KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT TOEGEPASTE WETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
Kasteelpark Arenberg 10, 3001 Heverlee

MULTI-USER SIGNAL AND SPECTRA CO-ORDINATION FOR DIGITAL SUBSCRIBER LINES

Promotor:
Prof. Dr. ir. M. Moonen

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de toegepaste wetenschappen

door

Raphael CENDRILLON

December 2004

KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT TOEGEPASTE WETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
Kasteelpark Arenberg 10, 3001 Heverlee

MULTI-USER SIGNAL AND SPECTRA CO-ORDINATION FOR DIGITAL SUBSCRIBER LINES

Jury:
Prof. Dr. ir. G. De Roeck, voorzitter
Prof. Dr. ir. M. Moonen, promotor
Prof. Dr. ir. G. Gielen
Prof. Dr. ir. S. McLaughlin (U. Edinburgh, U.K.)
Prof. Dr. ir. B. Preneel
Prof. Dr. ir. L. Vandendorpe (U.C.L.)
Prof. Dr. ir. J. Vandewalle
Prof. Dr. ir. S. Vandewalle

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de toegepaste wetenschappen

door

Raphael CENDRILLON

U.D.C. 621.391.827 December 2004

D/2004/7515/88

ISBN 90-5682-550-X

Acknowledgement

I would like to thank my family for all the support and love they have given me over the years. My parents have always believed in me and this taught me to believe in myself. To my brothers and sister I love you all very much.

To Prof. Marc Moonen, you've been a great supervisor over the past four years. I've learnt a lot and my experience here has been a very rewarding one. Thanks for your support, enthusiasm, and for believing in me.

To my colleagues at Alcatel: Tom Bostoen, Radu Suciu, Katleen Van Acker, Piet Vandaele, Etienne Van den Bogaert and Jan Verlinden, working with you has been a pleasure.

I would like to thank the reading committee: Prof. Georges Gielen, Prof. Bart Preneel and Prof. Joos Vandewalle, for their time, effort and continued support throughout my doctorate. I would also like to thank the jury members: Prof. Stephen McLaughlin, Prof. Luc Vandendorpe, Prof. Stefan Vandewalle and the chairman Prof. Guido De Roeck.

To my colleagues Wei Yu and George Ginis, your ideas, thoughts and energy have been enlightening and inspiring. Thank you for your company and good humour.

Many others have supported me during my Ph.D. and a few deserve special mention. To Prof. John Cioffi, thank you for hosting me at Stanford University. Your time, interest and continued support have been a blessing. To Dr. Michail Tsatsanis and Dr. Jacky Chow, thank you for hosting me and your support and advice over the years. To Prof. John Homer, you made it all possible, thanks mate!

To the group at the Katholieke Universiteit Leuven: Olivier Rousseaux, Thomas Klasen, Imad Barhumi, Sharon Gannot, Geert Van Meerbergen, Geert Leus, Geert Ysebaert, Koen Vanbleu, Gert Cuypers, Simon Doclo, Geert Rombouts, Ann Spriet, Koen Eneman, Hilde Vanhaute, Toon van Waterschoot, Jan Vangorp, Paschalis Tsiaflakis, Jan Schier, Matteo Montani, and Deepaknath Tan-

dur. Thank you for providing such a friendly, supportive workplace.

Finally to the Belgian people and all the great friends I have made here, Kristien, Jace, Audrey and Matteo, thank you for your hospitality and friendship. Living overseas can be a difficult time. Your companionship has made living here not only a rewarding experience, but an enjoyable one aswell.

Abstract

The appetite amongst consumers for ever higher data-rates seems insatiable. This booming market presents a huge opportunity for telephone and cable operators. It also presents a challenge: the delivery of broadband services to millions of customers across sparsely populated areas. Fully fibre-based networks, whilst technically the most advanced solution, are prohibitively expensive to deploy. Digital subscriber lines (DSL) provide an alternative solution. Seen as a stepping-stone to a fully fibre-based network, DSL operates over telephone lines that are already in place, minimizing the cost of deployment.

The basic principle behind DSL technology is to increase data-rate by widening the transmission bandwidth. Unfortunately, operating at high frequencies, in a medium originally designed for voice-band transmission, leads to crosstalk between the different DSLs. Crosstalk is typically 10-15 dB larger than the background noise and is *the* dominant source of performance degradation in DSL.

This thesis develops practical multi-user techniques for mitigating crosstalk in DSL. The techniques proposed have low complexity, low latency, and are compatible with existing customer premises equipment (CPE). In addition to being practical, the techniques also yield near-optimal performance, operating close to the theoretical multi-user channel capacity.

Multi-user techniques are based on the coordination of the different users in a network, and this can be done on either a spectral or signal level.

Spectra coordination, also known as dynamic spectrum management (DSM), minimizes crosstalk by intelligently setting the transmit spectra of the modems within the network. Each modem must achieve a trade-off between maximizing its own data-rate and minimizing the crosstalk it causes to other modems within the network. The goal is to achieve a fair trade-off between the rates of the different users in the network.

The first part of this thesis investigates the optimal design of transmit spectra for a network of crosstalking DSLs. This problem was previously considered

intractable since it requires the solution of a high-dimensional, non-convex optimization. This thesis uses a dual-decomposition to solve the optimization in an efficient, tractable way. The resulting algorithm, *optimal spectrum balancing*, achieves significant gains over existing spectra coordination algorithms, typically doubling or tripling the achievable data-rate.

The second part of this thesis investigates multi-user signal coordination. In the upstream, reception is done in a joint fashion; the signals received on each line are combined to cancel crosstalk whilst preserving the signal of interest.

Existing crosstalk cancelers are based on decision feedback, which leads to problems with error propagation, high complexity, and a long latency. To address this problem, this thesis presents a simple linear canceler based on the well known zero-forcing criterion. This technique has a low complexity, short latency, and operates close to the theoretical channel capacity.

In the downstream, transmission is done in a joint fashion; predistortion is introduced into the signal of each user prior to transmission. This predistortion is chosen such that it annihilates with the crosstalk introduced in the channel. As a result the customer premises (CP) modems receive a signal that is crosstalk free.

Existing precoder designs either give poor performance or require the replacement of CP modems, which raises a huge legacy issue. To address this problem, this thesis presents a simple linear precoder based on a channel diagonalizing criterion. This technique has a low complexity, does not require the replacement of CP modems, and operates close the the theoretical channel capacity.

Despite the low complexity of the techniques described, signal coordination is still too complex for current implementation. This problem is addressed in this thesis through a technique known as partial cancellation. It is well known that the majority of crosstalk experienced on a line comes from the 3 to 4 surrounding pairs in the binder. Furthermore, since crosstalk coupling varies dramatically with frequency, the worst effects of crosstalk are limited to a small selection of tones. Partial cancelers exploit these facts to achieve the majority of the performance of full cancellation at a fraction of the complexity.

Partial canceler and precoder design is discussed and shown to be equivalent to a resource allocation problem. Given a limited amount of available run-time complexity, a modem must distribute this across lines and tones such that the data-rate is maximized. This thesis presents the optimal algorithm for partial canceler design and several simpler, sub-optimal algorithms. These algorithms are shown to achieve 90% of the data-rate of full cancellation at less than 30% of the complexity.

Notation

Mathematical Notation

x	scalar x
$\mathbf{x}$	vector $\mathbf{x}$
$\mathbf{X}$	matrix $\mathbf{X}$
$[\mathbf{X}]_{\text{row } n}$	row n of matrix $\mathbf{X}$
$[\mathbf{X}]_{\text{col } m}$	column m of matrix $\mathbf{X}$
$\mathbb{X} \setminus \mathbb{Y}$	elements contained in set $\mathbb{X}$ and not in the set $\mathbb{Y}$
$\lvert\mathbb{X}\rvert$	cardinality of set $\mathbb{X}$
$\lvert x\rvert$	absolute value of scalar x
$[x]^+$	$\max(0, x)$
$[x]_a^b$	$\max(a, \min(x, b))$
$\lfloor\cdot\rfloor$	round down to nearest integer
$\lVert\cdot\rVert$	$L2$-norm
$(\cdot)^T$	matrix transpose
$(\cdot)^H$	matrix Hermitian transpose
$\overset{\text{qr}}{=}$	QR decomposition
$\overset{\text{svd}}{=}$	SVD decomposition
$\text{conj}(\cdot)$	complex conjugate
$\text{dec}(\cdot)$	decision operation
$\det(\cdot)$	matrix determinant
$\text{diag}\{\mathbf{x}\}$	diagonal matrix with vector $\mathbf{x}$ as diagonal
$\mathcal{E}\{\cdot\}$	statistical expectation
$I(x; y)$	mutual information between x and y
$\max(x, y)$	maximum of x and y
$\min(x, y)$	minimum of x and y
$\mathcal{O}(\cdot)$	order

Fixed Symbols

$\mathbb{A}^{(N)}$	set of strictly diagonally dominant matrices of size $N \times N$
b_k^n	bitloading of user n on tone k
$b_{k,\mathrm{bc}}^n$	BC single-user bound
$b_{k,\mathrm{mac}}^n$	MAC single-user bound
$\mathcal{F}_K$	DFT matrix of size K
f_s	DMT symbol rate
$\mathbf{H}_k$	crosstalk channel matrix on tone k
$\mathbf{h}_k^n$	column n of $\mathbf{H}_k$
$\overline{\mathbf{h}}_k^n$	row n of $\mathbf{H}_k$
$h_k^{n,m}$	channel from TX m to RX n on tone k
$\mathcal{I}_K$	IDFT matrix of size K
$\mathbf{I}_N$	identity matrix of size N
K	number of DMT-tones
L	Lagrangian dual function
L_k	Lagrangian dual function on tone k
$\mathbb{M}_k^n$	crosstalkers cancelled when detecting user n on tone k
$\underline{\mathbb{M}}_k^n$	crosstalkers not cancelled when detecting user n on tone k
N	number of lines within the binder
$\mathbf{P}_k$	crosstalk precoding matrix on tone k
P_n	transmit power available to modem n
R_n	data-rate on line n
R_n^{target}	target data-rate for line n
s_k^{mask}	PSD mask on tone k
$\widetilde{s}_k^n$	PSD of symbol intended for receiver n on tone k, $\widetilde{x}_k^n$
s_k^n	PSD of TX n on tone k
$\mathbf{s}_n$	length K vector containing PSD of TX n on all tones
$\overline{\mathbf{s}}_k$	length N vector containing PSDs of all TXs on tone k
$\mathbf{U}_k$	left singular-vectors of $\mathbf{H}_k$
$\mathbf{V}_k$	right singular-vectors of $\mathbf{H}_k$
w_n	weight for user n in weighted rate-sum
x_k^n	signal sent by TX n on tone k
$\hat{x}_k^n$	estimate of user n's symbol on tone k
$\widetilde{x}_k^n$	symbol intended for user n on tone k prior to precoding
$\mathbf{x}_k$	transmitted vector on tone k
y_k^n	received signal of line n on tone k
$\mathbf{y}_k$	received vector on tone k
z_k^n	noise of line n on tone k
$\mathbf{z}_k$	noise vector on tone k
α_k	degree of diagonal dominance on tone k
β_k	precoder scaling factor on tone k
Δ_f	inter-tone spacing
Γ	SNR-gap to capacity
Λ_k	singular values of $\mathbf{H}_k$

λ_n	Lagrange multiplier of line n
μ_n	proportion of run-time complexity allocated to user n
σ_k^n	noise power of RX n on tone k
$\widetilde{\sigma}_k^n$	noise power of RX n on tone k after cancellation filter
$\mathbf{0}_{x\times y}$	zeros matrix of size $x \times y$

Acronyms and Abbreviations

ADC	Analog to Digital Converter
ADSL	Asymmetric Digital Subscriber Line
AFE	Analog Front-end
AWG	American Wire Gauge
AWGN	Additive White Gaussian Noise
BC	Broadcast Channel
CDMA	Code Division Multiple Access
CO	Central Office
CLEC	Competitive Local Exchange Carrier
CP	Customer Premises
CPE	Customer Premises Equipment
CWDD	Column-wise Diagonal Dominance
DFC	Decision Feedback Canceler
DFE	Decision Feedback Equalizer
DFT	Discrete Fourier Transform
DMT	Discrete Multi-tone
DP	Diagonalizing Precoder
DS	Downstream
DSLAM	Digital Subscriber Line Access Multiplexer
DSM	Dynamic Spectrum Management
EFM	Ethernet in the First Mile
FDMA	Frequency Division Multiple Access
FEQ	Frequency-domain Equalizer
FFT	Fast Fourier Transform
IC	Interference Channel
IDFT	Inverse Discrete Fourier Transform
ILEC	Incumbent Local Exchange Carrier
ISI	Inter-symbol Interference
IW	Iterative Waterfilling
KKT	Karush Kuhn Tucker
LAN	Local Area Network
MAC	Multi-access Channel
MIMO	Multi-input Multi-output
ONU	Optical Network Unit
PBO	Power Back-off

PSD	Power Spectral Density
RFI	Radio Frequency Interference
RLCG	Resistance Inductance Capacitance Conductance
RT	Remote Terminal
RWDD	Row-wise Diagonal Dominance
RX	Receiver
SIC	Successive Interference Cancellation
SINR	Signal to Interference plus Noise Ratio
SMC	Spectrum Management Centre
SNR	Signal to Noise Ratio
SVD	Singular Value Decomposition
THP	Tomlinson-Harashima Precoder
TX	Transmitter
UMTS	Universal Mobile Telecommunications System
US	Upstream
USD	United States Dollar
VDSL	Very high-speed Digital Subscriber Line
ZF	Zero Forcing
ZFP	Zero Forcing Precoder

Contents

II Multi-user Signal Coordination

4 Receiver Coordination

5 Transmitter Coordination

Chapter 1

Introduction

1.1 Digital Subscriber Lines

Digital communication has undergone a revolution in the last decade. Typical connections speeds have increased from 14.4 kbps in 1994, to 1.5 Mbps today, a hundred-fold improvement. This revolution is being driven by the explosion of the Internet and new high-speed applications like video-streaming, file-sharing of music and movies, teleworking and video-conferencing. The appetite amongst consumers for ever higher data-rates seems insatiable, and will continue to grow as new technologies like *high definition television* (HDTV) take hold.

Sales of broadband access today exceed $22 billion worldwide[72]. This will grow substantially as countries like China and India industrialize. This booming market presents a huge opportunity to telephone and cable operators. It also presents a challenge: the delivery of broadband services to millions of customers, across sparsely populated areas.

Whilst technically the most advanced solution, fully fibre-based networks are prohibitively expensive to deploy. Optical terminal equipment, and the trenching of fragile fibres is extremely costly. The expected recovery period for the initial investment on a fully fibre network is 7.5 years, time that companies do not have in today's volatile market[78, 55].

Digital subscriber lines (DSL) provide an alternative solution. Seen as a stepping-stone to a fully fibre-based network, DSL provides connectivity in the last mile between the *customer premises* (CP) and the fibre-network core. DSL operates over telephone lines that are already in place, minimizing the cost of deployment.

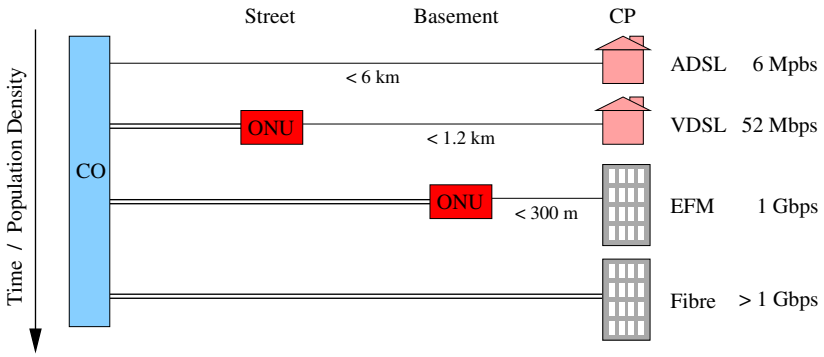

Figure 1.1: DSL Network Evolution

With DSL the fibre network grows through evolution rather than revolution. Instead of replacing the entire network with fibre in one operation, an extremely expensive option, with DSL the fibre network grows according to customer demand. In the beginning, fibre is used to connect the *central offices* (CO) to the network core. ADSL provides connectivity from the CO to the CP, providing *downstream* (DS) rates of up to 6 Mbps.

As demand increases, fibre can be laid to the end of each street where an *optical network unit* (ONU), also known as a *remote terminal* (RT), is installed, as shown in Fig. 1.1. VDSL provides connectivity from the ONU to the CP, increasing rates to 52 Mbps. In high density housing and office buildings, fibre can be extended to the basement. *Ethernet in the First Mile* (EFM), a technology based on DSL, then connects each office to an ONU in the basement, providing symmetrical rates of up to 1 Gbps[2].

Following this evolutionary approach, operators can deploy their fibre networks as demand grows. Expenditure on extra infrastructure is fueled using revenue from existing services. This leads to a fast return on investment and a lower risk for operators. With DSL, fibre can be deployed in a heterogeneous fashion, and scaled to match demand. Fibre can be deployed to all basements in the central business district, to the end of the street in urban areas, and to the CO in suburban and rural areas.

One of the main drives behind the development of DSL technology, was a desire by *telephone network operators* (telcos) to enter the broadband consumer market. Until recently, broadband access in many countries was dominated by *cable network operators* (cablecos) who provide Internet access over the same coaxial cable they use to provide television service.

ADSL was originally developed in 1987 with the goal of providing television

services over the phone network. These plans failed, but ADSL did not. The Internet boom, that began with the first commercial Internet service provider in 1989, created a massive demand for broadband access. ADSL technology was developed, and initial field trials began in 1995.

Today the original dream of television service over DSL is being revisited, with operators deploying triple-play services, a combination of video, high-speed Internet and voice. This is driving demand for still higher data-rates, and new DSL technologies such as ADSL2+ and *very high bit-rate digital subscriber line* (VDSL) are being developed in response.

VDSL is now being deployed in Korea and Japan where high density housing makes fibre-to-the-basement economically feasible. Access rates up to 70 Mbps are currently provided and demand continues to grow.

Cable modems present the biggest threat to DSL as a competing technology for broadband access. At the same time wireless and satellite systems are being developed that threaten to take a share of the broadband market. Satellite technology has a natural advantage in rural areas where the population density is too low to justify installing an RT. For a low number of subscribers wireless and satellite solutions are much cheaper since they do not require heavy investment in infrastructure.

In developing countries such as India and China there is often no telephone infrastructure in place. Most citizens do not own a fixed line telephone and rely on mobile phones instead. Here DSL loses its main benefit, which is the use of existing telephone infrastructure. So wireless and satellite systems will find a large potential market in these places.

Despite these specific cases, for conventional broadband access DSL and coaxial cable will continue to dominate the market. The primary reason behind this is that wireless is an inherently more expensive delivery means, in terms of bits/second /Hz/user, than wireline. This higher cost results from a number of fundamental differences between wireline and wireless transmission, which we now describe.

