{N}_1$ in terms of power to the propagating electromagnetic fields. The sum of G_1P_0 and $\mathcal{N}_1$ forms the input to the second amplifier: not only will this stage add its own noise $\mathcal{N}_2$, it will also amplify the noise of the preceding stage, $G_2\mathcal{N}_1$. The total noise power increases monotonically with the number of amplifiers, and this phenomenon along such an optical amplifier chain is known as *amplified* spontaneous emission (ASE).

To refine our picture, consider the situation in Figure 2-1. Each gain element G_i is followed by a loss element L_i (an important special case is $L_i = 1/G_i$). The signal $\mathcal{P}_N$ and ASE noise $\mathcal{N}_N$ terms at the output are

$$\mathcal{P}_N = \prod_{k=1}^{N} G_k L_k\, P_0 \tag{2.4}$$

$$\mathcal{N}_N = \mu\, h\nu\, B_o \sum_{k=1}^{N} (G_k - 1) L_k \prod_{j=k+1}^{N} G_j L_j \tag{2.5}$$

2.1.4 Signal to noise ratio

Next, we consider the terms that define the mean-squared current at the output of an APD. The signal-to-noise ratio will involve only ratios of these quantities, and so

our simple definition is reasonable.

Following the discussion in [6], identifying the signal and noise contributions in the photocurrent is most convenient in terms of B_e, the electronic bandwidth of the APD, and B_o, the optical bandwidth of the signal, and it's useful to think in the frequency domain in terms of power spectral densities that can be integrated over the relevant bandwidth to yield a quantity that has the dimensions of power. We assume that the detector has a uniform frequency response[1]. In terms of the mean-square components of $i(t)$, the noise terms are

$$\begin{aligned} <{i_N}^2> &= <{i_{sh}}^2> + <{i_{s-ASE}}^2> + <{i_{ASE-ASE}}^2> + <{i_{th}}^2> \\ &= \bar{g}^{2+x}\, e\eta\, 2B_e(I_S + 2m\, I_N) + 2\bar{g}^2\,\eta^2 I_S\, I_N \frac{2B_e}{B_o} \\ &\quad + m\,\bar{g}^2\,\eta^2 I_N^2 \frac{2B_e}{B_o^2}\left(B_o - \frac{B_e}{2}\right) + \frac{4k_B\, T_e B_e}{R} \end{aligned} \tag{2.6}$$

where T_e is the effective temperature of the detector and R is its resistance. The number of ASE modes is given by m (for single-mode EDFAs without a polarizer at the output $m = 2$).

The four noise terms in [2.6] are: the shot noise due to the signal and the ASE, the beat noise of the signal and ASE fields, the beat noise among the ASE components themselves, and the thermal noise of the detector.

Typically, the noise current component due to ASE-ASE beat noise can be neglected if the signal power $\mathcal{P}_N$ is not allowed to drop too far. Optical filtering of the broadband ASE noise may also help to cut down this component. For similar reasons, the shot noise due to ASE can also be ignored.

The photodetector SNR, using [2.2] and [2.6] is

$$\boxed{\text{SNR} = \frac{<{i_s}^2>}{<{i_N}^2>}} \tag{2.7}$$

[1]For a model similar to the one we've discussed that specifically accounts for the filtering characteristics of typical optical detectors, see [20] and [28].

2.1.5 Bit error rate

In digital communication systems, the optical signals are usually amplitude modulated (on-off keyed, OOK). The random nature of the detection process leads to a finite probability of error for any signal-processing decision rule. Though we will be concerned only with [2.7], we provide a highly simplified discussion here of how that definition is related to the more familiar concept of a bit error rate (BER).

For a simple binary hypothesis (H_0, H_1) testing problem in a Bayesian framework [17], we identify a sufficient statistic y and compute its distributions $p_{\mathsf{y}|H}(y|H_0)$ and $p_{\mathsf{y}|H}(y|H_1)$ under each of the hypotheses. Choosing between H_0 and H_1 is equivalent to comparing the likelihood ratio $\mathcal{L}(\mathsf{y})$ to a threshold η

$$\mathcal{L}(\mathsf{y}) = \frac{p_{\mathsf{y}|H}(y|H_1)}{p_{\mathsf{y}|H}(y|H_0)} \underset{\hat{H}(\mathsf{y})=H_0}{\overset{\hat{H}(\mathsf{y})=H_1}{\gtrless}} \eta \tag{2.8}$$

where $\hat{H}(\mathsf{y})$ is our decision based on the statistic y. To form a measure of how "good" our decision is, we define the following probabilities

$$P_D = \Pr\left\{\hat{H}(\mathsf{y}) = H_1 \mid H = H_1\right\} \tag{2.9}$$

$$P_F = \Pr\left\{\hat{H}(\mathsf{y}) = H_1 \mid H = H_0\right\} \tag{2.10}$$

In optical communications, H_1 represents the presence of a pulse, and H_0 the absence of a pulse, and so the above terms refer to the probabilities of (successful) detection and of 'false-alarm', respectively. Obviously, P_F defines an error event, and its *a priori* probability is $\Pr\{H = H_0\}$.

There is a second type of error that can occur: we may declare $\hat{H}(\mathsf{y}) = H_0$ when, in reality, $H = H_1$. The probability of this event is $1 - P_D$, and its associated *a priori* probability is $\Pr\{H = H_1\}$.

Combining the two independent cases, the overall probability of error is

$$P_e = P_F \Pr\{H = H_0\} + (1 - P_D) \Pr\{H = H_1\} \tag{2.11}$$

There are many levels at which this formalism can be applied to optical signals. We'll adopt one of the simplest, originally due to Personick [20]. The conditional distributions in [2.8] are assumed to be Gaussian, and the Personick metric is the Q-factor,

$$Q = \frac{< i_{s_1} > - < i_{s_0} >}{\sqrt{< {i_{N_1}}^2 >} + \sqrt{< {i_{N_0}}^2 >}} \tag{2.12}$$

where $< i_{s_0} >$ and $< i_{s_1} >$ are the mean photocurrents associated with the signals under the two hypotheses, and $< {i_{N_0}}^2 >$ and $< {i_{N_1}}^2 >$ are the corresponding mean-squared noise components (variances).

The decision threshold for this detection scheme is the weighted average of the mean signal component of the photodetector output current,

$$\frac{\sqrt{< {i_{N_1}}^2 >} < i_{s_0} > + \sqrt{< {i_{N_0}}^2 >} < i_{s_1} >}{\sqrt{< {i_{N_0}}^2 >} + \sqrt{< {i_{N_1}}^2 >}} \tag{2.13}$$

and the corresponding bit error rate, using the standard notation for Gaussian integrals, is

$$\text{BER} = \frac{1}{\sqrt{2\pi}} \int_Q^\infty \exp\left(-\frac{x^2}{2}\right) dx = \frac{1}{2}\text{erfc}\left(\frac{Q}{\sqrt{2}}\right) \tag{2.14}$$

A lower BER implies that fewer errors occur over a given time interval, and a higher quality-of-service (QoS) can be obtained. Usually, achieving a low BER implies a trade-off in some other system resource, e.g. if the noise added grows monotonically with increasing propagation distance, a BER threshold sets a limit on how far apart the transmitter-receiver pair can be. Typical values of acceptable BER range from 10^{-12} to 10^{-9} for on-off keyed systems without error-correction coding.

Personick's model is a rather crude one: the filtered photodetector output is better expressed in terms of components along an orthonormal basis over the pulse interval T [11]. The dimensionality of the space of finite energy signals with a bandwidth B_o and time spread T is about $2B_oT + 1$ which, for convenience, is assumed to be even $= 2\mathcal{M}$. The filtered noise process in each polarization at the photodetector over the interval T can be written as $\sum_{i=1}^{2\mathcal{M}} n_i\phi_i(t)$ where we take n_i as i.i.d zero-mean Gaussian with variance $N_0/2$. The signal can also be expanded in this orthonormal basis as $\sum_{i=1}^{2\mathcal{M}} s_i\phi_i(t)$. In units such that the photodetector has unit gain, the integral of the photodetector output (without considering shot noise) is

$$\mathsf{x} = \int_0^T \left(\sum_{i=1}^{2\mathcal{M}} (s_i + n_i)\, \phi_i(t) \right)^2 dt \tag{2.15}$$

$$= \sum_{i=1}^{2\mathcal{M}} (s_i + n_i)^2 \tag{2.16}$$

Further accounting for the shot noise process, the photodiode generates electrons following an inhomogeneous Poisson process with rate equal to the square of the field envelope [11, 18]. The total number of electrons y generated over a bit time by the photodiode follows a Laguerre distribution, and the conditional error probabilities may then be evaluated explicitly.

We may approximate y as Gaussian, and a target error probability $P_e = Q(q)$ results in a necessary signal to noise ratio [11]

$$\boxed{E/N_0 \approx q^2 \left(1 + \frac{1}{2N_0}\right) + q\sqrt{\mathcal{M}\left(1 + \frac{1}{N_0}\right)}} \tag{2.17}$$

where $N_0/2 = \mathcal{N}/2B_o$ from [2.3] is the power spectral density of this additive white Gaussian noise (AWGN) source and $2E$ is the signal of the "on" pulses in OOK modulation.

The probability of error (which is the same as the "bit error rate" BER) under

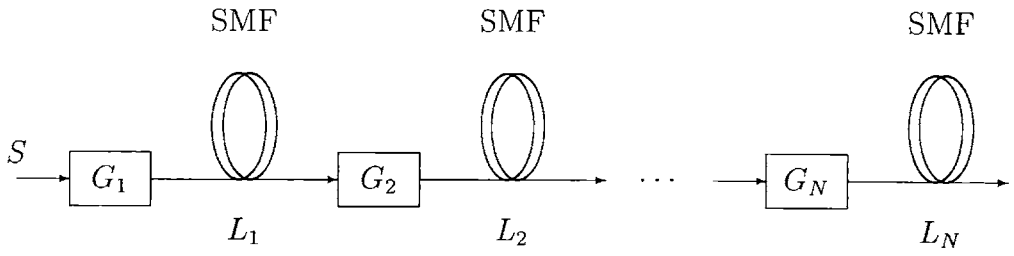

Figure 2-1: An amplifier chain consisting of N repeated sections. Section $i = 1, 2, \ldots, N$ is characterized by a gain $G_i \geq 1$ and followed by a loss (a section of single mode fiber, SMF) $L_i \leq 1$.

this assumption is

$$P_e \approx Q\left(\frac{2E/N_0}{\sqrt{\mathcal{M}(1+1/N_0)} + \sqrt{\mathcal{M}(1+1/N_0) + 4E/N_0(1+1/2N_0)}}\right) \tag{2.18}$$

where terms such as $\mathcal{M}N_0^2$ represent the ASE-ASE beat noise and $2EN_0$ represents the signal-ASE noise beat noise.

Both Personick's expression and the one given above are related to the physically intuitive concept of the signal-to-noise ratio [2.7], which shall be our figure-of-merit.

2.2 Optimizing gain and placement

A typical optical bus network consists of alternating sections of amplifying and loss elements. In particular, if we're using EDFAs as the gain elements, the length of the doped sections and the doping concentration are the typical parameters that control the gain. The length of the undoped section between EDFAs determines the loss; and conversely, specifying the loss (e.g. in dB) uniquely determines the propagation length in a given fiber. In this section, we investigate the conditions under which the signal-to-noise ratio at the end of such an optical amplifier chain is maximized.

We'll assume throughout this discussion that the effect of fiber nonlinearities can be ignored. Furthermore, the amplification is linear in field amplitudes, and the noise is additive and white as discussed in the previous section. The signal-to-noise ratio (SNR) at any point along this amplifier chain—and in particular after the N^{th} stage—is then completely determined by the individual gains and losses of these elements. There are several ways that we can formulate our optimization problem: we consider a very general approach that will prove quite powerful in subsequent more physically intuitive reasonings. The optimization problem that we consider first is[2]

$$\begin{aligned} \max \quad & \text{SNR}\,(G_1, G_2, \ldots, G_N,\, L_1, L_2, \ldots, L_N) \\ & \text{subject to} \quad \prod_{k=1}^{N} G_k = G_{max}, \quad G_k \geq 1 \,\forall\, k, \\ & \qquad\qquad \prod_{k=1}^{N} L_k = L_{min}, \quad 0 < L_k \leq 1 \,\forall\, k. \end{aligned} \tag{2.19}$$

In a given practical implementation, the constraints are frequently tighter—for example, the entire domain $[1, G_{max}]$ may not be available to each G_i. We solve this particular problem as indicative of the upper bound on the achievable, with the caveat that this bound will probably be a loose one. Correspondingly, our solution to this problem will be *indicative* of the correct design approach, rather than specify explicitly what the various system parameters will be: the usefulness of the solution to this problem in the network design context is to tell us how to improve performance in a given situation, if possible, and what the upper bounds of such improvement methods are. But the approach we take is quite general, and we expect other problems—besides the one of immediate concern—can also be analyzed in a similar fashion. A particular problem with a more direct physical interpretation is analyzed in Section 3.4 after we have developed our model for amplification in erbium-doped fibers.

We'll first consider the case $N = 2$, when there are only two sections. As we've

[2]See Appendix A for a discussion on the form of the constraints.

seen in Section 2.1, the signal and ASE noise powers at the output of this two-element chain are

$$\mathcal{P} = G_2 L_2 G_1 L_1 P_0 \tag{2.20}$$

$$\mathcal{N} = \underbrace{\mu h\nu B_o}_{\alpha}[(G_1 - 1)L_1 G_2 L_2 + (G_2 - 1)L_2] \tag{2.21}$$

In this section, we focus on optimization arguments, and the numerical values of the various parameters are largely unimportant. We make the following assumption

$$\frac{e\eta}{h\nu} \approx 1 \tag{2.22}$$

so that the optical powers [2.20, 2.21] carry over directly to currents generated in response to them by a photodiode of quantum efficiency η.

Further, we assume that the thermal noise of the detector is dominated by either the signal-ASE beat noise or the signal shot noise: this is the case of most practical interest when a high-gain preamplifier is used at the detector, or when the optical amplifier chain comprises of more than a few elements.

Using these assumptions, we can write the signal-to-noise ratio (SNR) function defined by these two components

$$\mathrm{SNR}\,(G_1, G_2, L_1, L_2) = \frac{P_0}{2\alpha} \frac{G_2 L_2 G_1 L_1}{K + (G_1 - 1)L_1 G_2 L_2 + (G_2 - 1)L_2} \tag{2.23}$$

in terms of the following parameters

$$\alpha = \mu h\nu B_o \tag{2.24}$$

$$K = eB_o/2\alpha$$

$$= 1/2\mu\eta \tag{2.25}$$

The variables that appear in this continuous, real-valued function are defined over

the closed intervals on the real axis $1 \leq G_i \leq G_{max}$, and $L_{min} \leq L_i \leq 1$. Usually, G_{max} is large—40 to 60 dB—and L_{min} is correspondingly small $\approx 1/G_{max}$. The main point is that these domains are all convex, compact subsets of $\mathbf{R}^1$.

Let us define the convex set S as the Cartesian product of these intervals

$$S = [1, G_{max}] \times [1, G_{max}] \times [L_{min}, 1] \times [L_{min}, 1] \tag{2.26}$$

Since S is closed and bounded on $\mathbf{R}^4$, by the Borel covering theorem [27, pages 64–66], it is compact.

Now, we recall a useful concept [1, page 68],

Definition 1 (Quasi-concavity) *A real-valued function f defined on a convex set S in $\mathbf{R}$ is said to be quasi-concave if the set*

$$\{\, \mathbf{s} \mid \mathbf{s} \in S,\ f(\mathbf{s}) \geq c \,\}$$

is convex for all real numbers c.

Note that SNR(s) is a real number. As we will show later, SNR is monotonic in each of G_1, G_2, L_1 and L_2. Therefore [1, page 68], the SNR is individually quasi-concave in G_1, G_2, L_1 and L_2.

We model the selection of the optimum allocation—the optimum G_i and L_i—as a non-cooperative game (of dimension 4) in strategic form between the players G_1, G_2, L_1 and L_2. The extension to a cooperative game will be discussed later. The strategy set for each player consists of the domain over which the corresponding variable (G_1, G_2, L_1 and L_2) is defined in $\mathbf{R}^1$. S as defined above is the strategy set for the game. The reward function P_i is the same for all players, and is the signal-to-noise ratio function, SNR.

We select a strategy for this game by allocating a real number for each of the players to form a strategy vector $\mathbf{s} = \{s_1, s_2, s_3, s_4\} \in S$. These four numbers s_i

determine the SNR, and consequently the reward for each player. Our goal is to optimize the reward: choose the optimum G_i and L_i to maximize the SNR.

A strategy vector $\mathbf{s} \in S$ can be diminished by deleting the ith strategy to form $\mathbf{s}_{-i} \in S_{-i}$, the space of all strategies except for that of player i. Similarly, a diminished strategy vector may be augmented $\mathbf{s}_{-i}\backslash t$ to form the same vector as $\mathbf{s}$ but with a new strategy t for player i.

Definition 2 (Reaction Function) *The reaction function for player i is the point-to-set mapping [10] $\mathbf{r}_i : S_{-i} \mapsto S_i$ defined by*

$$\mathbf{r}_i(\mathbf{s}_i) = \{\, \mathbf{s}_i \mid P_i(\mathbf{s}_{-i}\backslash\mathbf{s}_i = \max_{\mathbf{t}\in S_i} P_i(\mathbf{s}_{-i}\backslash\mathbf{t}) \,\}$$

which defines player i's optimal strategy in reaction to a given set of strategies by the other players.

The Cartesian product set of all such point-to-set mappings is called the reaction function $\mathbf{R}(\mathbf{s})$ for the game. As we'll show shortly, the nature of the SNR function implies that this is nonempty for every $\mathbf{s} \in S$.

The following theorem states conditions for the existence of a particularly important $\mathbf{s}^*$, which is a fixed-point of $\mathbf{R}$, i.e. $\mathbf{s}^* \in \mathbf{R}(\mathbf{s}^*)$.

Theorem 1 (Nash Theorem [14, pages 267–268]) *In a non-cooperative game of dimension N, if the strategy sets S_i are convex compact subsets of a finite-dimensional space, and the reward functions P_i are continuous and individually quasi-concave, then there exists a point $\mathbf{s}^* \in S$ such that for every $i \in N$,*

$$P_i(\mathbf{s}^*\backslash\mathbf{s}_{-i}) \leq P_i(\mathbf{s}^*), \qquad \textit{for all } s_i \in S_i.$$

To see this, note that our reward functions, each equal to the SNR function, are individually quasi-concave. Therefore, the set of points $\mathbf{t}_i$ that maximize $P_i(\mathbf{s}_{-i}\backslash\mathbf{t})$ is convex. Since the SNR function is continuous, and the domains of definition for G_1,

G_2, L_1 and L_2 are compact, $\mathbf{t}_i$ is compact and nonempty, and so, $\mathbf{R}$ is convex and compact.

Since the reward functions (SNR) in continuous, and S is compact, it follows that $\mathbf{R}$ is upper semi-continuous[3]. The following theorem guarantees the existence of a fixed-point which is common to the ranges of each of the individual reaction functions.

Theorem 2 (Kakutani [13, page 256]) *Let* $\mathbf{R}$ *be an upper semi-continuous point-to-set mapping from a compact convex set* S *into itself such that for each* $\mathbf{s} \in S$, $\mathbf{R}(\mathbf{s})$ *is compact and convex. Then there exists a point* $\mathbf{s}^* \in S$ *(called the fixed point) such that* $\mathbf{s}^* \in \mathbf{R}(\mathbf{s}^*)$.

This fixed point is a Nash equilibrium for the problem.

□

The definition of our game satisfies the conditions for the existence of a Nash equilibrium: a strategy vector, consisting of picking a value in the appropriate domain of definition, for each of G_1, G_2, L_1 and L_2 such that no player's SNR can be increased by a unilateral strategy change. In other words, given the Nash equilibrium for our problem, changing G_1 or any of the other parameters individually will not improve the SNR.

In the next section, we use a simple argument to construct a Nash equilibrium, by examining the structure of the individual strategies and reaction functions. We'll see that this point has some special properties that will also enable us to characterize a different form of optimality: Pareto efficiency [14]. We can think of a strategy S as formed by an *allocation* X of the gain and loss constraints, G_{max}, L_{min} among the players so that the respective constraints are satisfied.

[3]Recall [1, page 68] that a point-to-set mapping

$$\mathbf{g}(\mathbf{x}) = \arg\max_{\mathbf{y} \in Y} f(\mathbf{x}, \mathbf{y})$$

is upper semi-continuous if f is continuous and Y is compact.

Definition 3 (Pareto Efficiency) *A feasible allocation X is said to be Pareto efficient if there exists no other feasible allocation X' such that all individuals (weakly) prefer X' to X and at least one individual strictly prefers X' to X.*

In our particular case, the benefit functions for all the players are the same: to acknowledge this explicitly, this is sometimes known as Pareto optimality.

2.2.1 Optimize over G_1, G_2

First, we consider the variation of SNR with respect to G_1 while keeping the other parameters fixed. This will allow us to define what the reaction function of G_1 might be, for instance. To highlight the dependence on G_1,

$$\frac{2\alpha}{P_0}\,\mathrm{SNR}(G_1) = \frac{A_1 G_1}{B_1 + C_1 G_1} \tag{2.27}$$

where

$$A_1 = G_2 L_2 L_1 \tag{2.28}$$

$$B_1 = K - G_2 L_2 L_1 + (G_2 - 1) L_2 \tag{2.29}$$

$$C_1 = G_2 L_2 L_1 \tag{2.30}$$

Note that SNR is a continuous and differentiable function of G_1, and in particular, since $L_i \leq 1$,

$$\frac{2\alpha}{P_0}\frac{d}{dG_1}\mathrm{SNR}(G_1) = \frac{B_1}{(B_1 + C_1 G_1)^2} \qquad \begin{cases} > 0 & \text{if } K > L_1 L_2 \\ < 0 & \text{if } K < L_1 L_2 \text{ and } G_1 = 1 \end{cases} \tag{2.31}$$

We sketch $\mathrm{SNR}(G_1)$ for the two cases $B_1 > 0$ and $B_1 < 0$ as shown in Figure [2-2], and we note that in the asymptotic limit $G_1 \to \infty$, the SNR approaches $A_1/C_1 = 1$.

Note that $K = 1/2\mu\eta$, where the inversion factor $\mu \approx 1$ and the photodiode quantum efficiency $\eta \approx 1$. Since $\min\{L_1, L_2\}$ is usually $\ll 1$, it's usually true that

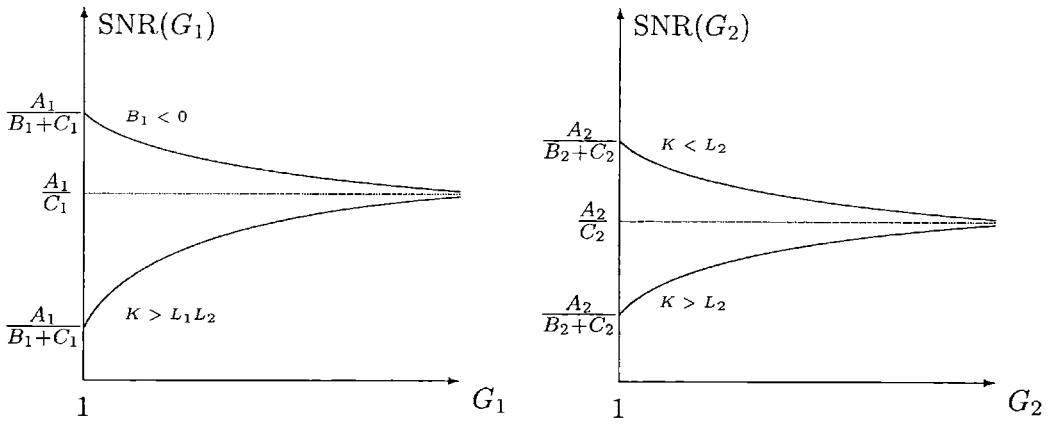

Figure 2-2: The dependence of gain on the signal-to-noise ratio SNR(G_1) and SNR(G_2) from [2.27]. As discussed in the text, the lower curve corresponds to most practical situations. Note that if $K > L_2$, then $K > L_1L_2$ since $L_1 < 1$.

$K > L_1L_2$. In this case, a reasonable strategy for the gain of this first stage would be to increase G_1 to the maximum possible, so that the SNR is maximized. This is consistent with our definition of a Nash equilibrium. However, given that the total gain that we can allocate (G_{max}) is bounded, this isn't the full story, as we'll see in the remainder of this section.

Using a similar argument, we consider the variation in SNR with G_2, which we write

$$\frac{2\alpha}{P_0}\text{SNR}(G_2) = \frac{A_2G_2}{B_2 + C_2G_2} \tag{2.32}$$

where

$$A_2 = G_1L_1\,L_2 \tag{2.33}$$

$$B_2 = K - L_2 \tag{2.34}$$

$$C_2 = (G_1 - 1)L_1L_2 + L_2 \tag{2.35}$$

By forming the first derivative with respect to G_2, we can easily show that $\mathrm{SNR}(G_2)$ is a monotonically increasing function, just as we observed for $\mathrm{SNR}(G_1)$.

$$\frac{2\alpha}{P_0}\frac{d}{dG_2}\mathrm{SNR}(G_2) = \frac{B_2}{(B_2 + C_2 G_2)^2} \qquad \begin{cases} > 0 & \text{if } K > L_2 \\ < 0 & \text{if } K < L_2 \end{cases} \tag{2.36}$$

The SNR sketch, shown in Figure [2-2], is of the same form as the one for G_1.

2.2.2 Cooperative solution

We now focus on the case $K > L_2$. As we will see later, this is indeed the optimal case. A similar argument to the one below, with complementary conclusions, will help us choose the right (G_1, G_2) for $K < L_2$.

In the absence of a constraint on G, a Nash equilibrium with respect to (G_1, G_2) for constant (L_1, L_2) would be the (non-cooperative) allocation $(G_1 \to \infty, G_2 \to \infty)$. However, this isn't possible—but G_1 and G_2 can cooperate, and form a joint strategy to optimize the SNR within the given constraints (i.e. find the Pareto optimal solution.)

We solve this problem by using the concept of arbitration.

Definition 4 (Arbitration Procedure [19]) *Given a payoff region P and a feasible status-quo allocation $(u_0, v_0) \in P$, an arbitration procedure Ψ produces a payoff pair $(u^*, v^*) = \Psi(P, (u_0, v_0))$ which is fair to both players.*

Note that the asymptotic limits for the two curves are different. In fact,

$$\frac{A_2}{C_2} = \frac{G_1 L_1}{(G_1 - 1)L_1 + 1} \leq 1 = \frac{A_1}{C_1} \tag{2.37}$$

Under the constraints of G_{max}, let's consider the arbitrated allocation $(G_1^* = G_{max}, G_2^* = 1)$, which is certainly Pareto superior to $(G_1 = 1, G_2 = G_{max})$. At this allocation, the SNR is A_1/C_1. An increase ΔG in G_2 cannot raise the SNR to more

than A_1/C_1: the asymptotic limit for the SNR by increasing G_2 is $A_2/C_2 \leq A_1/C_1$. In fact, the constraint $G_1G_2 = G_{max}$ implies that G_1 shall have to reduce, for small ΔG, by approximately $\Delta G\, G_{max}$, thereby reducing the SNR. This affects the reward function for G_2, and so moves away from Pareto optimality.

We make the following observations that result from this arbitration.

1. Consider any $S_G^{'} \in S$. If the optimal $\{P_1(G_1^*, G_2^*), P_2(G_1^*, G_2^*)\} \in S_G^{'}$ and there exists another $\{P_1(G_1, G_2), P_2(G_1, G_2)\} \in S_G^{'}$, then

$$\Psi\left(S_G^{'}, \{P_1(G_1, G_2), P_2(G_1, G_2)\}\right) = \{P_1(G_1^*, G_2^*), P_2(G_1^*, G_2^*)\}$$

   In words, the arbitration is independent of irrelevant alternatives.

2. The arbitration is independent of monotone order-preserving linear transformations to P_1 or to P_2, the payoff regions for G_1 and G_2 respectively. This follows from the definition of the reward function SNR.

By the Nash Theorem for cooperative games [19, page 138], the feasible Pareto efficient arbitrated allocation we've found is unique. Alternatively, if we take our status quo feasible point as $(G_1 = G_{max}, G_2 = 1)$, it follows from [2.31] and [2.36] that no other point is Pareto optimal. Having satisfied the conditions for the Nash theorem for cooperative games, we know that there exists exactly one Pareto optimal point. Therefore, the arbitration we've constructed in the preceding paragraphs is indeed the unique solution, in the Nash sense.

To summarize our argument in a different light, assuming G_{max} is sufficiently large ($G_{max} \to \infty$) that setting $G_{1,2} = G_{max}$ puts us in the asymptotic limits of either curve, we note that

$$\lim_{G_1 \to 1} \left\{ \lim_{G_2 \to \infty} \mathrm{SNR}\,(G_1, G_2) \right\} = L_1 \leq 1 \tag{2.38}$$

$$\lim_{G_2 \to 1} \left\{ \lim_{G_1 \to \infty} \mathrm{SNR}\,(G_1, G_2) \right\} = 1 \tag{2.39}$$

Therefore, the allocation $(G_1 = G_{max}, G_2 = 1)$ is the unique Pareto efficient Nash solution of the cooperative game for fixed L_1 and L_2.