To begin with, wireline media have a lower attenuation per unit distance than wireless media. This is natural since propagation of an electromagnetic signal through free space leads to more loss than along a waveguide, such as a telephone line or coaxial cable. Furthermore, wireline systems have channels that vary very slowly with time. This allows techniques such as bitloading and powerloading to be applied to increase spectral efficiency. Additionally the overhead required for synchronization and channel identification will be much lower in the slowly varying wireline channel, than in a wireless environment where the channel typically changes for every packet that is received.

Interference in a DSL network is suppressed to a large extend by the insulation

between the twisted pairs. This allows different lines to transmit data in the same frequency range at the same time. Since each customer has their own phone line, the total capacity of the network grows with the number of users. Hence a DSL system can potentially serve an unlimited number of users. In wireless systems users must share a common, limited bandwidth. There is no natural suppression of interference in the transmission medium. As a result, each user must employ time-division, frequency-division, code-division or some other orthogonal multi-access technique to prevent interference. The total capacity of the network is limited by the available bandwidth and as the number of subscribers increases the average data-rate of each subscriber decreases. To maintain the same data-rate as the number of subscribers grows, the operator must decrease cell size and increase the number of base-stations, an extremely expensive operation.

It should be kept in mind that base-stations themselves must be connected to the network backbone using some kind of wireline technology such as DSL, coaxial cable or fibre. So the use of a wireless access point simply shifts the wireline system design problem further back into the network. The problem however must still be solved.

In general wireline access technology will always be cheaper in terms of bits/second/Hz/user because it is technically an easier problem to solve. This is reflected in the cost of *customer premises equipment* (CPE), which in 2003 cost $400 USD for a wireless MAN terminal, and $50 USD for DSL[42, 84].

Despite the higher cost per user of wireless systems in high-density urban and sub-urban areas, they will continue to find application in niche markets such as rural areas. Here fixed wireless or satellite access may be a more economic solution. It should also be noted that with satellite access upstream connectivity must still be provided over a wireline network, e.g. DSL. Furthermore, with satellite systems low latency is difficult to achieve, which creates problems for voice-over-IP and video-conferencing applications.

Perhaps the biggest advantage of wireless access is the low initial investment required to roll out a network and begin serving customers. For example, with $4.2 million USD it is possible to deploy a network over 500 square km serving up to 6000 subscribers[42]. This is orders of magnitude lower than the cost of rolling out a DSL or coaxial network to serve the same area. An additional problem for new operators entering the market, the so-called *competitive local exchange carriers* (CLEC), is that the *incumbent local exchange carriers* (ILEC) currently have a monopoly on the twisted-pair network. This is unlikely to change in the near-future as recent economic problems with the dot-com bubble and the resulting effect on the telecoms industry has delayed plans in many countries for liberalization of local loop access (unbundling). This makes it difficult for CLECs to enter the DSL market and will lead to many of these companies moving to wireless access technologies instead. The lower

cost of entry into the wireless market and relative ease of deployment may lead to a more dynamic competitive environment and help break the telco/cableco duopoly that is developing in many countries.

So far we have considered wireless access a a competing technology to DSL. Wireless and wireline technologies are inherently different, and offer different trade-offs of mobility, convenience, ubiquity, data-rate and cost. The broadband networks of the future will not consist of either wireless or wireline technology alone, but a dynamic mixture of both. At home many users may prefer a high-speed, low-cost DSL line to provide connectivity, coupled with a wireless *local area network* (LAN) hub for convenient access. Away from home users may happily sacrifice some data-rate to have convenient, mobile access which may be delivered through UMTS, IEEE 802.11 LANs, IEEE 802.16 metropolitan area networks or some combination of all three[56, 77, 76]. An adaptive, intelligent network that can seamlessly switch users from one access technology to another is the goal of future access networks. Both wireline and wireless technology have an important and synergistic part to play in this future.

There are three challenges that limit the future growth of DSL services:

Rate

The demand for ever-higher connection rates continues to grow. This is driven by the desire for triple-play services, e.g. delivering two HDTV channels, at 12 Mbps per channel, plus high-speed Internet at 10 Mbps, plus a voice/music channel of 1 Mbps requires a 35 Mbps service. ADSL systems today offer 3 Mbps in high density urban areas. In suburban and rural areas the access rates are often 256 kbps or less.

Increasing access rate is a major challenge for telcos. This is particularly crucial due to the competition from cablecos, who continue to upgrade their networks to provide higher access rates. The coaxial cable is a superior medium to twisted pair, and cable networks today are only limited by the switching speed of CPE. Note that, since the cable network is a shared medium, all CP modems must switch at the full rate of the cable, which corresponds to the number of active users times the access rate of each user. As a result, CP modems for cable networks are more expensive to manufacture than for DSL. This slight advantage will soon change as Moore's law decreases the cost of computing power. Hence it is imperative that telcos increase access rates to remain competitive.

Reach

Customers in suburban and rural areas are typically situated far from the CO. Over such distances channel attenuation is high due to the poor quality of the twisted-pair medium. This limits the number of customers that can be reached

with DSL services.

This problem is particularly evident in geographically sparse countries like the USA and Australia where DSL penetration is less than 5%[50]. Compare this with countries like Korea, which has a penetration of 29%, and it is clear that telcos are missing out on a large opportunity for revenue.

Symmetry

Existing DSL technologies such as ADSL are asymmetric, providing a higher rate in the DS than in the *upstream* (US). Whilst this makes sense in conventional applications such as web-browsing and video-streaming, the growth of peer-to-peer file-sharing of music and movies, video conferencing and teleworking via virtual private LANs is increasing the demand for US data-rate. Providing high US and DS rates in the limited bandwidth available is a major challenge for DSL vendors and operators alike.

All three of these issues, rate, reach and symmetry, can be addressed by extending the fibre network closer to the customer. The DSL network then operates over shorter lines, leading to a lower channel attenuation and higher data-rates. However the deployment of remote, fibre-fed terminals at the end of each street is expensive. Computing power, on the other hand, is cheap and continues to go down in price. This motivates the use of signal processing techniques, rather than fibre deployment, to increase performance. The development of advanced coding, equalization and multi-user transmission techniques is essential for DSL to stay competitive with coaxial networks. This thesis focuses on the use of multi-user techniques to improve DSL performance.

1.2 The Crosstalk Problem

The twisted-pair medium was originally designed with voice-band communication in mind. Traditional voice band modems limit transmission to below 4 kHz and, as a result, are limited to a data-rate of 56 kbps.

The basic principle behind DSL technology is to increase the achievable data-rate by widening the transmission bandwidth. ADSL uses frequencies up to 1.1 MHz, which allows it to provide data-rates up to 6 Mbps. VDSL uses frequencies up to 12 MHz, which increases the maximum data-rate to 52 Mbps.

Unfortunately, operating at such high frequencies in a medium originally design for voice-band transmission leads to its own problems. The twisted pairs in the access network are bundled together within large binder groups, which typically contain 20 to 100 individual pairs. The high frequencies used in DSL give rise to electromagnetic coupling between the different twisted-pairs. This leads to

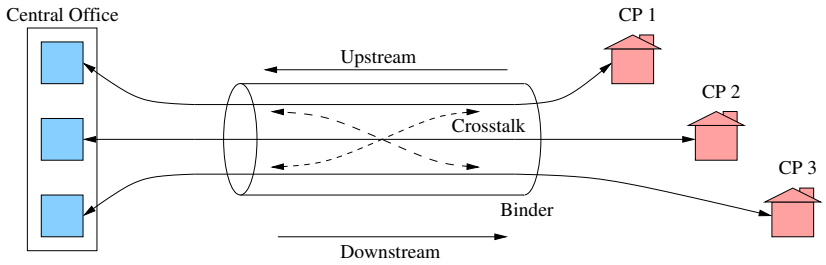

Figure 1.2: Crosstalk

interference or *crosstalk* between the different systems operating within the binder, as shown in Fig. 1.2. Crosstalk is typically 10-15 dB larger than the background noise and is *the* dominant source of performance degradation in DSL.

Crosstalk transforms the twisted-pair binder into a multi-user channel. Significant work has been done on multi-user communication techniques, typically motivated by wireless applications. These techniques can also be applied in DSL to mitigate crosstalk and this is the focus of this thesis.

Whilst the DSL environment shares some superficial similarities to the wireless environment, in many ways it is fundamentally different. For example the DSL channel is quite static, changing once every few hours, unlike the wireless channel, which varies continually. Power constraints are not an issue in DSL since modems use a mains power supply. The DSL channel has a much smaller attenuation than a typical wireless channel, and this makes design easier.

On the other hand, DSL modems typically operate at a much higher rate than wireless systems. An ADSL modem runs at 4000 symbols per second, and transmits over 256 tones, so a simple multiplication operation requires 1 million floating-point operations per second. This puts strict limitations on the complexity of any signal processing. As will be shown, considerable effort must be put into reducing the complexity of multi-user techniques in DSL.

1.3 State of the Art

Current modems operate operate in a single-user fashion. Crosstalk is treated as background noise; it decreases the receiver-side SNR and leads to a significant degradation in data-rate. Fig. 1.3 shows the data-rates achieved by a group of 25 VDSL modems. The modems are deployed in a common binder and suffer mutual crosstalk. Clearly there is a significant performance penalty

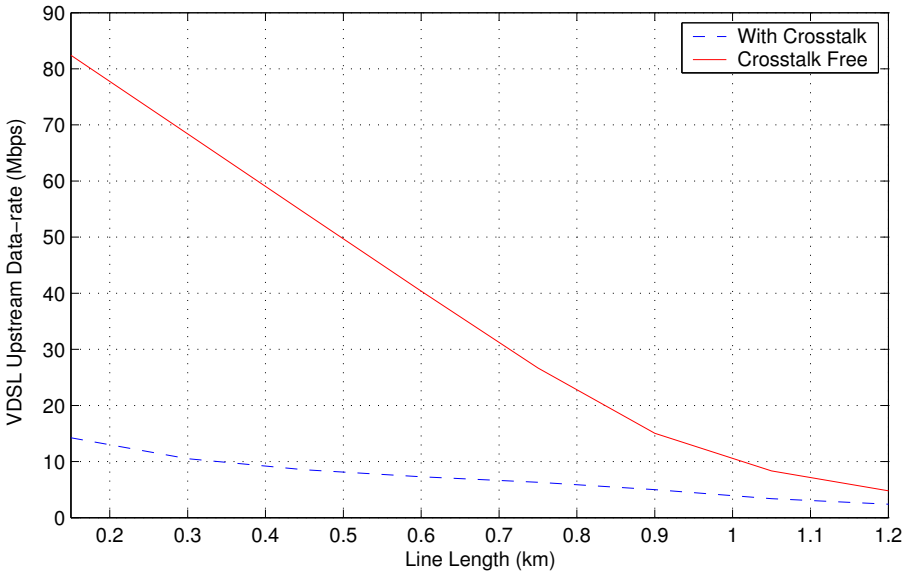

Figure 1.3: Data-rate loss due to Crosstalk in VDSL

as a result of crosstalk.

Using multi-user techniques such as multi-user spectra optimization can help minimize the effects of crosstalk. Existing modems are not capable of adjusting their transmit spectra, and instead employ fixed transmit masks. Whilst there is some provision in the new DSL standards for a programmable transmit mask, at present no DSL product makes use of this capability[4]. Fig. 1.4 shows the data-rates achieved by a group of 25 ADSL modems. The modems are deployed in a common binder and suffer mutual crosstalk. The achievable data-rates are shown with fixed transmit masks, and with optimized transmit spectra, according to the optimal spectrum balancing algorithm from Chapter 3. Clearly, existing modems suffer a significant performance penalty for using fixed transmit spectra.

Another multi-user technique, known as crosstalk cancellation, can completely remove crosstalk allowing operation on the *crosstalk free* line from Fig. 1.3. Unfortunately this technique is not available in existing modems due to its high complexity, long latency, and inability to work with existing customer premises equipment.

Many techniques have been proposed in literature for both crosstalk cancellation and multi-user spectra coordination. A detailed study of these techniques is deferred to the relevant chapters. The main problems with these techniques

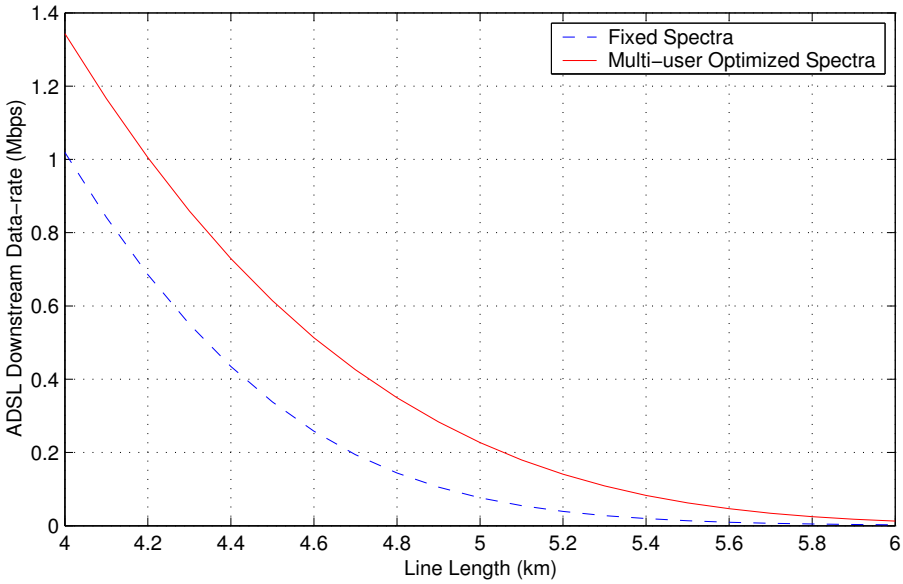

Figure 1.4: Data-rate loss due to Unoptimized Transmit Spectra in ADSL

are complexity, latency, and incompatibility with existing equipment.

The goal of this thesis is to develop practical multi-user techniques for DSL that can be applied in existing or near-future DSL platforms. In response, this thesis develops algorithms that have low complexity, low latency, and are compatible with existing *customer premises equipment* (CPE). In addition to being practical, the algorithms are also shown to yield near-optimal performance, operating close to the theoretical multi-user channel capacity.

1.4 Thesis Overview and Contributions

An overview of the thesis and its major contributions is now given. Multi-user techniques are based on the coordination of different users in a network. This can be done on a spectral or signal level.

Part I of this thesis investigates multi-user spectra coordination. With spectral coordination the transmit spectra of the modems within a network are limited in some way to minimize the negative effects of crosstalk. Each modem must achieve a trade-off between maximizing its own data-rate and minimizing the crosstalk it causes to other modems within the network. The goal is to achieve

a fair trade-off between the rates of the different users in the network.

Chapter 3 investigates the design of optimal transmit spectra for a network of crosstalking DSLs. This problem was previously considered intractable since it requires the solution of a high-dimensional, non-convex optimization. Chapter 3 shows how the application of a dual-decomposition solves the optimization in an efficient, tractable way. The resulting algorithm, which we name *optimal spectrum balancing*, achieves significant gains over existing spectra coordination algorithms, typically doubling or tripling the achievable data-rate. The material in Chapter 3 has been published as [40, 39, 110, 14, 94, 97], submitted for publication as [20, 95], and has been patented by Alcatel[32]. The *optimal spectrum balancing* algorithm was submitted to standardization as [36, 37, 38, 35] and is now part of the draft ANSI standard on Dynamic Spectrum Management[8].

Part II of this thesis investigates multi-user signal coordination. In a DSL network, the line-side transceivers are often co-located at the CO. This allows modems to be co-ordinated on a signal level.

In the US, signal coordination is used between co-located CO receivers. Reception is done in a joint fashion; the signals received on each line are combined to cancel crosstalk whilst preserving the signal of interest.

Chapter 4 discusses crosstalk canceler design. Existing techniques are based on decision feedback between the different users within the binder. To prevent error propagation decoding must be done before decisions are fed back, which leads to a high computational complexity and latency. To address this problem, a simple linear canceler is presented based on the well known ZF criterion. This technique has a low complexity and latency. It is shown that, due to a special property of upstream DSL channels, this design operates close to the theoretical channel capacity. A low complexity algorithm is proposed for spectra optimization when crosstalk cancellation is employed. This material has been published as [34, 22, 28, 23] and submitted for publication as [18].

In the downstream, signal coordination is used between co-located CO transmitters. Transmission is done in a joint fashion; predistortion is introduced into the signal of each user prior to transmission. This predistortion is chosen such that it annihilates with the crosstalk introduced in the channel. As a result the customer premises (CP) modems receive a crosstalk free signal.

This technique, known as crosstalk precoding, is discussed in Chapter 5. Existing precoder designs lead either to poor performance or require the replacement of CP modems. Millions of CP modems are currently in use, owned and operated by a multitude of customers. Replacing these modems presents a huge legacy issue. To address this problem a simple linear precoder is presented based on a channel diagonalizing criterion. The precoder has a low complexity

and works with existing CP modems. It is shown that, due to a special property of downstream DSL channels, this design operates close to the theoretical channel capacity. A low complexity algorithm is proposed for spectra optimization when crosstalk precoding is employed. This material has been published as [17, 29], submitted for publication as [19] and submitted to standardization as [33].

As a by-product, the work in Chapters 4 and 5 produced as set of bounds on the determinants and inverses of diagonally dominant matrices. These are listed in Appendix B.

Despite the low complexity of the techniques presented in Chapters 4 and 5, signal coordination still requires a much higher complexity than is available in existing DSL modems. Crosstalk cancellation and precoding have a complexity that scales quadratically with the number of lines within a binder. For typical binders, which contain anywhere from 20 to 100 lines, these techniques are outside the scope of present day implementation and may remain so for several years. Chapter 6 addresses this problem through a technique known as partial cancellation.

It is well known that the majority of crosstalk experienced on a line comes from the 3 to 4 surrounding pairs in the binder. Furthermore, since crosstalk coupling varies dramatically with frequency, the worst effects of crosstalk are limited to a small selection of tones. Partial cancelers exploit these facts to achieve the majority of the performance of full cancellation at a fraction of the complexity. Whilst the idea of partial cancellation has been discussed in literature, no work has specifically focused on partial canceler design.

Chapter 6 investigates partial canceler and precoder design, which is in essence a problem of resource allocation. Given a limited amount of available run-time complexity, a modem must distribute this across lines and tones such that the data-rate is maximized. Chapter 6 presents the optimal algorithm for partial canceler design and several simpler, sub-optimal algorithms. These algorithms are shown to achieve 90% of the data-rate of full cancellation at less than 30% of the complexity. This material has been published as [27, 25, 24, 26] and has been patented by Alcatel[30, 31].

Conclusions are drawn and interesting areas for further research are discussed in Chapter 7.

Chapter 2

Basic Concepts

2.1 Digital Subscriber Lines

2.1.1 Discrete Multi-tone Modulation

Modern DSL systems can be divided into two camps: single-carrier and *discrete multi-tone* (DMT) modulated systems. Before the development of DSL, all voiceband modems were based on single-carrier modulation. In voiceband transmission the lower complexity of single-carrier systems made them a more attractive option.

In broadband systems such as DSL, the transmission channel is frequency selective. This results in *inter-symbol interference* (ISI) which degrades performance significantly if left unaddressed. In single-carrier systems ISI can be removed through the use of a *decision feedback equalizer* (DFE) at the receiver. Whilst this improves performance it has a high run-time complexity and can suffer from error propagation.

An alternative is to use a Tomlinson-Harashima precoder at the transmitter to precompensate for ISI. This avoids problems with error propagation, however it requires accurate channel knowledge at the transmitter. Hence the receiver must measure the channel and communicate this to the transmitter, which results in a high transmission overhead and increased computational complexity.

DMT modulation was proposed to address the short-comings of single carrier systems. With DMT modulation the frequency selective channel is effectively divided into many parallel sub-channels, known as tones, as shown in Fig. 2.1. Within each sub-channel the channel response is approximately flat, so transmission over the sub-channels does not suffer from ISI. As a result a scalar

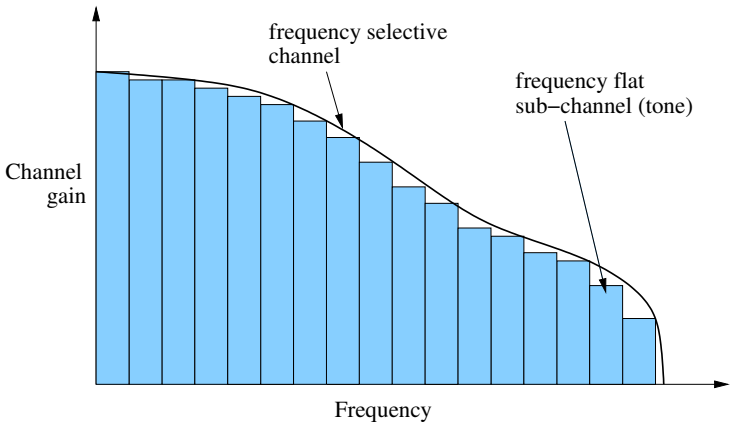

Figure 2.1: Discrete Multi-tone Transmission (Sub-channels)

multiplication is sufficient to equalize each sub-channel. Combined with efficient modulation through the *fast Fourier transform* (FFT), this leads to a much lower complexity than single-carrier systems with a DFE[13]. Furthermore, this approach does not suffer from error propagation.

Time-Domain Transmission

DMT modulation is now described in more detail. Consider transmission through a channel with ISI. Denote the transmit sequence x_i^{time} , which has a sampling rate $F_s = 1/T_s$. If the transmitter and receiver are synchronized, then the discrete-time signal after sampling at rate F_s at the receiver is

$$y_i^{\text{time}} = \sum_{l=0}^{L} h_l^{\text{time}} x_{i-l}^{\text{time}} + z_i^{\text{time}}, \tag{2.1}$$

where $h_l^{\text{time}} \triangleq h(lT_s)$, and $h(t)$ denotes the continuous-time impulse response of the channel. L is chosen such that $h_l^{\text{time}} = 0$ for all $l > L$. The term $z_i \triangleq z(iT_s)$, where $z(t)$ is continuous-time additive Gaussian noise at the receiver. This term will be used to capture thermal noise, *radio frequency interference* (RFI) and alien crosstalk.

Consider a block of symbols $\mathbf{x}^{\text{time}} \triangleq [x_K^{\text{time}}, \ldots, x_{1-L}^{\text{time}}]^T$ to be transmitted through the channel. Denote the corresponding received sequence as $\mathbf{y}^{\text{time}} \triangleq [y_K^{\text{time}}, \ldots, y_1^{\text{time}}]^T$. From (2.1) transmission can be modelled in matrix form as

$$\mathbf{y}^{\text{time}} = \mathbf{H}_{\text{toeplitz}} \mathbf{x}^{\text{time}} + \mathbf{z}^{\text{time}}, \tag{2.2}$$

where $\mathbf{z}^{\text{time}} \triangleq [z_K^{\text{time}}, \ldots, z_1^{\text{time}}]^T$ and the $K \times K + L$ Toeplitz channel matrix

$$\mathbf{H}_{\text{toeplitz}} \triangleq \begin{bmatrix} h_0^{\text{time}} & \cdots & h_L^{\text{time}} & 0 & \cdots & 0 \\ 0 & h_0^{\text{time}} & \cdots & h_L^{\text{time}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \cdots & 0 & h_0^{\text{time}} & \cdots & h_L^{\text{time}} \end{bmatrix}.$$

The Cyclic Prefix

In order to ensure that the DMT sub-carriers, known as *tones*, remain orthogonal after propagation through the ISI channel, a cyclic prefix is used[80, 102]. The cyclic prefix is a copy of the last L data-symbols, placed at the beginning of the transmitted block. A cyclic prefix can be incorporated into the vector $\mathbf{x}^{\text{time}}$ by setting

$$\mathbf{x}^{\text{time}} = \begin{bmatrix} \mathbf{x}_{\text{data}}^{\text{time}} \\ \mathbf{x}_{\text{cp}}^{\text{time}} \end{bmatrix},$$

where the data $\mathbf{x}_{\text{data}}^{\text{time}} \triangleq [x_K^{\text{time}}, \ldots, x_1^{\text{time}}]^T$ and the cyclic prefix

$$\mathbf{x}_{\text{cp}}^{\text{time}} \triangleq [x_K^{\text{time}}, \ldots, x_{K-L+1}^{\text{time}}]^T.$$

From (2.2), transmission can be modelled as

$$\begin{aligned} \mathbf{y}^{\text{time}} &= \mathbf{H}_{\text{toeplitz}} \begin{bmatrix} \mathbf{x}_{\text{data}}^{\text{time}} \\ \mathbf{x}_{\text{cp}}^{\text{time}} \end{bmatrix} + \mathbf{z}^{\text{time}}, \\ &= \mathbf{H}_{\text{circ}} \mathbf{x}_{\text{data}}^{\text{time}} + \mathbf{z}^{\text{time}}, \end{aligned} \tag{2.3}$$

where $\mathbf{H}_{\text{circ}}$ is the $K \times K$ circulant Toeplitz matrix with

$$\mathbf{h}^{\text{time}} \triangleq [h_0^{\text{time}}\ \mathbf{0}_{1\times K-L-1}\ h_L^{\text{time}}, \ldots, h_1^{\text{time}}]^T,$$

as its first column. So the effect of the cyclic prefix is to convert the linear convolution of the channel into a circular convolution. As will be shown in the following section, since circular convolution in time is equivalent to multiplication in frequency the CP ensures that the tones remain orthogonal after propagation through the channel.

Frequency Domain Transmission

Frequency-domain transmission is now examined in more detail. Define the frequency-domain symbol to be transmitted on tone k as x_k^{freq}, and the vector of frequency-domain symbols $\mathbf{x}^{\text{freq}} \triangleq [x_1, \ldots, x_K]^T$.These symbols are efficiently modulated using the IFFT. So

$$\mathbf{x}_{\text{data}}^{\text{time}} = \mathcal{I}_K \mathbf{x}^{\text{freq}}, \tag{2.4}$$

where $\mathcal{I}_K$ denotes the K-point IDFT matrix. At the receiver the signal $\mathbf{y}^{\text{time}}$ is efficiently demodulated using the FFT. So

$$\mathbf{y}^{\text{freq}} = \mathcal{F}_K \mathbf{y}^{\text{time}}, \tag{2.5}$$

where $\mathcal{F}_K$ denotes the K-point DFT matrix and $\mathbf{y}^{\text{freq}} \triangleq [y_1, \ldots, y_K]^T$. Combining (2.3), (2.4) and (2.5) yields

$$\mathbf{y}^{\text{freq}} = \mathcal{F}_K \mathbf{H}_{\text{circ}} \mathcal{I}_K \mathbf{x}^{\text{freq}} + \mathbf{z}^{\text{freq}},$$

where the frequency-domain noise vector

$$\mathbf{z}^{\text{freq}} \triangleq \mathcal{F}_K \mathbf{z}^{\text{time}} = [z_1, \ldots, z_K]^T.$$

Define the frequency-domain transfer function for the channel as

$$\mathbf{h}^{\text{freq}} \triangleq [h_1, \ldots, h_K],$$

where h_k is the channel response on tone k. The frequency-domain transfer function is

$$\mathbf{h}^{\text{freq}} = \mathcal{F}_K \mathbf{h}^{\text{time}}.$$

Circulant matrices are diagonalized by the DFT and IDFT matrices, so

$$\mathcal{F}_K \mathbf{H}_{\text{circ}} \mathcal{I}_K = \mathbf{H}_{\text{freq}},$$

where $\mathbf{H}_{\text{freq}} = \text{diag}\{h_1, \ldots, h_K\}$. Another way of interpreting this is that circular convolution in the time-domain corresponds to a multiplication in the frequency domain. Hence the received signal after demodulation is

$$\mathbf{y}^{\text{freq}} = \mathbf{H}_{\text{freq}} \mathbf{x}^{\text{freq}} + \mathbf{z}^{\text{freq}},$$

Since $\mathbf{H}_{\text{freq}}$ is diagonal, transmission now occurs independently on each tone. The received signal on tone k

$$y_k = h_k x_k + z_k.$$

Equalization of the channel can be implemented with low complexity by simply multiplying y_k with h_k^{-1} at the receiver. The estimate of the symbol on tone k is thus

$$\begin{aligned} \hat{x}_k &= h_k^{-1} y_k, \\ &= x_k + h_k^{-1} z_k. \end{aligned}$$

The overall complexity of DMT is $\mathcal{O}(2K \log_2 K + K)$ per transmitted block, which includes $K \log_2 K$ operations for modulation (demodulation) with the IFFT (FFT) and one multiplication per-tone for equalization. Recall that K denotes the number of DMT tones, whilst L denotes the length of the channel impulse response. For comparison, the DFEs employed in single-carrier systems have a complexity of $\mathcal{O}(LK)$. Typical values in VDSL are $K = 4096$ and $L = 320$. In this case DMT reduces complexity by a factor of 12, giving it a significant advantage over single-carrier systems.

Bitloading

Define the noise power on tone k as $\sigma_k \triangleq \mathcal{E}\{|z_k|^2\}$ and the transmit power as $s_k \triangleq \mathcal{E}\{|x_k|^2\}$, where $\mathcal{E}\{\cdot\}$ denotes the statistical expectation operation. On tone k the theoretical capacity with DMT is

$$c_k = \Delta_f \log_2 (1 + SNR_k),$$

where Δ_f denotes the tone-spacing and the *signal-to-noise ratio* (SNR) on tone k is defined

$$SNR_k \triangleq \sigma_k^{-1} |h_k|^2 s_k.$$

Most practical coding schemes are characterized by an SNR-gap to capacity Γ, which determines how closely the code comes to the theoretical capacity. Γ is a function of the coding gain, desired noise margin and target probability of error[89, 53]. So in practice the achievable data-rate is

$$c_k = \Delta_f \log_2 \left(1 + \Gamma^{-1} SNR_k\right). \tag{2.6}$$

In DMT systems the receiver measures the SNR on each tone and reports this back to the transmitter. The transmitter can then adaptively vary the number of bits used on each tone by choosing different constellation sizes, a technique known as bitloading. Bitloading allows DMT systems to achieve a high spectral efficiency. The bitloading on a tone is the number of bits transmitted per DMT-symbol. Using (2.6) the achievable bitloading on tone k is

$$b_k = f_s^{-1} \Delta_f \log_2 \left(1 + \Gamma^{-1} SNR_k\right),$$

where f_s denotes the DMT symbol-rate. The total data-rate of the modem is then

$$R = f_s \sum_k b_k.$$

Typically $f_s = \Delta_f$ and

$$b_k = \log_2 \left(1 + \Gamma^{-1} SNR_k\right).$$

Powerloading

DSL systems typically operate under a set of spectral masks which ensure that spectral compatibility is maintained with other communication systems that may exist within the same binder

$$s_k \leq s_k^{\text{mask}}, \forall k. \tag{2.7}$$

Modems also operate under a total transmit power constraint that arises from limitations on the analog front-end

$$\sum_k s_k \leq P.$$

A modem can vary the power allocated to each tone s_k, and will do so in an attempt to maximize its total data-rate subject to any spectral mask and total power constraints

$$\begin{aligned} s_k^{\text{opt}} &= \arg \max_{s_1,\ldots,s_K} R_k \\ &\text{s.t.} \quad \sum_k s_k \leq P \\ &\qquad s_k \leq s_k^{\text{mask}}, \forall k. \end{aligned} \tag{2.8}$$

This is referred to as powerloading. Since the objective function (2.8) is concave and the constraints form a convex set, the KKT conditions are sufficient for optimality. The *Karush-Kuhn-Tucker* (KKT) conditions imply

$$s_k^{\text{opt}} = \left[\frac{1}{\lambda} - \frac{\Gamma\sigma_k}{|h_k|^2}\right]_0^{s_k^{\text{mask}}}, \tag{2.9}$$

where $[x]_a^b \triangleq \max(a, \min(x, b))$. The *waterfilling level* $1/\lambda$ must be chosen such that either the power constraint is tight $\sum_k s_k = P$, or $\sum_k s_k < P$ and the modem transmits at mask on all tones $\lambda = 0$. Efficient algorithms exist to find the appropriate λ with complexity $\mathcal{O}(K \log K)$[12].

Provided a powerful enough error-correcting code is used, powerloading allows DMT systems to operate arbitrarily close to the theoretical channel capacity. The natural way in which DMT systems implement powerloading is one of their major advantages over single-carrier systems.

2.1.2 Multi-user Channels

So far the discussion has been restricted to DSL systems operating in isolation. This section considers the interaction of several DSL modems operating within the same binder. A multi-user channel model is developed that incorporates crosstalk effects.

Multi-user Transmission

Consider several modems operating within the same binder as depicted in Fig. 2.2. The modems are assumed to be synchronized and transmit simultaneously. The discrete-time signal after sampling at rate F_s at receiver n is

$$y_i^{\text{time},n} = \sum_{l=0}^{L} \left(h_l^{\text{time},n,n} x_{i-l}^{\text{time},n} + \sum_{m \neq n} h_l^{\text{time},n,m} x_{i-l}^{\text{time},m} \right) + z_i^{\text{time},n}, \tag{2.10}$$

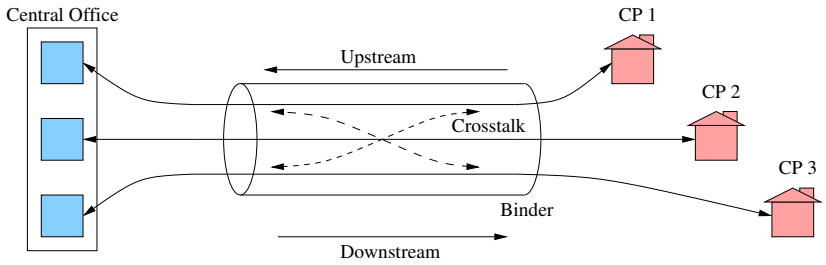

Figure 2.2: Multi-user Transmission

where $x_i^{\text{time},n}$ is the time-domain sequence transmitted by modem n. Here $h_l^{\text{time},n,m} \triangleq h^{n,m}(lT_s)$ where $h^{n,m}(t)$ denotes the continuous-time impulse response of the channel from transmitter m to receiver n. When $m = n$, $h^{n,m}(t)$ is a direct channel. When $m \neq n$, $h^{n,m}(t)$ is a crosstalk channel. The first term in (2.10) is the signal of interest for receiver n, whilst the second term is the crosstalk from all other transmitters.

The additive Gaussian noise sequence experienced by receiver n is denoted $z_i^{\text{time},n}$. L is now chosen such that $h_l^{\text{time},n,m} = 0$ for all n,m, and $l > L$. As before, DMT modulation converts the frequency-selective channel into several independent sub-channels, or tones. Denote the gain on tone k from transmitter m to receiver n as $h_k^{n,m}$. This can be found through the DFT of the corresponding impulse response

$$[h_1^{n,m}, \ldots, h_K^{n,m}]^T = \mathcal{F}_K \left[h_0^{\text{time},n,m} \; \mathbf{0}_{1\times K-L-1} \; h_L^{\text{time},n,m}, \ldots, h_1^{\text{time},n,m} \right]^T.$$

The signal at receiver n on tone k in the multi-user case is

$$y_k^n = \sum_{m=1}^{N} h_k^{n,m} x_k^m + z_k^n, \tag{2.11}$$

where N denotes the number of users in the binder. Equation (2.11) can be expressed in matrix form as follows. Define the vectors $\mathbf{x}_k \triangleq [x_k^1, \cdots, x_k^N]^T$, $\mathbf{y}_k \triangleq [y_k^1, \ldots, y_k^N]^T$ and $\mathbf{z}_k \triangleq [z_k^1, \ldots, z_k^N]^T$ which contain the transmitted, received and noise signals for all modems on tone k respectively. Define the multi-user channel matrix as $\mathbf{H}_k \triangleq [h_k^{n,m}]$. The diagonal elements of $\mathbf{H}_k$ contain the direct channels whilst the off-diagonal elements contain the crosstalk channels. Transmission on tone k can now be written as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k. \tag{2.12}$$

Empirical Channel Models

Table 2.1: RLCG Parameters

Cable Type	TP1	TP2
diameter (mm)	0.4	0.5
$r_{0c}\,(\Omega/\text{km})$	286.176	174.559
a_c	0.1476962	0.0530735
$l_0\,(\mu\text{H/km})$	675.369	617.295
$l_\infty\,(\mu\text{H/km})$	488.952	478.971
b	0.929	1.152
$f_m\,(\text{kHz})$	806.339	553.760
$c_\infty(\text{nF/km})$	49	50
$g_0\,(\text{n℧/km})$	43	0.00023487476
g_e	0.7	1.38

Exhaustive measurement campaigns have been made to model the direct and crosstalk channels in DSL networks. As a result the direct channel of a twisted-pair can be accurately estimated using an incremental RLCG model which defines the resistance, inductance, capacitance and conductance per kilometer of twisted pair. The models of R, L, C, and G for copper cable are

$$\begin{aligned} R_k &= \left(r_{0c}^4 + a_c f_k^2\right)^{1/4}, \\ L_k &= \left(l_0 + l_\infty (f_k/f_m)^b\right)\left(1 + (f_k/f_m)^b\right)^{-1}, \\ C_k &= c_\infty, \\ G_k &= g_0\,(f_k)^{g_e}, \end{aligned}$$

where $f_k \triangleq \Delta_f \cdot k$ is the frequency on tone k in Hz[6]. The models are frequency dependent. The parameters r_{0c}, a_c, l_0, l_∞, f_m, b, c_∞, g_0 and g_e depend on the cable diameter, materials and construction. Values of these parameters for the standard cable types TP1 and TP2 are listed in Tab. 2.1.