2.2.3 Optimize over L_1, L_2

This section runs essentially parallel to the previous one: instead of G_1 and G_2, we deal with L_1 and L_2. First, we consider the variation of SNR with respect to L_1 while keeping the other parameters fixed. To highlight the dependence on L_1,

$$\frac{2\alpha}{P_0}\,\mathrm{SNR}(L_1) = \frac{a_1 L_1}{b_1 + c_1 L_1} \tag{2.40}$$

where

$$a_1 = G_2 L_2 G_1 \tag{2.41}$$

$$b_1 = K + (G_2 - 1)L_2 \tag{2.42}$$

$$c_1 = (G_1 - 1)\, G_2 L_2 \tag{2.43}$$

Taking the derivative of the above with respect to L_1,

$$\frac{2\alpha}{P_0}\,\frac{d}{dL_1}\mathrm{SNR}(L_1) = \frac{b_1}{(b_1 + c_1 L_1)^2} \tag{2.44}$$

which is always positive, and so, the SNR is a monotonically increasing function of L_1 as shown in Figure [2-3].

Analysis of SNR versus L_2 follows along exactly the same lines.

$$\frac{2\alpha}{P_0}\,\mathrm{SNR}(L_2) = \frac{a_2 L_2}{b_2 + c_2 L_2} \tag{2.45}$$

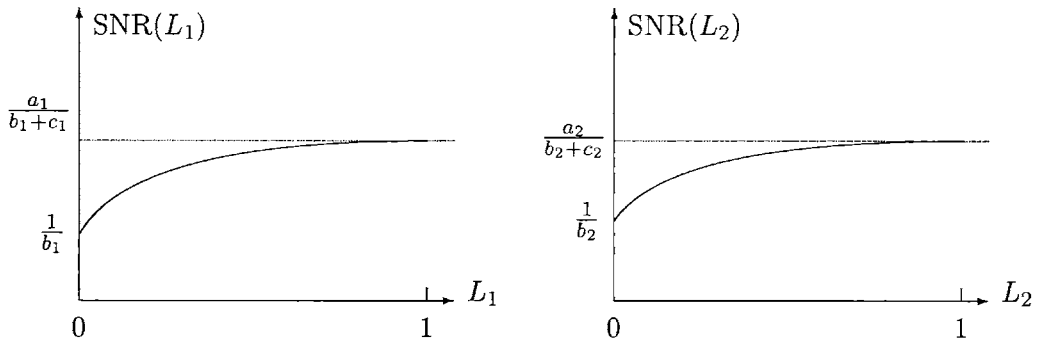

Figure 2-3: The loss dependence of signal-to-noise ratio SNR(L_1) and SNR(L_2) from [2.40] and [2.45].

where

$$a_2 = G_1 L_1 G_2 \tag{2.46}$$

$$b_2 = K \tag{2.47}$$

$$c_2 = (G_1 - 1)L_1 G_2 + (G_2 - 1) \tag{2.48}$$

Given the familiar nature of this functional relationship, it should come as no surprise by now that, once again, we optimize the SNR in the limit $L_2 = 1$. Taking the derivative as before,

$$\frac{2\alpha}{P_0}\frac{d}{dL_2}\text{SNR}(L_2) = \frac{b_2}{(b_2 + c_2 L_2)^2} \tag{2.49}$$

As in the case of G_1 and G_2, we have to choose which of the two variables L_1 or L_2

to maximize. From the figure, it's easy to see that the following is true for $L_{min} \approx 0$,

$$\lim_{L_1 \to 1} \left\{ \lim_{L_2 \to 0} \mathrm{SNR}\,(L_1, L_2) \right\} = \frac{1}{K} \tag{2.50}$$

$$\lim_{L_2 \to 1} \left\{ \lim_{L_1 \to 0} \mathrm{SNR}\,(L_1, L_2) \right\} = \frac{1}{K + (G_2 - 1)} \leq \frac{1}{K} \tag{2.51}$$

with equality in the last relation if and only if $G_2 = 1$.

Therefore, the allocation $(L_1 = 1, L_2 = L_{min})$ is a Nash equilibrium (for fixed G_1 and G_2). In connection with an earlier statement on page 34, note that $K > L_2$ is indeed the usual case.

As a reminder, our definition of L is a number between 0 and 1. To relate this to the length of fiber $d_{1,2}$ that corresponds to this loss, note that $d_i \propto \exp(-L_i)$. Our Nash equilibrium allocation of (L_1, L_2) is to set $d_1 \to 0$, and d_2 to take up the rest of the propagation distance (as defined by L_{min}).

2.2.4 Optimality

We assume that G_{max} and L_{min} can be distributed arbitrarily among the two stages: in the standard terminology of economics, this is the assumption of free distribution in an economy of goods. Then [14, page 171],

Theorem 3 *In an economy that supports free distribution, if preferences are continuous and strongly monotonic, then a feasible allocation X is Pareto efficient if and only if there is no other feasible allocation X' that is strictly preferred by all players.*

In the context of our game, the preferences are defined by the SNR function and are the same for all players. As we've seen in the previous section, this function is certainly continuous, and we've shown by taking the first derivative that it is strictly monotonic in each of the variables [2.31, 2.36, 2.44, 2.49]. Therefore, each of the individual preferences are strongly monotonic.

The Nash solution that we've found satisfies the criterion of the above theorem: as we've shown, a shift away from that allocation strictly decreases the SNR for each of

the players. Any other resulting allocation cannot be Pareto efficient: consequently, the Nash solution $\{G_1 = G_{max}, G_2 = 1, L_1 = 1, L_2 = L_{min}\}$ is also Pareto optimal.

The Nash solution of the optimum allocation co-operative game for the two-amplifier case *with unsaturated gain* is to arbitrate $G_1 = G_{max}$, $G_2 = 1$, and $L_1 = 1$, $L_2 = L_{min}$. No other feasible allocation is also Pareto optimal. This implies that all the gain is inserted at the first stage of this amplifier chain, and the distance between the two amplifier stages should be set to zero: the propagation distance is entirely defined by L_2.

We caution that, though a similar approach may be taken for more complicated nonlinear amplification problems, the results are likely to be different.

2.2.5 Optical amplifier chains

We present a simple constructive proof to generalize our optimality argument. Consider a chain of N amplifiers, each characterized by gain G_k, and each followed by a loss element L_k. The problem can be formulated more precisely in terms of the following definition,

Definition 5 (Reallocation) *If* $\mathbf{X} = \{G_1, G_2, \ldots, G_N, L_1, L_2, \ldots, L_N\}$ *is status quo feasible allocation, the following mapping where* $\mathbf{X}'$ *is also feasible defines a reallocation.*

$$\mathbf{X} \mapsto \mathbf{X}' = \{G_1', G_2', \ldots, G_N', L_1', L_2', \ldots, L_N'\}$$

In our problem, feasibility is defined by the constraints in [2.19]. The signal and ASE noise terms at the output of such a chain are

$$\mathcal{P} = P_0 \prod_{k=1}^{N} G_k L_k \tag{2.52}$$

$$\mathcal{N} = \alpha \sum_{k=1}^{N} (G_k - 1) L_k \prod_{j=k+1}^{N} G_j L_j \tag{2.53}$$

It's obvious from the above expression that if we can reduce the overall noise contribution from a subset $N' \leq N$ of the amplifiers, the overall signal-to-noise ratio improves. Consider $N' = 2$, and we write the ASE noise $\mathcal{N}$ with a subscript N for the number of stages,

$$\begin{aligned}\mathcal{N}_N &= \alpha \sum_{k=1}^{N} (G_k - 1) L_k \prod_{j=k+1}^{N} G_j L_j \\ &= \{\alpha(G_1 - 1)L_1 G_2 L_2 + \alpha(G_2 - 1)L_2\} \prod_{k=3}^{N} G_k L_k \\ &\quad + \mathcal{N}_{N-\{1,2\}} \end{aligned} \tag{2.54}$$

Recall that the arbitration between the first two stages in each of G and L is independent of irrelevant alternatives. The optimal strategy for allocating the gains and losses using $\mathcal{N}_N$ is also the Pareto optimal Nash solution, $\{G_1 = G_{max}, G_2 = 1, L_1 = 1, L_2 = L_{min}\}$ formed by setting G_1 and L_2 to their respective limits. This pair of amplifiers collapses to a single stage: $G_2 = 1$ implies a no-gain, no-loss ideal transmission line, and similarly $L_1 = 1$ implies the length of undoped fiber between the last two amplifiers $d_{12} = \log L_1 = 0$. In essence, we've concatenated the gains of these two stages, and located this new amplifier at the left-most edge of the region we're considering.

We recursively consider the combination of this resultant amplifier and the next one and apply the same argument. At each stage, the two amplifiers collapse to a single stage, located at the left-hand edge of the pair, with an increased gain at the first stage and an increased propagation distance beyond the second amplifier stage. Provided we don't run into a upper-bound limitation on the total gain allowed at a single stage, this procedure can be repeated *ad infinitum*. This is undoubtedly a Pareto optimal arbitration among the N players.

Similar to the Nash theorem for two-person games is the generalization to an N-person cooperative game [26, pages 309–314]: this Pareto optimal arbitration is

unique and is the Nash solution to the problem.

In a lumped amplifier chain modeled as a $2N$-person cooperative game, the unique Pareto optimal reallocation strategy that optimizes the SNR starting with $\{G_1, G_2, \ldots, G_N, L_1, L_2, \ldots, L_N\}$ is the Nash arbitration solution:

$$\begin{aligned} &\{G_1' \to \prod_{k=1}^{N} G_k, G_2' \to 1, \ldots, G_N' \to 1\} \\ &\{L_1' \to 1, L_2' \to 1, \ldots, L_N' \to \prod_{k=1}^{N} L_k\} \end{aligned} \tag{2.55}$$

As a footnote, we point out that this isn't always implementable: the gain that we can attain in a single stage is limited by the onset of non-linear optical effects, and that will modify our SNR expressions considerably.

2.3 Lumped and distributed amplifiers

In a typical EDFA, the length of the erbium-doper fiber is of the order of a few meters, and undoped fiber is used for the remainder of the distance between the source and the destination. Consider a section of fiber perhaps three orders of magnitude longer, but with a lighter doping concentration so that the net gain is the same. The same extent of fiber serves as the gain medium and as the transmission medium. This is called a distributed amplifier.

For our present discussion, we can view a distributed amplifier as a limiting case of a lumped amplifier chain. If a total gain of G_0 is to be distributed among N stages, one particular way of achieving this is to allocate to each stage a gain of G/N. As $N \to \infty$, the gain of each stage goes to zero and so does the inter-amplifier spacing, but the overall gain is, by definition, unchanged. In essence, we've taken a discrete distribution over to its continuous equivalent, while preserving the "area under the curve".

It's clear from our discussion that the SNR for such a system is worse than that for any lumped amplifier chain (with a gain element preceding a lossy fiber). Consider section k of length L_k in a chain of N amplifiers (N is large but finite). For this section, we've shown that it's optimal to maintain a single amplifier of gain $G_k = G_0/N$ at the front end of the fiber of length L_k. In particular, this arrangement is better than spreading G_k out evenly over L_k (i.e. converting the k^{th} section to a distributed amplifier). This holds for all sections $k = 1, 2, \ldots, N$.

The same result is derived from a different perspective in [6, pages 121–136]. But we will see that for our distribution network problem (maximizing the number of users), this sub-optimal solution is the more natural one.

Chapter 3

Characteristics of EDFAs

In constructing a deductive system
such as that contained in the present work,
there are two ... tasks that have to be ... performed.
On the one hand, we have to analyse existing mathematics
with a view to discovering what premises are employed,
and whether they are capable of reduction
to more fundamental premises.
—B. RUSSELL, Preface to *Principia Mathematica* (1910)

Optical (glass) fibers doped with lanthanides, e.g. erbium, can achieve lasing gain, and can be pumped with relatively low levels of light to perform amplification of optical signals. The first Erbium-doped fiber amplifiers (EDFAs) were developed by researchers in the University of Southampton and the AT&T Bell Laboratories in the late eighties. In this chapter, we cannot completely discuss the considerable research activity in this field over the past decade, but instead focus on those aspects relevant to applications of erbium-doped fibers in optical distribution networks.

To preview our development, we describe an unsaturated gain model, and the evolution of the pump through an erbium-doped fiber. We analyze the dependence of the signal gain as a function of the input pump power. As we aim to present analytical derivations where possible, models which must inherently be evaluated numerically in a given situation (e.g. high-gain EDFAs) are not discussed.

3.1 Rate Equations

An Er-doped fiber pumped by an optical beam can be described by the standard rate equations for a two or three level laser system, depending on the method (and therefore the optical frequency) of pumping. The theoretical analysis of Er-doped fibers is well known [6], and we will not discuss it in completeness here. Instead, the major part of our development will deal with a simpler model for the evolution of pump and signal beams derived from basic considerations and under certain simplifying considerations. This will allow us to get some useful answers to the size-of-the-network problem, without a great deal of symbolic manipulation

Firstly, we assume that the relevant laser transitions are homogeneously broadened. In the absence of saturation, this assumption changes none of the predictions of the main phenomena, e.g. ASE spectrum, as noted in [6]. Moreover, for the types of calculations we carry out, the insight gained from analytical solutions possible with this simplification is valuable.

On a related note, the ligand-field induced Stark effect in the laser system causes a splitting of the each of the (two or three) levels into a number of manifolds. We assume that because of thermalization, the populations within each individual manifold (for an energy level of total angular momentum J, there are now J+1/2 energy sublevels) follow Boltzmann's distribution. At thermal equilibrium, the same rate equations as before are satisfied by the *sum* of the population densities of each level. This is convenient because the overall pumping and emission rates can indeed be characterized by experiments, whereas transitions between individual Stark sub-levels would require further knowledge, and also question our assumption of homogeneous broadening.

3.1.1 Propagation of intensity

The concept of cross sections is an effective way to express the strength of atomic transitions and the effect of applied electric fields. Consider a thin slab of erbium-doped fiber of thickness Δz and transverse area S along which the erbium ion density

is ρ. Let $\mathcal{N}_1$ and $\mathcal{N}_2$ be the fractional densities, or populations, of atoms in the lower and upper levels of the atomic transition of interest. Assuming ions are neither created nor lost, the total population $\mathcal{N}_1 + \mathcal{N}_2 \equiv \mathcal{N} = \rho$.

When this slab of active medium is illuminated by an optical beam of power $P(z,t)$, each lower level atom behaves as if it has an effective area or cross section σ_a for power absorption. Similarly, each upper level atom is characterized by an effective area σ_e for emission. The change in power $\Delta P(z,t)$ for this optical wave upon passing through this slab is ([23], pages 286–287)

$$\Delta P(z,t) = [\mathcal{N}_2(z,t)\sigma_e - \mathcal{N}_1(z,t)\sigma_a]\ P(z,t)\,\Delta z \tag{3.1}$$

Before we consider the limiting case $\Delta z \to 0$, let's introduce two modifications to the above model. As discussed in Desurvire [6, page 33], we can capture the effect of incomplete overlap between the doped fiber core and the modal (electromagnetic field) distribution of the optical power by a multiplicative confinement factor, $\Gamma_s \leq 1$.

Secondly, let's normalize the populations to the laser ion density, so that, in terms of $N_{1,2} = \mathcal{N}_{1,2}/\mathcal{N}$, the above equation becomes

$$\frac{\partial P_s}{\partial z}(z,t) = \mathcal{N}\,\Gamma_s\,[(\sigma_e + \sigma_a)N_2(z,t) - \sigma_s]\,P_s(z,t) \tag{3.2}$$

To simplify the notation, we introduce the absorption and emission coefficients,

$$\alpha_s = \rho\,\Gamma_s\,\sigma_a \tag{3.3}$$

$$\gamma_s = \rho\,\Gamma_s\,\sigma_e \tag{3.4}$$

so that the equation for optical propagation becomes

$$\boxed{\frac{\partial P_s}{\partial z}(z,t) = [(\gamma_s + \alpha_s)N_2(z,t) - \alpha_s]\,P_s(z,t)} \tag{3.5}$$

We'll use this equation in a later chapter, generalizing to multiple optical beams:

among the minor changes will be a wavelength dependency of the cross sections for the various optical beams, which are usually called "channels" in a wavelength-division multiplexing (WDM) scenario.

We now have a equation describing the evolution of an optical beam, whether it be at the pump wavelength or at the signal frequencies, along an Er-doped fiber. Next, we derive the corresponding equation that describes the change in the populations of the two levels as a result of this optical power transfer.

3.1.2 Two-level system

Rate equations for laser systems are usually described in terms of pumping and transition rates, labeled R and W, which modify the population densities, N_1 and N_2, of the lower and upper atomic levels of Er-doped glass. In particular, using the standard two-level model [29, pages 192–193], without direct pumping to the levels,

$$\frac{dN_2}{dt} = -\frac{N_2}{t_2} - (\gamma_s N_2 - \alpha_s N_1) W_i \tag{3.6}$$

$$\frac{dN_1}{dt} = -\frac{N_1}{t_1} + \frac{N_2}{\tau} + (\gamma_s N_2 - \alpha_s N_1) W_i \tag{3.7}$$

where τ is the spontaneous lifetime of the upper state. We disregard the term dependent on t_1, which represents a further transition from the lower level to a ground level: in our case, the lower level is itself the ground level. Also, τ characterizes the upper-to-lower level transition: therefore, $t_2 = \tau$.

The emission rate $W_i(t)$, which we'll soon generalize to include a longitudinal coordinate z, can be defined for a specific transition (e.g. the $2 \rightarrow 1$ laser transition in Er-doped glass), in terms of the signal intensity $I_s(t)$ [29, page 193],

$$W_{21}(t) = \frac{{\lambda_s}^2}{8\pi n^2 h \nu_s \tau} \Gamma_s I_s(t)\, g(\lambda_s) \tag{3.8}$$

where $g(\lambda_s)$ is the lineshape function. and Γ_s is the multiplicative factor defined

earlier. From fundamental considerations, the emission cross section is defined [23, page 288],

$$g(\lambda_s) = 8\pi n^2 \tau \frac{\sigma_e(\lambda_s)}{{\lambda_s}^2} \tag{3.9}$$

resulting in the following definition for the emission rate

$$W_{21}(t) = \frac{\sigma_e(\lambda_s)}{h\nu_s}\Gamma_s I_s(t) = \frac{\sigma_e(\lambda_s)}{S}\frac{1}{h\nu_s}\Gamma_s P_s(t) \tag{3.10}$$

using the relation between power and intensity in terms of the cross-section area S of the fiber.

With this definition, the rate equation for the upper level population becomes

$$\frac{dN_2}{dt} = -\frac{N_2}{\tau} - \frac{1}{\rho S}\frac{1}{h\nu_s}P_s(t)\left[\gamma_s N_2 - \alpha_s N_1\right] \tag{3.11}$$

Next, we can consider the effects of pulse propagation, and the longitudinal coordinate z, on the above rate equations. First, the derivatives with respect to time t are replaced with partial derivatives.

Let the electromagnetic energy density in the optical pulse be $D(z,t)$, such that the intensity of the pulse is

$$I_s(z,t) = v_g\, D(z,t) \approx \frac{c}{n} D(z,t) \tag{3.12}$$

where we approximate the group velocity in Er-doped glass with the phase velocity.

In a short segment of length Δz, the rate of change of stored energy is given by the energy flux differential across this length, plus the rate of stimulated emission within this segment [23, pages 363–365],

$$\frac{\partial}{\partial t}\left[D(z,t)\,\Delta z\right] = I_s(z,t) - I_s(z+\Delta z,t) + \left[\gamma_s N_2(z,t) - \alpha_s N_1(z,t)\right] I_s(z,t)\,\Delta z \tag{3.13}$$

In the limit $\Delta z \rightarrow 0$, the above becomes

$$\frac{1}{c}\frac{\partial I_s(z,t)}{\partial t} + \frac{\partial I_s(z,t)}{\partial z} = \left[\gamma_s N_2(z,t) - \alpha_s N_1(z,t)\right] I_s(z,t) \tag{3.14}$$

We transform to moving coordinates defined by

$$t \rightarrow t - z/c \tag{3.15}$$

and, in this reference frame that moves with the forward-traveling pulse,

$$\frac{1}{\sigma_e \rho}\frac{\partial P_s(z,t)}{\partial z} = \left[\gamma_s N_2(z,t) - \alpha_s N_1(z,t)\right] P_s(z,t) \tag{3.16}$$

where we've divided both sides of [3.14] by S to convert from intensity to optical power.

Finally, using [3.11],

$$\boxed{\frac{\partial N_2(z,t)}{\partial t} = \frac{N_2(z,t)}{\tau} - \frac{1}{\rho S}\left\{\frac{1}{h\nu_s}\frac{\partial P_s(z,t)}{\partial z}\right\}} \tag{3.17}$$

These are essentially the same equations presented without derivation in [25], and we'll use this model in a later chapter to analyze a particular form of the Er-doped fiber amplifier. But first we present the model of Desurvire, derived using the same development as above, with some further assumptions.

3.1.3 Desurvire's model

Two other simplifying assumptions enable us to obtain a closed-form solution to the above coupled, non-linear equations. In particular, we assume that the Er-doing profile is confined to a small central region of the fiber core, and that the sum of all normalized signal+ amplified spontaneous emission (ASE) powers is dominated by the normalized pump power at each fiber coordinate z. This is the unsaturated

gain regime: once the signal input is reduced below a certain level, the EDFA gain is independent of the signal input or output powers.

We do not repeat the analysis in [6]: the rate equations reduce to

$$\frac{dq}{dz} = -\alpha_p \frac{q}{1+q} - \alpha'_p q \tag{3.18}$$

$$\frac{dp_k}{dz} = \alpha_k \frac{1}{1+q}\left[\left(\frac{\eta_k - \eta_p}{1+\eta_p}\, q - 1\right) p_k + \frac{\eta_k}{1+\eta_p} 2q\, p_{0k}\right] - \alpha'_s p_k \tag{3.19}$$

where p and q represent optical powers at the signal and pump wavelengths normalized to the corresponding saturation power (of the order of magnitude of 1 mW for the EDFA parameters considered in this chapter)

$$P_{sat}(\nu) = \frac{h\nu S}{\tau\Gamma[\sigma_e(\nu) + \sigma_a(\nu)]} \tag{3.20}$$

The equivalent ASE powers are also normalized in terms of the optical bandwidth,

$$p_{0k} = \frac{h\nu_k\, B_o}{P_{sat}(\nu_k)} \tag{3.21}$$

We've defined the pump and signal Er-doping absorption coefficients in terms of the cross-section parameters,

$$\alpha_p = \rho_0\, \sigma_{ap} \Gamma_p \tag{3.22}$$

$$\alpha_k = \rho_0\, \sigma_{ak} \Gamma_k \tag{3.23}$$

where ρ_0 is the peak value of the Er-doping concentration and the (power-dependent) overlap integral factors, in terms of the power mode size $\omega_{k,p}$ at the signal or pump

wavelength, and the appropriate waveguide dimension a_0, are

$$\Gamma_p = 1 - \exp\left[-\left(\frac{a_0}{\omega_p}\right)^2\right] \tag{3.24}$$

$$\Gamma_k = 1 - \exp\left[-\left(\frac{a_0}{\omega_k}\right)^2\right] \tag{3.25}$$

Typically, the overlap factors are $\ll 1$ for $a_0 \ll \omega_{p,k}$.

In the above equations, the absorption coefficients $\alpha'_{s,p}$ account for fiber background loss at the respective frequencies of the signal and pump beams, and become important for distributed amplifiers, where the length of the gain section can be of the order of kilometers. η_s and η_p are the ratios of the emission to absorption cross-sections at the signal and pump wavelengths respectively.

The new rate equations are still nonlinear in the pump power, but for a given fiber length, it is possible to derive a relationship between the input and output pump powers as a function of the fiber signal gain. We shall use this in Section 3.3 to obtain conditions under which the absorption of pump power at a given signal gain is minimized.

3.2 Material characteristics

So far, we've talked about the basic physical processes that govern amplification in Er-doped glass fibers. The absorption and emission cross-sections $\sigma_a(\lambda)$ and $\sigma_e(\lambda)$ figure prominently in the rate equations, and as we've mentioned earlier, these parameters are primarily of phenomenological origin—they have been measured and tabulated for a range of wavelengths.

Our goal in this short section is to point out a few references in the literature for the parameters that will appear repeatedly in our discussion. The experimental determination of these parameters in a given situation is by no means a trivial exercise, and the references we cite will also provide typical numerical values for these

parameters. These values will, in turn, enable us to get a numerical feel for the algebraic (analytical) answers we get by studying the questions of network scale using Er-doped fibers.

Absorption spectra can be measured directly from the fiber using a white light source, and the results are usually quite accurate for large doping concentrations (e.g. $\alpha_s > 10$ dB/m). Part of our work will deal with fibers which are doped far more lightly, and for these low concentration fibers, experimental procedures have to account for the background loss. Measurements can be calibrated using spectral loss measurements away from the 1.5 μm resonance as described in [4].

There are other approaches as discussed in [6], one of which is by using the Fuchtbauer-Ladenburg formula, relating the peak cross sections to the effective line widths of the absorption and emission line shapes. The emission and absorption cross sections are linked in theory via McCumber's relation developed for phonon-terminated optical masers [16]. Giles' method [8] measures the peak cross section ratio by relating the small-signal gain to the loss coefficient at the signal wavelength. This is called the cutback procedure: the curves for the small-signal gain and for the loss coefficients are obtained by cutting small lengths of the fiber at the output end and measuring the corresponding output signal power.

The saturation powers P_{sat} at any wavelengths can be directly measured without first having to find out the mode size, fluorescence lifetime and cross sections. The general procedure is to measure the fiber transmission $P_{out}(\lambda)/P_{in}(\lambda)$ and fit the experimental points using an analytical model (e.g. the models we're about to describe). The best fit yields a numerical value for $P_{sat}(\lambda)$ as described in [6, pages 274–277].

There are several interesting aspects to this topic, and a considerable amount of research has been carried out over the last decade, but we will refer the reader to [6] for an overview and further sources in the literature.

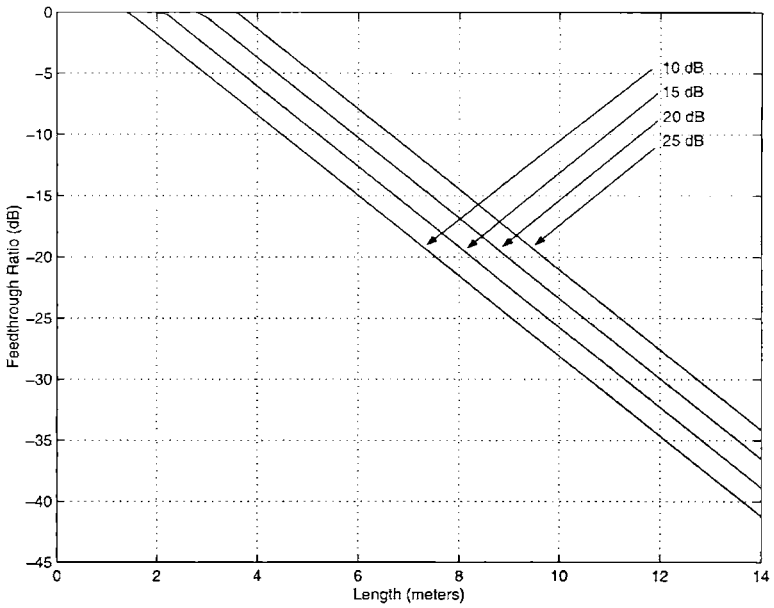

Figure 3-1: The ratio of normalized pump output to input q_l/q_0 (in dB) for the following typical 1480 nm pump parameters: $\alpha_p = 1.32\,\mathrm{m}^{-1}, \alpha_k = 1.50\,\mathrm{m}^{-1}, \eta_p = 0.23, \eta_k = 1.3, \alpha'_{p,k} = 0$

3.3 Pump propagation

We now investigate the evolution of the pump along an Er-doped fiber. In particular, we look for the conditions under which the absorption of the pump can be neglected. We'll see that complete pump transparency for a given signal gain G_k can be achieved only for a certain fiber length l_{opt} and only in the limit of infinite input pump power, However, even with finite, but large, input pump power, it's possible to achieve substantial gain for relatively short fiber lengths. We demonstrate the relationship between the input pump power and the resultant pump absorption for a desired signal gain.

Returning to [3.18] and [3.19], we can write an equation for the output of the pump power, q_l, as a function of the input pump power, q_0, the length of the Er-doped fiber, l and the signal gain, G_k,

$$\boxed{q_l = q_0 \exp\left(\underbrace{-\alpha_p \frac{1+\eta_p}{1+\eta_k}(1+\epsilon_k)C}_{B}\, l\right) \exp\left(\underbrace{\frac{\alpha_p}{\alpha_k}\frac{1+\eta_p}{1+\eta_k}}_{D} \log_e G_k\right)} \tag{3.26}$$

where $\epsilon_{p,k} = \alpha'_{p,k}/\alpha_{p,k}$ is the ratio of background to ionic absorption loss at the respective wavelengths, and

$$C = \epsilon_p - \frac{1+\epsilon_p}{1+\epsilon_k}\left(\frac{\eta_p - \eta_k}{1+\eta_p} + \epsilon_k\right) \tag{3.27}$$

We note that the above relationship, in the limit of negligible background losses, provides a correction to that stated in [2]. The C and D parameters defined in that paper should read

$$C = \alpha_s \frac{\eta_s - \eta_p}{1+\eta_s} \tag{3.28}$$

$$D = \frac{\alpha_p}{\alpha_s}\frac{1+\eta_p}{1+\eta_s} \tag{3.29}$$

The relationship between output and input pump powers at a given signal gain and for a particular length of Er-doped fiber is demonstrated in Figure 3-1 for a typical set of parameters at a 1480 nm pump. Note that $q_l/q_0 \leq 1$ with equality implying that the pump is not attenuated.

Further, the required input pump power, q_0, for a fiber of length l and gain G_k is given by

$$\boxed{q_0 = b\frac{e^{A_1 l} - 1}{e^{A_2 l} - 1} e^{A_2 l}} \tag{3.30}$$

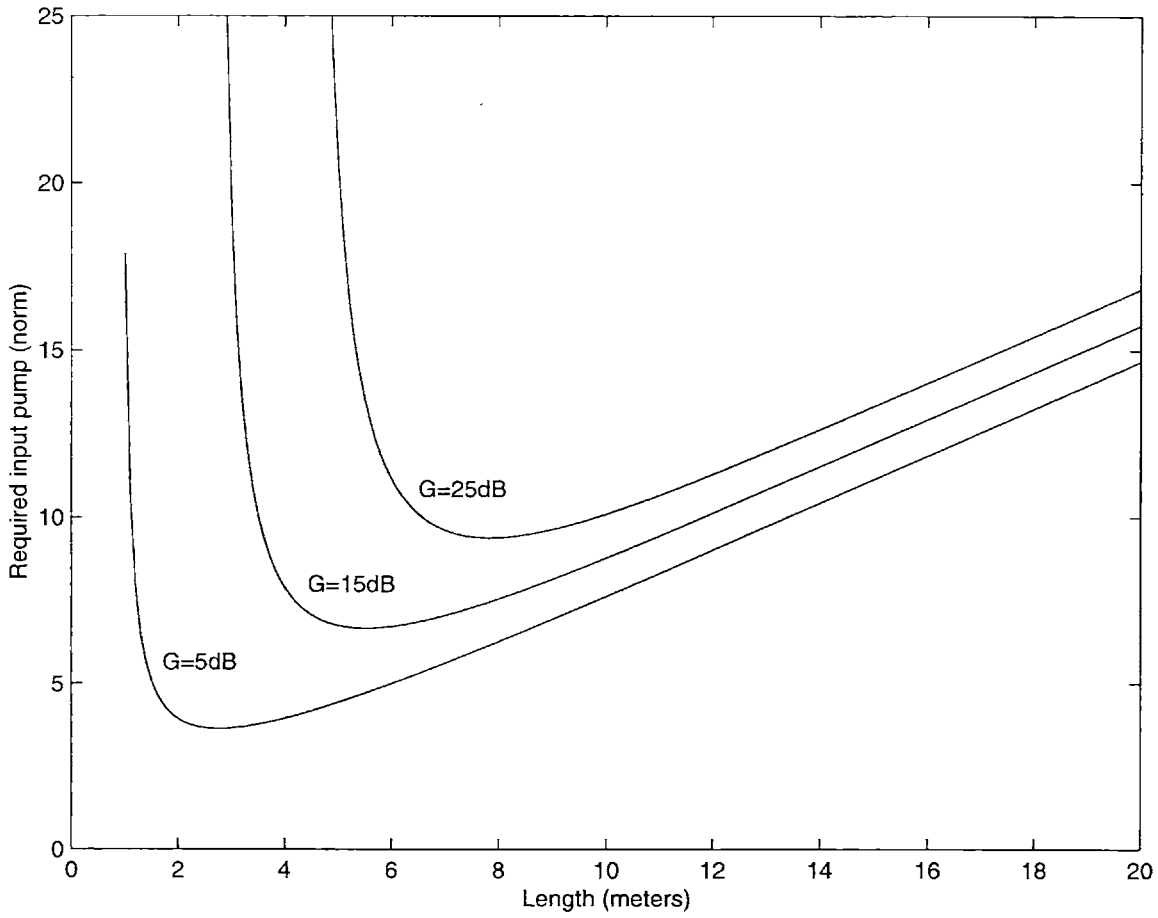

Figure 3-2: The minimum normalized input pump required to achieve a certain gain, G, as a function of the Er-doped fiber length, L, for the following typical 1480 nm pump parameters: $\alpha_p = 1.32\,\mathrm{m}^{-1}, \alpha_k = 1.50\,\mathrm{m}^{-1}, \eta_p = 0.23, \eta_k = 1.3, \alpha'_{p,k} = 0$

where

$$A_1 = \epsilon_p \alpha_p \frac{1+\eta_p}{1+\eta_k}\left(1+\epsilon_k+\frac{G_k}{\alpha_k L}\right) \tag{3.31}$$

$$A_2 = \alpha_p(1+\epsilon_p)\frac{1+\eta_p}{1+\eta_k}\left(\frac{\eta_k-\eta_p}{1+\eta_p}-\epsilon_k-\frac{G_k}{\alpha_k l}\right) \tag{3.32}$$

$$b = \frac{A_2}{\epsilon_p(A_1+A_2-\alpha'_p)} \tag{3.33}$$

From Figure 3-2, we note that in order to be able to achieve a certain level of gain (from any length of Er-doped fiber), the normalized input pump power must exceed a certain threshold, which can be found from [3.30] by setting $dq_0(l)/dl \mid_{l=l^*} = 0$ and then using [3.30] again to find $q_0(l^*)$.

3.3.1 Pump transparency

Dividing both sides of [3.26] by q_0 and setting the left-hand side equal to unity gives a condition for complete pump transparency,

$$Bl = D \log_e G_k \tag{3.34}$$

where B and D are defined in [3.26]. We can easily show that this corresponds to the asymptotic limit of the curves in Figure 3-2, i.e. for the particular $\hat{l}$ that solves [3.34], $q_o \to \infty$. To see this in the simplified case where we ignore background losses, we follow Desurvire in simplifying [3.30] to

$$q_0 = \alpha_p l \frac{Q_k}{1 - \exp[\alpha_p l(Q_k - 1)]} \tag{3.35}$$

where

$$Q_k = \frac{1 + \eta_p}{1 + \eta_k} \left(1 + \frac{G_k}{\alpha_k l}\right) \tag{3.36}$$

and for the condition of [3.34], $Q_k \to 1$ as $l \to \hat{l}$.

3.3.2 Finite input pump power

For a finite q_0, we can solve the transcendental [3.35] for the gain that a particular length of fiber achieves. The results of the numerical solution are plotted in Figure 3-3.

Note that for high input pump power, the gain increases exponentially (and linearly in dB) with the length of the Er-doped fiber: we expect this for situations where

the pump "bleaches" the fiber absorption at that wavelength. For lower pumps, the gain falls off at increased lengths because a significant portion of the pump is absorbed.

Starting with an available pump power budget, and a desired signal gain, we use Figure 3-3 to get the desired length of Er-doped fiber, and then use Figure 3-1 to evaluate the absorption in the pump for this length of fiber. For $q_0 = 100$, we see that the feedthrough ratio is typically less than -0.3 dB for G$\leq$25 dB. Note that for higher signal gain, our assumption of unsaturated gain is itself subject to revision.

While perfect pump transparency can't be achieved, the loss in the pump power, with an input pump power $q_0 = 100$, *can be limited.* A plot of the dB loss in pump power (called the feedthrough ratio, or FT) is plotted as a function of the signal gain for a family of q_0 curves in Figure 3-4.

For high input pump powers, we can achieve high gain with small pump loss, whereas to maintain the same level of pump loss (i.e. feedthrough ratio) for lower pump input levels, the gain must be reduced.

For remotely pumped amplifier chains, the optimum length of an EDFA doesn't follow the traditional definition. In particular, we don't wish to extract the maximum power from the pump, which leads to the"optimum length" being defined as that at which the output normalized pump power is $(1 + \eta_p)/(\eta_s - \eta_p) \simeq O(1)$.

Rather, we aim for the minimum length possible to achieve a certain amount of gain, given that the input pump power is many times the saturation pump power. This enables operation near the $l \to \hat{l}$ point, and the gain that we can achieve for a given length is given by the input pump power.

As Figure 3-3 shows, the gain at a particular fiber length converges quickly even as the input pump power increases asymptotically, and this limit is indeed the well-known $G(\infty)$ parameter.

Even with arbitrarily low signal powers, this (unsaturated) maximum EDFA gain is limited by both amplifier self-saturation and laser oscillation due to ASE. The

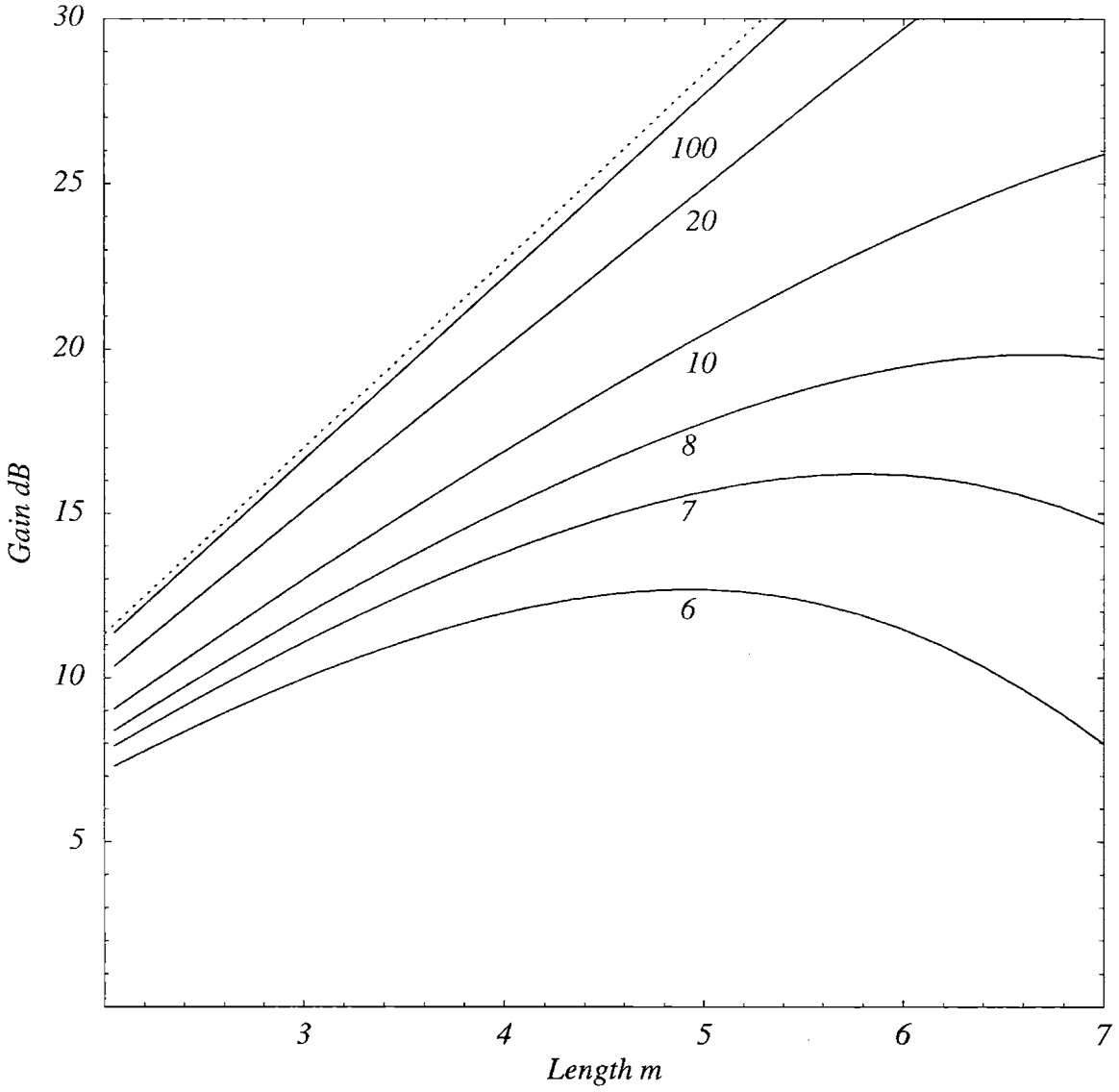

Figure 3-3: The maximum gain possible as a function of the Er-doped fiber length, given a normalized input pump q_0 (6, ..., 100) for the following typical 1480 nm pump parameters: $\alpha_p = 1.32\,\mathrm{m}^{-1}, \alpha_k = 1.50\,\mathrm{m}^{-1}, \eta_p = 0.23, \eta_k = 1.3, \alpha'_{p,k} = 0$. The dotted line corresponds to the gain at 'infinite q_0', $G(\infty)$

second limiting factor is particularly sensitive to backscattering from optical elements e.g. in chains of optical amplifiers.

As the pump propagates along the undoped fiber, it is attenuated just as the signal is, and therefore, the input pump to Er-doped sections further down the line will be less than the original input pump power. Figure 3-4 shows that in order to let a sufficient amount of the pump propagate to a subsequent Er-doped section, we may have to reduce the gain by *reducing the length of the Er-doped section* so as not to attenuate the signal too much.

This is particularly true of distribution networks, where the distance between stations is small (e.g. 5–10 km), and a loss of 1–2 dB in the Er-doped section is comparable to the propagation loss of the pump in the undoped region.

3.4 Alternate form of the SNR optimization problem

In Chapter 2 we derived the optimum allocation of gain and loss along a chain of optical amplifiers under quite general constraints [2.19] and the form of those constraints was further discussed in Appendix A. In this section, we use the physical model of the erbium-doped fiber amplifier that we have developed to show the equivalence between the problem that was solved earlier and one that has a more obvious physical interpretation.

We consider again a chain of alternating gain and loss sections, as shown in Figure 3-5, but now specify that the chain of amplifiers is remotely pumped. The constraint that arises naturally in this formulation is that of limited pump power: we assume that the total available normalized pump power (at the input to the chain) is q_0. We will show that the previous solution i.e. increase the gain G_1 of the first stage to the maximum allowed, is still the optimum one.

As before, we do not discuss in this report *what* the maximum allowed gain is:

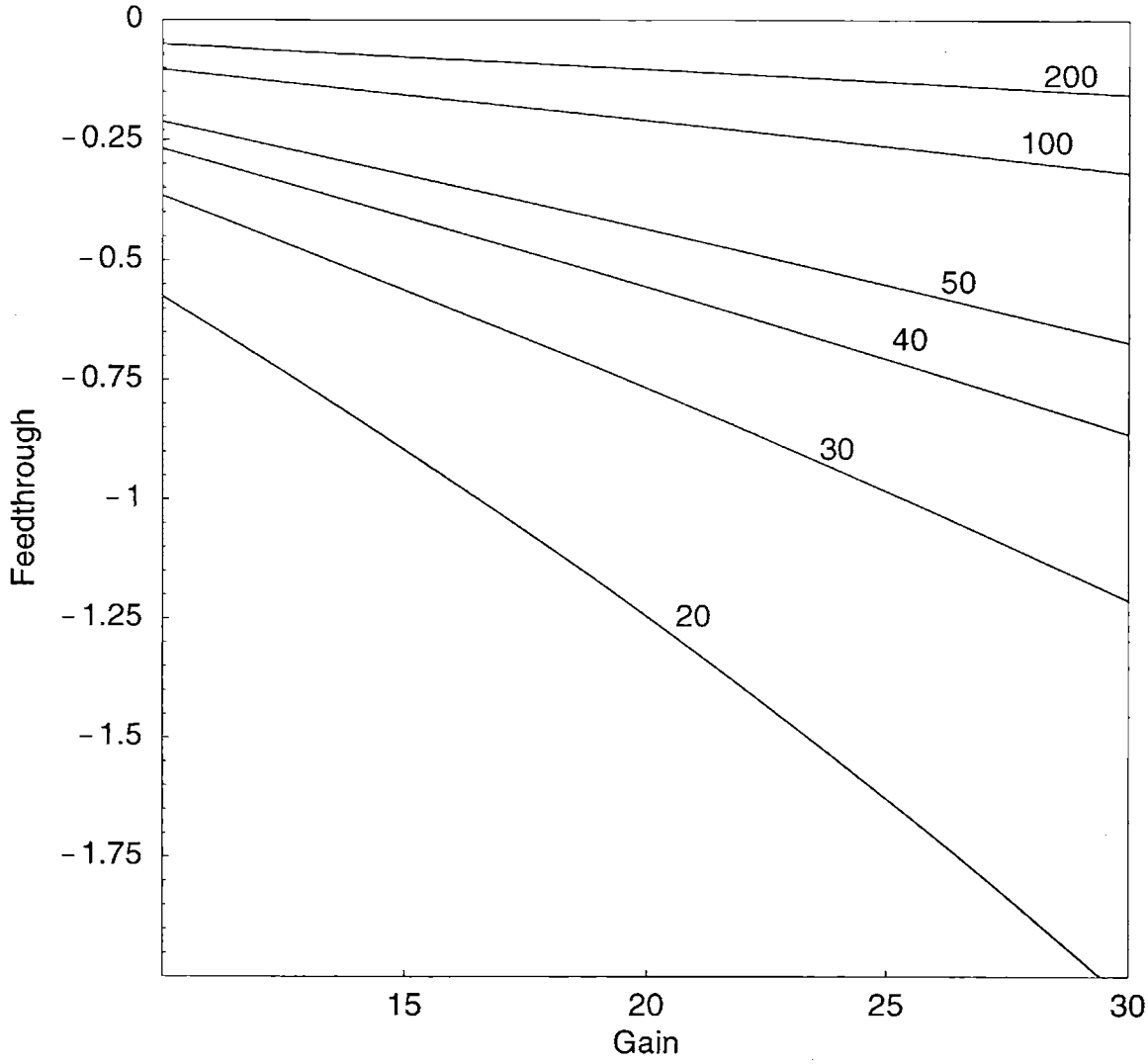

Figure 3-4: Feedthrough (FT), defined as the reduction in the pump power (in dB) as a function of the signal gain. The curves are plotted for a family of normalized input pump powers q_0 (20, ... ,200) and for the following typical 1480 nm pump parameters: $\alpha_p = 1.32\,\mathrm{m}^{-1}, \alpha_k = 1.50\,\mathrm{m}^{-1}, \eta_p = 0.23, \eta_k = 1.3, \alpha'_{p,k} = 0$.

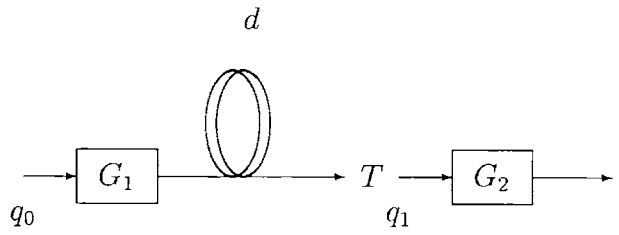

Figure 3-5: A chain of two amplifiers G_1 and G_2 connected by undoped fiber of length d with excess loss T between the elements

typically, nonlinear effects in propagation and in amplification must be considered for a complete treatment, which substantially complicates the problem. For our purposes, it's sufficient to explicitly acknowledge the presence of such a constraint. Our discussion will consider a chain of two amplifiers—the results can be easily extended to more elements as was done in the earlier optimization problem.

We assume, naturally, that the pump is not fully exhausted in the first amplifier itself, so that a second gain stage can be supported. Let the spacing between the stages be d, and the excess loss between the stages be T (e.g. coupling between the doped and undoped sections of fiber). Since the pump is not fully exhausted in the first gain stage, the (normalized) pump power at the output of the first amplifier is $\mathrm{FT}\, q_0$ where FT is the feedthrough ratio defined by [3.26] as

$$\mathrm{FT} = \frac{q_l}{q_0} = e^{-Bl} G_1^D \tag{3.37}$$

where both B and D are positive and l is the length of erbium-doped fiber that provides G_1.

The pump that is input to the second stage, accounting for propagation and excess

losses, is

$$q_1 = e^{-\alpha_p' d} T \, \mathrm{FT} \, q_0 \tag{3.38}$$

$$= \underbrace{e^{-\alpha_p' d} T e^{-Bl}}_{P} G_1^D \, q_0 \tag{3.39}$$

where α_p' is the background absorption coefficient at the pump wavelength in the undoped sections of fiber.

We assume that this remaining pump is fully utilized at the second (and final) stage to provide the maximum possible gain. From Figure 3-6, we see that the relationship between the desired gain G and the minimum required (normalized) pump power for such a gain to be possible (for any length of fiber) is linear. The exact dependency is not important: we model the relationship as phenomenological rather than carrying out tedious algebraic manipulation

$$\hat{q} = \alpha G + \beta \tag{3.40}$$

where, for the particular values used in Figure 3-6, $\alpha = 1.08$ and $\beta = 3.30$. Physically, β represents the pump power necessary to overcome absorption in the erbium-doped fiber at the signal wavelengths.

The gain that we can obtain at the second stage is

$$G_2 = \frac{1}{\alpha} \left[P G_1^D q_0 - \beta \right] \tag{3.41}$$

and the product $\Pi = G_1 G_2$ used in the constraints of [2.19] is

$$\Pi = G_1 \frac{1}{\alpha} \left[P G_1^D q_0 - \beta \right] \tag{3.42}$$

Consider what happens if we decrease the gain of the first stage to

$$G_1' = G_1 - \Delta G_1 \tag{3.43}$$

where, for our analysis, ΔG_1 is small compared to G_1. The pump power now available to feed the second gain stage is

$$\begin{aligned} q_1' &= P(G_1 - \Delta G_1)^D q_0 && (3.44) \\ &\simeq P G_1^D \left(1 - D\frac{\Delta G_1}{G_1}\right) q_0 && (3.45) \end{aligned}$$

where we note, in passing, that

$$\left|\frac{q_1'}{G_1'}\right| > \left|\frac{q_1}{G_1}\right| \tag{3.46}$$

The gain that we can achieve with this pump power is

$$G_2' = \frac{1}{\alpha}\left[P G_1^D \left(1 - D\frac{\Delta G_1}{G_1}\right) q_0 - \beta\right] \tag{3.47}$$

Consider $\Pi' = G_1' G_2'$,

$$\begin{aligned} \Pi' &= \frac{1}{\alpha}(G_1 - \Delta G_1)\left[P G_1^D \left(1 - D\frac{\Delta G_1}{G_1}\right) q_0 - \beta\right] && (3.48) \\ &\simeq \Pi - \frac{1}{\alpha}\left[(D+1) P G_1^D q_0 - \beta\right] \Delta G_1 && (3.49) \end{aligned}$$

where we neglect terms $O(\Delta G^2)$.

As typical numerical values, $q_0 = O(100)$, $P = O(0.1)$, $D = 0.47$, $\alpha = 1.08$ and $\beta = 3.3$. Since $G_1 \geq 1$ by definition, we have shown that $\Pi' < \Pi$.

As discussed in Appendix A, reducing the product $\Pi \rightarrow \Pi'$ is suboptimal (assuming that the signal shot noise is not dominant) when both Π and Π' are feasible alternatives. In physical terms, it's better to use as much of the pump as possible at

the first stage. In the optimal case (if possible) $G_2 = 1$ and so

$$G_1 = \left[\frac{e^{Bl}e^{\alpha'_p d}}{Tq_0}(\alpha + \beta)\right]^{1/D} \tag{3.50}$$

Subsequent optimization over d (again, if realizable) will result, as in Chapter 2, in $Te^{-\alpha'_p d} \rightarrow 1$ i.e. $d = 0, T = 1$. Now we realize that if we can physically alter the geometry of the problem to use length l of erbium-doped fiber to provide this new gain $\hat{G}_1$, then

$$\hat{G}_1 = \frac{1}{\alpha}(q_0 - \beta) \tag{3.51}$$

This result shows the generality of our earlier approach, which was derived independently of the physical considerations of pumping in amplifier chains. Of course, we have specifically disallowed any nonlinear effects, which will probably determine the maximum gain allowed at any one stage.

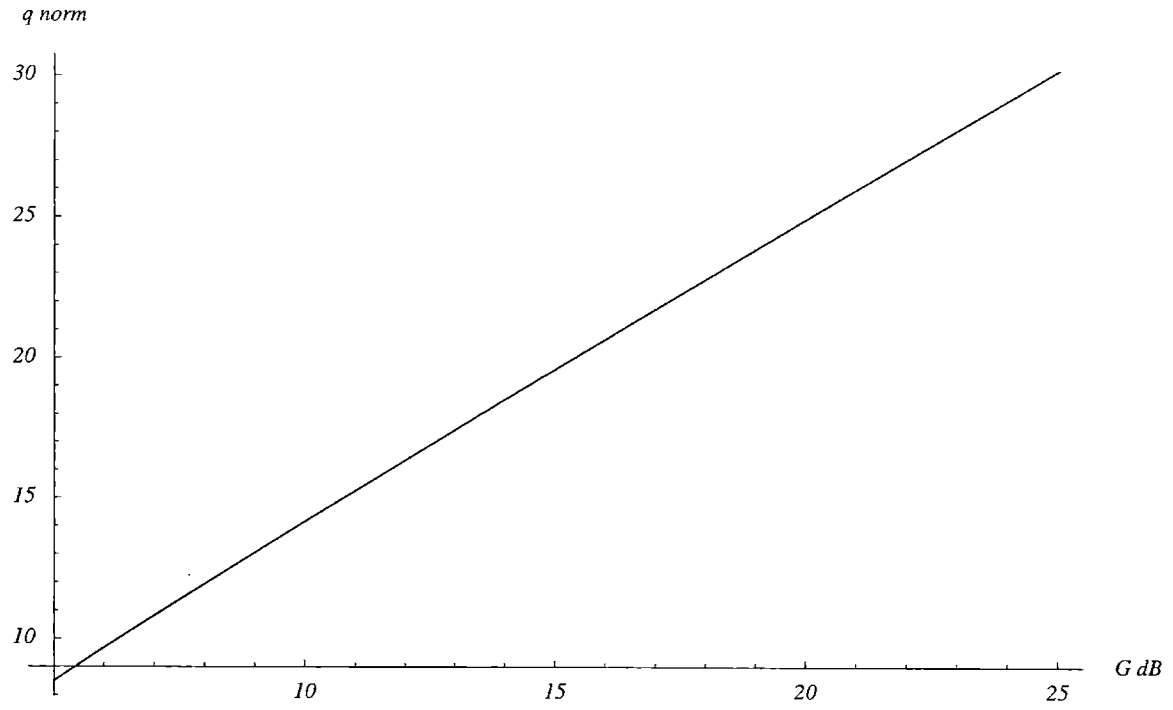

Figure 3-6: The minimum required (normalized) input pump power for a desired gain, assuming that the EDFA length is optimized for minimum required pump power. We use the following typical 1480 nm pump parameters: $\alpha_p = 1.32\,\mathrm{m}^{-1}, \alpha_k = 1.50\,\mathrm{m}^{-1}, \eta_p = 0.23, \eta_k = 1.3, \alpha'_{p,k} = 0$. The slope is 1.08 and the intercept on the ordinate is 3.31 (calculations in Mathematica$^{\mathrm{R}}$)

Chapter 4

Parametric design of distribution networks

On the other hand, when we have decided upon our premises,
we have to build up again as much as may seem necessary
of the data previously analyzed,
and as many other consequences of our premises
as are of sufficient general interest to deserve statement.
—B. RUSSELL, Preface to *Principia Mathematica* (1910)

We use the results of the previous chapters to evaluate the capabilities of remotely-pumped Er-doped fiber segments in an optical network. We consider distribution networks with one signal and one co-propagating pump beam, and derive bounds on the number of receivers that can be supported, for typical systems parameters. The central assumption is that a single pump beam is input at the head of the distribution bus, and once the utility of this pump is exhausted, we have reached the limit of our network. It's simple to extend our results if a second pump beam is then injected.

In the first section, we discuss the numerical values of the signal-to-noise ratio along a typical chain of optical amplifiers. In particular, we introduce the set of parameters that describes the APD receiver, the representative of a user along this optical network. Using these parameters, we discuss the bus distribution network, and various extensions to this model, with varying degrees of realism and optimality. Our goal is to evaluate how many receivers we can support using a bus distribution network. As will be obvious in retrospect, these results are highly dependent on the numerical values of the parameters.

4.1 Signal-to-noise ratio

We assume that a 1480 nm pump is input at one end of a chain of optical amplifiers (see Figure 2-1), each of gain G_k, where a passive section of (single-mode) fiber resulting in attenuation L_k connects amplifiers k and $k+1$, or, in the final section, the last amplifier to a detector.

As discussed earlier, if the input signal power is P_0, the signal and ASE noise powers at the output of such a chain (i.e. after the N^{th} element) are

$$\mathcal{P}_N = P_0 \prod_{k=1}^{N} G_k L_k \tag{4.1}$$

$$\mathcal{N}_N = \mu\, h\nu\, B_o \sum_{k=1}^{N} (G_k - 1) L_k \prod_{j=k+1}^{N} G_j L_j \tag{4.2}$$

and we will specifically account for a particular contribution to L_k in a subsequent section. If the gain of each stage is uniform and balances the loss in the following section of fiber i.e. $G_k = 1/L_k \equiv G$ then $\mathcal{N}_N$ grows linearly with the number of amplifiers N,

$$\mathcal{N}_N = \mu\, h\nu\, B_o\, (1 - 1/G) N \tag{4.3}$$

Such a "transparent bus distribution network" is not necessarily the optimum allocation of gains and losses (as analyzed in earlier chapters) but is simple to analyze in closed form—other situations may be analyzed numerically as appropriate.