The propagation constant per unit length for the twisted pair at tone k is

$$\gamma_k = \sqrt{(R_k + j2\pi f_k L_k)\,(G_k + j2\pi f_k C_k)}.$$

The characteristic impedance of the line on tone k is defined as

$$Z_{0,k} \triangleq \sqrt{\frac{R_k + j2\pi f_k L_k}{G_k + j2\pi f_k L_k}}.$$

The direct channel transfer function for a twisted-pair of length d km can now be modelled as

$$h_k(d) = \frac{Z_L + Z_S}{Z_L \cosh(\gamma_k d) + Z_{0,k} \sinh(\gamma_k d) + Z_S Z_L Z_{0,k}^{-1} \sinh(\gamma_k d) + Z_S \cosh(\gamma_k d)}, \tag{2.13}$$

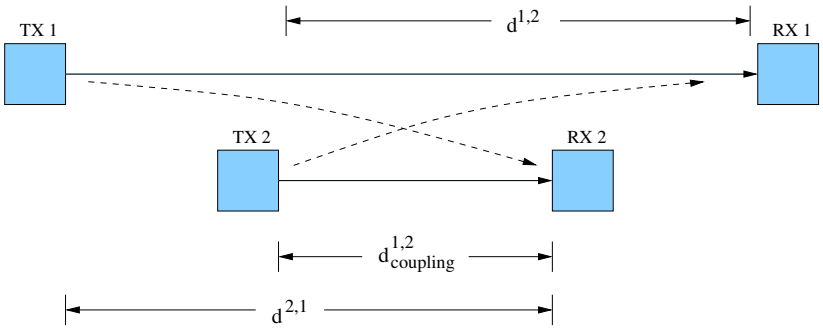

Figure 2.3: Coupling distances

where Z_S is the source impedance of the transmitting modem and Z_L is the load impedance of the receiving modem.

Empirical models for crosstalk channels are based on 1% worst-case analysis. So in 99% of cases the crosstalk is less severe than the empirical models suggest. Such worst-case models are used to ensure that DSL modems operate for the majority of customers.

In the 1% worst-case models, the crosstalk channel gain between two lines is

$$h_k^{n,m} = \alpha_{k,n,m} \left| h_k(d^{n,m}) \right| ,$$

where

$$\alpha_{k,n,m} \triangleq K_{\text{xf}} \cdot (f_k/f_0) \sqrt{d_{\text{coupling}}^{n,m}} \tag{2.14}$$

and $f_0 = 1$ MHz and $K_{\text{xf}} = 0.0056$[7]. As shown in Fig. 2.3, $d_{\text{coupling}}^{n,m}$ is the length of the binder segment over which coupling between line m and line n occurs, and is measured in kilometers. Note that

$$d_{\text{coupling}}^{n,m} \leq \min(d_m, d_n)$$

where d_n is the length of line n. The entire distance from the crosstalk source (transmitter m) to the crosstalk victim (receiver n) is $d^{n,m}$. The term $h_k(d^{n,m})$ denotes the transfer function for a channel of length $d^{n,m}$ as defined in (2.13).

Measured Channels

Measurements of direct and crosstalk channels have also been made on real cables for a limited number of cable lengths. These can be used to obtain a more realistic evaluation of DSL system performance.

Shown in Fig. 2.4 is the direct channel transfer function from a 1 km line of diameter 0.5 mm. The empirical transfer function is included for comparison.

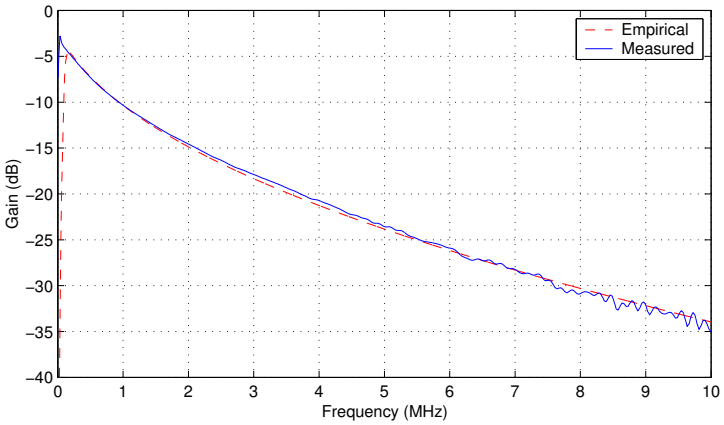

Figure 2.4: Direct Channel Transfer Functions (1 km cable, 0.5 mm pairs)

It is clear that the empirical and measured transfer function match quite well for the direct channels. This is generally the case.

Shown in Fig. 2.5 is a crosstalk channel transfer function from another 1 km line into the 1 km line just described. As can be seen, the empirical model is quite poor at predicting the transfer function of the crosstalk channel. There are several periodic dips in the measured transfer function. These result from the rotation of the different twisted-pairs around one-another within the binder, an effect not included in the empirical models[52]. Despite this the empirical models are still useful for worst-case analysis. They allow the performance of DSL systems to be guaranteed in 99% of deployments, since they are based on 1% worst-case statistics.

More advanced empirical models have been proposed which take the rotation of twisted-pairs into account[52]. This work is still at an early stage and requires more thorough verification before it can be used for accurately predicting DSL system performance.

This thesis uses a combination of empirical models and actual channel measurements to evaluate performance.

2.2 Multi-user Information Theory

Information theory is a useful tool for characterizing the achievable capacity of a communication channel. It can also yield insight into the design of optimal communication systems.

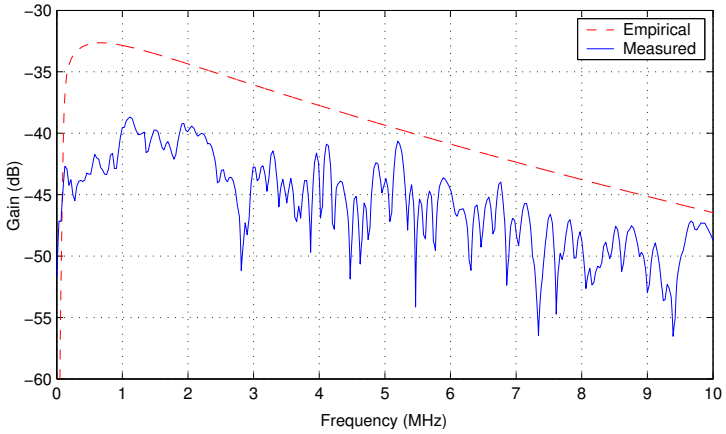

Figure 2.5: Crosstalk Channel Transfer Functions (1 km cable, 0.5 mm pairs)

Multi-user information theory is concerned with the analysis of multi-user channels. Since DSL systems operate in the presence of crosstalk, the DSL network is a multi-user channel. Multi-user information theory is then a valuable tool for the analysis and design of DSL systems.

2.2.1 Rate Regions

In multi-user channels there is an inherent trade-off between the rates of different users. Increasing the rate of one user, by increasing his transmit power, causes more interference to the other users in the network, and their rate is subsequently decreased. Similarly, there may be a limitation on the total amount of transmit power. Allocating more power to one user may preclude the allocation of power to another user.

Due to this inherent trade-off, it is not possible to characterize the capacity of a multi-user channel with a single number. Rather, capacity must be characterized through a *rate region*, a set of all possible rate combinations that can be achieved by the users in a channel. An example rate region is shown in Fig. 2.6. The operating point a is achievable, the operating point b is not.

The rate region depends on the type of channel under consideration. There are many different types of multi-user channel; each type is characterized by the degree of co-ordination available between transmitters or receivers. The most relevant to DSL will now be described.

2.2.2 Interference Channel

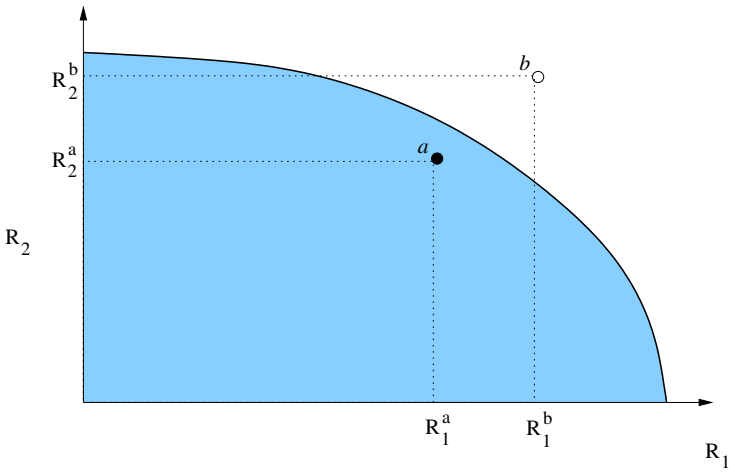

Figure 2.6: Example Rate Region

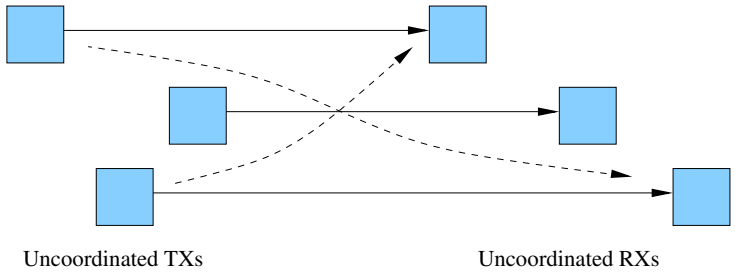

Figure 2.7: Interference Channel

In the *interference channel* (IC) no signal level co-ordination is possible between transmitters or receivers. That is, neither joint encoding at the transmitters nor decoding at the receivers is possible. Each receiver decodes its signal independently and in the presence of the interference from other users as depicted in Fig. 2.7.

The capacity region of the IC is unknown and has been an important problem in information theory since it was first introduced by Shannon[86]. Despite this in a few special cases the capacity region is known. For example, Carleil and Sato showed that very strong interference is equivalent to no interference at all[16, 82]. The strong interference assumption is

$$\frac{|h_k^{n,m}|^2}{\sigma_k^n} \geq \frac{|h_k^{m,m}|^2}{\sigma_k^m}, \forall n, m \neq n. \tag{2.15}$$

In Carleil's scheme a receiver first detects the interference from the other users, treating its own signal of interest as noise. The interfering signals can be detected without error as a result of the strong interference assumption. The interference can then be removed, allowing the receiver to detect its signal of interest as if interference were not present.

In DSL the crosstalk channels are typically weaker than the direct channels; the strong interference condition does not hold, and the interference subtraction scheme just described is inapplicable. Furthermore, these schemes are computationally complex. For this reason current DSL systems treat crosstalk as noise. The bitloading of modem n on tone k is then limited to

$$b_k^n = I\left(x_k^n; y_k^n\right),$$

where $I(a;b)$ denotes the mutual information between a and b, and we assume $f_s = \Delta_f$. As the number of crosstalkers becomes large the interference tends to a Gaussian distribution[75], and the bitloading of modem n on tone k becomes

$$b_k^n = \log_2\left(1 + \frac{\left|h_k^{n,n}\right|^2 s_k^n}{\sum_{m\neq n}\left|h_k^{n,m}\right|^2 s_k^m + \sigma_k^n}\right). \tag{2.16}$$

The total rate of modem n is thus $R_n = f_s \sum_k b_k^n$. Each modem has a total power constraint. Denote the total power constraint of modem n as $\overline{P}_n$. So

$$\Delta_f \sum_k s_k^n \leq \overline{P}_n,$$

where $\overline{P}_n$ denotes the total power that modem n can transmit. This arises from limitations on each modem's analog front-end. For convenience this is reformulated as

$$\sum_k s_k^n \leq P_n, \forall n, \tag{2.17}$$

where $P_n \triangleq \overline{P}_n/\Delta_f$. So assuming that interference is treated as noise, the capacity region of the IC is

$$\mathcal{C}_{\text{IC}} = \bigcup_{\sum_k s_k^n \leq P_n, \forall n} \left\{(R_1, \ldots, R_N) \,:\, R_n \leq f_s \sum_k I\left(x_k^n; y_k^n\right)\right\}.$$

Here the union is taken across all possible transmit spectra in order to characterize the capacity region. In practice this is prohibitively complex and a more efficient search algorithm is required. This is discussed further in Chapter 2.7.

2.2.3 Multi-access Channel

In the *multi-access channel* (MAC) co-ordination is possible between receivers, and they can jointly decode the signals from the different transmitters. No co-ordination is possible between transmitters. This is depicted in Fig. 2.8.

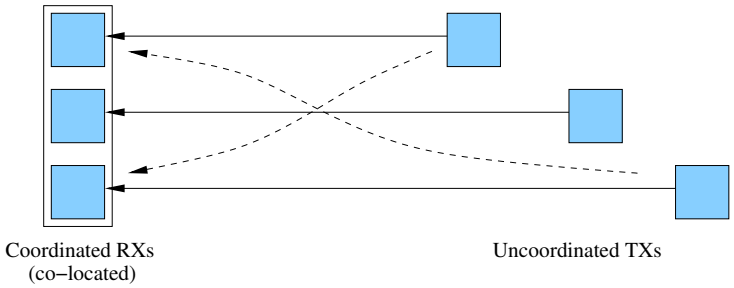

Figure 2.8: Multi-access Channel

An example of a MAC is the uplink of a wireless LAN, where many laptops transmit to a single base-station. Another example is the upstream DSL channel, where many CP transmitters communicate to a set of co-ordinated CO receivers that use joint decoding to cancel crosstalk. This is discussed further in Chapter 4.

Let us start by considering the so-called *single-user bound*, which is the capacity achieved when only one user (CP modem) transmits and all receivers (CO modems) are used to detect that user. Since only one user transmits the received signal at the CO is

$$\mathbf{y}_k = \mathbf{h}_k^n x_k^n + \mathbf{z}_k,$$

where $\mathbf{h}_k^n \triangleq [\mathbf{H}_k]_{\text{col}\,n}$. Using the single-user bound the achievable bitloading of user n on tone k is limited to

$$\begin{aligned} b_k^n &\leq I(x_k^n; \mathbf{y}_k), \\ &= b_{k,\text{mac}}^n, \end{aligned} \tag{2.18}$$

where $I(a;b)$ denotes the mutual information between a and b. Here

$$b_{k,\text{mac}}^n \triangleq \log_2 \left(1 + s_k^n \mathbf{h}_k^n \mathbf{S}_{z,k}^{-1} \mathbf{h}_k^n\right),$$

where the noise correlation is defined $\mathbf{S}_{z,k} \triangleq \mathcal{E}\left\{\mathbf{z}_k \mathbf{z}_k^H\right\}$. With spatially white background noise, $\mathbf{S}_{z,k} = \sigma_k \mathbf{I}_N$, the single-user bound simplifies to

$$b_{k,\text{mac}}^n = \log_2 \left(1 + \sigma_k^{-1} s_k^n \left\|\mathbf{h}_k^n\right\|_2^2\right).$$

In the single-user case with spatially white noise, the single-user bound can be achieved by applying a matched filter to the received vector $\mathbf{y}_k$. The estimate of the transmitted symbol is then

$$\hat{x}_k^n = \left\|\mathbf{h}_k^n\right\|_2^{-2} \left(\mathbf{h}_k^n\right)^H \mathbf{y}_k,$$

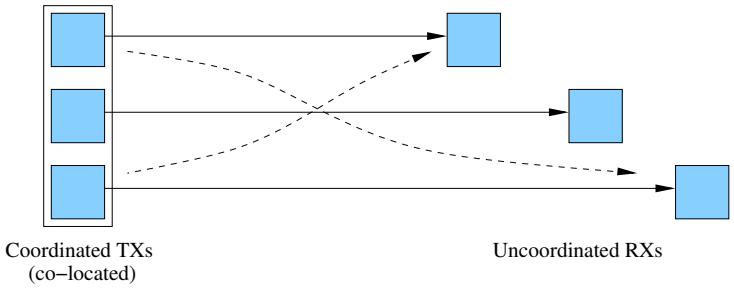

Figure 2.9: Broadcast Channel

$$= \quad x_k^n + \|\mathbf{h}_k^n\|_2^{-2} \left(\mathbf{h}_k^n\right)^H \mathbf{z}_k,$$

which leads to a data-rate of $b_{k,\mathrm{mac}}^n$. Here $(\cdot)^H$ denotes the Hermitian transpose. In the multi-user case, the single-user bound can be achieved by detecting a user last in a *successive interference cancellation* (SIC) structure[59, 96].

From (2.18), the total rate of user n can be bounded

$$R_n \leq f_s \sum_k b_{k,\mathrm{mac}}^n.$$

Assuming that a total power constraint (2.17) applies to each modem, the MAC capacity region can be bounded

$$\mathcal{C}_{\mathrm{MAC}} \subset \bigcup_{\sum_k s_k^n \leq P_n,\, \forall n} \left\{ (R_1, \ldots, R_N) \,:\, R_n \leq f_s \sum_k b_{k,\mathrm{mac}}^n \right\}.$$

In Chapter 4 it is shown that this bound is tight for DSL channels. The bound is then sufficient for the evaluation of multi-user techniques in DSL. An exact characterization of the MAC capacity region is possible and can be useful for other applications, such as wireless communications, where the single-user bound is not tight. The interested reader is directed to [93, 108, 100, 105].

2.2.4 Broadcast Channel

In the *Broadcast Channel* (BC) co-ordination is possible between transmitters, and they can jointly encode the signals intended for different receivers. No co-ordination is possible between receivers. This is depicted in Fig. 2.9.

An example of a BC is the downlink of a wireless LAN, where a single base-station transmits to several laptops. Another example is the downstream DSL channel, where a set of co-ordinated CO transmitters communicate to multiple

CP receivers. The CO transmitters jointly encode their signals to precompensate for the effects of crosstalk. This is discussed further in Chapter 5.

Considering the single-user bound, which is the capacity achieved when all transmitters (CO modems) are used to communicate to a single receiver (CP modem). In this case the received signal on the CP modem is

$$y_k^n = \overline{\mathbf{h}}_k^n \mathbf{x}_k + \mathbf{z}_k,$$

where $\overline{\mathbf{h}}_k^n \triangleq [\mathbf{H}_k]_{\text{row } n}$. Using the single-user bound the achievable bitloading of user n on tone k is limited to

$$\begin{aligned} b_k^n &\leq I(\mathbf{x}_k; y_k^n), \\ &= \log_2 \left(1 + (\sigma_k^n)^{-1} \overline{\mathbf{h}}_k^n \mathbf{S}_{x,k} \left(\overline{\mathbf{h}}_k^n\right)^H\right), \end{aligned} \tag{2.19}$$

where the transmit correlation matrix is defined $\mathbf{S}_{x,k} \triangleq \mathcal{E}\left\{\mathbf{x}_k \mathbf{x}_k^H\right\}$. Define the elements of the correlation matrix $s_k^{n,m} \triangleq [\mathbf{S}_{x,k}]_{n,m}$, and the diagonal elements $s_k^n \triangleq [\mathbf{S}_{x,k}]_{n,n}$. Since $\mathbf{S}_{x,k}$ is positive semi-definite, it follows that

$$s_k^{n,m} \leq \sqrt{s_k^n s_k^m}, \ \forall n, m. \tag{2.20}$$

Now consider the inner-term of (2.19), which is

$$\begin{aligned} \overline{\mathbf{h}}_k^n \mathbf{S}_{x,k} \left(\overline{\mathbf{h}}_k^n\right)^H &= \sum_v h_k^{n,v} \sum_m s_k^{v,m} \text{conj}\left(h_k^{n,m}\right), \\ &\leq \sum_v |h_k^{n,v}| \sum_m \sqrt{s_k^v} \sqrt{s_k^m} \, |h_k^{n,m}|, \\ &= \sum_v |h_k^{n,v}| \sqrt{s_k^v} \sum_m |h_k^{n,m}| \sqrt{s_k^m}, \\ &= \left(\sum_m |h_k^{n,m}| \sqrt{s_k^m}\right)^2, \end{aligned}$$

where conj(.) denotes the complex conjugate operation, and (2.20) is used in the second line. This allows a looser bound to be formed

$$b_k^n \leq b_{k,\text{bc}}^n, \tag{2.21}$$

where

$$b_{k,\text{bc}}^n \triangleq \log_2 \left(1 + (\sigma_k^n)^{-1} \left(\sum_m |h_k^{n,m}| \sqrt{s_k^m}\right)^2\right).$$

In the single-user case the single-user bound can be achieved with a matched transmit filter

$$x_k^m = \text{conj}\left(h_k^{n,m}\right) |h_k^{n,m}|^{-1} \sqrt{s_k^m} \widetilde{x}_k^n,$$

where $\widetilde{x}_k^n$ denotes the *quadrature amplitude modulated* (QAM) symbol intended for user n. Without loss of generality, we assume that the power of $\widetilde{x}_k^n$ is set to unity. This ensures that the PSD level is correct

$$\mathcal{E}\left\{|x_k^m|^2\right\} = s_k^m.$$

The received signal at modem n on tone k is then

$$y_k^n = \left(\sum_m |h_k^{n,m}| \sqrt{s_k^m}\right) \widetilde{x}_k^n + z_k^n.$$

At receiver n an estimate of the transmitted symbol $\widetilde{x}_k^n$ can be formed

$$\begin{aligned} \hat{x}_k^n &= \left(\sum_m |h_k^{n,m}| \sqrt{s_k^m}\right)^{-1} y_k, \\ &= \widetilde{x}_k^n + \left(\sum_m |h_k^{n,m}| \sqrt{s_k^m}\right)^{-1} z_k^n, \end{aligned}$$

which leads to a data-rate of $b_{k,\mathrm{bc}}^n$. In the multi-user case the single-user bound can be achieved through dirty paper coding[106, 101].

From (2.21) the total rate of user n can be bounded

$$R_n \leq f_s \sum_k b_{k,\mathrm{bc}}^n.$$

Assuming that a total power constraint (2.17) applies to each modem, the BC capacity region can be bounded

$$\mathcal{C}_{\mathrm{BC}} \subset \bigcup_{\sum_k s_k^n \leq P_n, \forall n} \left\{(R_1, \ldots, R_N) \,:\, R_n \leq f_s \sum_k b_{k,\mathrm{bc}}^n\right\}.$$

Chapter 5 shows that this bound is tight in DSL channels. The bound is then sufficient for the evaluation of multi-user techniques in DSL. An exact characterization of the BC capacity region is possible and can be useful for other applications, such as wireless communications, where the single-user bound is not tight. The interested reader is directed to [106, 101, 62, 104].

Part I

Multi-user Spectra Coordination

Overview

Crosstalk is a major problem in modern DSL systems such as ADSL and VDSL. Crosstalk can be mitigated through the coordination of DSL modems. This can be done either on a spectral or signal level. Spectral coordination is discussed in this part of the thesis. Signal coordination is discussed in Part II.

Signal level coordination leads to maximum performance. However, for signal level coordination to be used either the transmitters or receivers must be co-located. In some situations this is not possible, for example the mixed deployment shown in Fig. 2.10. Here one DSL service is deployed from the CO and another from a *remote terminal* (RT). Since neither the head-end modems nor the CP modems are co-located, it is impossible to coordinate transmission or reception on a signal level. As a result, the only way to mitigate crosstalk is through spectral coordination.

Signal coordination increases the run-time complexity of DSL modems significantly. Spectral coordination, on the other hand, does not increase run-time complexity. So when the cost of DSL equipment must be kept low, spectral coordination is preferable.

With spectral coordination the transmit spectra of the modems within a network are limited to minimize the negative effects of crosstalk. Each modem must achieve a trade-off between maximizing its own data-rate and minimizing the crosstalk it causes to other modems within the network. The goal is to achieve a fair trade-off between the rates of the different users.

A classical scenario is shown in Fig. 2.10 where a binder carries a mixture of CO and RT distributed DSLs. Since the RT is located further downstream than the CO, it has a relatively strong crosstalk channel into CP1. In some cases the crosstalk channel from the RT to CP1 can be even stronger than the direct channel from the CO to CP1. If the RT transmits at full power it will induce a large amount of crosstalk on the CO distributed line, significantly reducing its data-rate. This is referred to as the near-far scenario since the near-end transmitter (RT) causes a huge amount of crosstalk to the far-end

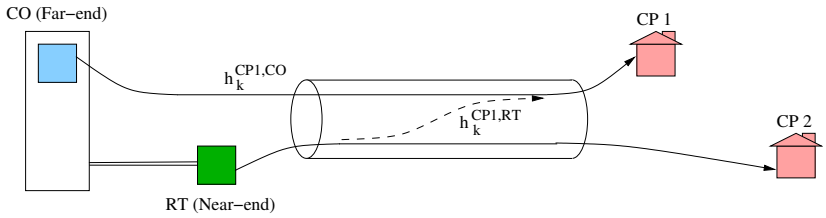

Figure 2.10: Mixed Deployment Scenario

receiver (CO). Clearly some power-backoff is necessary on the RT transmitter to ensure that a fair rate is achieved by the CO line.

From an information theory perspective the DSL network is an interference channel since signal coordination is not possible. Our goal is to characterize the capacity region of this interference channel, and the corresponding optimal transmit spectra.

Chapter 3 investigates the design of optimal transmit spectra for a network of interfering DSLs. This problem was previously considered intractable since it requires the solution of a high-dimensional, non-convex optimization. It is shown that, through the use of a dual-decomposition, the optimization can be solved in an efficient, tractable way. The resulting algorithm, which we name *optimal spectrum balancing*, gives significant gains over existing spectral coordination techniques, typically doubling or tripling data-rates.

The material in Chapter 3 has been published as [40, 39, 110, 14, 94, 97], submitted for publication as [20, 95], and has been patented by Alcatel[32]. The *optimal spectrum balancing* algorithm was submitted to standardization as [36, 37, 38, 35] and is now part of the draft ANSI standard on Dynamic Spectrum Management[8].

Chapter 3

Optimal Spectrum Balancing

3.1 Introduction

This chapter investigates the design of transmit spectra for a network of interfering DSLs[1]. Static spectrum management is the traditional approach and employs identical spectral masks for all modems. To ensure widespread deployment, these masks are based on worst case scenarios[6]. As a result they can be overly restrictive and lead to poor performance.

Dynamic spectrum management (DSM), a new paradigm, overcomes this problem by designing the spectra of each modem to match the specific topology of the network[47, 88, 44]. These spectra are adapted based on the direct and crosstalk channels seen by the different modems. They are customized to suit each modem in each particular situation.

A DSM algorithm known as *iterative waterfilling* was recently proposed and demonstrates the spectacular performance gains that are possible[107]. An unanswered question at this point is: How much more can be achieved?

The goal of this chapter is to address this question. The focus is on centralized spectrum management where a *spectrum management center* (SMC) is responsible for setting the spectra of the modems within a network. The chapter will present an algorithm for *optimal spectrum balancing* in the DSL interference channel. Assuming that all modems employ *discrete multi-tone* (DMT) mod-

[1]The work in this chapter was done in close collaboration with Prof. Wei Yu, University of Toronto, Canada.

ulation this algorithm achieves the best possible balance between the rates of the different modems in the network, allowing operation at any point on the rate region boundary.

The algorithm is suitable for direct application when a SMC is available. Note that one disadvantage of centralized algorithms is that when a new line is activated, or if a line is deactivated, re-optimisation of the modem transmit spectra is necessary to ensure optimal performance. This is one disadvantage of centralized algorithms with respect to more autonomous algorithms such as iterative waterfilling. Furthermore centralized spectrum management requires a SMC that may not be available in the unbundled case, where multiple operators share the same binder. In these cases an autonomous algorithm may be preferred. Optimal spectrum balancing can still be useful here since it provides an upper bound on performance of all DSM algorithms, both centralized and autonomous. Furthermore, the spectra generated by the proposed algorithm provide valuable insight that can be used in autonomous algorithm design.

One may argue, if centralized control is available (via a SMC), why is it that crosstalk cancellation, enabled through signal coordination, is not implemented? Although crosstalk cancellation leads to greater performance gains, it is more complex to implement and is not feasible when head-end modems are not co-located in the same *central office* (CO) or *remote terminal* (RT). Furthermore, since crosstalk cancellation uses signal coordination, it requires an entirely new design of both the *DSL access multiplexer* (DSLAM) and *customer premises* (CP) modems. Spectral coordination, on the other hand, only involves setting the transmit PSD levels of the modems. This can be done without any change to the modem hardware currently deployed in the field and is feasible to implement right now. We also note that in several specific scenarios crosstalk cancellation is possible even without signal coordination[41, 113]. Whilst performance gains are possible, these techniques are highly complex. The rest of the chapter assumes that crosstalk cancellation is not performed, and each modem treats crosstalk as additive noise.

The multi-user DSL channel with no signal coordination is an example of an interference channel in multi-user information theory. The capacity region and the optimal code design for the interference channel are long-standing open problems in information theory. This chapter considers an achievable rate region for the interference channel within the context of currently deployed DSL modems in the field. In this case, interference must be treated as noise, and the optimization of the achievable rate region is reduced to the optimization of the joint spectra amongst all the users. Hence the solution obtained using the optimal spectrum balancing algorithm proposed in this chapter, although not the best possible for the interference channel, is optimal within the current capabilities of the DSL modems already developed.

The main difficulty in the optimal design of the multi-user spectrum in the

DSL context is the computational complexity associated with the optimization problem. The constrained optimization problem is non-convex, and a naive exhaustive search leads to an exponential complexity in the number of tones K in the system. In ADSL $K = 256$ whilst in VDSL $K = 4096$. This leads to a computationally intractable problem.

The algorithm presented in this chapter overcomes the exponential complexity in K through the use of a technique called dual decomposition. The computational complexity of the proposed algorithm, although linear in K, is still exponential in the number of users. Nevertheless, it gives a practical way to compute the achievable rate regions for channels with a small number of users. Doing so was not possible prior to this work.

Despite the large reduction in complexity that the optimal spectrum balancing algorithm achieves, in some scenarios it may still be too complex for practical implementation. To address this issue, a simpler algorithm is developed based on an iterative approach. This algorithm has a quadratic complexity in the number of users, and is applicable to existing DSL modems. As will be shown in Section 3.6, the algorithm exhibits near-optimal performance, yielding significant improvement over existing state-of-the-art.

The rest of the chapter is organized as follows. The system model for a network of interfering DSL modems is given in Section 3.2. The problem is then to characterize the achievable rate region and the corresponding transmit spectra. This problem is formulated in Section 3.3. Section 3.3.2 describes existing solutions, which are typically heuristic and sub-optimal. In Section 3.3.4 it is shown that trying to find the optimal solution directly through an exhaustive search is computationally intractable. Section 3.3.5 and 3.3.6 show that the spectrum balancing problem has an equivalent dual problem. This can be decomposed into separate sub-problems that are then solved independently on each tone. The resulting algorithm, which we name optimal spectrum balancing, is presented in Section 3.4 and gives an efficient solution to the spectrum balancing problem. The complexity of the algorithm is discussed in Section 3.4.3. Section 3.5 describes a simpler algorithm that solves the spectrum management problem through an iterative approach. Section 3.6 compares the performance of the proposed algorithms to existing spectrum balancing techniques. Conclusions are drawn in Section 3.7.

3.2 System Model

This chapter only considers DSM as applied to DMT modulated modems. Whilst some form of DSM can also be applied to single carrier modems it often leads to inferior performance since dynamic shaping of the transmit spectra is not possible. As such it is assumed that any non-DMT systems form part

of the background noise.

It is assumed that each modem treats the interference from other modems as noise. This is an interference channel, and the achievable bitloading of modem n on tone k is given by (2.16).

We denote the maximum bitloading that a modem can support as $b_{\max}$, which lies in the range 8-15 in current standards[9][3][7]. Since a practical error correction coding scheme will be employed, the system will experience an SNR-gap to capacity, denoted as Γ. Modifying (2.16) to incorporate the maximum bitloading limitation and the SNR-gap to capacity leads the following bitloading for modem n on tone k

$$b_k^n \triangleq \min\left(b_{\max}, \log_2\left(1+\frac{1}{\Gamma}\frac{\left|h_k^{n,n}\right|^2 s_k^n}{\sum_{m\neq n}\left|h_k^{n,m}\right|^2 s_k^m+\sigma_k^n}\right)\right). \tag{3.1}$$

The data-rate on line n is then

$$R_n = f_s \sum_k b_k^n.$$

In practice the relationship between the received *signal-to-interference-plus-noise ratio* (SINR) and the bitrate may be more complex and is in fact dependent on the coding scheme employed within the modem. In particular the maximum bitloading will apply to the encoded data-rate. In this chapter (3.1) will be used for simplicity however the algorithms presented here can be applied to any arbitrary function that relates the bitloading to the SINR on each tone.

3.3 The Spectrum Management Problem

We restrict our attention to the two user case for ease of explanation. Extensions to more than two users will be discussed in Section 3.4.2. The spectrum management problem for the two user case is defined as

$$\max_{\mathbf{s}_1,\mathbf{s}_2} R_2 \quad \text{s.t.} \quad R_1 \geq R_1^{\text{target}}, \tag{3.2}$$

where R_n^{target} denotes the target data-rate of user n, and the PSD vector of user n is defined $\mathbf{s}_n \triangleq [s_1^n, \ldots, s_K^n]$. As described in Section 2.2.1, the rate region is a plot of all possible operating points, or rate combinations that can be achieved in a multi-user channel. Operating points on the boundary of the region are said to be optimal. These points and their corresponding PSD combinations can be characterized by solving the spectrum management problem (3.2) for a range of values of R_1^{target}. This is the goal of this chapter.

3.3.1 Constraints

The optimisation (3.2) is typically subject to a total power constraint on each modem

$$\sum_k s_k^n \leq P_n,\ n = 1, 2. \tag{3.3}$$

Spectral mask constraints may also apply

$$s_k^n \leq s_k^{\text{mask}},\ \forall k, n. \tag{3.4}$$

Naturally the spectra must also be non-negative.

$$s_k^n \geq 0,\ \forall n, k. \tag{3.5}$$

3.3.2 Existing Solutions

This Section will give a review of existing solutions to the spectrum management problem. These solutions are typically heuristic and result in sub-optimal solutions.

Flat Power Back-off

With the flat power back-off method a modem transmits the same PSD on all tones[73]. This PSD is set to the minimum possible value that still allows the modem to achieve its target data-rate. The PSD of user n with flat power back-off is set to s_{flat}^n on all tones, where s_{flat}^n is chosen such that

$$f_s \sum_k b_k^n(s_{\text{flat}}^1, s_{\text{flat}}^2) = R_n^{\text{target}},\ \forall n.$$

Here $b_k^n(s_k^1, s_k^2)$ denotes the bitloading of user n corresponding to the PSD combination (s_k^1, s_k^2) as calculated by (3.1). Flat power back-off cannot vary the degree of power back-off with frequency. Since crosstalk coupling varies significantly with frequency this is a major disadvantage.

Reference PSD Method

In the reference PSD method each modem sets its transmit PSD such that the corresponding received PSD is equal to the *reference PSD*[79, 83, 73]. Consider modem n and its transmit PSD on the kth tone, s_k^n. The corresponding received PSD will be

$$s_{k,\text{rx}}^n = |h_k^{n,n}|^2\, s_k^n. \tag{3.6}$$

The reference PSD method requires that

$$s_{k,\text{rx}}^n = s_{k,\text{ref. PSD}}^n,\ \forall n, k.$$

Combining this with (3.6) implies that the transmit PSD must be

$$s_k^n = |h_k^{n,n}|^{-2}\, s_{k,\text{ref PSD}}^n.$$

The reference PSD method has been adopted by standardization groups for use in VDSL[2]. In the standards the reference PSD is specified as

$$s_{k,\text{ref PSD}}^n = \begin{cases} -60 - 22\sqrt{f_k}\ \text{dBm/Hz}, & \text{in US band 1 } (k < 1972); \\ -60 - 17.18\sqrt{f_k}\ \text{dBm/Hz}, & \text{in US band 2 } (k \geq 1972), \end{cases}$$

where f_k denotes the frequency in MHz on tone k[9, 7]. The logic behind this choice of reference PSD is as follows. First note that the attenuation on an L km line can be well approximated by

$$18.0975L\sqrt{f_k}\,\text{dB}.$$

So $22\sqrt{f_k}$ is the attenuation experienced by a 1216m line. If a modem transmits at the spectral mask, which is -60 dBm/Hz, then the received PSD on a 1216 m line will be $-60-22\sqrt{f_k}$ dBm/Hz. So the reference PSD method forces each line to adjust its transmit PSD, such that the corresponding received PSD is equal to the received PSD of a 1216 m line. This forces all lines to perform in a similar way, there-by limiting the crosstalk that lines cause one-another and assuring a fair rate-allocation for all lines. This approach is also known as the *reference length method*; here the *reference length* is set to 1216m in US band 1[73]. Unfortunately all lines will now achieve the same data-rate as a 1216 m line, in US band 1, regardless of line length. So shorter lines cannot exploit their lower channel attenuation to achieve higher data-rates.