We convert the optical powers given above to the photocurrents they generate when incident upon an APD—the expected value and variance of this photocurrent defines the signal-to-noise ratio. Using our simple definition of the Personick Q-factor, or the more sophisticated [2.17], we can then obtain the corresponding probability of error (i.e. bit error rate). Note that the variances of the photocurrent are not the same for the mark and space symbols. The ASE-signal beat noise typically dominates

in the presence of an optical pulse, but obviously not when a pulse is absent.

For purposes of our calculation, we assume a transparent bus distribution network and that the following parameters describe the system,

Electronic bit rate (data)	B	10 Gbits/s
Optical bandwidth	B_o	$2\,B$
Receiver electronic bandwidth	B_e	B
Inversion factor	μ	1
Detector quantum efficiency	η	0.8
APD avalanche gain	$\bar{g}$	17
APD excess noise factor	x	0
Effective receiver temperature	T_e	$4 \times 290\,K$
Detector Resistance	R	$50\,\Omega$

which yields a receiver sensitivity of about -28 dBm (see Figure 4-2).

Defining the various mean-square current components at the output of an APD, we can evaluate the SNR as a function of the number of amplifiers, for a given signal input power. This implies that we can also evaluate the BER, e.g. using Personick's approximation that would be observed by a detector based on this APD after each amplifier. The results are shown in Figure 4-1 for a single signal channel with input power $P_0 = 0.02$ mW as would be typical in such a network.

Alternatively, we can consider P_0 as the variable of interest, and ask what the required P_0 is for a given SNR target: this defines the receiver sensitivity as shown in Figure 4-2. For example, consider two cases: one with $N = 10$ amplifiers and the second with $N = 100$. As the ASE noise increases with the number of amplifiers, we would expect that a higher input signal power would be required to attain the same SNR threshold in the second case.

As expected, the growing ASE component results in a higher P_0 requirement with increasing numbers of amplifiers, and the increase is quite linear in dB for $N > 10$. For the first few stages, the receiver thermal noise dominates the ASE components.

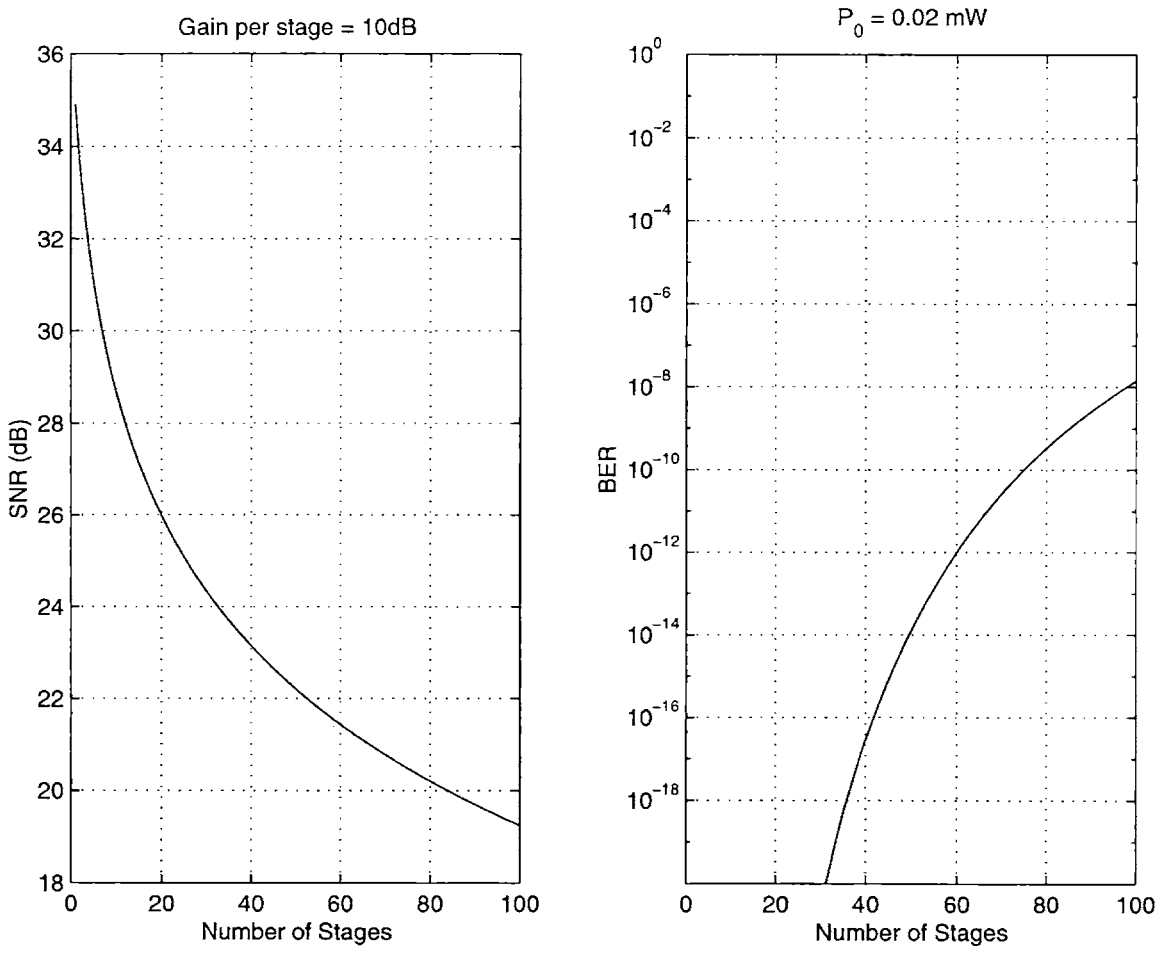

Figure 4-1: The signal-to-noise ratio (SNR) and Personick's approximation to the bit-error rate (BER) as measured by an APD ($\bar{g}$=17, x=0) as a function of the number of repeated amplifiers. The gain of each stage is 10 dB and balances the attenuation in the signal in traversing the next section of fiber. The input signal power is 0.02 mW

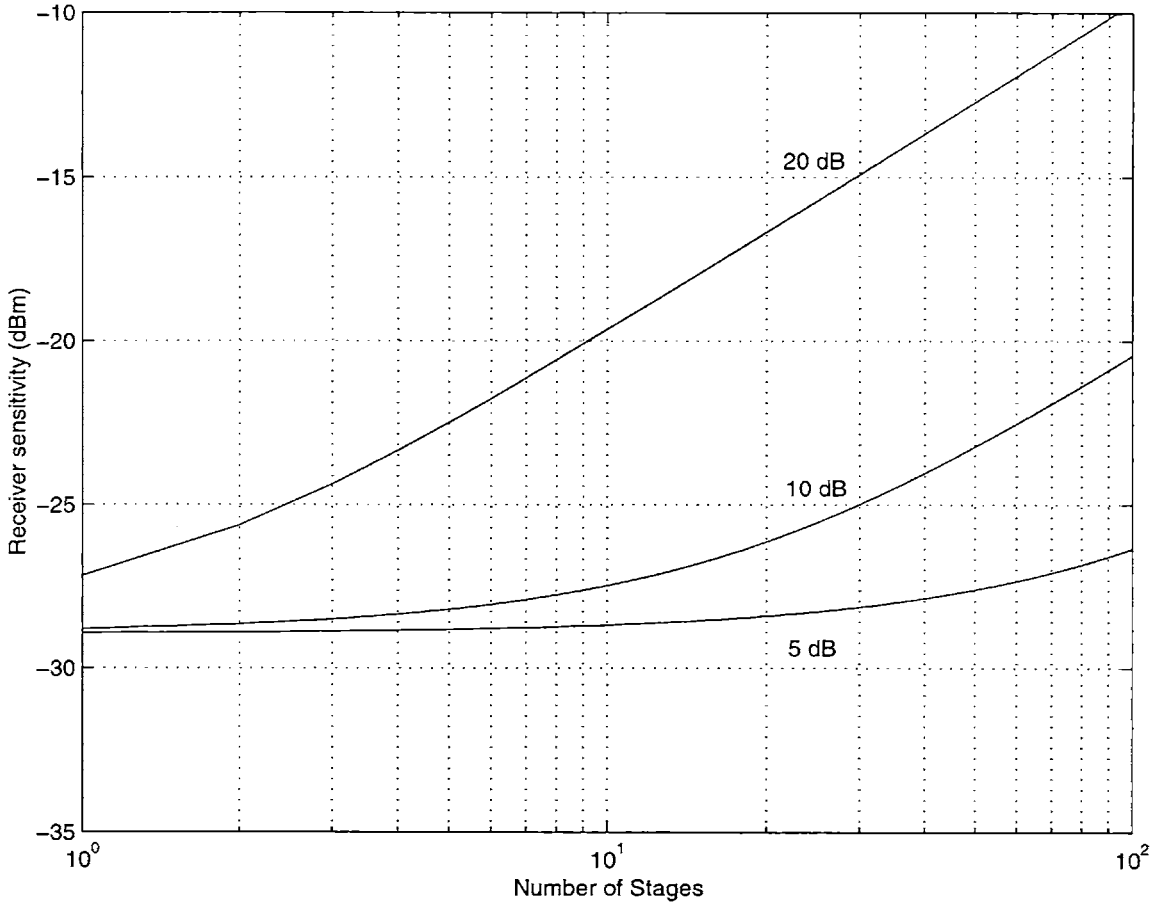

Figure 4-2: The required incident power on the APD to maintain a target SNR (receiver sensitivity) as a function of the number of repeated amplifiers for three different conditions: the gain of each stage is 5 dB, 10 dB or 20 dB and in each case, balances the attenuation in the signal in traversing the next section of fiber.

Plots such as this are common in the literature, and our use of them will be limited to two simple interpretations. Along an amplifier chain for a particular G, this plot tells us what the required signal level is for a sufficient SNR. Correspondingly, if we specify the acceptable minimum receiver sensitivity, we can determine how many stages of amplification are allowed i.e. the length of the network.

4.2 Bus distribution

It's convenient to relate the power that is incident on the APD to the signal power in the optical fiber. In an optical bus network, the receivers stationed along the bus each divert (or 'tap') a small fraction of the signal power e.g. by using a weakly-coupled resonator. This tap fraction is then incident on the APD for that receiver station.

Since the gains and losses are balanced throughout the network, the power required by Figure 4-2 incident on the APD, expressed as a fraction of the input signal power, tells us what fraction of the signal power needs to be diverted from the bus to be incident on the APD: this number is in fact the tap fraction. For example, the receiver sensitivity after twenty amplifiers in a chain with G=10 dB is about -26 dBm. If the input signal level is -5 dBm, we require a weakly-coupled resonator with a tap fraction no smaller than -21 dB so that the SNR constraint is met.

Also, from an earlier chapter, we've seen that for a sufficiently high (normalized) input pump power q_0, the absorption of the pump in the amplifier can be neglected. Nevertheless, there will be some loss resulting from the coupling between the fiber and the tapping resonator etc. We assume, as we did in our analysis of the passive distribution network, that at each stage, the pump suffers a loss T.

The propagation equation for the pump can be written as

$$q_N = q_0 \prod_{k=1}^{N} L_k T \tag{4.4}$$

where L_k is the absorption loss of the pump in the SMF sections between amplifiers.

We can write this as $L_k = \exp(-\alpha'_p\, d_k)$ where d_k is the length of the SMF section.

For a 1480 nm pump, $\alpha'_p \simeq 0.5$ dB/km, but is much higher (about 1.2 dB/km) for a 980 nm pump. Since we neglect pump absorption in the EDF as compared to T, the additional loss in the pump is almost entirely from these SMF sections, a 980 nm pump can travel a far shorter distance than a 1480 nm pump. At the same time, we've seen from Figure 3-3 that beyond $q_0 > 100$, the gain doesn't increase very much: we're already pretty close to the asymptotic limit of infinite q_0. So the additional pump power that we can inject with present-day 980 nm sources doesn't really help very much in the early stages either.

Let's assume that the stages are evenly spaced: $d_k = d$, and so, the pump that is input to the $N + 1^{\text{th}}$ stage is

$$q_N = q_0 T^N e^{-N\alpha'_p d} \tag{4.5}$$

which implies that

$$\ln\left(\frac{q_N}{q_0}\right) = N(\ln T - \alpha'_p d) \tag{4.6}$$

In our parametric design of the network, we specify that the gain of each stage should be G, and this implies, according to Figure 3-3 that, at any stage, we need a minimum q so that this gain can be achieved. Therefore, defining $\hat{q}$ as the minimum (normalized) pump power required to be able to provide gain G, the number of taps that we can support is

$$\aleph = \left\lfloor \frac{\ln(\hat{q}/q_0)}{\ln T - \alpha'_p d} \right\rfloor \tag{4.7}$$

As a numerical example, assume $G = 10$ dB, and from Figure 3-3, $\hat{q} = 5$. Suppose that at the input to this chain, we're given $q_0 = 175$, and the attenuation at each stage $T = -0.5$ dB. Assume that the inter-amplifier spacing $d = 50$ m, and $\alpha'_p = 0.5$ dB/km for a 1480 nm pump. Then, $\aleph = 20$ i.e. we can permit twenty taps along this fiber.

There is another factor to consider: the signal gain at each stage is 10 dB of which we expend a negligble fraction in propagation (assuming that $\alpha_s' = 0.5$ dB/km) between amplifiers. In order to maintain a constant signal level, we have to distribute the 10 dB of signal gain at each tap among as many users as possible. We show two of the many choices available in Figure 4-3: the first is rather too optimistic to be of practical value, and we are quite generous in allowing for nonidealities in the second case. A practical implementation in a wide variety of configurations should have performance between that evaluated for our examples.

Assuming an ideal star distribution network at each tap (which we call a star subnetwork), we have from Figure 4-2 that the minimum detectable power at the twentieth tap is -26 dBm. For simplicity, we'll use the same figure as a conservative estimate for each of the previous taps as well. Let us assume that the input signal power is -5 dBm (which is about as high as we can let it be, if we want to avoid the nonlinear regime, given that the signal level rises by 10 dB after amplification). Then, we can support no more than $10^{3.1} \approx 1250$ receivers at this stage. Since there are twenty stages in all, the theoretical upper limit on the total number of receivers that can be supported with this set of parameters (using a uniform tap fraction) is 25000.

We can right away see that this is quite impractical. How does one divide a -5 dBm signal among 1250 APDs at one tap location without incurring any losses? A slightly better, but still quite optimistic, scenario is to assume that at each tap location, we allocate the expendable gain ($10 - 0.025 \simeq 10$ dB in the above example) by a simple distribution tree (which we call a tree subnetwork) of branching factor b. In other words, all the photodiodes are located at the terminal vertices of the tree, and at each split, the tree opens up into b branches, with an excess loss of Δ (in dB) for an imperfect split. Counting the tap from the bus as an edge, the graph is b-regular except, of course, for its terminal vertices. At each node, therefore, we attenuate the dropped signal power by $(10 \log_{10} b + \Delta)$ dB. We note that Δ can account for both

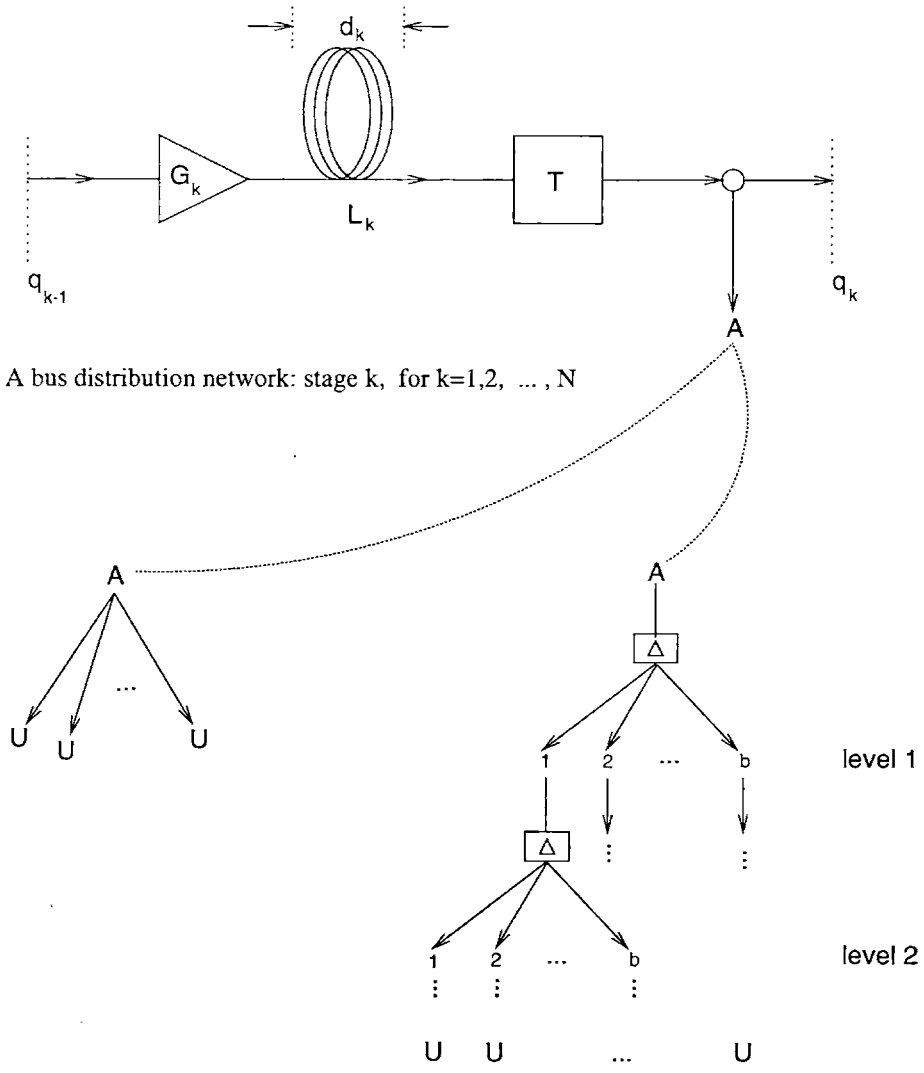

Figure 4-3: A bus distribution network, comprised of a main bus with remotely-pumped EDFAs, whose k^{th} stage is shown, with input pump power q_{k-1}, and output pump power q_k. Gain G_k is followed by a length d_k of undoped fiber resulting in attenuation L_k and is further followed by an "excess loss" T. At each stage, a sub-network is introduced to connect to users U: two examples are an ideal star and a b-regular tree of α levels with an excess loss of Δ per stage.

coupling and absorption losses e.g. a small but finite propagation distance between successive levels of splitting, as would be necessary in any practical device.

Let P_d be the signal level (in dBm) that is tapped [i.e. input signal power (dBm) + expendable gain (dB)] and P_{min} the receiver sensitivity (in dBm). Then the number of receivers we can support at each tap is b^α where

$$\alpha = \left\lfloor \frac{P_d - P_{min} - \zeta}{10 \log_{10} b + \Delta} \right\rfloor \tag{4.8}$$

where ζ is the excess loss suffered by the signal encountered at each tap (e.g. at a WDM demux).

For example, if we assume $b = 2$, $\Delta = 1$ dB, $P_d = +5$ dBm, $P_{min} = -25$ dBm, and $\zeta = -1$ dBm, then $\alpha = 7$. The number of receivers that we can support by this seven-level (binary) tree is $2^7 = 128$. Accounting for each of the taps, the total number of receivers that can be supported is $20 \times 128 = 2560$ which is quite a bit more conservative. Nevertheless, our assumptions have been quite simple, and a network designer who wishes to use this method will be well-advised to consider any major sources of nonideal behavior into account.

We consider a parametric approach to designing a distribution network based on a fiber bus. We stipulate that the gain of each amplifier stage is 10 dB, the inter-amplifier spacing is 5 km, and that the signal level at each amplifier input remains constant. Consequently, 10 dB of the signal is dropped at each tap and we can support twenty such taps before the pump is too weak to provide another gain stage (of 10 dB gain).

At each tap location, the best we can do is to distribute the dropped signal power by a lossless star subnetwork among about 1250 users. But this is highly impractical: not only will any real distribution device have some inherent loss, there are few applications of a distribution network that requires 1250 users to be situated at exactly the same location.

To remedy this, we consider a tree subnetwork to distribute the 10 dB dropped signal power while allowing a modest loss at each splitting stage. The total number of receivers that can be supported this way (about 2500) is substantially increased over that possible with a passive (unamplified) bus distribution network, which can typically support about 40 detectors under the same assumptions.

A more careful approach is to account for the different detection thresholds for each of the eight stages, using Figure 4-2. This implies that we use different tap fractions at each of the receivers, though this may not be practical in a commercial implementation. Considering an ideal star subnetwork, the total number of users across all the 20 stages is over 36000 (up from 25000) wheras for the tree subnetwork approach, we can suport about 4200 users (up from 2560). The floor operation in the above formulae can provide substantial advantage to the use of non-uniform taps for the tree subnetwork approach: another level of splitting results in a factor of b increase in the number of users for that subnetwork.

As we mentioned at the start of this chapter, this is essentially an *ad-hoc* approach to designing a bus distribution network, with little attempt at overall optimization. We've picked $G = 10$ dB merely as representative of the highest gain we can achieve without entering the nonlinear regime, in accordance with our earlier results on the optimum allocation of gains and losses in an optical amplifier chain. And we haven't optimized the use of the pump power for all stages: in the above example, even though the pump power after the twentieth stage is too weak to provide a gain of 10 dB, it can nevertheless provide a somewhat smaller gain, and a still smaller gain at a stage after that. Obviously, increasing the number of taps will increase the number of receivers even if we cannot support a gain of 10 dB at each of the added stages.

What complicates optimization arguments applied directly to the above expressions is the floor operation in several of the formulae. It's inviting to play games

with the parameters, adjusting them a little one way or the other to squeeze in an additional stage, but such a procedure offers no insight into what the upper limits of bus distribution networks are.

Moreover, it's usually the case that each tap corresponds to only one receiver. Assuming that successive receivers are 50–100 m apart, it isn't very reasonable to stipulate 10 dB of gain and then have to couple out nearly all of the amplified signal so that the signal doesn't encounter the (high-power) nonlinear regime. networks, the inter-receiver spacing is quite a bit closer than 5 km.

What we would like is to add only enough gain to offset the propagation losses and the tap fraction for a single receiver: both of these may be quite a bit less than 1 dB for receivers spaced apart by 10–20 meters, as in a typical computer network. The Desurvire model isn't quite suited to calculations of this sort, and we use instead the model we've developed from first principles in the previous chapter. As we will see, it's simplest to account for these taps explicitly while modeling the characteristics of the entire network as one single distributed amplifier.

4.3 Two-level example

In this section, we apply the above formulation to determine the number of users that a particular network architecture[1] can support. The system we consider is a hybrid of the models discussed so far, and as shown in Figure 4-4, has two levels of bus distribution lines from the access node to the end users. Amplifiers are present only along the main bus, and the subsidiary buses are passive.

We represent the fraction of signal power tapped from the main bus to start the j^{th} subsidiary bus by t_j and the fraction of signal power further tapped from that subsidiary bus to feed an end user by α_j. For signal power P_0 input at the access node (i.e. the head of the distribution network), the signal and ASE powers after the

[1] architecture due to V.W.S. Chan, MIT

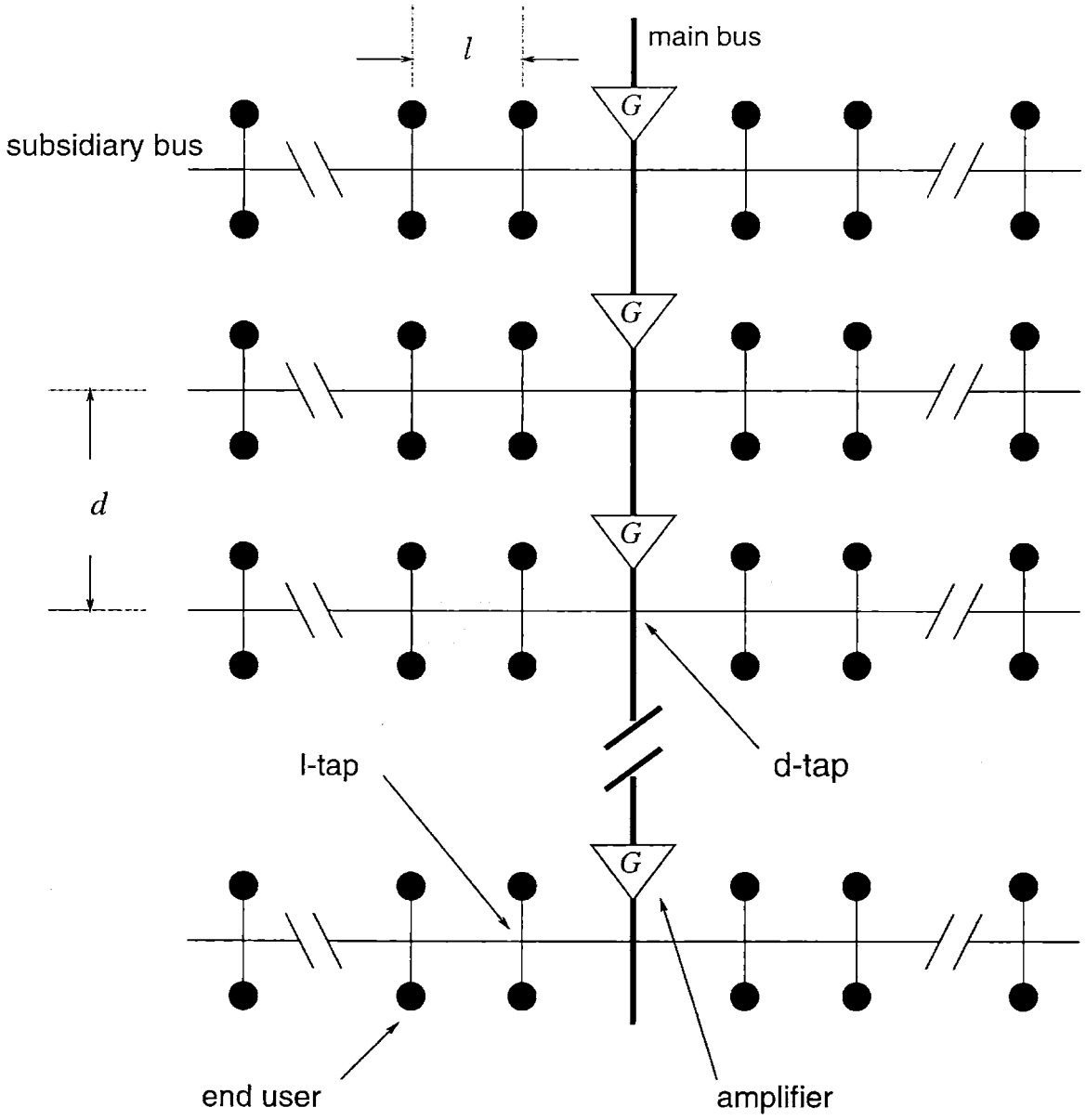

Figure 4-4: A two-level distribution network, comprised of a main bus with remotely-pumped EDFAs and passive subsidiary buses at regular intervals of d km. End users are situated along the subsidiary buses at regular intervals of l km. The signal and pump beams are input in the main bus at the top of the network, and the network is assumed to be unidirectional for purposes of analysis.

N^{th} stage are

$$\mathcal{P}_N = P_0 \prod_{k=1}^{N} (1-t_k)(1-\zeta)G_k L_k \tag{4.9}$$

$$\mathcal{N}_N = \mu\, h\nu\, B_o \sum_{k=1}^{N} (1-t_k)(G_k-1)L_k \prod_{j=k+1}^{N} (1-t_j)G_j L_j \tag{4.10}$$

We assume for our model that the taps come after the gain and the ζ-loss stage, but before the L attenuation so that the corresponding fractions that enter each of the pair of N^{th}-stage passive subsidiary buses are

$$p_N = \frac{1}{2}\mathcal{P}_N \left(\frac{t_N}{1-t_N}\right) 1/L_N \tag{4.11}$$

$$n_N = \frac{1}{2}\mathcal{N}_N \left(\frac{t_N}{1-t_N}\right) 1/L_N \tag{4.12}$$

For a "transparent" distribution network, we use only enough amplification to overcome the attenuation in the signal,

$$(1-t_k)G_k L_k = 1 \qquad \forall\, k = 1, 2, \ldots, N \tag{4.13}$$

and dropping the indices from G_k and L_k and generalizing the index N to the more conventional i, the above expressions simplify to

$$p_i = P_0\, G\, t_i \tag{4.14}$$

$$n_i = \mu h\nu\, B_o\, N(G-1)t_i \tag{4.15}$$

Equation [4.7] tells us how many subsidiary buses we can support from the main bus. If d is the spacing between two consecutive subsidiary buses (assumed uniform), and each gain stage provides uniform gain G, then the transparency condition dictates

$$t = 1 - \frac{1}{G} e^{-\alpha_s' d} \tag{4.16}$$

where α'_s is the background absorption coefficient at the signal wavelength. From this equation, the (uniform) tap fraction can be determined.