Longer lines are generally not active in US band 2 since their direct channel attenuation is too high. For this reason less PBO is required in US band 2 than in US band 1. To take this into account, the reference length in US band 2 is decreased to 950 m, allowing modems to transmit with a higher PSD in that band. The attenuation of a 950 m line is approximately $17.18\sqrt{f_k}$, which motivates the choice for the reference PSD in US band 2. The use of different reference lengths in different frequency ranges is also known as the *multiple reference length method*[73]. The choice of 1216 m and 950 m as the reference lengths was based on an optimization over several representative scenarios. The goal of the optimization was to ensure that the reach of several representative services does not decrease by more than 10% as a result of crosstalk[79, 83].

The short-fall of the reference PSD method is that it forces all lines to adopt the same pair of reference lengths for each US band. Furthermore, the transmit PSD is forced to be of the form $-60-18.0975L\sqrt{f_k}$ dBm/Hz. These constraints reduce flexibility in the allocation of transmit spectra and can lead to poor performance.

[2] In ADSL a spectrum management technique has not yet been standardized[1].

Reference Noise Method

The reference noise method works with the weighted rate-sum, which will be described in Section 3.3.5. Essentially maximizing the weighted rate-sum,

$$wR_1 + (1-w)R_2,$$

is equivalent to the original spectrum management problem (3.2) provided a correct weight w is chosen[87].

Since the weighted rate-sum optimization is non-convex, the reference noise method makes the assumption of high SINR[87]. Under this assumption the bitloading of user n on tone k can be approximated as

$$b_k^n \simeq \log_2 \left(\frac{1}{\Gamma} \frac{\left|h_k^{n,n}\right|^2 s_k^n}{\sum_{m \neq n} \left|h_k^{n,m}\right|^2 s_k^m + \sigma_k^n} \right).$$

Total power constraints (3.3) are not considered and the optimization is done only under spectral mask (3.4) and non-negativity (3.5) constraints. Continuous bitloading is assumed and the maximum bitloading constraint is neglected. In the 2 user case this leads to a convex optimization that can be solved in closed form. The resulting optimal PSDs for tone k are

$$\begin{aligned} s_k^1 &= \left[\frac{w}{1-2w} \sigma_k^2 \left| h_k^{2,1} \right|^{-2} \right]_0^{s_k^{mask}}, \\ s_k^2 &= \left[\frac{1-w}{2w-1} \sigma_k^1 \left| h_k^{1,2} \right|^{-2} \right]_0^{s_k^{mask}}, \end{aligned} \tag{3.7}$$

where $[x]_a^b \triangleq \min(\max(x,a),b)$. Note that with such PSDs the crosstalk experienced by line 1 will be $\frac{1-w}{2w-1}\sigma_k^1$, which is a scaled version of the noise experienced by line 1. The same can be said of line 2. Note that if $w > 0.5$, then $s_k^1 = 0$, and if $w < 0.5$, then $s_k^2 = 0$. This is quite counter-intuitive, since a higher w should place more importance on the data-rate of user 1. Furthermore, there is a singularity at $w = 0.5$. Using (3.7) with values of $w \neq 0.5$ will result in only one active user on each tone. This *frequency division multiple access* (FDMA) approach typically leads to poor performance in DSL.

In the general N user case no closed form solution for maximizing the weighted rate-sum exists. Despite this (3.7) can still be applied as a somewhat heuristic approach. The weighting term $\frac{w}{1-2w}$ in (3.7) is set to unity which leads to the following transmit spectra for user n on tone k

$$s_{k,\text{rnoise}}^n = \left[\sigma_k^{\text{ref}} \left| h_k^{\text{ref},n} \right|^{-2} \right]_0^{s_k^{\text{mask}}}.$$

Here σ_k^{ref} denotes the so-called *reference noise*. This is the noise experienced by a theoretical *reference line*. The transfer function $h_k^{\text{ref},n}$ is the crosstalk

channel from transmitter n into the reference line. Both the reference noise and reference line are parameters that must be chosen by the modem designer. Typically the reference line is the longest line in the binder, and the reference noise is chosen as a representative noise PSD.

Note that with the reference noise method, the crosstalk that each line induces on the reference line is equal to the reference noise. This is the defining characteristic of this method.

Section 3.6 will show that the reference noise method is near-optimal at low frequencies, where the high-SINR assumption is valid. Unfortunately at high frequencies the assumption is invalid and the reference noise method gives poor performance.

Iterative Waterfilling

Iterative waterfilling is a power allocation algorithm based on the concept of greedy optimization[107, 109]. Each line tries to selfishly maximize its own data-rate under a given power constraint P_n,

$$\begin{aligned} s^n_{k,\text{iw}} \quad &= \quad \arg\max_{\mathbf{s}_n} \sum_k \log_2\left(1+\frac{1}{\Gamma}\frac{|h^{n,n}_k|^2 s^n_k}{\sum_{m\neq n}|h^{n,m}_k|^2 s^m_k+\sigma^n_k}\right), \\ \text{s.t.} \quad &\sum_k s^n_k \leq P_n, \\ &s^n_k \geq 0,\ \forall k. \end{aligned}$$

In the continuous bitloading case this leads to a waterfilling PSD similar to that described in (2.9). In iterative waterfilling spectral masks are not applied. Furthermore, each user must waterfill against both the background noise and interference of other users. The resulting PSD for user n on tone k is

$$s^n_{k,\text{iw}} = \left[\frac{1}{\lambda_n} - \Gamma\frac{\sum_{m\neq n}|h^{n,m}_k|^2 s^m_k+\sigma^n_k}{|h^{n,n}_k|^2}\right]^+, \tag{3.8}$$

where $[x]^+ \triangleq \max(0,x)$ and λ_n is chosen such that $\sum_k s^n_k = P_n$[107]. In iterative waterfilling the modem sets its total power P_n to the minimum possible value that still allows the modem to achieve its target data-rate[107]. Put another way, the waterfilling level, λ_n, is chosen such that

$$f_s\sum_k \log_2\left(1+\frac{1}{\Gamma}\frac{|h^{n,n}_k|^2 s^n_{k,\text{iw}}(\lambda_n)}{\sum_{m\neq n}|h^{n,m}_k|^2 s^m_k+\sigma^n_k}\right) = R^{\text{target}}_n,$$

where $s^n_{k,\text{iw}}(\lambda_n)$ denotes the waterfilling PSD with waterfilling level λ_n, as calculated by (3.8).

Algorithm 3.1 Iterative Waterfilling

repeat
  for $n = 1, \ldots, N$
    repeat
      $s_k^n = s_{k,\mathrm{iw}}^n(\lambda_n)$, $\forall k$, found using (3.8)
      if $f_s \sum_k b_k^n < R_n^{\mathrm{target}}$ and $\sum_k s_k^n \leq P_n$, then increase λ_n, else decrease λ_n
    until convergence
  end
until convergence

With iterative waterfilling the users adjust their PSDs one-by-one according to (3.8). The process iterates through all the users until convergence is reached. The algorithm is listed as Alg. 3.1.

CO distributed lines are typically long and experience large attenuation at high frequencies. As a result they are only active on low frequencies. The RT distributed lines, on the other hand, are typically short and experience a constant attenuation with frequency. The best strategy is then for the RT to switch off its lower tones and transmit only in the high frequencies, where the CO lines are not active. The RT still achieves a good data-rate, and the lower frequencies are left free for the CO lines to use.

Unfortunately iterative waterfilling does not follow this approach. RT distributed lines are typically short and experience little crosstalk from CO distributed lines. As a result the transmit PSDs with iterative waterfilling are flat on RT distributed lines. The low frequencies are not switched off to protect the CO lines and this leads to poor performance.

Several studies have been made on the stability[94, 97], implementation[14] and expected gain[98] of iterative waterfilling.

Other Techniques

In [43] an optimal solution was formulated to the spectrum balancing problem based on simulated annealing. Unfortunately, simulated annealing cannot guarantee convergence to the global optimum, and the convergence speed can also be slow. Furthermore, due to the high complexity of this algorithm, the transmit PSD is constrained to be flat within each transmission band. Whilst this reduces complexity, it significantly limits the search space of the algorithm and generally leads to poor performance.

An attempt to formulate an optimal solution to the spectrum balancing problem was also made in [85]. In the proposed algorithm, several symmetry constraints were imposed on the transmit spectra to ensure a convex optimization. These constraints limit the search space of the algorithm and generally lead to

poor performance.

In [45] the optimal solution to the spectrum balancing problem was considered for the 2 user case. It was shown that if the crosstalk channels are very weak the weighted rate-sum optimization (3.11) is approximately convex. This allows the optimal transmit spectra to be found efficiently through the use of convex optimization techniques. Unfortunately this approach does not extend to more than 2 users. Furthermore, in VDSL or in RT distributed ADSL, the crosstalk channels are not sufficiently weak and the proposed algorithm is inapplicable.

In [46] the same authors considered the spectrum balancing problem under a strong interference condition (2.15). Unfortunately in many cases the crosstalk is not strong enough to satisfy (2.15), but is too strong to satisfy the weak interference condition required for the algorithm in [45]. Furthermore, the algorithm in [46] assumes that the strong interference is detected and subtracted prior to the detection of the signal of interest. Such advanced interference subtraction techniques are highly complex and are not available in current DSL systems.

3.3.3 Bitloading

Continuous Bitloading

Let us now return our attention to the spectrum management problem (3.2) presented earlier. Consider the case where the modems can support any possible bitloading. Denote the accuracy with which modems can control their transmit PSD as Δ_s. In current standards Δ_s is set to 0.5 dBm/Hz[5]. The total power (3.3) and spectral mask constraints (3.4) make it possible to upper bound the transmit power on any tone

$$s_k^n \leq s_k^{\mathrm{n,max}},$$

where $s_k^{n,\max} \triangleq \min(P_n,\ s_k^{\mathrm{mask}})$. Hence the range of s_k^n can be limited to

$$s_k^n \in \{0, \Delta_s, \ldots, s_k^{n,\max}\}.$$

So for the 2 user case on tone k there are $q_k = \prod_{n=1}^{2} \left(s_k^{n,\max}\Delta_s^{-1} + 1\right)$ possible PSD combinations.

Discrete Bitloading

In practice DSL modems can only support a fixed set of discrete bitloadings. So the search space can be reduced to the PSDs corresponding to these exact bitloadings, reducing complexity considerably without affecting optimality. Define the vector $\mathbf{b}_k \triangleq [b_k^1, \ldots, b_k^N]^T$, which contains the bitloading of all users on tone k, and $\overline{\mathbf{s}}_k \triangleq [s_k^1, \ldots, s_k^N]^T$, which contains the PSD of all users on tone

k. Provided that $b_k^n \leq b_{\max}$, $\forall n$, then (3.1) can be used to find the PSD combination $\overline{\mathbf{s}}_k$ corresponding to a particular bitloading combination $\mathbf{b}_k$ as is now shown. First define $\mathbf{A}_k \triangleq a_k^{n,m}$ where

$$a_k^{n,m} \triangleq \begin{cases} 0 & n = m, \\ \Gamma |h_k^{n,m}|^2 & n \neq m. \end{cases}$$

Also define $\boldsymbol{\sigma}_k \triangleq \Gamma[\sigma_k^1, \ldots, \sigma_k^N]^T$, $\mathbf{D}_k \triangleq \text{diag}\{|h_k^{1,1}|^2, \ldots, |h_k^{N,N}|^2\}$ and $\Lambda_k \triangleq \text{diag}\{2^{b_k^1} - 1, \ldots, 2^{b_k^N} - 1\}$. Since $b_k^n \leq b_{\max}$, (3.1) can be rewritten as

$$|h_k^{n,n}|^2 \, s_k^n - \Gamma(2^{b_k^n} - 1) \sum_{m \neq n} |h_k^{n,m}|^2 \, s_k^m = \Gamma(2^{b_k^n} - 1)\sigma_k^n, \, \forall n. \tag{3.9}$$

Note that taking (3.9) for each n forms a set of n linear equations in $\overline{\mathbf{s}}_k$. These can be written in matrix form as

$$(\mathbf{D}_k - \Lambda_k \mathbf{A}_k)\overline{\mathbf{s}}_k = \Lambda_k \boldsymbol{\sigma}_k.$$

The PSD combination required to support a particular bitloading combination $\mathbf{b}_k$ is then

$$\overline{\mathbf{s}}_k = (\mathbf{D}_k - \Lambda_k \mathbf{A}_k)^{-1} \Lambda_k \boldsymbol{\sigma}_k. \tag{3.10}$$

In the remainder of the chapter for the 2 user case, $s_k^n(b_k^1, b_k^2)$ is used to denote the PSD of user n corresponding to the bitloadings b_k^1, b_k^2 as calculated by (3.10). Hence the range of PSD combinations (s_k^1, s_k^2) can be limited to

$$(s_k^1, s_k^2) \in \left\{ \left(s_k^1(b_k^1, b_k^2), s_k^2(b_k^1, b_k^2) \right) \mid b_k^n \in \{0, \ldots, b_{\max}\}, \, \forall n \right\}.$$

So on tone k there are $q_k = (b_{\max} + 1)^2$ possible PSD combinations.

3.3.4 Exhaustive Search

At this point a simplistic algorithm could be proposed to find the optimal PSDs based on an exhaustive search. On tone k there are q_k possible PSD combinations. Taking all possible PSD levels across all tones results in $\prod_k q_k$ possible PSD combinations. The feasibility of each PSD combination is determined based on any power constraints as described in Section 3.3.1, and on the target rate constraint for user 1. The PSD combination that maximizes the data-rate of user 2 is then selected.

Unfortunately whilst this algorithm is simple to implement, its complexity in the discrete bitloading case is $\mathcal{O}((b_{\max} + 1)^{2K})$. With $K = 256$ in ADSL and $K = 4096$ in VDSL, this results in a computationally intractable problem. In the continuous bitloading case the complexity is even higher.

So an exhaustive search for the optimal PSDs leads to a computationally intractable problem. The fundamental reason behind this is as follows. The

target rate constraint for the first line and the total power constraint associated with each line couples the power allocation problem across frequency. As such the PSD combination must be searched in a joint fashion across all tones. This results in an exponential complexity in K and an intractable problem.

The following sections will show that the dual decomposition approach can be used to transform this problem into an equivalent one that has a linear complexity in K and is tractable. Since the development has many stages a brief overview is given here before proceeding with a detailed explanation in sections 3.3.5, 3.3.6, and 3.4.

To start with, Section 3.3.5 replaces the original optimisation problem (3.2), with a weighted rate-sum maximization. With a correctly chosen weight w, maximizing the weighted rate-sum implicitly enforces the target rate constraint on user 1.

In Section 3.3.6 the power constrained optimisation is replaced by an equivalent dual problem. This dual problem consists of an unconstrained optimisation of a Lagrangian. In the Lagrangian the total power constraints are enforced through the use of Lagrangian multipliers which form part of the objective function. When the Lagrangian multipliers are chosen correctly, maximizing the Lagrangian will implicitly enforce the power constraints. The power constraints need not be explicitly enforced and the problem can be decoupled across frequency.

After this decoupling the optimisation can be solved by maximizing the Lagrangian independently on each tone, an approach known as dual decomposition. This leads to a complexity that is linear, rather than exponential, in K and the problem becomes computationally tractable. This is the main innovation in this chapter.

The dual decomposition method is a commonly used approach in convex optimization theory for solving constrained optimization problems through an equivalent unconstrained dual-problem. The dual problem is chosen so that it can be decomposed into several simpler sub-problems. The dual decomposition method has been applied in other communication problems with convex objective functions such as joint routing and resource allocation[103] and power allocation in the vector multiple access channel[110]. This work shows that the dual decomposition method can also be applied to non-convex optimizations.

3.3.5 The Weighted Rate-sum

We start by considering the following optimization problem where the objective is to maximize the weighted rate-sum

$$\max_{\mathbf{s}_1,\mathbf{s}_2} wR_1 + (1-w)R_2. \tag{3.11}$$

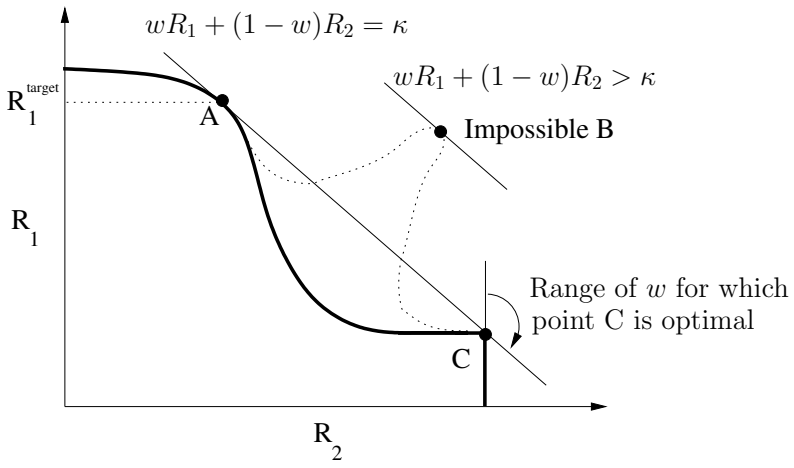

Figure 3.1: Optimality of A in the weighted rate-sum (3.11) implies optimality in the spectrum management problem (3.2)

The following theorem shows that solving this problem is equivalent to solving the original spectrum management problem (3.2).

Theorem 3.1 *For any $0 \leq w \leq 1$, there exists at least one R_1^{target} for which the weighted rate-sum optimization (3.11) is equivalent to the original spectrum management problem (3.2).*

Proof: The proof is made by illustration. As shown in the rate region in Fig. 3.1, for any $w \in [0,1]$ there will be at least one point that maximizes the weighted rate-sum. If there are multiple optimal points the optimization search will need to explore each point in turn. In this case there are two points A and C. Consider one of these points $A \triangleq (R_1^a, R_2^a)$. Assume that there exists some other point in the rate region $B \triangleq (R_1^b, R_2^b)$ such that $R_1^b > R_1^a$ and $R_2^b > R_2^a$. This would imply that point B leads to a larger weighted rate-sum than point A, but this is contradicted by the optimality of point A in the weighted rate-sum (3.11). So no such point B can exist. Hence R_2^a is the highest rate for line 2 which will allow a target rate of R_1^a be achieved on line 1. This implies that point A is optimal in terms of the original spectrum management problem (3.2) for the target rate $R_1^{\text{target}} = R_1^a$. ■

In Section 3.4 the optimal spectrum balancing algorithm is described. It finds the optimal solution to (3.11) for any particular w. Theorem 3.1 implies that solving (3.11) is equivalent to solving (3.2) for some particular R_1^{target}. So the proposed algorithm is guaranteed to always yield an optimal solution to the

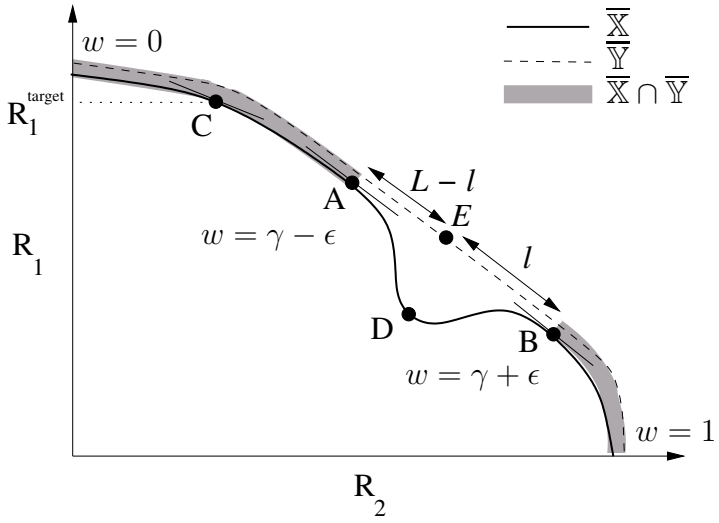

Figure 3.2: Operating points in $\mathbb{X} \cap \mathbb{Y}$ can be found through a weighted rate-sum optimization

spectrum management problem (3.2). The full proof is deferred to Theorem 3.4 and Appendix A.

By sweeping through different values of w a large portion of the rate region can be characterized. Unfortunately points which lie on the *interior* of the *convex hull* of the rate region, e.g. point D in Fig. 3.2, cannot be found with the proposed algorithm. This is because these points are not optimal in terms of a weighted rate-sum (3.11). For example in Fig. 3.2 both A and B are superior to point D. This is one of the problems inherent to the use of the weighted rate-sum as an optimization metric, however the weighted rate-sum appears difficult to avoid since trying to solve (3.2) directly leads to an exponential complexity in K and an intractable problem.

Fortunately all achievable points on the *convex hull* of the rate region *can* be characterized using a weighted rate-sum and hence can be found using optimal spectrum balancing. This statement is formalized in the following theorem.

Theorem 3.2 *For any rate region $\mathbb{X}$, define $\overline{\mathbb{X}}$ as the boundary of $\mathbb{X}$, $\mathbb{Y}$ as the convex hull of $\mathbb{X}$, and $\overline{\mathbb{Y}}$ as the boundary of $\mathbb{Y}$. Consider any operating point $C \triangleq (R_1^c, R_2^c)$ which is achievable $C \in \mathbb{X}$ and on the boundary of the convex hull of the rate region $C \in \overline{\mathbb{Y}}$ as depicted in Fig. 3.2. There exists some w such that the PSDs at point C can be found through a weighted rate-sum maximization.*

Proof : C is on the boundary of the convex set $\mathbb{Y}$. So there exists no point $D \triangleq (R_1^d, R_2^d) \in \mathbb{Y}$ such that $R_1^d > R_1^c$ and $R_2^d > R_2^c$. This implies that for some w

$$wR_1^c + (1-w)R_2^c \geq wR_1^d + (1-w)R_2^d, \ \forall (R_1^d, R_2^d) \in \mathbb{Y}.$$

Now since $\mathbb{X} \subset \mathbb{Y}$

$$wR_1^c + (1-w)R_2^c \geq wR_1^d + (1-w)R_2^d, \ \forall (R_1^d, R_2^d) \in \mathbb{X}.$$

So the point C gives the maximum weighted rate-sum of all points within the rate region $\mathbb{X}$ for some particular weight w. Hence the point C is optimal in the weighted rate-sum (3.11) for that w and can be found through a weighted rate-sum maximization. ■

Corollary 3.3 *For any convex rate-region, all optimal operating points can be found through a weighted rate-sum optimization.*

Proof : In a convex rate region, the boundary of the convex hull $\overline{\mathbb{Y}}$, contains the entire boundary of the rate region and $\overline{\mathbb{X}} = \overline{\mathbb{Y}}$. All optimal operating points in terms of the original spectrum management problem (3.2) lie on the boundary of the rate region $\overline{\mathbb{X}}$. Hence Theorem 3.2 implies that all optimal operating points can be found through a weighted rate-sum optimization. ■

Theorem 3.2 implies that *any achievable operating point on the boundary of the convex hull of the rate region can be found through a weighted rate-sum optimization*. If the rate region is close to being convex, then the majority of the optimal operating points can be found. Thankfully this is the case in DSL channels as is now explained.

In the wireline medium there is some correlation between the channels on neighbouring tones. If the channel is sampled finely enough then neighbouring tones will see almost the same channels (both direct and crosstalk).

Imagine that the tone spacing is fine enough such that $h_k^{n,m} \simeq h_{k+l}^{n,m}$, $0 \leq l \leq L-1$. Consider two points in the rate region, $A = (R_1^a, R_2^a)$ and $B = (R_1^b, R_2^b)$ and their corresponding PSDs $(s_k^{1,a}, s_k^{2,a})$ and $(s_k^{1,b}, s_k^{2,b})$. It is possible to operate at a point $E = (\frac{l}{L}R_1^a + \frac{L-l}{L}R_1^b, \frac{l}{L}R_2^a + \frac{L-l}{L}R_2^b)$ for any $0 \leq l \leq L-1$ as depicted in Fig. 3.2. This is done by setting the PSDs to $(s_k^{1,a}, s_k^{2,a})$ on tones $k \in \{pL+1, \ldots pL+l\}$ for all integer values of p, and to $(s_k^{1,b}, s_k^{2,b})$ on all other tones.

For example, to operate at a point E that is 2/3 between A and B (on the side closer to A), it is required that $l = 2$ and $L = 3$. Thus the PSDs are set to $(s_k^{1,a}, s_k^{2,a})$ on tones $k \in \{1, 2, 4, 5, 7, 8, \ldots, K-1\}$ and to $(s_k^{1,b}, s_k^{2,b})$ on

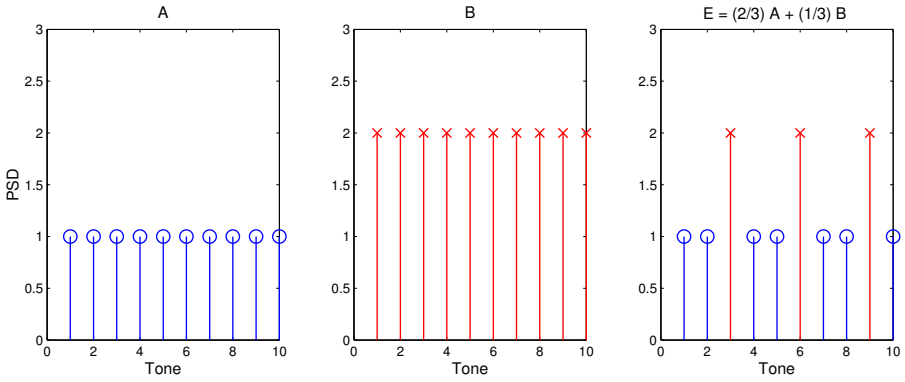

Figure 3.3: Frequency-sharing

tones $k \in \{3, 6, 9, \ldots, K\}$. This is depicted in Fig. 3.3. For this to work the tone spacing must be small enough such that the channel is approximately flat over $L = 3$ neighbouring tones. That is, it is necessary that $h_k^{n,m} \simeq h_{k+1}^{n,m} \simeq h_{k+2}^{n,m}$, $\forall\, k \in \{1, 4, \ldots, K\}$.

For large L (small tone spacing), practically any operating point between A and B can be achieved. Thus for any two points in the rate region, any point between them is also within the rate region. This is the definition of a convex set. As such the rate region is approximately convex in DMT systems with sufficiently small tone spacings. This approximation becomes exact as the inter-tone spacing approaches zero.

In ADSL and VDSL the tone spacing Δ_f is 4.3125 kHz. In both measured and empirical wireline channels this tone-spacing is found to be small enough such that the rate regions are convex or nearly-convex.

It should be made clear that even when the rate region is non-convex, the PSD combinations returned by the optimal spectrum balancing algorithm are provably optimal (see Theorem 3.4), resulting in an operating point on the boundary of the rate region. The convexity of the rate region affects only the ability of the proposed algorithm to explore all optimal operating points. It does *not* affect the optimality of the points found by the algorithm.

Note that one should not confuse convexity of the rate-region with convexity of the objective function (3.11). In practice the rate regions are seen to be nearly-convex, however the optimisation problem is highly non-convex, exhibiting many local maxima. For this reason conventional convex optimisation techniques cannot be applied and an exhaustive search is required on each tone.

3.3.6 Dual Decomposition

In the previous section it was shown that the spectrum management problem (3.2) can be solved through a weighted rate-sum optimization (3.11). It was also shown that in DSL the rate region is approximately convex, allowing almost all optimal operating points to be found. This section will show how the weighted rate-sum can be solved in a computationally tractable way.

The total power constraints (3.3) can be incorporated into the optimization problem by defining the Lagrangian

$$L \triangleq wR_1 + (1-w)R_2 + \lambda_1(P_1 - \sum_k s_k^1) + \lambda_2(P_2 - \sum_k s_k^2). \tag{3.12}$$

Here λ_n is the Lagrangian multiplier for user n and is chosen such that either the power constraint on user n is tight $\sum_k s_k^n = P_n$ or $\lambda_n = 0$[3]. The constrained optimization (3.11) can now be solved via the unconstrained optimization

$$\max_{\mathbf{s}_1,\mathbf{s}_2} L(s_k^1, s_k^2, w, \lambda_1, \lambda_2). \tag{3.13}$$

Define the Lagrangian on tone k

$$L_k \triangleq wb_k^1 + (1-w)b_k^2 - \lambda_1 s_k^1(b_k^1, b_k^2) - \lambda_2 s_k^2(b_k^1, b_k^2).$$

Clearly the Lagrangian (3.12) can be decomposed into a sum across tones of L_k and a term which is independent of s_k^1 and s_k^2

$$L = f_s \sum_k L_k + \lambda_1 P_1 + \lambda_2 P_2.$$

This is known as the dual decomposition. As a result the optimization can be split into K per-tone optimizations that are coupled only through w, λ_1 and λ_2. This will lead to a linear, rather than exponential complexity in K and a computationally tractable search.

3.4 Optimal Spectrum Balancing

3.4.1 2-User Algorithm

The optimal spectrum balancing algorithm is listed as Alg. 3.2. If discrete bitloading is employed then the maximization in the function *optimize_s* is limited to the PSD combinations corresponding to valid bitloading combinations, as calculated by (3.10).

[3]Note that when $\lambda_n = 0$, the power constraint on user n is implicitly enforced through the target rate constraints on the other users. In this case the power constraint need not be explicitly enforced, and the corresponding Lagrangian becomes zero.

The algorithm operates as follows. It is necessary to search through both λ_1 and λ_2 to find values which place sufficient importance on the total power constraint terms within the Lagrangian (3.12). Variation of w makes it possible to map out the optimal, achievable points on the convex hull of the rate region.

The algorithm contains three loops, an outer loop that varies w, an intermediate loop that searches for λ_1 and an inner loop that searches for λ_2. Bisection is used in each search.

When searching for λ_n, it is first necessary to find a value of λ_n which ensures that the power constraint of user n is satisfied. This value is stored in $\lambda_n^{\max}$. Note that a larger λ_n places more emphasis on the power constraint of user n in the Lagrangian. As a result, using a larger λ_n will result in a lower total power for user n.

Once $\lambda_n^{\max}$ is found the algorithm proceeds to bisection. Note that after the algorithm has completed, for each user either $\sum_k s_k^n = P_n$ or the corresponding Lagrangian multiplier is driven to zero, $\lambda_n = 0$. Thus the Lagrangian and the original objective become equivalent. More rigorously,

Theorem 3.4 *For each w Alg. 3.2 returns a PSD combination that is optimal for the spectrum management problem (3.2). That is, for some R_1^{target}*

$$\begin{aligned} \mathbf{s}_1, \mathbf{s}_2 \quad = \quad & \arg\max_{\mathbf{s}_1, \mathbf{s}_2} R_2, \\ \text{s.t.} \quad & R_1 \geq R_1^{\text{target}}, \\ & \sum_k s_k^n \leq P_n, \ \forall n, \\ & 0 \leq s_k^n \leq s_k^{n,\max}, \ \forall n, k. \end{aligned} \tag{3.14}$$

Here R_1^{target} is in fact the rate of user 1 at convergence of the algorithm. Varying w from 0 to 1 allows all optimal operating points that lie on the convex hull of the rate region to be found. If the rate region is convex then all optimal operating points are found.

Proof : See Appendix A. ∎

Note that the cost function on each tone L_k is still non-convex. Hence the optimization of L_k must be solved through exhaustive search, which has an exponential complexity in N. The important observation is that since the optimization on each tone is solved independently the algorithm has a linear, rather than exponential, complexity in K. This results in a computationally tractable algorithm.

If the function *optimize_s* finds multiple PSD combinations that yield the same value for the Lagrangian L_k, then all PSD combinations are stored. This

Algorithm 3.2 Optimal Spectrum Balancing (2 Users)

Main Function

for $w = 0 \ldots 1$
  $\mathbf{s}_1, \mathbf{s}_2 = \text{optimize\_}\lambda_1(w)$
end

Function $\mathbf{s}_1, \mathbf{s}_2 =$ ***optimize***$\_\lambda_1(w)$
$\lambda_1^{\max} = 1$, $\lambda_1^{\min} = 0$
while $\sum_k s_k^1 > P_1$
  $\lambda_1^{\max} = 2\lambda_1^{\max}$
  $\mathbf{s}_1, \mathbf{s}_2 = \text{optimize\_}\lambda_2(w, \lambda_1^{\max})$
end
repeat
  $\lambda_1 = (\lambda_1^{\max} + \lambda_1^{\min})/2$
  $\mathbf{s}_1, \mathbf{s}_2 = \text{optimize\_}\lambda_2(w, \lambda_1)$
  if $\sum_k s_k^1 > P_1$, then $\lambda_1^{\min} = \lambda_1$, else $\lambda_1^{\max} = \lambda_1$
until convergence

Function $\mathbf{s}_1, \mathbf{s}_2 =$ ***optimize***$\_\lambda_2(w,\lambda_1)$
$\lambda_2^{\max} = 1$, $\lambda_2^{\min} = 0$
while $\sum_k s_k^2 > P_2$
  $\lambda_2^{\max} = 2\lambda_2^{\max}$
  $\mathbf{s}_1, \mathbf{s}_2 = \text{optimize\_}s(w, \lambda_1, \lambda_2^{\max})$
end
repeat
  $\lambda_2 = (\lambda_2^{\max} + \lambda_2^{\min})/2$
  $\mathbf{s}_1, \mathbf{s}_2 = \text{optimize\_}s(w, \lambda_1, \lambda_2)$
  if $\sum_k s_k^2 > P_2$, then $\lambda_2^{\min} = \lambda_2$, else $\lambda_2^{\max} = \lambda_2$
until convergence

Function $\mathbf{s}_1, \mathbf{s}_2 =$ ***optimize***$\_s(w, \lambda_1, \lambda_2)$
for $k = 1 \ldots K$
  $s_k^1, s_k^2 = \arg\max_{s_k^1, s_k^2} L_k(s_k^1, s_k^2, w, \lambda_1, \lambda_2)$
    s.t. $0 \le s_k^n \le s_k^{n,\max}$, $\forall n$
  (solve by 2-D exhaustive search)
end

ensures that if multiple points in the rate region are optimal for the same weight w, see for example points A and C in Fig. 3.1, then each of these points is discovered.

3.4.2 N-User Algorithm

In the previous section the optimal spectrum balancing algorithm and optimality proof was only given for 2 user channels. The extension to more than 2 users is now described. In the general case of N users there will be a target rate constraint on the first $N-1$ users and the goal is to maximize the rate of the Nth user. The spectrum management problem is then

$$\begin{aligned} \max_{\mathbf{s}_1,\ldots,\mathbf{s}_N} R_N \quad \text{s.t.} \quad & R_n \geq R_n^{\text{target}},\ \forall n < N; \\ & \sum_k s_k^n \leq P_n,\ \forall n; \\ & 0 \leq s_k^n \leq s_k^{n,\max},\ \forall n,k. \end{aligned} \tag{3.15}$$

Using a similar approach to Theorem 3.2, it can be shown that (3.15) is equivalent to maximizing a weighted rate-sum

$$\begin{aligned} \max_{\mathbf{s}_1,\ldots,\mathbf{s}_N} \sum_n w_n R_n \quad \text{s.t.} \quad & \sum_k s_k^n \leq P_n,\ \forall n; \\ & 0 \leq s_k^n \leq s_k^{n,\max},\ \forall n,k; \end{aligned} \tag{3.16}$$

where the weights $w_1,\ldots,w_{N-1}$ are chosen such that the target rate constraints on users $1,\ldots,N-1$ are met. The weight for the Nth user is arbitrarily defined as $w_N \triangleq 1-\sum_{n=1}^{N-1} w_n$. To enforce the total power constraints on all users, N Lagrangian multipliers are required $\lambda_1,\ldots,\lambda_N$. The constrained optimization (3.16) is solved through the unconstrained optimization

$$\max_{s_k^1,\ldots,s_k^N} L_k \quad \text{s.t.} \quad 0 \leq s_k^n \leq s_k^{n,\max},\ \forall n;$$

where the Lagrangian on tone k is defined

$$L_k(s_k^1,\ldots,s_k^N,w_1,\ldots,w_N,\lambda_1,\ldots,\lambda_N) \triangleq \sum_n w_n b_k^n(s_k^1,\ldots,s_k^N) - \lambda_n s_k^n.$$

In the 2-user case bi-section is only done on λ_1 and λ_2. In the N user case bi-section must be done on $\lambda_1,\ldots,\lambda_N$, which leads to an exponential complexity in the number of users N. It is possible to overcome this exponential complexity in the bisection of λ-space by replacing the bisection with a sub-gradient approach. The resulting algorithm for the N-user case is listed as Alg. 3.3[111][4]. This algorithm is now described.

[4]The sub-gradient descent version of optimal spectrum balancing was developed in collaboration with Prof. Wei Yu, University of Toronto, Canada. Prof. Yu is given primary credit for the development of this version of the algorithm.

The algorithm consists of an inner loop and an outer loop. The inner loop updates the Lagrangian multipliers λ_n and w_n. The update rule for w_n, based on sub-gradient descent, is

$$w_n = \left[w_n + \epsilon \left(R_n^{\text{target}} - \sum_k b_k^n \right) \right]^+ ,$$

where ϵ is an appropriately chosen step-size. Constraints are added to ensure w_n remains positive. One can interpret w_n as the priority given to user n in the optimization. If the data-rate of user n is below its target, then w_n is increased to allocate more priority to user n. The process is repeated until user n achieves its target rate, or $w_n = 0$. Effectively user n chooses the least possible priority w_n required to achieve his target rate, thereby minimizing the disturbance caused to the other modems in the network.