Following the notation of Figure 4-4, let the inter-user spacing along a subsidiary bus be l, and the minimum detectable power for any user along the i^{th} subsidiary bus (dictated by [4.15]) be $\hat{P}_i$. Along this simple passive distribution bus, we reach the limit on the number of pairs of users $\hat{m}_i$ when

$$2\,\hat{P}_i = \frac{1}{2} p_i \left[e^{-\alpha'_s l}(1 - 2\alpha_i) \right]^{\hat{m}_i} \tag{4.17}$$

where p_i is obtained from [4.14].

Rewriting the above expression, the total numbers of users that can be supported per "d-tap" along the main bus (recognizing that there are two end users per "l-tap", along the subsidiary bus and two subsidiary buses per d-tap) is

$$\hat{M}_i = \left\lfloor \frac{\log 4\dfrac{\hat{P}_i}{p_i}}{\log\left(1 - 4\dfrac{\hat{P}_i}{p_i}\right) - \alpha'_s l} \right\rfloor \tag{4.18}$$

At each end user along the i^{th}, the tap fraction is

$$\alpha_i = \frac{\hat{P}_i}{\frac{1}{2} p_i} \tag{4.19}$$

since there are two subsidiary buses for each tap (contributing p_i in total power) along the main line where $\hat{P}_i$ is determined by the signal-to-noise ratio criterion.

If we are to account for an excess loss δ per l-tap along the subsidiary bus, analogous to the role of T in [4.4], [4.17] becomes

$$2\hat{P}_i = \frac{1}{2} p_i \left[e^{-\alpha'_s l}(1 - 2\alpha_i)\delta \right]^{\hat{m}_i} \tag{4.20}$$

and therefore, the number of users per tap [4.18] is

$$\hat{M}_i = \left\lfloor \frac{\log 4\dfrac{\hat{P}_i}{p_i}}{\log\left(1 - 4\dfrac{\hat{P}_i}{p_i}\right) - \alpha_s' l + \log\delta} \right\rfloor \tag{4.21}$$

4.3.1 Numerical calculations

We use the same parameters to describe the system as tabulated earlier. We assume that the normalized input pump power q_0 is 250 and the input signal power $P_0 = 0.1$ mW, and we use the results of the previous chapter to approximate $\hat{q}$, the required normalized input pump power to achieve a certain desired gain. We account for excess losses using $T = -1.5$ dB, $10\log_{10}\delta = -0.5$ dB and assume ζ to be compensated for by a slight increase in the appropriate G. We illustrate the results of numerical evaluation for two particular instances.

It is simple to modify the calculations to take the effect (i.e. added noise power resulting in a slightly higher detection thereshold) of ζ explicitly into account, but it's effect for the number of d-taps $\aleph \simeq 10$ is quite small anyway.

Since we assume that the subsidiary buses are equally spaced along the main bus, it's simple to calculate the ASE power at each of the d-taps (along the main bus). The number of d-taps is determined by the given input pump power. Once we know the ASE power that is tapped into any given subsidiary bus, we calculate the minimum signal power that is necessary to meet a simple SNR test at a bit-rate of 10 Gbits/s: we use the approximate from of Personick's Q-factor test as discussed earlier.

For $d = 50$ meters and $l = 10$ meters, the following table represents the total number of users that can be supported (using a uniform tap fraction). Subsidiary buses farther from the access node suffer from a lower SNR and therefore, the corresponding $\hat{M}_i$ is lower.

G	$\hat{q}$	$\aleph$	$\sum_i^{\aleph} \hat{M}_i$	t_i
gain (dB)	req. pump	# d-taps	# total users	tap fraction
5	4	14	1128	0.665
10	5	13	1000	0.894
15	7	12	928	0.967
20	8	12	912	0.989
25	9	11	868	0.997

For such short inter-user and inter-bus distances, the critical parameter in these calculations is the excess loss per stage (both in the main bus and in the subsidiary buses). We find that it should be possible to support about a thousand users with such a distribution tree, but caution that the calculations may need to be adapted to fit particular situations, e.g. if the signal has already suffered ASE accumulation before the access node (as is likely). The MATLAB source code provided in Appendix C may be modified accordingly.

Note that for spatially compact networks such as the one considered here, it's advantageous to keep the gain per stage as low as possible to minimize the SNR degradation due to amplified spontaneous emission. Of course, if the inter-user spacing is to be increased, it may become *necessary* to use higher levels of amplification so that the tap fraction (which we can think of as extracting the component of the gain added in excess of propagation losses) remains positive.

Also, this agrees with our analysis of the optimum allocation of gains and losses along an optical amplifier chain: recall that our solution was to set $G_2 = G_3 = \cdots = G_N = 1$ so that the gain stage G_1 is as far away from the detector as possible. If G_1 represents the gain that is encountered before the start of the distribution network (e.g. in the long-haul backbone network), the number of users that can be supported is indeed higher when we lower $G = G_2 = G_3 = \cdots = G_N$ to the minimum necessary to overcome attenuation and simultaneously avoid the nonlinear regime.

We assume uniform G for each of the stages. While it is true that we want to pick G as low as possible, the floor operation in the formulae involves some searching for the optimum G. For any particular value of G that we choose:

1. Find the minimum normalized input power $\hat{q}_0$ needed to generate G for the particular EDF characteristics, which is given by

$$\hat{q}_0 = \min_L q_0$$

where q_0 is given by [3.30]. In some situations, this step may be simplified e.g. by assuming $\hat{q}_0 = q_0/10$.

2. The number of d-taps that are allowed is given by [4.7]

3. The pair of equations [4.14] and [4.15], with p_i as the variable of interest, results in $\hat{P}_i$, the minimum tapped signal power for the i^{th} subsidiary bus, that satisfies an appropriate SNR or BER constraint.

4. Using this value of $\hat{P}_i$ and p_i as given by [4.14], we calculate the number of users per d-tap from [4.21].

For an upstream signal input at the end users in Figure 4-4, we typically require that the headroom at the head (access node of the feeder network) be sufficient to support a subsequent downstream distribution network. This implies that we need to overcome the worst-case loss $L_{max} = 1/\hat{M}_i + (\hat{M}_i/4)(\alpha_s' l + \delta)$ in the signal, e.g. by inserting an amplifier of the appropriate gain $G = 1/L_{max}$ just before the d-tap. In other respects, our unsaturated amplifier model is reciprocal, but a practical implementation will need to carefully consider additional elements as appropriate.

Chapter 5

Distributed EDFAs and bus networks

If I wanted to fell a tree
I would not start with tearing the leaves
and then breaking the twigs and cutting the branches,
but apply the axe to the lowest part
—N.C. CHAUDHURI, Three Horsemen of the New Apocalypse (1997)

This chapter discusses a particular form of the erbium-doped fiber amplifier that is more appropriate, at least from an analytical viewpoint, to a bus distribution network where a single user is served at each drop point. From earlier discussions on the minimum detectable signal power, we know that the fraction of signal power in the bus that needs to be diverted to a detector, the "tap fraction" can be quite small. For example, if the minimum required signal power for the desired SNR (or BER as appropriate) is -25 dBm, and if the signal power along the distribution bus is -5 dBm, the necessary tap fraction is -20 dB.

At each tap along the bus, the signal level drops by the tap fraction, which we can offset by amplifying the signal before the next tap so that the signal power remains constant. But when we service only one receiver per tap, the gain necessary to overcome this drop is small compared to or of the same magnitude as the gain necessary to overcome propagation losses. It's reasonable to consider an extended section of lightly-doped fiber, which simultaneously serves as the transmission and the gain medium. We model the entire bus distribution line as a single gain Er-doped fiber (amplifier), with the gain spread out along the entire length, hence the name

'distributed' EDFA. This may, of course, model a real, physical system (distributed EDFAs do exist) but also serves as a mean-field approximation for the case of lumped gain elements. Since the inter-stage gain, whether lumped or distributed, is small, we expect the approximation to be a good one.

The rate equations which we've derived in an earlier chapter, and are also used in [25], describe the evolution of the population density of the excited state in a two-level laser system and of the optical power along the Er-doped fiber. We characterize the taps along the distribution bus by a (spatial) density function in the optical power evolution equations: the fraction of the optical power (in each channel) that is diverted away from the bus to a receiver. To keep our equations simple, we'll consider a two-level model, with pumping at 1480 nm (background loss coefficient 0.5 dB/km). This makes sense from a practical standpoint—the background absorption coefficient at 980 nm is significantly higher (1.2 dB/km), and for distribution spans in the tens of kilometers, this additional loss can be quite costly.

As we will see, the fraction of the optical power that is tapped at successive detectors need not be constant. Since increased propagation length implies a greater total gain, and hence greater ASE, we need greater signal levels incident on the detector to maintain the same BER. If we assume that the signal level in the bus is maintained constant, this means the tap fractions form an increasing sequence. This observation will allow us to form a bound on the total number of receiver stations that can be supported.

5.1 Rate equations

In an earlier chapter, we've derived the basic forms of the rate equations that we'll use. We consider a generalization of that model to include more than one optical beam traversing the erbium-doped fiber (EDF).

Consider a section of EDF of length L along which N optical channels propagate. Channel k, at wavelength λ_k carries optical power $h\nu_k P_k(z,t)$ with a confinement

factor Γ_k. Let the fraction of atoms in the excited state be $N_2(z,t)$. The rate equation that describes the change in the upper level population is

$$\frac{\partial N_2(z,t)}{\partial t} = -\frac{N_2(z,t)}{\tau} - \frac{1}{\rho S}\sum_{k=1}^{N}\frac{\partial P_k(z,t)}{\partial z} \tag{5.1}$$

where τ is the spontaneous lifetime of the upper level, ρ is the number density of active Er atoms and S is the fiber core cross section.

We define a 'tap function', $t_k(z)$ which represents the fraction of power in channel k at distance z from the input end of the fiber that would be tapped (i.e. diverted from the transmission bus to a receiver) if a receiver were to be stationed there. For a total of M receiver stations along the fiber,

$$f_k(z) = \sum_{m=1}^{M}\delta(z - z_m)I_k(m)\,t_k(z) \tag{5.2}$$

represents the taps along the entire EDF length for channel k, where $I_k(m) = \{1, 0\}$ is a vector of indicator variables representing whether a fraction $t_k(z_m)$ of channel k is tapped at $z = z_m$ or not.

Each of the m receivers is characterized by an N-vector of indicator variables, $[I_1(m), I_2, (m), \ldots, I_N(m)]^T$, representing whether channel k is tapped or not. While this $I_{N\times M}$ matrix is useful for simulations of channel loading and dynamics, the case of $I_k(m) \equiv 1$ represents the scenario a where each channel is tapped at each receiver.

The vector $[z_m]$ for $m = 1, 2, \ldots, M$ represents the z coordinate of each of the taps. This allows us to consider arbitrary spacing of the taps, particularly for numerical simulations, but we shall first consider the case

$$z_m = \frac{m}{M}L \qquad m = 1, 2, \ldots, M \tag{5.3}$$

representing M taps that are uniformly spaced by L/M.

As before, the propagation equation for the power in any channel (including the

pump) can be written as

$$\frac{\partial P_k(z,t)}{\partial z} = u_k \left[(\gamma_k + \alpha_k)N_2(z,t) - \alpha_k - \alpha'_k - f_k(z)\right] P_k(z,t) \tag{5.4}$$

where

$$\alpha_k = \rho \Gamma_k \sigma_a(\lambda_k) \tag{5.5}$$

$$\gamma_k = \rho \Gamma_k \sigma_e(\lambda_k) \tag{5.6}$$

$$= \eta_k \alpha_k \tag{5.7}$$

are the absorption and emission coefficients as defined earlier, and α'_k is the background loss coefficient at λ_k, and $u_k = \pm 1$ indicates forward or backward propagation relative to the z axis.

The meaning of the definition [5.2] now becomes clear. Specifically, consider the case of uniformly spaced taps. By integrating [5.4] in a small neighborhood Δz of $z = m^* L/M$,

$$\Delta P_k(z,t)|_{z=m^*\frac{L}{M}} = \ldots - P_k(m^*\frac{L}{M}, t) \int_{m^*\frac{L}{M}-\frac{\Delta z}{2}}^{m^*\frac{L}{M}+\frac{\Delta z}{2}} \delta\left(z - m^*\frac{L}{M}\right) t_k(z)\, dz \tag{5.8}$$

$$= \ldots - P_k\left(m^*\frac{L}{M}, t\right) t_k\left(m^*\frac{L}{M}\right) \tag{5.9}$$

which represents the fraction of the power in channel k that is dropped (by a power splitter) at $z = m^* L/M$.

We require that $\tau, \rho S$ are independent of z, and assume that we can exchange the order of operations d/dt and $\int dz$[1]. In doing so, we specifically disallow a z-dependence of ρ, but still allow non-uniform taps. Integrating [5.1] over the total

[1]See Appendix B for a discussion of the validity of this.

length of the EDF from $z = 0$ to L,

$$\left(\frac{d}{dt} + \frac{1}{\tau}\right) \int_0^L N_2(z,t)\, dz = -\frac{1}{\rho S} \sum_{k=1}^{N} P_{out}(\lambda_k, t) - P_{in}(\lambda_k, t) \tag{5.10}$$

We define the path-averaged upper level fraction, $\bar{N}_2(t)$, as

$$\bar{N}_2(t) = \frac{1}{L} \int_0^L N_2(z,t)\, dz \tag{5.11}$$

and the path-averaged exponential gain coefficient (for channel k), $\bar{g}_k(t)$, as

$$\bar{g}_k(t) = \frac{1}{L} \log \frac{P_{out}(\lambda_k, t)}{P_{in}(\lambda_k, t)} \tag{5.12}$$

Therefore, [5.10] can be rewritten as

$$\boxed{\left(\frac{d}{dt} + \frac{1}{\tau}\right) \bar{N}_2(t) = -\frac{1}{\tau L \zeta} \sum_{k=1}^{N} P_{in}(\lambda_k, t)[\exp(\bar{g}_k(t)L) - 1]} \tag{5.13}$$

where the saturation parameter $\zeta = \rho S/\tau$ is defined as the number density of Er atoms divided by the spontaneous lifetime in the upper level [25].

Also, from [5.4] after dividing both sides by $P_k(z,t)$ and integrating over the total length of the EDF from $z = 0$ to L,

$$\frac{1}{L} \int_0^L \frac{1}{P_k(z,t)} \frac{\partial P_k(z,t)}{\partial z} = (\gamma_k + \alpha_k)\bar{N}_2(t) - \tilde{\alpha_k} \tag{5.14}$$

where we define

$$\tilde{\alpha_k} = \alpha_k + \alpha_k' + \frac{1}{L} \sum_{m=1}^{M} t_k(z_m) \tag{5.15}$$

Recognizing that the left-hand side of [5.14] is given by [5.12], the average expo-

nential gain constant can be written as

$$\bar{g}_k(t) = (\gamma_k + \alpha_k)\bar{N}_2(t) - \tilde{\alpha}_k \tag{5.16}$$

$$= \frac{\zeta}{P_{sat}(\lambda_k)}\bar{N}_2(t) - \tilde{\alpha}_k \tag{5.17}$$

where the saturation power (expressed in photons) is

$$P_{sat}(\lambda_k) = \frac{\rho S}{\tau(\gamma_k + \alpha_k)} = \frac{\zeta}{\gamma_k + \alpha_k} \tag{5.18}$$

Combining [5.13] and [5.17], we can describe the time-evolution of the upper-level fraction, given the input channel powers.

$$\boxed{\left(\frac{d}{dt} + \frac{1}{\tau}\right)\bar{N}_2(t) = -\frac{1}{\tau L \zeta}\sum_{k=1}^{N} P_{in}(\lambda_k, t)\left\{\exp\left[\left(\frac{\zeta}{P_{sat}(\lambda_k)}\bar{N}_2(t) - \tilde{\alpha}_k\right)L\right] - 1\right\}} \tag{5.19}$$

Once the above ordinary differential equation for $\bar{N}_2(t)$ is solved, we can obtain the path-averaged gain for each channel from [5.17], and thereby determine the output powers for each channel.

Given the non-algebraic nature of the above equation, finding an explicit solution for the dynamics of $\bar{N}_2(t)$ usually has to be carried out using numerical techniques. In the next section, we focus on a special case of [5.19] which will yield some important insights.

5.2 Steady state

We consider the special case of [5.19] when all the involved quantities are independent of time. By definition, this is the steady-state solution, and $\bar{N}_2$ satisfies a transcen-

dental equation

$$\bar{N}_2 = -\frac{1}{L\zeta}\sum_{k=1}^{N} P_{in}(\lambda_k)\left\{\exp\left[\left(\frac{\zeta}{P_{sat}(\lambda_k)}\bar{N}_2 - \tilde{\alpha_k}\right)L\right] - 1\right\} \tag{5.20}$$

Since we're mainly concerned with the limitations on the size of the distribution network, rather than its dynamics, we'll assume the steady state conditions for the remainder of the chapter.

5.3 Uniform taps

To simplify the following analysis, we assume uniform and uniformly spaced taps along the bus, but allow for differences in the taps for the signal and the pump.

$$t_k\left(m\frac{L}{M}\right) \equiv t_{s,p} \tag{5.21}$$

If we further assume, as we did in an earlier chapter, that the pump power dominates the signal powers, then we can reduce the right-hand side of [5.20] to a single contribution: from the pump,

$$\bar{N}_2 = -\frac{1}{L\zeta}P_{in}(\lambda_p)\left\{\exp\left[\left(\frac{\zeta}{P_{sat}(\lambda_p)}\bar{N}_2 - \tilde{\alpha_p}\right)L\right] - 1\right\} \tag{5.22}$$

where, depending upon the physical geometry of the network,

$$\tilde{\alpha_p} = \begin{cases} \alpha_p + \alpha'_p + \frac{M}{L}t_p & \text{if pump is tapped along with the signal} \\ \alpha_p + \alpha'_p & \text{otherwise} \end{cases} \tag{5.23}$$

The first definition is applicable if power splitters with non-zero drop response at the pump wavelength (e.g. Y-junctions, or weakly coupled resonant structures) are used to tap a fraction of the signal power. Instead, if a bandpass WDM mux-demux is

used to separate the pump before the tap, the second definition applies. The solution of the simple transcendental equation, [5.22], enables characterization of each of the channel gains, using [5.17].

Consider the case of perfectly uniform signal channels, or equivalently, only one signal channel, indexed by s rather than k. The transparency pump power defines unity net gain for the signal channel, or equivalently, the (steady-state) path-averaged exponential gain constant at the signal wavelength $\bar{g}_s = 0$. We can derive a simple condition on the required $\bar{N}_2$, using [5.17],

$$\bar{N}_2 = \frac{\tilde{\alpha}_s}{\gamma_s + \alpha_s} \tag{5.24}$$

$$= \frac{\alpha_s + \alpha'_s + Mt_s/L}{\gamma_s + \alpha_s} \tag{5.25}$$

Since this represents the *fraction* (≤ 1) of atoms in the excited state, the following inequality must be satisfied:

$$\boxed{\frac{M}{L} \leq \frac{\gamma_s - \alpha'_s}{t_s} = \frac{\eta_s \alpha_s - \alpha'_s}{t_s}} \tag{5.26}$$

which defines a bound on the number of stations per unit length that can be supported for a given tap fraction.

Furthermore, the pump power required to achieve transparency can be found by substituting [5.24] into [5.22], and using [5.18],

$$P_{in}(\lambda_p)/P_{sat}(\lambda_p) = L \frac{(\gamma_p + \alpha_p)(\alpha_s + \alpha'_s + Mt_s/L)/(\gamma_s + \alpha_s)}{1 - \exp\left[\left(\frac{\gamma_p + \alpha_p}{\gamma_s + \alpha_s}\alpha_s - \alpha_p - \alpha'_p - I_{p-tap}\frac{Mt_p}{L}\right) L\right]} \tag{5.27}$$

where $I_{p-tap} \in \{1, 0\}$ is an indicator variable that takes on values depending on whether the pump is tapped along with the signal or not. This equation can also be used to define the maximum serviced length of Er-doped fiber L, for a given pump input power, $P_{in}(\lambda_p)$.

Equations [5.26] and [5.27] can be combined to describe a bound on the number

of receivers that can be supported. We assume the condition in [5.26] to be satisfied with equality, and substitute in [5.27] with the assumption that L is large so that the denominator of [5.27] ≈ 1,

$$\boxed{L^* = \frac{P_{in}(\lambda_p)/P_{sat}(\lambda_p)}{\gamma_p + \alpha_p}} \tag{5.28}$$

and consequently,

$$\boxed{M^* = \frac{P_{in}(\lambda_p)/P_{sat}(\lambda_p)}{t_s}\left(\frac{\gamma_s - \alpha'_s}{\gamma_p + \alpha_p}\right)} \tag{5.29}$$

The validity of this approximation depends, of course, on the numerical values of the various parameters. We'll see that for a representative set of numerical values, this is indeed valid. Using the same approximation in [5.27], if we are given L or M, we can solve for the other

$$L = \frac{1}{\alpha_s + \alpha'_s}\left\{\left(\frac{\gamma_s + \alpha_s}{\gamma_p + \alpha_p}\right)\frac{P_{in}(\lambda_p)}{P_{sat}(\lambda_p)} - Mt_s\right\} \tag{5.30}$$

$$M = \frac{1}{t_s}\left\{\left(\frac{\gamma_s + \alpha_s}{\gamma_p + \alpha_p}\right)\frac{P_{in}(\lambda_p)}{P_{sat}(\lambda_p)} - (\alpha_s + \alpha'_s)L\right\} \tag{5.31}$$

Since both L and M must be positive, we can derive an upper bound

$$\boxed{\tilde{M} = \frac{P_{in}(\lambda_p)/P_{sat}(\lambda_p)}{t_s}\left(\frac{\gamma_s + \alpha_s}{\gamma_p + \alpha_p}\right)} \tag{5.32}$$

which gives the maximum number of users that can be supported (we've not dealt with noise-related bounds yet), and

$$\boxed{\tilde{L} = \frac{P_{in}(\lambda_p)/P_{sat}(\lambda_p)}{\gamma_p + \alpha_p}\left(\frac{\gamma_s + \alpha_s}{\alpha_s + \alpha'_s}\right)} \tag{5.33}$$

is the maximum length of erbium-doped fiber that this level of pump input power can support.

The given conditions will determine which form of the constraint is more applicable: if M/L is the starting point, then M^* and L^* are the appropriate bounds. Note that $M^* < \tilde{M}$ and $L^* < \tilde{L}$. However, if we are given either M or L and can trade off a lower receiver density for increased propagation length or number of users, then $\tilde{L}$ or $\tilde{M}$ is what we seek.

It's also possible, in practice, to design networks with different receiver densities along successive sections of the EDF. A particular span of the the fiber may serve a particularly high concentration of users, and may be followed by a essentially a transmission line without a significant number of taps. This "clustered" model is difficult to analyze theoretically, but the results of the parametric approach of the previous chapter can be combined with those presented here. In a later section, we'll consider nonuniform taps (albeit for a different rationale), and the mathematics are uniformly applicable to a wider variety of situations than we will point out explicitly.

At signal transparency, we've found two bounds on L and hence on M: one given by the simple fact that L and M must both be positive, and the second which bounds the receiver density M/L by a characteristic of the upper-level population. In a later section, we'll evaluate these bounds for typical numerical values.

Before we do so, we investigate another bound: that due to the signal-to-noise (SNR) ratio required to maintain a particular bit-error rate (BER).

5.4 Noise power

Following Desurvire [6, pages 76–77], the amplifier noise, related to the photon statistics master equation, is defined as

$$N(z) = G(z) \int_0^z \frac{a(z')}{G(z')}\, dz' \tag{5.34}$$

where

$$a(z) = \sigma_e(\lambda_s)\rho\Gamma_s N_2(z) \tag{5.35}$$

$$G(z) = \exp\left[\int_0^z \{\sigma_e \rho\Gamma_s N_2(z') - \sigma_a \rho\Gamma_s N_1(z') - \alpha_s' L - M t_s\}\, dz'\right] \tag{5.36}$$

In our analysis, we have dealt with path-averaged quantities, and so an evaluation of $N(L)$ is not possible, particularly in the case of transparency. However, if we assume uniform and complete medium inversion along the fiber length (typically achieved for negligible absorption of a high-power 980 nm pump), $N_1(z) \equiv 0$ in [5.36] implies that

$$\sigma_e \rho\Gamma_s N_2 = \alpha_s L + M t_s \tag{5.37}$$

and using this fact along with $G(z) \equiv 1$ in [5.34]

$$N(L) = \alpha_s' L + M\, t_s \tag{5.38}$$

and the noise power in bandwidth B_o is

$$P_N(L) = N(L)\, h\nu\, B_o \tag{5.39}$$

For uniform, but incomplete inversion (negligible absorption of a high-power 1480 nm pump), we have to account for the non-zero lower level population density where

$$N_2|_{max} = \frac{\tilde{\alpha_s}}{\gamma_s + \alpha_s} \tag{5.40}$$

since the gain coefficient is always negative at the pump wavelength. Since $N_1 + N_2 = 1$, and we have normalized the population densities by ρ, the doping concentration

along the fiber (number density),

$$P_N(L) = \left\{ \sigma_a(\lambda_s)\, \rho \Gamma_s \frac{\gamma_s + \alpha_s - \tilde{\alpha}_s}{\gamma_s + \alpha_s} + \alpha'_s \right\} L + M\, t_s \tag{5.41}$$

As compared to a distributed amplifier without periodic taps along the signal path, the added gain needed to offset the drop in signal power along the length of the EDF increases the noise power by the factor $M\, t_s$, and also lowers the corresponding SNR.

For the case of maximum M attained when [5.26] is satisfied with equality,

$$P_N^{MAX}(L) = \left\{ \frac{\alpha_s}{\gamma_s + \alpha_s} (\gamma_s + \alpha_s - \tilde{\alpha}_s) + \gamma_s \right\} L\, h\nu\, B_o \tag{5.42}$$

$$= \gamma_s L\, h\nu\, B_o \tag{5.43}$$

which is, in general, an upper bound for the ASE noise power.

5.5 SNR constraint

A given level of input pump power $P_{in}(\lambda_p)$ limits the length of an EDF that can provide signal transparency using the results of the previous section and the determines the number of stations that can be supported. Here, we can derive closed-form expressions for the important and practical case of very high pump power, so that we can assume uniform medium inversion along the entire fiber length.

Let us assume that the optical power in the signal is $P_S(L)$ and the optical power in the noise is $P_N(L)$. For input signal power P_0,

$$P_S(L) = P_0 \tag{5.44}$$

$$P_N(L) = (\alpha'_s L + M t_s)\, h\nu\, B_o \tag{5.45}$$

where, for simplicity, we use [5.38] to account for the noise power. The same ar-

guments can be applied to [5.41] as well—the only difference is that of an additive constant increasing the noise power in the second case.

We define the following terms which have the dimensions of currents (as generated by a photodiode in response to incident optical fields),

$$I_S(L) = P_S(L)\,\frac{e}{h\nu_s} \tag{5.46}$$

$$I_N(L) = P_N(L)\,\frac{e}{h\nu_s} \tag{5.47}$$

Assume that we detect the fields by an avalanche photodiode with gain factor $\bar{g}$, excess noise factor x and quantum efficiency η, where B_e is the electronic bandwidth (Gbits/s, related to the bit rate of communications) and B_o is the optical bandwidth (nanometers, related to the passband of the optical window of the detector). We write down the mean-square current terms that define the signal-to-noise ratio:

$$\begin{aligned}
<{i_S}^2> &= <\bar{g}>^2\eta^2{I_S}^2 && \text{signal} \\
<{i_{N_1}}^2> &= 2\eta^2 I_S\, I_N(L)\tfrac{2B_e}{B_o}<\bar{g}>^2 && \text{signal-ASE beat noise} \\
<{i_{N_2}}^2> &= 2\eta\, eB_e[I_S + 2I_N(L)]<\bar{g}>^{2+x} && \text{signal and ASE shot noise} \\
<{i_{N_3}}^2> &= \eta^2 I_N(L)^2\,\tfrac{B_e}{B_o}(2B_o - B_e)<\bar{g}>^2 && \text{ASE-ASE beat noise} \\
<{i_{N_4}}^2> &= 4\,k_B\tfrac{T_e}{R}B_e && \text{detector thermal noise}
\end{aligned} \tag{5.48}$$

where k_B is Boltzmann's constant, T_e is the effective temperature (Kelvin) and R is the resistance (Ω) of the detector.

The signal to noise ratio (SNR) is defined in terms of the above mean-squared currents as

$$\boxed{\mathrm{SNR}(L) = \frac{<{i_S}^2>}{<{i_{N_1}}^2> + <{i_{N_2}}^2> + <{i_{N_3}}^2> + <{i_{N_4}}^2>}} \tag{5.49}$$

and, given a target SNR and input signal power P_0, [5.49] is a quadratic equation in L. In fact, if we ignore the ASE-ASE beat noise term, [5.49] reduces to a simple

linear equation in L.

There is another way to analyze this equation: for a given L, [5.49] is a quadratic equation in P_0, the signal power required to achieve the target SNR. Alternatively, using [2.17], we can calculate the necessary signal power for a target bit-error-rate (BER).

At the end of this section of transparent fiber, the signal power is the same as at the input, by definition of transparency. However, we have added ASE noise along the fiber. We can then determine the number of stations that can be supported by a *passive* distribution bus, with the receiver thresholds set by taking into account the SNR at the output of the EDF. As we've seen in Chapter 1, this is typically a small number (about 40) which we will ignore.

5.6 Numerical example

We continue to assume that the taps along the fiber are uniform, though we'll see in the next section that this is not the optimum that can be achieved.

We assume the following numerical values for the various parameters, [5]:

signal absorption coefficient ($\lambda_s = 1.55\,\mu$m)	α_s	4.0 dB/km
pump absorption coefficient ($\lambda_p = 1.48\,\mu$m)	α_p	1.6 dB/km
background loss coefficient	$\alpha'_{s,p}$	0.5 dB/km
Ratio of σ_e to σ_a ($\lambda_s = 1.55\,\mu$m)	η_s	1.42
Ratio of σ_e to σ_a ($\lambda_p = 1.48\,\mu$m)	η_p	0.37

and we assume that $t = 0.01$, implying that the tap fraction is -20 dB. Also, we assume that the pump is not tapped along with the signal at each receiver along the fiber, so that the maximum possible utility is gained from a given pump input power.

The first bound, given by [5.26], then implies that the number of receiver stations per kilometer $M/L \leq 244$.

In an earlier section, we've analyzed the limit imposed by the limited input pump power available for amplification of the signal as given by [5.28]. First, we verify that the approximation we made in deriving that relationship, and the ones that followed it, is indeed valid. We want

$$1 - \exp\left[\left(\frac{\gamma_p + \alpha_p}{\gamma_s + \alpha_s}\alpha_s - \alpha_p - \alpha'_p\right) L\right] \approx 1 \tag{5.50}$$

and substituting in numerical values, we want $1 - \exp(-1.75\,L) \approx 1$ which is satisfied with about 1% or less error if $L \geq 2.5$ km. Since the span of our distribution networks will turn out to be quite a bit longer than this, our approximation is self-consistent.

Substituting in the appropriate numerical values, we see that by ignoring noise constraints, the maximum length $L^* = 0.505 \times P_{in}(\lambda_p)/P_{sat}(\lambda_p)$. If the normalized input pump power $q \equiv P_{in}(\lambda_p)/P_{sat}(\lambda_p) = 100$, we have $L^* = 50.5$ km, and therefore, $M^* = 244 \times L^* \approx 12{,}300$ receivers.

If we'd rather deal with a fixed number of receivers rather than a receiver density (number of receivers per kilometer), then we can use [5.30] to evaluate the maximum possible transmission length for a given number of receivers, or [5.31] for the converse.

The upper bounds [5.32,5.33] can be evaluated for e.g. $q = 100$, $t_s = 10^{-2}$, yielding $M < 30{,}700$ and $L < 84$ km. Note that this exceeds the receiver density bound (M^*/L^*), and so the earlier bound is tighter.

We can design the length of our network to suit a given number of users, or the other way around. Upper bounds on each of the parameters are given by simple relationships in terms of the input power and tap fraction. Also, we can use the receiver density M/L as the starting parameter instead, which may be more appropriate in some applications.

The tradeoff between propagation length and the receiver density will be explored in the next section.

In evaluating the signal-to-noise ratio bound [5.49], we've shown the results for two SNR thresholds, -25 dB and -30 dB, assumed uniform for each receiver station, and the following parameters (as in earlier chapters) for detectors:

Electronic bit rate (data)	B	10 Gbits/s
Optical bandwidth	B_o	$2\,B$
Receiver electronic bandwidth	B_e	B
APD avalanche gain	$\bar{g}$	50
APD excess noise factor	x	0
Effective receiver temperature	T_e	$4 \times 290\,K$
Detector Resistance	R	$50\,\Omega$

Further, we use the ASE noise power defined for uniform but incomplete inversion (from 5.41),

$$P_N(L) = \left[\left\{\frac{\alpha_s(\gamma_s + \alpha_s - \tilde{\alpha_s})}{\gamma_s + \alpha_s} + \alpha_s'\right\} L + M\,t_s\right] h\nu\,B_o \tag{5.51}$$

$$= \frac{\gamma_s}{\gamma_s + \alpha_s}\left[(\alpha_s + \alpha_s')L + M\,t_s\right] h\nu\,B_o \tag{5.52}$$

The results of the numerical solution of [5.49] are shown in Figure 5-1. We've plotted the input signal power (in dBm) along the ordinate (though this is the parameter that determines the corresponding abscissa). Conversely, if we wish to design an EDF of a given length, this figure indicates the required input signal power. We have consistently assumed that the EDF is not in the nonlinear regime of amplification, and this will limit the maximum input signal power we can tolerate.

The results are encouraging! While we must caution that these are theoretical, and therefore, rather optimistic calculations, it's evident that a distributed EDF can support a large number of stations: in the thousands of receivers, over a distance in the tens of kilometers. This is precisely suited for the application we have in mind: a distribution network.

In carrying out this calculation, we've assumed that the tap fraction at each

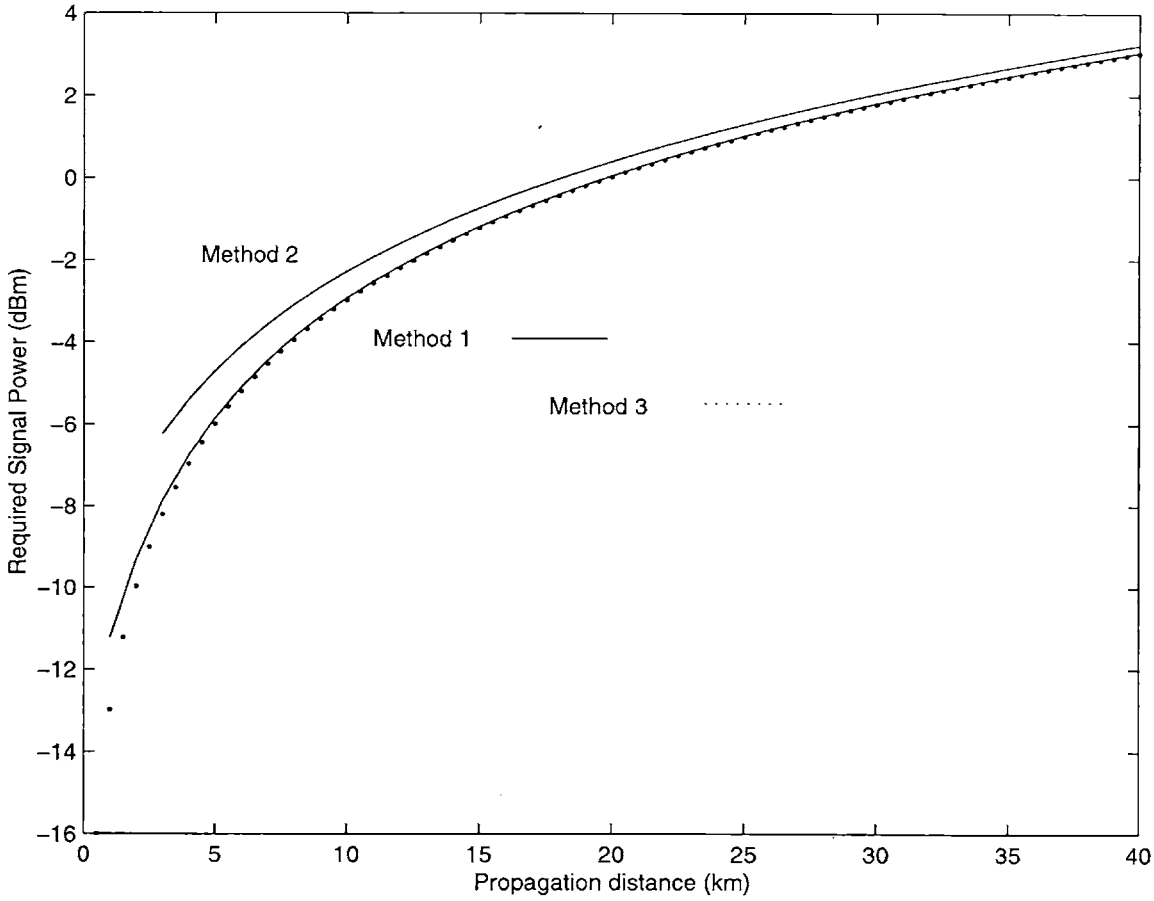

Figure 5-1: The maximum length of EDF that can be supported by a given normalized input signal power (or vice versa), as dictated by the SNR constraint. The curves for Methods 1 and 3 are almost coincident. Three algorithms are considered as described in the text. The receiver density is 100 receivers per kilometer and the tap fraction = 0.01

receiver along the fiber is uniform. Two algorithms have been used to evaluate the required signal power in Figure 5-1. Personick's Q factor accounts for the different APD output variances (noise terms) for the presence and absence of a pulse, and the exact calculation forms Method 2. A simpler algorithms assumes that the variances are equal—this is common in the literature—and then we can solve a simple quadratic equation to find the required signal power: this is Method 1. The more careful analysis of Humblet and Azizoglu considered in [2.17] with an assumed $\mathcal{M} = 36$ forms the basis of Method 3. For network lengths more than a few kilometers, the results of Method 3 coincide with those of Method 1. Appendix C lists the MATLAB code for this plot.

We've found two types of bounds that constrain the span of a distributed EDFA. Numerically evaluating these bounds using the parameters we've used in this section, we see, firstly, that the number of receivers per kilometer cannot exceed 244. Next, for a given input pump power (e.g. 60 times the saturation power at 1480 nm), and assuming 3000 receivers along the fiber, the span of the fiber cannot exceed about 40 kilometers. This implies that the receiver density is 75 detectors per km, well within our bound of 244. And then, assuming that this level of pump is high enough to cause uniform (but incomplete) inversion along the fiber, the required signal levels are about 0 dBm for a receiver threshold of -25 dB, and -2.5 dBm for a receiver threshold for -30 dB.

In a later section, we show how we can obtain the limiting performance of such a system. At each receiver, we tap no more than necessary to meet a certain SNR. At the input end of the fiber, the ASE power is less than at the output. And so, we'd expect a lower signal level to satisfy the SNR criterion. Now consider the situation we've been describing this far—the signal level is maintained constant along the length of the fiber. A lower signal level requirement for the receiver at the input end of this fiber implies that the tap fraction at this end can be *lower* than at the output end. In other words, the limiting performance is achieved in the case of *non-uniform taps.*

The effect of lowering the receiver density from e.g. 244 to 75 receivers per kilometer will be analyzed in the next section.

5.7 Tradeoff pump input for receiver density

In the context of the numerical values we've used in the previous section, the maximum receiver density is 244 users per kilometer. In certain situations, we may not need such a high density of users—it's possible to trade receiver density for maximum propagation length as we've seen in our calculations. In this section, we formulate those results algebraically.

Let the normalized input pump power be $q = P_{in}(\lambda_p)/P_{sat}(\lambda_p)$. Assuming the highest possible receiver density [5.26], the propagation length L^* is given by

$$L^* = \frac{q}{\gamma_p + \alpha_p} \tag{5.53}$$

Let's consider a particular receiver density M/L so that

$$L = \frac{q}{\gamma_p + \alpha_p} \frac{\gamma_s + \alpha_s}{\alpha_s + \alpha_s' + Mt_s/L} \tag{5.54}$$

where $\gamma_s > \alpha_s' + Mt_s/L$.

The increase in propagation length $\Delta L = L - L^*$ is

$$\boxed{\Delta L = \frac{q}{\gamma_p + \alpha_p} \frac{\gamma_s - \alpha_s' - Mt_s/L}{\alpha_s + \alpha_s' + Mt_s/L}} \tag{5.55}$$

Using numerical values for the parameters from the previous section, for an input $q = 100$ and $M/L = 75$ receivers per kilometer instead of the theoretical limit of 244, we have $\Delta L = 19.5$ km. Similarly, if we were to assume $q = 60$, then $\Delta L = 11.7$ km, and $L = L^* + \Delta L = 42$ km as discussed in the previous section.

5.8 Non-uniform taps

We'll continue to assume a single channel for the signal, but now relax the constraint of uniform taps. We can also relax the assumption that the taps are uniformly spaced. As we've discussed earlier, this is a closer picture of ideality: as L increases, a larger fraction of the signal is necessary to satisfy [5.49], whose denominator grows monotonically with L.

Consider the original definition of the tap function, $f_s(z)$

$$f_s(z) = \sum_{m=1}^{M} \delta(z - z_m) I_s(m)\, t_s(z) \tag{5.56}$$

where $[z_m]$ is the vector of tap locations, as before. In a network design problem, $[z_m]$ represents the locations of the receiver stations, and is a given parameter.

Under the assumption of complete and uniform medium inversion, we can consider the noise power at the end of a section of EDF of length z,

$$P_N(z) = \left[\alpha'_s z + \sum_{m=1}^{M(z)} t_s(z_m) \right] h\nu\, B_o \tag{5.57}$$

where $M(z)$ is the number of taps in $(0, z)$. If, for any z that we consider, we assume that inversion is uniform but incomplete,

$$\begin{aligned} P_N(z)^{NC} &= \left[\left(\alpha'_s + \alpha_s \frac{\gamma_s - \alpha_s - \frac{1}{z} \sum_{m=1}^{M(z)} t_s(z_m)}{\gamma_s + \alpha_s} \right) z + \sum_{m=1}^{M(z)} t_s(z_m) \right] h\nu\, B_o \\ &= \frac{\gamma_s}{\gamma_s + \alpha_s} \left[(\alpha_s + \alpha'_s) z + \sum_{m=1}^{M(z)} t_s(z_m) \right] h\nu\, B_o \end{aligned} \tag{5.58}$$

but we will deal with the simpler notation of complete inversion for the remainder of this section.

For example, if we restrict ourselves to uniformly spaced taps,

$$z_m = \frac{m}{M}L \qquad m = 1, 2, \ldots, M \tag{5.59}$$

$$M(z) = \left\lfloor \frac{z}{L}M \right\rfloor \tag{5.60}$$

For complete and uniform medium inversion, signal power excursion is minimal, and we assume that transparency is maintained, i.e. [5.44] is assumed valid for all $z \in (0, L)$. We can now obtain $t_s(z)$ by the following recursive process.

For $0 < z \leq z_1$, we can find the noise power from [5.57]

$$P_N^1(z) = (\alpha_s' z) h\nu\, B_o \tag{5.61}$$

since there are, by definition, no taps before z_1. Using this value in the SNR constraint calculation [5.49], yields P_1^*, the minimum detectable signal power, given that the noise power is [5.61]. Therefore, the tap fraction at $z = z_1$ is $t_s(z_1) = P_1^*/P_0$.

For $z_1 < z \leq z_2$, the definition of the noise power must now account for the tap at z_1, which we have just computed,

$$P_N^2(z) = [\alpha_s' z + t_s(z_1)]\, h\nu\, B_o \tag{5.62}$$

and, again, the SNR constraint calculation [5.49] yields P_2^*, the minimum detectable signal power at the second stage. Therefore the tap fraction at $z = z_2$ is $t_s(z_2) = P_2^*/P_0$.

We can proceed in this recursive fashion until we reach a constraint that we'll discuss in a moment. In particular, for $z_{m-1} < z \leq z_m$, the noise power at z is

$$P_N^m(z) = [\alpha_s' z + \sum_{i=1}^{m-1} t_s(z_i)] h\nu\, B_o \tag{5.63}$$

and using this value in the SNR constraint calculation, with signal level P_0, yields

P_m^*, the minimum detectable signal power at the m^{th} stage and the corresponding tap fraction, $t_s(z_m) = P_m^*/P$. At each stage, the vector of the preceding taps, $t_s(z_i)$ is available to us.

It's easy to see that the $t_s(z_k)$ form an increasing sequence, but there is a physical limit on any of the terms: we cannot tap more than 100% of the signal power! This bound imposed by the accumulated ASE is evident when the following condition is met:

$$\boxed{t_s(z_m) \geq 1 \quad \text{for some } m < \infty} \tag{5.64}$$

More formally, since the sequence t_s represents the sequence of tap *fractions*, the bound on the number of stations M^* that can be supported given a vector of tap locations $[z_m], m = 1, 2, \ldots, M$ is

$$\boxed{M^* = \min\left\{ \min_{1 \leq m \leq M} \{ m \mid t_s(z_m) \geq 1 \}, M \right\}} \tag{5.65}$$

where we assume that $\min[x] = \infty$ if x is an empty set.

The upper level population (fraction) N_2 must be less than 1, and using [5.24],

$$\sum_{m=1}^{M(z)} t_s(z_m) \leq L(\gamma_s - \alpha_s') \tag{5.66}$$

When this condition is satisfied with equality, substituting into [5.58], the noise power at the output for complete inversion is

$$P_N^{MAX}(L) = \gamma_s L \, h\nu \, B_o \tag{5.67}$$

which is the same as for the case of uniform taps, as expected.

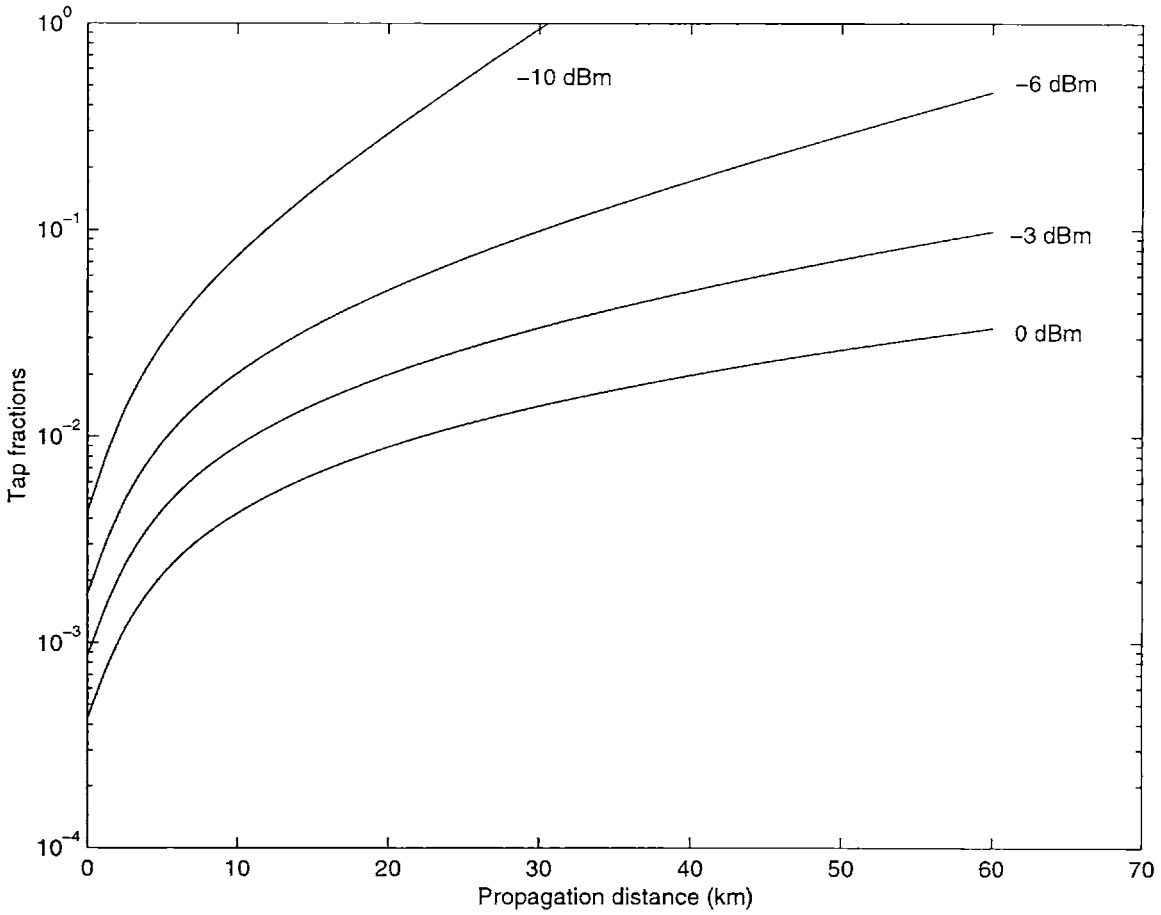

Figure 5-2: Tap fractions along the EDF, for receivers spaced apart by 10 meters, and input signal power $P_0 = -10$ dBm to $P_0 = 0$ dBm.

5.9 Numerical example (contd.)

The numerical values for our parameters are the same as before. The receiver sensitivity is not a rational function of the propagation length, and consequently, neither are the tap fractions. It's simplest to evaluate the algorithm numerically, e.g. in MATLAB. We've provided the code listing in Appendix C: it's simple to modify the parameters to address a special circumstance. For quicker execution, we've used the approximation to Personick's Q-factor sufficient statistic described earlier: for large L, the two methods give near-identical results.

Figure 5-2 plots the tap fractions t_k for receivers spaced apart by only 10 meters, whereas in Figure 5-3, we've increased the spacing to 40 meters.

In all cases, the tap fractions form an increasing sequence, and moreover, depend critically on the input signal power P_0. The results show that signal powers in the sub-dBm range are sufficient for typical LAN distances, particularly if the tap fractions are selected according to this procedure. As compared to the case of uniform taps, designing the taps optimally increases the propagation distance e.g. by an order an magnitude for $P_0 = -10$ dBm.

Another parameter that affects the growth rate of the tap fraction sequence is the receiver density M/L. Receivers spaced apart by 10 meters should be sufficiently generous for most applications. As discussed for the case of uniform taps, a lower receiver density increases the span of the network for the same signal input power level P_0, whereas for a still higher density (but within the $N_2|_{MAX}$ bound described earlier), we will reach the tap fraction bound $t_M \leq 1$ sooner.

We've assumed, in our analysis so far, that the tap fraction represents the small fraction of signal power that is necessary for detection. In a later chapter, we will see that the same sequence of mathematical steps can be applied to a different interpretation: t_k now represents a division of the signal (and pump) power into two or more equal parts. The physical structure that a sequence of such operations results in is called a distribution tree, which we analyze in the same framework as the bus network, but with a higher order-of-magnitude scale for the tap fractions.

Non-uniform taps permit a lower input signal power to serve the same distribution network (i.e. same length and number of users). Similarly, the number of users, overall length or receiver density can be increased for the same input signal power. As expected, the tap fractions form an increasing sequence.

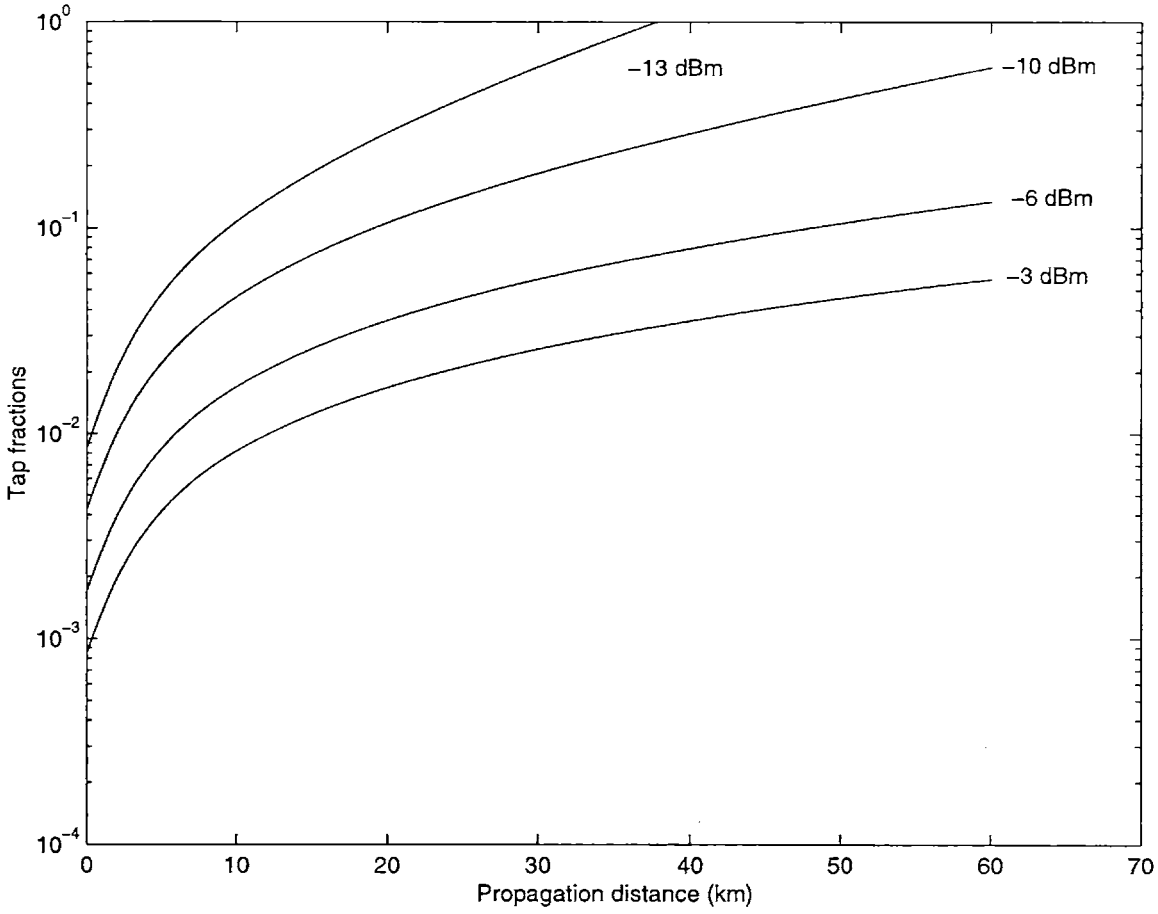

Figure 5-3: Tap fractions along the EDF, for receivers spaced apart by 40 meters, and input signal power $P_0 = -13$ dBm to $P_0 = -3$ dBm.

Chapter 6

Time-varying tap function

I don't know if you have ever seen one of those old maps
where they mark a spot with a cross and put "Here be dragons"
or "Keep ye eye skinned for hippogriffs,"
but I had always felt that some such kindly warning
might well be given to pedestrians and traffic
with regard to this Steeple Bumpleigh.
—P.G. WODEHOUSE, Jeeves in the Morning (1971)

In this chapter, we consider an important extension to our model for distributed EDFAs. So far, we've assumed that the tap function $t_k(z)$ is independent of time i.e. once we set the tap locations and the tap fractions, $f_k(z)$ does not affect the gain dynamics of the EDFA.

It's interesting to ask what the effect of a small (dynamic) perturbation in $t_k(z)$ is on the steady-state solution that we've found in the previous chapter. What we're most interested in is the resulting change, if any, in the output power of each channel $P_{out}(\lambda_k, t)$. We will show that perturbing the tap function of any channel indeed affects the remaining channels: this is called crosstalk, and present a simple closed-form expression to quantify this.

We consider the following perturbation

$$t_k(z,t) = t_k(z) + \Delta t_k(t) \qquad |\Delta t_k(t)| \ll t_k(z) \text{ for all } z,t \tag{6.1}$$

where $t_k(z)$ is the time-independent steady-state tap function defined in the previous

chapter.

This perturbation redefines our effective absorption coefficient [5.15]

$$\tilde{\alpha_k}(t) = \tilde{\alpha_k} + \underbrace{\frac{1}{L}\Delta t_k(t)}_{x_k(t)} \qquad |\Delta t_k(t)| \ll t_k(z) \text{ for all } z, t \tag{6.2}$$

This results in a perturbation of the path-averaged upper-level population density, path-averaged exponential gain constant and output power for each channel, which we define as

$$\bar{N}_2(t) = \bar{N}_2[1 + \Delta\bar{N}_2(t)] \qquad |\Delta\bar{N}_2(t)| \ll 1 \tag{6.3}$$

$$\bar{g}_k(t) = \bar{g}_k[1 + \Delta\bar{g}_k(t)] \qquad |\Delta\bar{g}_k(t)| \ll 1 \tag{6.4}$$

$$P_{out}(\lambda_k, t) = P_{out}(\lambda_k)[1 + y_k(t)] \qquad |y_k(t)| \ll 1 \tag{6.5}$$

Our goal is to relate $y_k(t)$ to $\Delta t_k(t)$, or equivalently to $x_k(t)$. The first step is to substitute these definitions into [5.13]. Separating the steady-state terms from those that depend explicitly on t, we get two equations:

$$\bar{N}_2 = -\frac{1}{L\zeta}\sum_{k=1}^{N} P_{in}(\lambda_k)[\exp(\bar{g}_k L) - 1] \tag{6.6}$$

$$\bar{N}_2\left(\frac{d}{dt} + \frac{1}{\tau}\right)\Delta\bar{N}_2(t) = -\frac{1}{\tau L\zeta}\sum_{k=1}^{N} P_{in}(\lambda_k)e^{\bar{g}_k L}\left[\exp\left\{\bar{g}_k\,\Delta\bar{g}_k(t)L\right\} - 1\right] \tag{6.7}$$

For small x, we can write $\exp(x) \approx 1 + x$, and so,

$$\bar{N}_2\left(\frac{d}{dt} + \frac{1}{\tau}\right)\Delta\bar{N}_2(t) \simeq -\frac{1}{\tau L\zeta}\sum_{k=1}^{N} P_{out}(\lambda_k)\,\bar{g}_k\,\Delta\bar{g}_k(t)L \tag{6.8}$$

Similarly, by substituting the new definitions into [5.16], we obtain the pair of

equations,

$$\bar{g}_k = (\gamma_k + \alpha_k)\bar{N}_2 - \tilde{\alpha}_k \tag{6.9}$$

$$\bar{g}_k\,\Delta\bar{g}_k(t) = (\gamma_k + \alpha_k)\bar{N}_2\,\Delta\bar{N}_2(t) - x_k(t) \tag{6.10}$$

Finally, by using the definitions in [5.12], we get

$$P_{out}(\lambda_k, t) = P_{in}(\lambda_k)e^{\bar{g}_k L}e^{\bar{g}_k\,\Delta\bar{g}_k(t)} \tag{6.11}$$

$$\simeq P_{out}(\lambda_k)\{1 + \bar{g}_k\,\Delta\bar{g}_k(t)L\} \tag{6.12}$$

which implies that

$$y_k(t) = \bar{g}_k\,\Delta\bar{g}_k(t)\,L \tag{6.13}$$

Substituting [6.10] into [6.8], we get

$$\bar{N}_2\left(\frac{d}{dt} + \frac{1}{\tau}\right)\Delta\bar{N}_2(t) = -\frac{1}{\tau L\zeta}\sum_{k=1}^{N} P_{out}(\lambda_k)\,L\,\left\{(\gamma_k + \alpha_k)\bar{N}_2\,\Delta\bar{N}_2(t) - x_k(t)\right\} \tag{6.14}$$

Recall the definition of $P_{sat}(\lambda_k)$ from [5.18]. Also, we define a saturation factor [25],

$$F_s = \sum_{k=1}^{N} P_{out}(\lambda_k)/P_{sat}(\lambda_k) \tag{6.15}$$

which represents the total output power, with each channel's power normalized to the saturation power at the corresponding wavelength. Then,

$$\bar{N}_2\left(\frac{d}{dt} + \frac{1 + F_s}{\tau}\right)\Delta\bar{N}_2(t) = \frac{1}{\tau\zeta}\sum_{k=1}^{N} P_{out}(\lambda_k)\,x_k(t) \tag{6.16}$$

$$= \frac{1}{\rho S}\sum_{k=1}^{N} P_{out}(\lambda_k)\,x_k(t) \tag{6.17}$$

Now, we take the Laplace transform of [6.17], and defining the following Laplace transform pairs

$$\mathcal{L}[\Delta\bar{N}_2(t)] = \Delta\bar{N}_2(s) \tag{6.18}$$

$$\mathcal{L}[x_k(t)] = X_k(s) \tag{6.19}$$

$$\mathcal{L}[y_k(t)] = Y_k(s) \tag{6.20}$$

$$\mathcal{L}[\Delta t_k(t)] = T_k(s) \tag{6.21}$$

we can write

$$\bar{N}_2\,\Delta\bar{N}_2(s) = \frac{\tau}{\rho S(1+F_s+\tau s)}\sum_{k=1}^{N} P_{out}(\lambda_k)\,X_k(s) \tag{6.22}$$

Also, taking the Laplace transform of [6.13] and using [6.10]

$$Y_k(s) = (\gamma_k+\alpha_k)\bar{N}_2\,\Delta\bar{N}_2(s)\,L - x_k(t)\,L \tag{6.23}$$

which, combining with [6.22] and using the definition of $x(t)$ results in

$$\boxed{Y_k(s) = \frac{1}{P_{sat}(\lambda_k)}\,\frac{1}{1+F_s+\tau s}\sum_{j=1}^{N} P_{out}(\lambda_j)\,\Delta T_j(s) - \Delta T_k(s)} \tag{6.24}$$

This represents, in the frequency domain, the solution to our perturbation problem. Note that it depends only on the output channel powers, and not on their path-averaged quantities, which is convenient from an experimental perspective. For example, in a communications network, a certain call might involve altering the tap fractions for a number of receivers. Given the output channel powers at that time and a bound on how much perturbation of the channel output powers are allowed, a network manager can decide whether or not to allow that call to go through. Similarly, the time-domain solution also indicates the time-scale over which perturbations resulting from a dynamic change in the tap fractions are damped below a threshold,

and this can be used in scheduling algorithms.

6.1 Single-channel perturbation

To gain some insight into [6.24], assume that the tap function for only a single channel p is perturbed:

$$\Delta t_k(t) = 0 \qquad \forall k \neq p \tag{6.25}$$

From [6.24], the output on channel p suffers a perturbation $y_p(t)$ whose Laplace transform is

$$Y_p(s) = -\Delta T_p(s) \left[1 - \frac{1}{1 + F_s + \tau s} \frac{P_{out}(\lambda_p)}{P_{sat}(\lambda_p)}\right] \tag{6.26}$$

whereas for all the remaining channels $q = 1, 2, \ldots, p-1, p+1, \ldots, N$

$$Y_q(s) = \frac{1}{1 + F_s + \tau s} \frac{P_{out}(\lambda_p)}{P_{sat}(\lambda_q)} \Delta T_p(s) \tag{6.27}$$

In physical terms, the effect of a small increase in the tap fraction for a particular channel p has two kinds of effects. Firstly, the power output for that channel is attenuated, as represented by the first term of [6.26]. Increasing the tap fraction results in a further increase in the magnitude of $y_p(t)$ and hence of the change in the power output.

There is also a second order effect—a singularity in the domain of the problem: a decrease in the power output for any given channel affects the upper-level population. Recall that the rate of change of $\bar{N}_2$ is proportional to the channel powers. Reducing the channel power reduces the rate of change of $\bar{N}_2$, and dampens the change in output power. This is shown by the second term of [6.26], which reflects the low-pass filtering. Since a change in $\bar{N}_2$ affects all channels, not just channel p, this term affects (with the saturation power for each channel acting as a normalizing constant) in the

output of each of the other channels, leading to crosstalk.

If we further assume that the perturbation is an impulse at t_0,

$$\Delta t_p(t) = \epsilon\,\delta(t - t_0) \quad \longleftrightarrow \quad \Delta T_p(s) = \epsilon \tag{6.28}$$

then the time-domain perturbation of channel p is

$$y_p(t) = -\epsilon + \frac{\epsilon}{\tau}\frac{P_{out}(\lambda_p)}{P_{sat}(\lambda_p)}\,e^{-t/t_c} \tag{6.29}$$

where t_c represents the characteristic time scale of this perturbative effect

$$t_c = \frac{\tau}{1 + F_s} \tag{6.30}$$

This clearly shows the two effects described above. Note that the characteristic time scale is the same as that defined in [25] for signal-tone perturbations in the input signal. This is expected: the first-order low-pass filter response is a characteristic of the EDFA model.

Chapter 7

Distributed EDFAs and tree networks

We do it now by writing analytic symbols on the blackboard,
but for your entertainment and interest,
I want you to ride in a buggy for its elegance,
instead of in a fancy automobile.
So we are going to derive this fact by purely geometrical arguments
—well, by essentially geometrical arguments.
—R.P. FEYNMAN (1964)

Our model for distributed EDFAs with taps along the length of the fiber can be very simply extended to a particular form of the tree distribution network. The problem we consider is that of a "fully-extended" tree: the users are the lowest-level leaves of a symmetric tree, and so the distance from each user to the head of the tree is the same.

For bus networks, the tap fractions t_k represent the the signal power coupled out of the bus transmission line at each receiver, and are typically small numbers $\approx 10^{-2}$. Now, we redefine t_k to be the splitting fraction (degree -1) at each node of the tree. For example, if a tree splits into two branches at each node, the splitting fraction of node k is $t_k \equiv 1/\gamma = 1/2$.

Each time we tap a fraction t_k, we begin another branch of the tree. If the taps are uniform, the degree of each of the nodes is the same and the tree is called "regular", but it's also possible to consider non-uniform taps. There is no special significance of a signal-to-noise ratio at each of the stages: all the receivers are at the lowest level, and so have identical SNRs. Rather, physical considerations can dictate what the

sequence of tap fractions should be: e.g. if we require no more than twelve users, a three-way split followed by two two-way splits will better use the available resources than a simple two-way splitting structure with four levels.

The mathematical analysis proceeds exactly as before, but since the tap fractions are now much larger in magnitude than before, far fewer taps can be allowed. If the total number of taps along the direct route from the head of the tree to any one user is M, the number of users the tree serves is γ^M.

There are two other minor interpretative changes we have to make. To obtain M from the equations of the previous chapter, we set $I_{p-tap} \equiv 1$, since the pump is also "tapped" along with the signal. This only strengthens the assumption that the denominator of [5.27] ≈ 1 Also, L now represents the distance of each user from the head of the tree, and we relate it to the total length of fiber used in the next section.

The bound on receiver density [5.26] now reads

$$M/L \leq (\eta_s - \alpha'_s)\gamma \tag{7.1}$$

For a tree that splits into two at each node, $\gamma = 2$ and for the numerical values of the EDFA parameters used in the previous chapter, we cannot support more than four such splits per kilometer along any particular branch. In other words, the number of receivers increases by a factor of $2^4 = 16$ for each kilometer of signal and pump propagation.

If we further assume that $M = 12$ is sufficient (which allows 4096 users at the lowest level), the maximum propagation length for a normalized input pump power $q = 100$ is [5.33] $\tilde{L} = 82.8$ km. Note that this is higher than for the bus network, even with non-uniform taps.

As before, we can now use the SNR constraint to obtain the minimum signal input power that is required at the head of the tree e.g. for $M = 12$.

7.1 Length of tree networks

As mentioned before, L now represents the distance of each of the users from the head of the tree. For simplicity, we represent the length of fiber between successive nodes as 1 unit. Let d_m be the total length of fiber used to form a tree with m splits along any head-to-user branch.

If we were to allow another γ-way split along this line, we would multiply the total number of users by γ, and increase

$$d_{m+1} = \gamma\, d_m + 1 \tag{7.2}$$

For a total of M splits along any one head-to-user branch,

$$d_M = \gamma^M d_0 + \sum_{m=0}^{M-1} \gamma^m \tag{7.3}$$

where $d_0 = 1$.

The summation is easily carried out:

$$d_M = \gamma^M + \frac{\gamma^M - 1}{\gamma - 1} = \frac{\gamma^{M+1} - 1}{\gamma - 1} \tag{7.4}$$

If we define the efficiency of fiber utilization η as the number of users served divided by the total length of fiber (in units)

$$\eta = \gamma^M \Big/ \frac{\gamma^{M+1} - 1}{\gamma - 1} \tag{7.5}$$

$$= \frac{\gamma - 1}{\gamma - (1/\gamma^M)} \tag{7.6}$$

For large M, we can further simplify

$$\boxed{\eta \simeq 1 - \frac{1}{\gamma}} \tag{7.7}$$

which shows that the efficiency increases for higher degrees of splitting. In the limit $\gamma \to \infty$, $\eta \to 1$, which is what we expect from a star distribution network. In any event, this is worse than for a bus distribution network where each successive fiber section adds 1 to the total number of users.

Distribution tree networks are simple to analyze if we redefine the tap fraction t_k to represent the splitting fraction at each node. We assume that all the users are situated at the lowest level leaves, and have shown that the efficiency of fiber usage is strictly less than that of a bus distribution network. But the performance of such a network is better than that of a bus network—increased number of users, or propagation distance or lower input signal power requirement, depending on which of these parameters we vary.

Nevertheless, a tree network in which all the users are situated at the lowest level isn't always practical. A particular implementation may call for a hybrid solution—bus distribution for the major part, with a tree at the terminus. The possibilities that a network designer can explore are quite extensive, and the framework we have provided can be easily adapted, and our code modified, as appropriate for a given problem.

Chapter 8

Conclusion

Let your body stir, if your mind is still becalmed. Swim.
No, now, like her, I can't cope with a crowd.
But you do, don't you, when you play as an extra fiddle in an orchestra.
How about walking? Walk to where you can walk to.
Walk around if you have nowhere to go.
It is five in the morning, but this is wintry London, there is no Venetian dawn.
The drifters out of the night pass those who are drifting into the day.
—V. SETH, An Equal Music (1999)

A section of erbium-doped fiber (EDF) can be pumped at 1480 nm or at 980 nm to provide gain for a number of signal channels in the 1530–1560 nm frequency band. A typical source for the pump is a semiconductor diode laser. The common erbium-doped fiber amplifier (EDFA) package integrates the pump diode and the EDF, but there is no physical barrier on isolating these two elements, and using a section of undoped fiber to connect them. In fact, the same section of signal-carrying single mode fiber (SMF) that couples to the EDF can be used to feed in the pump beam as well. In an optical network, the pumping sources are the components of the fiber amplifier that require management and control—the gain-providing section of EDF can be treated just like any other section of fiber from a maintenance perspective. The separation of these two components considerably simplifies network design.

In transmitting a signal from a source to a receiver, the most important channel model is that of an optical amplifier chain. The users are situated along the length of this chain, and at each receiver, a fraction of the signal power is coupled out of the

fiber ('tapped') and used for detection. We need repeated amplification to offset the losses incurred by absorption along the SMF and by the taps. It's quite intuitive that the signal-to-noise ratio (SNR) of an optical amplifier chain is maximized when all the gain is situated at the input end of the fiber—as far away from the detector as possible. We've formulated this result precisely using standard methods in optimization and results from game theory. The principal concept is that of a Nash solution for a cooperative game, and uses the fact that the strategy sets are compact and the individual reward functions are quasi-concave.

A well-known model [6] is used to demonstrate a particular method for constructing such remotely-pumped EDFA chains. The results clearly show that the number of receivers that can be supported with amplification is substantially higher than is possible without amplification. But we've seen that this method is characterized by large parametric dependencies: changing the numerical values of the system parameters by a small amount results in very large changes in the final answers. Given the nature of the formulae involved, it's not easy to optimize these results, and anyway, the basic principles on which we've built this model aren't really relevant for bus networks which service a single user at each tap location.

Since both the tap fractions and the inter-user distance in typical bus distribution networks are small, we assume that the entire bus is a lightly-doped EDF. The fiber serves simultaneously as the propagation (source-to-receiver) and gain medium. In order to characterize the performance of these types of networks, we've derived a simple model for the rate equations that describes an EDFA from first principles. We explicitly account for the taps by using a singular density function that represents the fraction of signal power that is coupled out at each user, and have solved the model in the steady-state conditions to identify the upper bounds on the number of receivers that can be supported.

We emphasize that the mathematical technique that enables us to get simple answers to this problem is the method of integrating under the differential in the rate

equation for the upper level population [25]. The principle on which this analysis rests is the well-known theorem of dominated convergence, and we've taken advantage of this by *constructing* a singular density function that captures the essential aspects of the detection mechanism while still satisfying the dominated convergence theorem. Other phenomena in EDFAs may be conveniently modeled this way, and Appendix B provides a starting point.

Investigating the dynamic behavior of EDFAs can be tricky or impractical. One particular case has important implications for network designers—the effects of changing the tap function. Using the Laplace transform and perturbation theory, we've solved the system of differential equations that model the EDFA. The low-pass filter nature of the EDFA as pointed out in a different context in [25] is once again evident, and we've identified the time scale over which these perturbative effects are felt. Furthermore, the parameters on which this model depend are precisely those that are supplied by typical network management & control services.

Appendix A

Inequality constraints

The Massachusetts Institute of Technology is committed to the principle of equal opportunity in education and employment. The Institute does not discriminate against individuals on the basis of race, color, sex, sexual orientation, religion, disability, age, veteran status, ancestry, or national or ethnic origin in the administration of its educational policies, admissions policies, scholarship and loan programs, and other Institute administered programs and activities, but may favor US citizens or residents in admissions and financial aid.

—MIT Nondiscrimination Policy (1999)

A generalization of the SNR optimization problem [2.19] is

$$\begin{aligned} \max \quad & \mathrm{SNR}\left(G_1, G_2, \ldots, G_N, L_1, L_2, \ldots, L_N\right) \\ & \text{subject to} \quad \prod_{k=1}^{N} G_k \leq G_{max}, \quad G_k \geq 1 \,\forall\, k, \\ & \qquad\qquad \prod_{k=1}^{N} L_k \leq L_{min}, \quad 0 < L_k \leq 1 \,\forall\, k. \end{aligned} \tag{A.1}$$

where we've replaced the equality constraints with inequalities. In physical terms, we require that the overall gain of the optical amplifier chain be *no more than* a certain threshold G_{max}. Since $L_{min} \propto e^{-\alpha d}$ for propagation distance d, the second constraint implies that the overall propagation distance should be *no less than* a threshold corresponding to L_{min}.

In this section, we demonstrate that, for all practical purposes, this problem re-

duces to the earlier problem,

$$\begin{aligned} \max \quad & \mathrm{SNR}\,(G_1, G_2, \ldots, G_N,\, L_1, L_2, \ldots, L_N) \\ \text{subject to} \quad & \prod_{k=1}^{N} G_k = G_{max}, \quad G_k \geq 1\,\forall\, k, \\ & \prod_{k=1}^{N} L_k = L_{min}, \quad 0 < L_k \leq 1\,\forall\, k. \end{aligned} \tag{A.2}$$

A.1 Sub-optimal in G

First, consider the gain constraint in [A.1]. Let

$$G_0 \triangleq \prod_{k=1}^{N} G_k < G_{max} \tag{A.3}$$

Writing out the SNR and isolating the dependence on G_1,

$$\mathrm{SNR}(G_0) = \frac{P_0}{2\alpha} \frac{G_1 L_1 \left(\prod_{k=2}^{N} G_k L_k\right)}{K + (G_1 - 1) L_1 \left(\prod_{k=2}^{N} G_k L_k\right) + \mathcal{N}_{2+}} \tag{A.4}$$

where $\mathcal{N}_{2+}$ represents the ASE noise contribution from all stages after the first in the optical amplifier chain.

Consider the arbitrated reallocation

$$G_1 \mapsto G_1' = \beta G_1 \qquad \beta > 1 \tag{A.5}$$

$$G_2 \mapsto G_2' = G_2 \tag{A.6}$$

$$\vdots \qquad \vdots \tag{A.7}$$

$$G_N \mapsto G_N' = G_N \tag{A.8}$$

where we choose β so that $\prod_{k=1}^{N} G_k' = G_{max}$. This is certainly feasible, and the new

SNR is

$$\mathrm{SNR}(G_{max}) = \frac{P_0}{2\alpha} \frac{\beta G_1 L_1 \left(\prod_{k=2}^{N} G_k L_k\right)}{K + (\beta G_1 - 1) L_1 \left(\prod_{k=2}^{N} G_k L_k\right) + \mathcal{N}_{2+}} \tag{A.9}$$

Taking the ratio of [A.9] and [A.4],

$$\frac{\mathrm{SNR}(G_{max})}{\mathrm{SNR}(G_0)} = \beta \frac{K + (G_1 - 1) L_1 \left(\prod_{k=2}^{N} G_k L_k\right) + \mathcal{N}_{2+}}{K + (\beta G_1 - 1) L_1 \left(\prod_{k=2}^{N} G_k L_k\right) + \mathcal{N}_{2+}} \tag{A.10}$$

Simplifying this expression, $\mathrm{SNR}(G_{max}) > \mathrm{SNR}(G_0)$ if (and only if)

$$\boxed{K - \left(\prod_{k=2}^{N} G_k\right) \underbrace{\left(\prod_{k=1}^{N} L_k\right)}_{L_{min}} + \mathcal{N}_{2+} > 0} \tag{A.11}$$

As before, $K \leq 1/2$ and $L_{min} \ll 1$. Since our optimal solution is to set $G_2 = G_3 = \cdots = G_N = 1$, this condition is usually satisfied in practice. In physical terms, if we assume that the signal shot noise doesn't dominate the ASE-signal beat noise, it's better to raise the overall allocatable gain to the highest attainable limit.

The Nash solution that we've found is consistent with the above assumption [A.11], and the inequality constraint for G_i in [A.1] can usually be replaced with an equality constraint.

A.2 Sub-optimal in L

Next, we consider the L constraint in [A.1]. Let

$$L_0 \triangleq \prod_{k=1}^{N} L_k < L_{min} \tag{A.12}$$

Writing out the SNR to isolate the dependence on L_N,

$$\mathrm{SNR}(L_0) = \frac{P_0}{2\alpha} \frac{G_N L_N \left(\prod_{k=1}^{N-1} G_k L_k\right)}{K + \mathcal{N}_{-1} G_N L_N + (G_N - 1) L_N} \tag{A.13}$$

where $\mathcal{N}_{2+}$ represents the ASE noise contribution from all stages until the last in the optical amplifier chain.

Consider the arbitrated reallocation

$$L_1 \mapsto L_1' = L_1 \tag{A.14}$$

$$L_2 \mapsto L_2' = L_2 \tag{A.15}$$

$$\vdots \qquad \vdots \tag{A.16}$$

$$L_N \mapsto L_N' = \gamma L_N \qquad \gamma > 1 \tag{A.17}$$

where we choose γ so that $\prod_{k=1}^{N} L_k' = L_{min}$. This is a feasible reallocation, and the new SNR is

$$\mathrm{SNR}(L_{min}) = \frac{P_0}{2\alpha} \frac{\gamma G_N L_N \left(\prod_{k=1}^{N-1} G_k L_k\right)}{K + \mathcal{N}_{-1} \gamma G_N L_N + \gamma (G_N - 1) L_N} \tag{A.18}$$

Taking the ratio of [A.18] and [A.13],

$$\frac{\mathrm{SNR}(L_{min})}{\mathrm{SNR}(L_0)} = \gamma \frac{K + \mathcal{N}_{-1} G_N L_N + (G_N - 1) L_N}{K + \mathcal{N}_{-1} \gamma G_N L_N + \gamma (G_N - 1) L_N} \tag{A.19}$$

Since $\gamma > 1$ and $K > 0$, $\mathrm{SNR}(L_{min}) > \mathrm{SNR}(L_0)$ in all cases: the SNR defined by a strict inequality constraint is always sub-optimal. In physical terms, an attenuator cannot improve the SNR.

The inequality constraint for L_i in [A.1] can always be replaced with an equality constraint.

Appendix B

The validity of equation 5.10

The use of improper functions thus does not involve any lack of rigor in the theory, but is merely a convenient notation, enabling us to express in a concise form certain relations which we could, if necessary, rewrite in a form not involving improper functions, but only in a cumbersome way which would tend to obscure the argument.
—P.A.M. DIRAC, The Principles of Quantum Mechanics (1930)

Starting from the rate equation [5.1], we integrate both sides with respect to z to get [5.10]. The main mathematical point we address here is the exchange of the order of differentiation and integration: is such an operation valid for our model? Specifically, we ask the question: is it true that

$$\frac{d}{dt}\int_{-\infty}^{\infty} N_2(z,t)\,dz = \int_{-\infty}^{\infty} \frac{\partial}{\partial t} N_2(z,t)\,dz \tag{B.1}$$

We'll use a particular version of the following well-known theorem [21],

Theorem 4 (The Dominated Convergence Theorem) *Consider a family* $\{f_h\}$ *of real-valued functions on* $\mathbf{R}$, *where* $0 < |h| < H$, *(H constant). Assume that*

1. f_h *is integrable for each* h,
2. *there exists a function* f *such that* $f_h \to f$ *almost everywhere as* $h \to 0$,
3. *there exists an integrable function* G, *independent of* h, *such that* $|f_h| \leq G$ *for all* h.

Then f is integrable, and

$$\lim_{h\to 0}\int f_h = \int \lim_{h\to 0} f_h = \int f.$$

Using this result, we can derive a sufficient (but not necessary) condition for our result [21, pages 153–154],

Theorem 5 *Let N_2 be a real-valued function defined on $\mathbf{R}\times J$ where J is an open interval in $\mathbf{R}$. Assume that for each fixed $t\in J$,*

1. *$z\mapsto N_2(z,t)$ is integrable,*
2. *$\frac{\partial}{\partial u}N_2(z,u)\Big|_{u=t}$ exists for almost all z,*
3. *there exists an integrable function $G(z)$, independent of t, such that*

$$\left|\frac{\partial}{\partial t}N_2(z,t)\right| \le G(z)$$

for almost all z and for all $t\in J$.

Then,

$$\frac{d}{dt}\int_{-\infty}^{\infty} N_2(z,t)\,dz = \int_{-\infty}^{\infty}\frac{\partial}{\partial t}N_2(z,t)\,dz \qquad (t\in J).$$

The first condition implies that the upper level population density, N_2, should be integrable: of course, without this assumption, we cannot define the path-averaged upper level fraction [5.11] etc.

Next, the set of tap locations $\{z_1, z_2, \ldots, z_M\}$ is the union of a collection of disjoint points (and so is a countable null set), and provided $N_2(z,t)$ is continuous over the rest of the Er-doped fiber, the second condition is also satisfied.

Finally, we'll construct a dominating function $G(z)$, independent of t and inte-

grable, so that the third condition is satisfied. From [5.1]

$$\begin{aligned}\left|\frac{\partial N_2(z,t)}{\partial t}\right| &= \frac{N_2(z,t)}{\tau} + \frac{1}{\rho S}\sum_{k=1}^{N}\frac{\partial P_k(z,t)}{\partial z} \\ &= \frac{N_2(z,t)}{\tau} + \frac{1}{\rho S}\sum_{k=1}^{N}\left[(\gamma_k+\alpha_k)N_2(z,t) - \tilde{\alpha}_k\right]P_k(z,t) \qquad \text{(B.2)}\end{aligned}$$

Since N_2 is the fraction of the laser ion density in the upper state, it must satisfy $N_2 \leq 1$, so that

$$\begin{aligned}\left|\frac{\partial N_2(z,t)}{\partial t}\right| &\leq \frac{1}{\tau} + \frac{1}{\rho S}\sum_{k=1}^{N}\left[(\gamma_k - \alpha_k' - f_k(z)\right]P_k(z,t) \qquad \text{(B.3)} \\ &\leq \frac{1}{\tau} + \frac{1}{\rho S}\sum_{k=1}^{N}\left[(\gamma_k - \alpha_k' - f_k(z)\right]P_{max} \qquad \text{(B.4)} \\ &\equiv G(z)\end{aligned}$$

where we've used [5.4] with $u_k \equiv 1$ for simplicity, and assumed that the optical powers in the fiber are always upper-bounded by P_{max}, independent of wavelength. The least upper bound of P_k in an erbium-doped fiber will usually be the input pump power: note that the pump is always absorbed, and the signal powers are usually far weaker than the pump.

Notice that each of the terms that define $G(z)$ is integrable: the only term that requires inspection is [5.2]

$$f_k(z) = \sum_{m=1}^{M}\delta(z - z_m)I_k(m)\,t_k(z) \qquad \text{(B.5)}$$

where $\delta(z)$ is the Dirac delta "function". By construction, the integral of the product $\delta(z - z_0)t_k(z)$ is indeed well-defined [7], and since the number of taps is finite, we can interchange summation and integration so that the integral of $f_k(z)P_k(z)$ exists, particularly when $P_k(z) \equiv P_{max}$.

As a footnote to this development, we have constructed this particular form of

the tap function with precisely this condition in mind. Exchanging the order of differentiation and integration as in the original EDFA model of [25] is very convenient and we would like to retain that facility while accounting for the taps along the fiber. We've identified the mathematical underpinnings for the validity of that model, and then constructed a density function effectively describing the taps so that we can still carry out the same mathematical procedures as before.

Equation [5.10] is valid provided

1. $z \mapsto N_2(z,t)$ is integrable,
2. $\frac{\partial}{\partial t} N_2(z,t)$ is continuous almost everywhere,
3. the optical power in the each channel is bounded.

Appendix C

Source code

The procedures and applications presented in this book
have been included for their instructional value.
They have been tested with care
but are not guaranteed for any particular purpose.
The publisher does not offer any warranties or representations,
nor does it accept any liabilities
with respect to the programs or applications.
—LaTeX: A Document Preparation System (1994)

C.1 Two-level parametric design

```
% MATLAB Source file
% Specific example: Parametric Design
%
% Name:                  specex.m
%
% Two-level parametric design of a specific distribution
% network. Plots \hat{M}_i for i=1, 2, \aleph.
% Calculates \aleph from \qhat (to be specified)
% Title of plot gives total number of users and of pairs
% of subsidiary buses. Ordinate specifies number of users
% for each pair of subsidiary buses.
%
%  Based on network architecture of V.W.S. Chan, MIT
%
```

```
% Author:                Shayan Mookherjea
% Last modified:         February 5, 2000
%
clear

                         % Computational parameters
R=50;                                    % detector resistance
F=4;                                     % noise figure of detector
T=290*F;                                 % effective detector temp
c=3*10^(8);                              % speed of light in vacuum
l=1.55*10^(-6);                          % center wavelength
nu=c/l;                                  % center frequency
eta=0.8;                                 % quantum efficiency of detector
h=6.634*10^(-34);                        % Planck's constant
e=1.6*10^(-19);                          % charge of electron
k=1.38*10^(-23);                         % Boltzmann's constant
Be=10*10^(9);                            % Electronic Bandwidth
Bo=2*Be;                                 % Optical Bandwidth

                         % EDFA Characteristics
AlphaP = 1.445;                          % pump absorption coefficient
AlphaS = 2.512;                          % signal absorption coefficient
EtaP = 0.37;                             % to get pump emission coefficient
EtaS = 1.42;                             % to get signal emission coefficient
BAlpha = 1.122;                          % background absorption coefficient
GammaP = EtaP*AlphaP;                    % pump emission coefficient
GammaS = EtaS*AlphaS;                    % signal emission coefficient

AvG = 50;                                % Avalanche Gain = 50
Q = 6;                                   % target SNR
const=e*eta/(h*nu);                      % optical power to  photodiode current

% Initialization
DeltaL = 0.5;                            % spacing (km) between bus taps
```

```
Deltad = 0.05;                          % spacing between sub-taps

G0 = 2.0;                               % default gain along bus (10 dB)
ExLossB = 0.7943;                       % excess loss per bus stage (-1.0 dB)
ExLossS = 0.8913;                       % excess loss per sub stage (-0.5 dB)

SigIn = 1e-4;                           % input signal power - 10 dBm

qhat = 3;                               % required norm. input pump power
q0 = 250;                               % supplied norm. input pump power
                                        % calculate number of bus taps allowed
notaps = floor(log(qhat/q0)/(log(ExLossB)-BAlpha*DeltaL));
pmin = zeros(1,notaps);                 % will store min reqd. power
mhat = zeros(1,notaps);                 % will store number of users per bus tap

psig = zeros(1,notaps);                 % signal power tapped after each amp
pase = zeros(1,notaps);                 % ase power tapped after each amp
tapfr= zeros(1,notaps);

tapfr0 = 1 - exp(BAlpha*DeltaL)/G0;     % the (default) tap fraction

for i=1:notaps,                         % you have the option of non-uniform taps
% tapfr(i) = tapfr0+i*(0.9999 - tapfr0)/notaps;
        tapfr(i) = tapfr0;              % uniform tap fractions
        psig(i)  = SigIn *tapfr(i)*G0;
        pase(i)  = h*nu*Bo*i*(G0-1)*tapfr(i);
                                        % the ase power tapped

      a = AvG;                                  % the preamplifier gain
      id2=a*const*pase(i);                      % ase current
      b = 2*Q*Q*(id2*2*Be/Bo+e*Be)*a;           % Linear in I_s
      c = Q*Q*(2*id2*2*Be*e + id2*id2*Be*(2*Bo-Be)/(Bo*Bo) + 4*k*T*Be/R);
      rt = b/(2*a*a) + (sqrt(b*b+4*c*a*a))/(2*a*a);
      pmin(i)=rt/const;
```

```
                                        % no. of users along a d-tap
mhat(i)=4*floor(log(4*pmin(i)/psig(i)) / (log(1 - 4*pmin(i)/psig(i))
          - BAlpha*Deltad + log(ExLossS)));
end

totalusers = sum(mhat);                 % total number of users
plot(mhat)
       xlabel('At Tap No.')
       ylabel('No. of users')
ttlstr = strcat('Number of users :',num2str(totalusers),' with :',
            num2str(notaps),' taps');
       title(ttlstr)
```

C.2 Equal taps: Figure 5-1

```
% MATLAB Source file
% Equal tap fractions
%
% Name:          eqtap.m
%
% This program calculates the input signal power necessary
% along an optical amplifier bus network. The tap locations
% are uniformly spaced, and the receiver density
% is supplied
%
% External Calls:   One of the following three.
%   SNRClc2a.m      Solves quadratic equation directly
%                       approximate but fast.
%   SNRClc2b.m      Use the Q factor correctly
%                       more accurate but slower.
%   SNRClc2c.m      According to (2.17) by Humblet
%                       and Azizoglu
% Author:           Shayan Mookherjea
% Last modified:    February 20, 2000
```

```
%
clear
global L rxden                  % used by the function that
                                % calculates the SNR threshold
% Definitions
global R F T const nu eta h e k Be Bo
global AlphaP AlphaS EtaP EtaS BAlpha GammaP GammaS AvG Q tapfr

          % Computational parameters
R=50;                           % detector resistance
F=4;                            % noise figure of detector
T=290*F;                        % effective detector temp
c=3*10^(8);                     % speed of light in vacuum
l=1.55*10^(-6);                 % center wavelength
nu=c/l;                         % center frequency
eta=0.8;                        % quantum efficiency of detector
h=6.634*10^(-34);               % Planck's constant
e=1.6*10^(-19);                 % charge of electron
k=1.38*10^(-23);                % Boltzmann's constant
Be=10*10^(9);                   % Electronic Bandwidth
Bo=2*Be;                        % Optical Bandwidth

          % EDFA Characteristics
AlphaP = 1.445;                 % pump absorption coefficient
AlphaS = 2.512;                 % signal absorption coefficient
EtaP = 0.37;                    % to get pump emission coefficient
EtaS = 1.42;                    % to get signal emission coefficient
BAlpha = 1.122;                 % background absorption coefficient
GammaP = EtaP*AlphaP;           % pump emission coefficient
GammaS = EtaS*AlphaS;           % signal emission coefficient

AvG = 50;                       % Avalanche Gain = 50
Q = 6;                          % target SNR
const=e*eta/(h*nu);             % optical power to  photodiode current
```

```
tapfr = 0.01;                % uniform tap fraction

% Initialization
indx=0;                      % counts how many taps
Lmax = 40;                   % calculation ends at 60 km
DeltaL = 1;                  % plotting intervals
spreq=zeros(1, (Lmax/DeltaL));  % Signal Power REQuired
L=DeltaL;                    % start off the first tap at (km)
rxden = 100;                 % receiver density (per km)

format compact;              % remove this line for speed (1/3)
% the WHILE loop
while (L<=Lmax),             % stop at max length
    indx=indx+1;             % also count how many taps
    disp(indx)               % remove this line for speed (2/3)
%   spreq(indx)=SNRClc2a/tapfr;
spreq(indx)=abs(fzero('SNRClc2b',2e-5,optimset('disp', 'off')))/tapfr;
%   spreq(indx)=SNRClc2c/tapfr;

    L=L+DeltaL;              % update the total length
end
format loose;                % remove this line for speed (3/3)

% Plot the required signal power
xax=linspace(DeltaL, DeltaL*indx, indx);
insigp=30+10*log10(spreq);                  % convert to dBm
plot(xax, insigp)                           % plot with labels
    xlabel('Propagation distance (km)')
    ylabel('Required Signal Power (dBm)')
```

C.2.1 Equal taps: SNR calculation: Method 1

```
% MATLAB Source file
% Minimum signal power required to meet Q
%
```

```
% Provides: SNRClc2a(x)
% Needs:  Global variables as defined below
% Algorithm: Directly solve the quadratic equation
%
% Author: Shayan Mookherjea
% Last Modified: January 14, 2000

function [DThr] = SNRClc2a(x)
global L rxden

% Computational parameters
global R F T const nu eta h e k Be Bo
global AlphaP AlphaS EtaP EtaS BAlpha GammaP GammaS AvG Q tapfr

% Power relationships
pase=(GammaS/(GammaS+AlphaS))*(AlphaS+BAlpha+(rxden*tapfr))*L*h*nu*Bo;
% noise power due to ASE

% We solve the quadratic equation for the minimum
% detectable power with a preamplifier in front.

   a = AvG; % the preamplifier gain
   id2=a*const*pase; % ase current
   b = 2*Q*Q*(id2*2*Be/Bo+e*Be)*a;  % Linear in I_s
   c = Q*Q*(2*id2*2*Be*e + id2*id2*Be*(2*Bo-Be)/(Bo*Bo) + 4*k*T*Be/R);
   rt = b/(2*a*a) + (sqrt(b*b+4*c*a*a))/(2*a*a);
   pmin=rt/const;
DThr = pmin; % This is the detection threshold (W)
```

C.2.2 Equal taps: SNR calculation: Method 2

```
% MATLAB Source file
% Minimum signal power required to meet Q
%
% Provides: SNRClc2b(x)
```

```
% Needs:  Global variables as defined below
% Algorithm: Find the zero of SNR- Q^2 using 'fzero'
%
% Author: Shayan Mookherjea
% Last Modified: January 14, 2000

function [SNRT] = SNRClc2b(x)
global L rxden

% Computational parameters
global R F T const nu eta h e k Be Bo
global AlphaP AlphaS EtaP EtaS BAlpha GammaP GammaS AvG Q tapfr

% Power relationships
psig=x; % signal power along bus
pase=(GammaS/(GammaS+AlphaS))*(AlphaS+BAlpha+(rxden*tapfr))*L*h*nu*Bo;
% noise power due to ASE

% Convert to currents and find SNRT
ISg = psig*(e*eta)/(h*nu);
INs = pase*(e*eta)/(h*nu);
Sig = (AvG*ISg)^2; % MS signal component
IN1 = 2*INs*2*Be/Bo *AvG *AvG*ISg; % signal-ASE beat noise
IN2 = 2*e*Be*(ISg+2*INs)*AvG; % shot noise
IN3 = (INs*AvG)^2*(Be/(Bo^2))*(2*Bo-Be); % ASE-ASE beat
IN4 = 4*k*T*Be/R; % thermal noise
Qform = Sig/(((sqrt(IN1+IN2+IN3+IN4))+sqrt(IN4))^2);% Personick suff. stat.
SNRT = Qform - Q^2; % what we need to find fzero of
```

C.2.3 Equal taps: SNR calculation: Method 3

```
% MATLAB Source file
% Minimum signal power required to meet Q
%
% Provides:                    SNRClc2a(x)
```

```
% Needs:                          Global variables as defined below
% Algorithm:                      Humblet-Azizoglu
%
% Author:                         Shayan Mookherjea
% Last Modified:                  February 20, 2000

function [DThr] = SNRC1c2c(x)
global L rxden

                                  % Computational parameters
global R F T const nu eta h e k Be Bo
global AlphaP AlphaS EtaP EtaS BAlpha GammaP GammaS AvG Q tapfr

M0=36;                            % M in equation (2.17)
% Power relationships
pase=(GammaS/(GammaS+AlphaS))*(AlphaS+BAlpha+(rxden*tapfr))*L*h*nu*Bo;
                                  % noise power due to ASE
pmin = pase*Q*Q*(1+h*nu/(2*pase))+pase*Q*sqrt(M0*(1+h*nu/pase));
DThr = pmin;
```

C.3 Inequal taps: Figure 5-2

```
% MATLAB Source file
% Inequal tap fractions
%
% Name:         ineqtap.m
%
% This program calculates the tap fractions necessary
% along an optical amplifier bus network. The tap locations
% are uniformly spaced by DeltaL, and at each tap location,
% no more power is tapped than is necessary to detect the
% signal given an SNR test.
%
% The tap fractions affect the accumulated ASE power.
%
% External Calls:   SNRCalc.m
%
% Author:           Shayan Mookherjea
% Last modified:    January 9, 2000
%                   Moved definitions to main file, and GLOBAL
%                   Simpler display (user feedback) (quicker)
%
clear
global L tapsum                 % used by the function that
                                % calculates the SNR threshold
% Definitions
global R F T const nu eta h e k Be Bo
global AlphaP AlphaS EtaP EtaS BAlpha GammaP GammaS AvG Q

            % Computational parameters
R=50;               % detector resistance
F=4;                % noise figure of detector
T=290*F;            % effective detector temp
c=3*10^(8);         % speed of light in vacuum
l=1.55*10^(-6);     % center wavelength
```

```
nu=c/l;              % center frequency
eta=0.8;             % quantum efficiency of detector
h=6.634*10^(-34);    % Planck's constant
e=1.6*10^(-19);      % charge of electron
k=1.38*10^(-23);     % Boltzmann's constant
Be=10*10^(9);        % Electronic Bandwidth
Bo=2*Be;             % Optical Bandwidth

          % EDFA Characteristics
AlphaP = 1.445;      % pump absorption coefficient
AlphaS = 2.512;      % signal absorption coefficient
EtaP = 0.37;         % to get pump emission coefficient
EtaS = 1.42;         % to get signal emission coefficient
BAlpha = 1.122;      % background absorption coefficient
GammaP = EtaP*AlphaP;   % pump emission coefficient
GammaS = EtaS*AlphaS;   % signal emission coefficient

AvG = 50;                % Avalanche Gain = 50
Q = 6;                   % target SNR
const=e*eta/(h*nu);      % convert optical power to  photodiode current

% Initialization
tapvecs=0;                      % will contain tap fraction sequence
tapsum=0;                       % the running sum of tap fraction
indx=0;                         % counts how many taps
Lmax = 60;                      % calculation ends at 60 km
PInput = .001;                  % Input power (W)
DeltaL = 0.04;                  % spacing between taps (km)
L=0;                            % start off the first tap at (km)
newtap= 0;                      % the latest tap fraction

format compact                  % suppress extra line feeds
% the WHILE loop
while (newtap <1)&(L<=Lmax),    % stop at max length
```

```
    indx=indx+1;                % also count how many taps
    newtap=SNRCalc/PInput;
    tapvecs(indx)=single(newtap);   % ...is added to the list
                                    % Update the global variables
    tapsum=tapsum+newtap;           % and the running sum is updated
    L=L+DeltaL;                     % as is the total length
end
format loose                        % back to usual format
% Plot the tap fractions
xax=linspace(DeltaL, DeltaL*indx, indx);
semilogy(xax, tapvecs)                      % plot with labels
    xlabel('Propagation distance (km)')
    ylabel('Tap fractions')
ttlstr=strcat('Input signal power:',num2str(round(30+10*log10(PInput))),' dBm');
    title(ttlstr)
```

C.3.1 Inequal taps: SNR calculation

```
% MATLAB Source file
% Calculates minimum signal power needed (mW)
% to meet SNR criterion
%
% Provides: SNRCalc(x)
% Needs:  Global variables as defined below
%
% Author: Shayan Mookherjea
% Last Modified: January 13, 2000

function [DThr] = SNRCalc(x)
global L tapsum

% Computational parameters
global R F T const nu eta h e k Be Bo
global AlphaP AlphaS EtaP EtaS BAlpha GammaP GammaS AvG Q
```

```
% Power relationships
pase=GammaS/(GammaS+AlphaS)*((AlphaS+BAlpha)*L+tapsum)*h*nu*Bo;
% noise power due to ASE

% We solve the quadratic equation for the minimum
% detectable power with a preamplifier in front.

   a = AvG; % the preamplifier gain
   id2=a*const*pase; % ase current
   b = 2*Q*Q*(id2*2*Be/Bo+e*Be)*a;  % Linear in I_s
   c = Q*Q*(2*id2*2*Be*e + id2*id2*Be*(2*Bo-Be)/(Bo*Bo) + 4*k*T*Be/R);
   rt = b/(2*a*a) + (sqrt(b*b+4*c*a*a))/(2*a*a);
   pmin=rt/const;
DThr = pmin; % This is the detection threshold
```

Appendix D

Further remarks on the optimization problem

Any linear ordering is a dense sum of scattered linear orderings.
—F. HAUSDORFF, *Math. Ann.* **65**, 435–505 (1908)

We have developed our optimization arguments in Chapter 2 along physical lines, for the final result is more physically meaningful than mathematically profound. Also, there are a number of other factors of physical origin which will affect any practical implementation of the results, e.g. finite input pump power to the EDFA. In this concluding section, we briefly highlight in a more mathematical setting the principles and our motivation for the game-theoretic analysis presented earlier.

We first see that altering our figure-of-merit slightly makes the analysis of optical amplifier chains quite a bit simpler. This is further explored in the second section in the context of the general theory of linear orderings.

D.1 Negative noise-to-signal ratio

Our figure of merit in the original optimization problem is the signal-to-noise ratio of the N^{th} amplifier along a chain of optical amplifiers. We can identify the associated reward functions with quantities that our hypothetical players may actually, in reality, want to maximize: the relationship between signal-to-noise ratio and communications performance is well-known. Nevertheless, the mathematics are somewhat

more elegant if we introduce the equivalent concept of a negative noise-to-signal ratio as follows:

Definition 6 (Negative noise-to-signal ratio: NNSR) *We define the negative noise-to-signal ratio in relation to our original definition of the signal-to-noise ratio SNR as*

$$NNSR = -1/SNR$$

Using the same approximations as in Chapter 2, i.e. the dominant noise sources are the signal-ASE beat noise and the signal shot noise, we can write the NNSR after the N^{th} stage as

$$\text{NNSR}_N = -\frac{2\alpha}{P_0}\frac{K + \sum_{k=1}^{N}(G_k - 1)L_k \prod_{j=k+1}^{N} G_j L_j}{\prod_{k=1}^{N} G_k L_k} \tag{D.1}$$

where $K = eB_o/2\alpha$ with $\alpha = \mu h\nu B_o$ as before and P_0 is the signal power input at the head of the chain.

We analyze our optimization problem in terms of NNSR rather than SNR. Firstly, we can now consider, if necessary, the noiseless case very simply: NNSR remains finite, while SNR $\to \infty$.

Specializing to the two-amplifier, two-attenuator chain as we did earlier, the new figure of merit is

$$\text{NNSR}\,(G_1, G_2, L_1, L_2) = -\frac{2\alpha}{P_0}\frac{K + (G_1 - 1)L_1\, G_2 L_2 + (G_2 - 1)L_2}{G_2 L_2\, G_1 L_1} \tag{D.2}$$

Our arguments proceed exactly as before, with appropriate modifications. For example, we can redraw Figure 2-2 as shown in Figure D-1.

The main advantage of this increased complexity in formulism is that chains of amplifiers can be analyzed quite easily. If we perform a linear transformation on

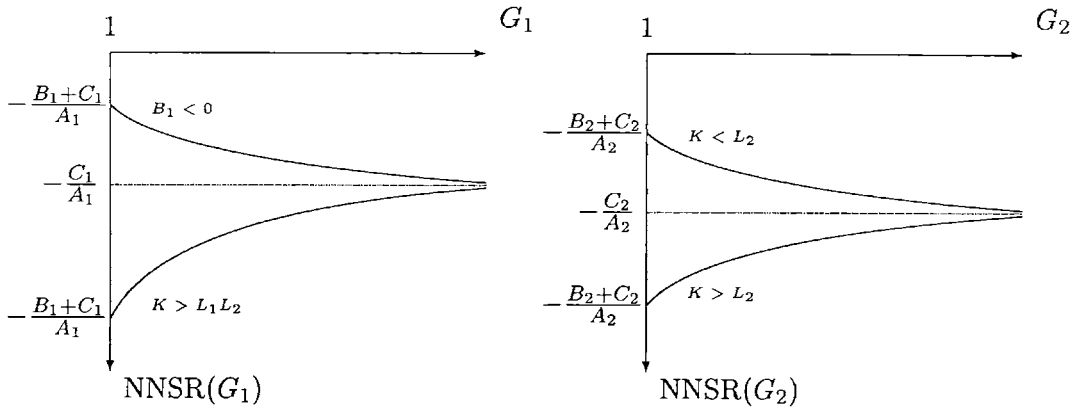

Figure D-1: The dependence of gain on the noise-to-signal ratio NNSR(G_1) and NNSR(G_2).

NNSR,

$$\text{NNSR}' = p\,\text{NNSR} + q, \qquad p \geq 0 \tag{D.3}$$

then both curves in Figure D-1 shift and scale appropriately, but it's still true that

$$p(-C_1/A_1) + q > p(-C_2/A_2) + q \tag{D.4}$$

which implies that our arbitration is still the Pareto optimal Nash solution, which is unique. This is merely a demonstration of the invariance of the arbitration under monotone order-preserving transformations, and we will return to this concept in the next section. First, we interpret this simple mathematical operation in a way that has substantial physical significance.

Let's consider the practically important case $G_k L_k = 1$, so that along a chain of N stages,

$$\overline{\text{NNSR}}_N = \frac{2\alpha}{P_0}\left[K + \sum_{k=1}^{N}(1 - 1/G_k)\right] \tag{D.5}$$

If we add another stage, the resultant $\overline{\text{NNSR}}$ can be written as

$$\overline{\text{NNSR}}_{N+1} = \left[1 - \frac{1}{G_{N+1}}\right]\frac{2\alpha}{P_0}\left[K + \sum_{k=1}^{N}(1 - 1/G_k)\right] + \frac{1}{G_{N+1}}\frac{2\alpha}{P_0}K \tag{D.6}$$

which, since $G_{N+1} \geq 1$ can be written as a linear transformation

$$\overline{\text{NNSR}}_{N+1} = p\,\overline{\text{NNSR}}_N + q \tag{D.7}$$

where $p \geq 0$.

Adding a stage has two implications. First, we change the N player game to an $N + 1$ player game, but this is a trivial modification in nomenclature only: our arbitration procedure can be carried out for any N. Then, we notice that by our definition of the game, we have changed the reward function for **all** the players by a monotone order-preserving **linear** transformation. We know that under precisely such transformations, our Nash solution is still valid and moreover, unique. The use of the negative noise-to-signal ratio makes it simple to demonstrate the correctness of our solution for arbitrary N-stage chains, of course, under the assumptions of feasibility.

D.2 Linear orderings

We have so far considered "free-market economies": in many practical situations, such an assumption may be too liberal. For example, not all G_k may be able to attain G_{max} even if we set all other $G_j = 1, j \neq k$. While there are countless possible

restrictions reality may impose, we briefly discuss a methodology that allows us to understand in which situation our analysis still holds valid.

Consider the set A of all feasible allocations

$$A = \left\{ a = (a_1, a_2, \ldots, a_N) \,\middle|\, a_k = \begin{pmatrix} G_k \\ L_k \end{pmatrix}, \quad k = 1, 2, \ldots, N \right\} \tag{D.8}$$

such that

$$\prod_k \begin{pmatrix} 1 & 0 \end{pmatrix} a_k \leq G_{max} \tag{D.9}$$

$$\prod_k \begin{pmatrix} 0 & 1 \end{pmatrix} a_k \leq L_{min} \tag{D.10}$$

and $1 \leq G_k \leq G_{max}$, $L_{min} \leq L_k \leq 1 \quad \forall k = 1, 2, \ldots, N$.

Our analysis using the signal-to-noise ratio (SNR) establishes a binary relation called a ranking R of A where

$$\forall a_1, a_2 \in A, \quad <a_1, a_2> \in R \quad \text{iff} \quad \mathrm{SNR}_N(a_1) < \mathrm{SNR}_N(a_2) \tag{D.11}$$

and we call $<A, R>$ a linear ordering.

Consider a mapping f of A onto B and a ranking S of B, which we are in general free to choose different from R. As an example, let f be the identity map and S be given by the NNSR. If we can write

$$\forall b_1, b_2 \in B, \quad <b_1, b_2> \in S \quad \text{iff} \quad \mathrm{NNSR}_N(a_1) < \mathrm{NNSR}_N(a_2) \tag{D.12}$$

then the linear orderings are isomorphic, $<A, R> \simeq <B, S>$. In our example, which represents our conversion from SNR to NNSR, f is in fact an automorphism.

For the identity map f, we can choose a different S according to any preferred criterion—for example, the bit error rate (BER) rather than NNRS—such that the orderings are isomorphic, i.e. have the same order type. It's obvious that having the

same order type is an equivalence relation on the class of all linear orderings.

Next, we can generalize f such that the linear orderings $< A, R >$ and $< B, S >$ again have the same order type: a simple example is

$$f(a) = f\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 10\log_{10}(a_1) - 10\log_{10}(\hat{a}_1) \\ 10\log_{10}(a_2) - 10\log_{10}(\hat{a}_2) \end{pmatrix} \tag{D.13}$$

which, for any choice of $\hat{a}_1$ and $\hat{a}_2$, expresses G_i and L_i in dBr rather than absolute units.

In fact, if order type τ_R is the representative of the equivalence class of $< A, R >$, we are free to choose any linear ordering S of order type τ_S such that τ_R is embeddable is τ_S, $\tau_R \preceq \tau_S$.

An added degree of complication arises if B is not Dedekind complete. We restrict our attention to $S = R$. Let $< I, S >$ be a linear ordering and for each $i \in I$ let $< A_i, S_i >$ be a linear ordering. For the moment, let's assume that A_i are disjoint. Define the genralized sum $\sum\{A_i \mid i \in I\}$ to be the linear ordering $< C, S >$ where $C = \cup\{A_i \mid i \in I\}$. If A_i are not disjoint, we modify our definition [22] by first replacing each $< A_i, S_i >$ by an isomorphic copy $< A_i', S_i' >$ so that $\{A_i' \mid i \in I\}$ is pairwise disjoint.

It follows trivially that our optimality conclusions hold: C is a linear ordering with the same binary relation S (i.e. NNSR) as each of its constituents. Moreover, we can further adapt A_i to a suitable form by the following lemma [22, pages 19–20]

Theorem 6 *Let f be an isomorphism of $< I, S >$ onto $< I', S' >$. For each $i \in I$ let $< A_i, S_i >$ be a linear ordering and for each $i' \in I'$ let $< A_i', S_i' >$ be a linear ordering; and assume that for each $i \in I$,*

$$< A_i, S_i > \simeq < A_{f(i)}, S_{f(i)} >$$

Then

$$\sum \{A_i \mid i \in I\} \simeq \sum \{A_{i'} \mid i' \in I'\}$$

This gives us considerable freedom in our choice of the isomorphic mappings.

Finally, we note that the constructive process we used in analyzing the N-amplifier chain in Chapter 2—starting at the left-hand end of the chain and analyzing groups of two stages—is a simple example of iterated condensations of the linear ordering, each iteration of which yields a linear ordering. An appropriate choice of condensation maps can considerably simplify the analysis in a given situation

References

[1] C. Berge and A. Ghouila-Houri. *Programming, Games and Transportation Networks.* Methuen, London, 1965.

[2] D.N. Chen, K. Motoshima, and E. Desurvire. A transparent optical bus using a chain of remotely pumped erbium-doped fiber amplifiers. *IEEE Photonics Technology Letters*, 5(3):351–353, March 1993.

[3] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg. *Atom-Photon Interactions.* John Wiley & Sons, paperback edition, 1998.

[4] S.P. Craig-Ryan, B.J. Ainslie, and C.A. Millar. Fabrication of long lengths of low excess loss erbium-doped optical fibre. *Electronics Letters*, 26(3):185, 1990.

[5] E. Desurvire. Analysis of distributed erbium-doped fiber amplifiers with fiber background loss. *IEEE Photonics Technology Letters*, 3(7):625–628, July 1991.

[6] E. Desurvire. *Erbium-doped Fiber Amplifiers.* John Wiley & Sons, 1994.

[7] I.M. Gelfand and G.E. Shilov. *Generalized Functions*, volume 1. Academic Press, 1964.

[8] C.R. Giles and D. DiGiovanni. Spectral dependence of gain and noise in erbium-doped fiber amplifiers. *IEEE Photonics Technology Letters*, 2(11):797, 1990.

[9] C.H. Henry. Theory of spontaneous emission noise in open resonators and its application to lasers and optical amplifiers. *Journal of Lightwave Technology*, LT-4(3):288–297, 1986.

[10] P Huard, editor. *Point-to-set Maps and Mathematical Programming*, volume 10 of *Mathematical Programming*. North-Holland, Amsterdam, 1979.

[11] P.A. Humblet and M. Azizoglu. On the bit error rate of lightwave systems with optical amplifiers. *Journal of Lightwave Technology*, 9(11):1576–1582, 1991.

[12] B. Huttner, S. Serulnik, and Y. Benaryeh. Quantum analysis of light-propagation in a parametric amplifier. *Physical Review A*, 42(9):5594–5600, November 1990.

[13] A.J. Jones. *Game Theory: Mathematical Models of Conflict.* Ellis Horwood, Chichester, 1980.

[14] D.G. Luenberger. *Microeconomic Theory.* McGraw-Hill, 1995.

[15] L. Mandel and E. Wolf. *Optical Coherence and Quantum Optics.* Cambridge, 1995.

[16] D.E. McCumber. Theory of phonon-terminated optical masers. *Physical Review*, 134(2A):299, 1964.

[17] D. Middleton. *An Introduction to Statistical Communication Theory.* McGraw-Hill, 1960.

[18] S. Mookherjea. Analysis of a photonic switching architecture. Bachelor of Science thesis at the Electrical Engineering department, California Institute of Technology, 1999.

[19] P. Morris. *Introduction to Game Theory.* Springer-Verlag, 1994.

[20] S.D. Personick. Receiver design for digital fiber optic communication systems. *Bell Systems Technical Journal*, 52:843–874, 1973.

[21] H.A. Priestley. *Introduction to Integration.* Clarendon Press, Oxford, 1997.

[22] J.G. Rosenstein. *Linear Orderings.* Academic Press, New York, 1982.

[23] A.E. Siegman. *Lasers.* University Science Books, Sausalito, California, 1986.

[24] R.E. Slusher, P. Grangier, A. Laporta, B. Yurke, and M.J. Potasek. Pulsed squeezed light. *Physical Review Letters*, 59(22):2566–2569, November 1987.

[25] Y. Sun, J.L. Zyskind, and A.K. Srivastava. Average inversion level, modeling and physics of erbium-doped fiber amplifiers. *IEEE Journal of Selected Topics in Quantum Electronics*, 3(4):991–1007, August 1997.

[26] J. Szép and F. Forgó. *Introduction to the Theory of Games.* D. Reidel, Dordrecht, 1985.

[27] A.E. Taylor. *General Theory of Functions and Integration.* Dover Publications, New York, 1985.

[28] Y. Yamamoto. Noise and error rate performance of semiconductor laser amplifiers in PCM-IM optical transmission systems. *IEEE Journal of Quantum Electronics*, QE-16(10):1073–1081, 1980.

[29] A. Yariv. *Quantum Electronics.* John Wiley & Sons, third edition, 1989.

[30] A. Yariv. *Optical Electronics in Modern Communications.* Oxford, fifth edition, 1997.

Index