Similarly the update rule for λ_n is

$$\lambda_n = \left[\lambda_n + \epsilon \left(\sum_k s_k^n - P_n \right) \right]^+ .$$

Constraints are added to ensure λ_n remains positive. One can interpret λ_n as the price for power. If user n is below his total power budget, then the price for power is decreased and user n will be allocated more power. The process is repeated until user n reaches his power budget, or $\lambda_n = 0$. This algorithm yields an optimal power allocation, allowing the modems to operate on the boundary of the N-D rate region. The optimality proof for this algorithm follows naturally from the 2-user proof given in Appendix A.

Note that the cost function on each tone L_k is still non-convex. Hence the optimization of L_k must be solved through an N-D exhaustive search, which has exponential complexity in N. So the overall complexity of the algorithm is still exponential in N. Due to the non-convexity of the cost function, it seems difficult to derive an optimal algorithm that does not make use of an exhaustive search on each tone. Hence, if optimality is required, an exponential complexity in N seems unavoidable.

In practice the computing power available at the SMC is often limited, and it may be preferable to find a near-optimal solution with lower complexity. Algorithms can be formulated to do this, based on the insight gained through the optimal spectrum balancing algorithm. These algorithms are described in Section 3.5.

3.4.3 Complexity

This section discusses the complexity of optimal spectrum balancing and shows that a significant complexity reduction can be achieved over the K-tone exhaus-

Algorithm 3.3 Optimal Spectrum Balancing (N users)[111]

repeat
  for each k: $s_k^1, \ldots, s_k^N = \arg\max_{s_k^1,\ldots,s_k^N} L_k$
    (solve by N-D exhaustive search)
  for each n: $w_n = [w_n + \epsilon\,(R_n^{\text{target}} - f_s \sum_k b_k^n)]^+$
  for each n: $\lambda_n = [\lambda_n + \epsilon\,(\sum_k s_k^n - P_n)]^+$
until convergence

tive search described in Section 3.3.4.

We first consider the 2-user case, described by Alg. 3.2. In the outer loops of the algorithm bisection is done on λ_1 and λ_2 such that the power constraints on both users are tight. Assume that an accuracy of ϵ_λ is required in each λ. In the function *optimize_*λ_1 the bisection will require $\log_2(1/\epsilon_\lambda)$ iterations to achieve an accuracy of ϵ_λ in λ_1. Each iteration will result in the function *optimize_*λ_2 being called. Similarly the function *optimize_*λ_2 will require $\log_2(1/\epsilon_\lambda)$ iterations to achieve an accuracy of ϵ_λ in λ_2. Each iteration will result in the function *optimize_s* being called.

The function *optimize_s* solves the weighted rate-sum optimization independently on each tone through exhaustive search. Hence it requires $K(b_{\max}+1)^2$ evaluations of L_k in the 2 user, discrete bitloading case. A similar expression can be written for the continuous bitloading case. So the total complexity of the proposed algorithm is $\mathcal{O}\left(K(b_{\max}+1)^2 \log_2(1/\epsilon_\lambda)^2\right)$. Typically setting ϵ_λ to 1×10^{-10} is sufficient to achieve an accuracy of 1% in the total power constraints (3.3). This leads to a complexity

$$V_{\text{OSB}} = \mathcal{O}\left(K(b_{\max}+1)^2\right).$$

For comparison solving the problem through a exhaustive search across all tones requires the evaluation of $(b_{\max}+1)^{2K}$ bitloading combinations. In most cases this is computationally intractable.

We now consider the N-user case. If bisection is used, the inner loop of the algorithm must be called 33^N times. Since the objective function on each tone is non-convex, the optimization requires $K(b_{\max}+1)^N$ evaluations of L_k in the discrete bitloading case. Evaluating L_k requires a weighted rate-sum of N users and so the total complexity of the algorithm is

$$\mathcal{O}\left(KN(b_{\max}+1)^N 33^N\right). \tag{3.17}$$

For comparison, the exhaustive search across all tones in the N user case requires the evaluation of $(b_{\max}+1)^{NK}$ bitloading combinations. For each bitloading combination the total rate must be calculated for each user across all tones. Hence the exhaustive search has a complexity

$$V_{\text{exhaustive}} = \mathcal{O}\left(KN(b_{\max}+1)^{KN}\right).$$

Algorithm 3.4 Iterative Spectrum Balancing

repeat
  for $n = 1 \dots N$
    repeat
      for each k: fix s_k^m, $\forall m \neq n$, then
        $s_k^n = \arg\max_{s_k^n} L_k$
        (solve by 1-D exhaustive search)
      $w_n = [w_n + \epsilon\,(R_n^{\text{target}} - f_s \sum_k b_k^n)]^+$
      $\lambda_n = [\lambda_n + \epsilon\,(\sum_k s_k^n - P_n)]^+$
    until convergence
  end
until convergence

Comparing this with (3.17) shows that the optimal spectrum balancing algorithm leads to a complexity reduction of

$$\Delta V = \mathcal{O}\left((b_{\max}+1)^{N(K-1)} 33^{-N}\right).$$

The first term $(b_{\max}+1)^{N(K-1)}$ can be interpreted as the benefit of replacing the NK dimensional non-convex optimisation with K separate N dimensional optimisations. The second term 33^{-N} is the penalty of searching through λ-space.

Typically $b_{\max} = 14$. In ADSL $K = 256$ and so the overall complexity reduction with the proposed algorithm is $\mathcal{O}(10^{298N})$. In VDSL $K = 4096$ and the overall complexity reduction is even higher at $\mathcal{O}(10^{4815N})$. To give an example, optimizing a 2 user ADSL system with Alg. 3.2 takes approximately 30 seconds on a modern PC. Using an exhaustive search would take 10^{590} years.

3.5 Iterative Spectrum Balancing

Despite the large reduction in complexity that the optimal spectrum balancing algorithm achieves, at large N it can still be too complex for practical implementation. Changing line conditions and the frequent addition of new users to a binder mean that practical DSM algorithms need to re-optimize the modem spectra in a matter of minutes. In this case it may be more interesting to find low complexity algorithms with near-optimal performance.

In recent work a simplified spectrum balancing algorithm was developed based on the optimal spectrum balancing approach[20]. Like optimum spectrum balancing this algorithm is based on a weighted rate-sum, which allows the selfish-optimum of iterative waterfilling to be avoided. However, the weighted rate-sum optimization is implemented in an iterative fashion leading to a quadratic,

rather than exponential, complexity in N. Whilst this is relatively early work, we include a description of it here for completeness.

In optimal spectrum balancing the transmit PSDs are searched jointly (3.16), which leads to an exponential complexity in N. An alternative approach is to search the PSDs of each user in an iterative fashion. The PSD of each user is updated one at a time. When updating the PSD of user n, the PSDs of all other users are fixed at their present values. The optimization is then

$$\max_{\mathbf{s}_n} \sum_n w_n R_n \quad \text{s.t.} \quad \sum_k s_k^n \leq P_n, \qquad 0 \leq s_k^n \leq s_k^{n,\max}, \ \forall k. \tag{3.18}$$

Note that the optimization is only done over the PSD of a single user. The algorithm iterates through the users, optimizing the PSD of each user in turn. The complete algorithm, which we name *iterative spectrum balancing,* is listed as Alg. 3.4.

The algorithm consists of an outer loop and an inner loop. In the inner loop the PSD of user n is optimized. In a similar fashion to optimal spectrum balancing, the update of each user's PSD is based on a weighted rate-sum, which allows the selfish-optimum of iterative waterfilling to be avoided. However, unlike optimal spectrum balancing, the optimization is only done on the PSD of a single user. So the N-dimensional exhaustive search is replaced by a 1-dimensional exhaustive search. This leads to a complexity which is quadratic, rather than exponential, in N.

This algorithm reduces complexity considerably and yields optimal performance in the broad range of scenarios that we have studied. Unfortunately have no proof for the optimality of iterative spectrum balancing. However we strongly suspect that such a proof exists, and that iterative spectrum balancing is in fact optimal for spectrum management in DSL channels. Perhaps of more practical interest is the fact that iterative spectrum balancing yields significant gains over the current state-of-the-art and is directly applicable to existing modems. Further of iterative spectrum balancing is an important area for future work.

3.6 Performance

This section compares the performance of the proposed algorithms with existing spectrum management techniques. For all simulations the line diameter is 0.5 mm (24-AWG). Direct and crosstalk channel transfer functions are calculated using the empirical models described in Section 2.1.2. The target symbol error probability is 10^{-7} or less. The coding gain and noise margin are set to 3

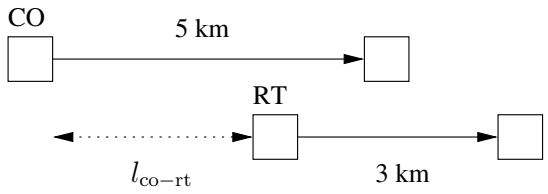

Figure 3.4: Downstream ADSL 2 User Scenario

dB and 6 dB respectively. Continuous bitloading is used and Δ_s is set to 0.1 dBm/Hz. The maximum bitloading is not constrained. As per the ADSL and VDSL standards the tone spacing Δ_f and DMT symbol rate f_s are set to 4.3125 kHz and 4 kHz respectively[9][3][7].

3.6.1 Remote Terminal Distributed ADSL

This section evaluates the performance of optimal spectrum balancing in downstream ADSL. A maximum transmit power of 20.4 dBm is applied to each modem[3]. The usual PSD constraint is not applied in the optimal spectrum balancing or iterative waterfilling algorithms. A spectral mask is applied to the flat PBO and reference noise method and is set at -40 dBm/Hz[3]. Background noise includes crosstalk from 10 ISDN, 4 HDSL, and 10 non-DSM ADSL disturbers which transmit at the spectral mask. In the reference noise method the reference length is set to 5 km and the reference noise to the background noise described above.

2-User Scenario

This scenario consists of a 5 km CO distributed line and a 3 km RT distributed line. The RT is located 4 km from the CO. This is depicted in Fig. 3.4 where $l_{\text{co-rt}} = 4$ km.

Fig. 3.5 shows the rate regions corresponding to the various spectrum management algorithms. For comparison the rate regions with iterative waterfilling, the reference noise method and flat PBO are shown. No PBO method for RT distributed ADSL modems has been defined in standardization and this is still an open issue[6]. A method for reducing the downstream transmit PSD known as the power cutback method is currently implemented in ADSL modems to prevent the receiver ADCs from being overloaded[3]. However on the 3 km RT distributed line this technique does not cause any reduction in the transmit PSD.

The PSDs corresponding to a 1 Mbps service on the CO distributed line are

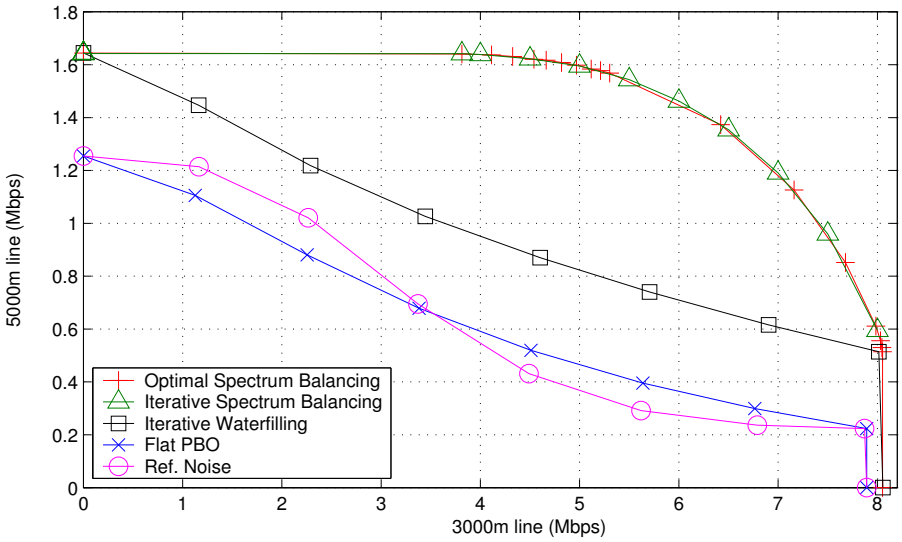

Figure 3.5: Rate Regions in Downstream ADSL

depicted in Fig. 3.6 and Fig. 3.7. The optimal PSD on the RT line decreases with frequency to reflect the increase in crosstalk coupling. This continues until 440 kHz where the CO line becomes inactive due to its low channel SNR above this frequency. Once the CO line becomes inactive a sudden increase in the optimal PSD on the RT line can be observed.

With the flat PBO algorithm, the RT line must employ a large amount of PBO to protect the CO line. This occurs because, unlike in optimal spectrum balancing, the flat PBO algorithm cannot vary the degree of PBO with frequency.

The iterative waterfilling algorithm gives similar results. Slightly less PBO is required since the CO line PSD has been boosted on the active tones as shown in Fig. 3.6. However the amount of PBO required is still much larger than with optimal spectrum balancing. The iterative waterfilling algorithm does not exploit the fact that crosstalk coupling is low at low frequencies. It also does not exploit the fact that the CO line is inactive above 440 kHz. Both of these facts could have been used to increase the transmit PSD on the RT line at low and high frequencies, leading to increased performance without a large degradation in the performance of the CO line. Due to this the iterative waterfilling algorithm gives poor performance.

It has been shown that the reference noise method is near-optimal when the SINR is high[87]. This is the case in low frequencies. For this reason the reference noise PSD matches the optimal PSD quite closely at frequencies below

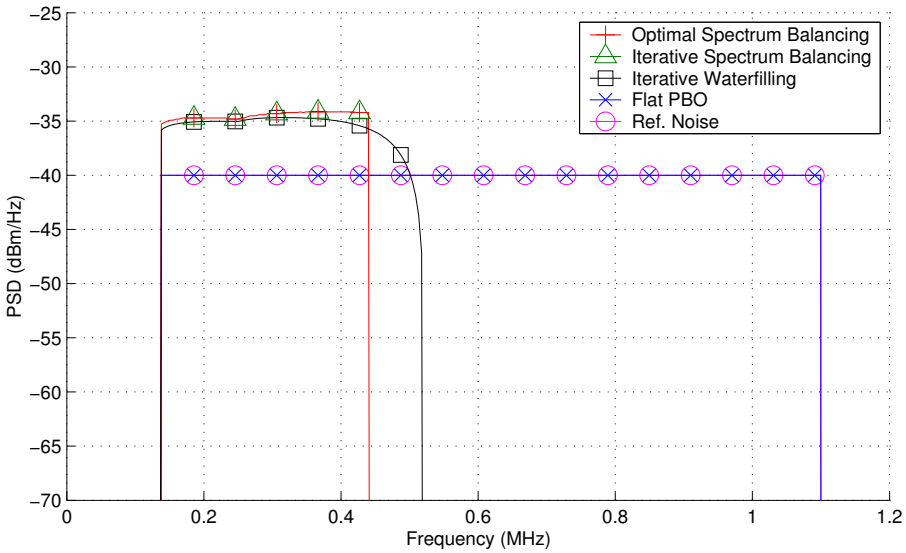

Figure 3.6: PSDs on CO line in Downstream ADSL (CO Line @ 1 Mbps)

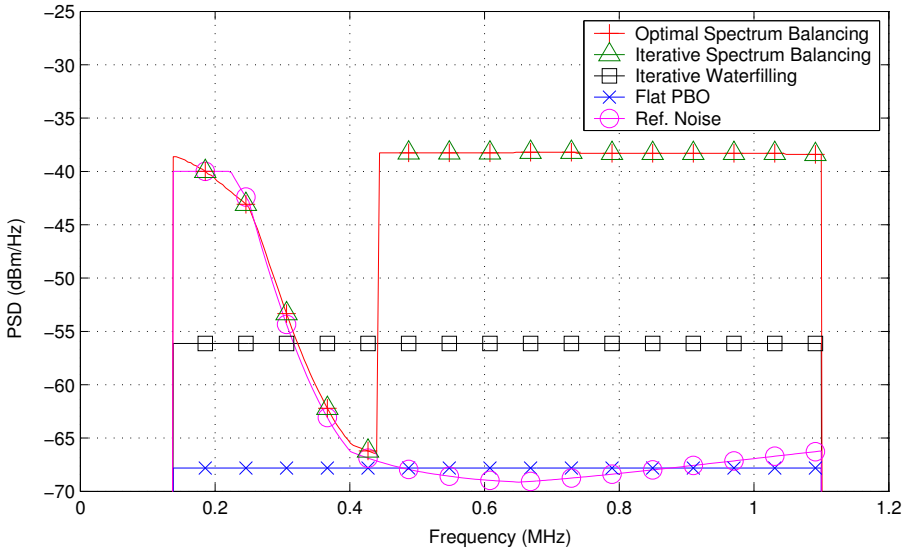

Figure 3.7: PSDs on RT line in Downstream ADSL (CO Line @ 1 Mbps)

Table 3.1: Achievable Rates in Downstream ADSL

Scheme	CO Rate	RT Rate
Flat PBO	1.0 Mbps	1.7 Mbps
Ref. Noise	1.0 Mbps	2.3 Mbps
Iterative Waterfilling	1.0 Mbps	3.6 Mbps
Iterative Spectrum Balancing	1.0 Mbps	7.4 Mbps
Optimal Spectrum Balancing	1.0 Mbps	7.4 Mbps

440 kHz. This could be exploited to create a low-complexity, near-optimal spectrum balancing algorithm.

As shown in Tab. 3.1 using optimal spectrum balancing instead of iterative waterfilling allows the data-rate on the RT distributed line to be increased from 3.6 Mbps to 7.4 Mbps whilst still maintaining a 1 Mbps service on the CO distributed line. So the data-rate is doubled through the use of optimal spectrum balancing.

Iterative spectrum balancing yields identical spectra and rate regions to optimal spectrum balancing, resulting in optimal performance. After simulating iterative spectrum balancing in a broad range of scenarios, it appears to be near-optimal in general. A detailed study of why iterative spectrum balancing yields near-optimal performance is an important area for future work. We postulate that this is due to the hierarchal structure of crosstalk, by which we mean that far-end users cause negligible crosstalk to near-end users. For example, in this scenario the RT causes significant crosstalk to the CO, but the CO has negligible impact on the RT. This appears to enable an iterative, user-by-user line-search to converge to the globally optimal solution.

To investigate the performance of optimal spectrum balancing in a broad range of scenarios we varied the distance from the CO to the RT $l_{\mathrm{co-rt}}$. Simulations were run with values of $l_{\mathrm{co-rt}}$ from 0 km to 5 km in 100 m increments. The target data-rate on the RT line was set to 7 Mbps. The resulting rate on the CO line with the different algorithms is then shown in Fig. 3.8. As can be seen, optimal spectrum balancing leads to a significant increase in data-rate over a broad range of scenarios.

4 User Scenario

This scenario consists of one 5 km CO distributed line, and 3 RT distributed lines: RT1, RT2 and RT3 as depicted in Fig. 3.9. The RTs are located at 2 km, 3 km and 4 km from the CO respectively. The corresponding line lengths are 4 km, 3.5 km and 3 km.

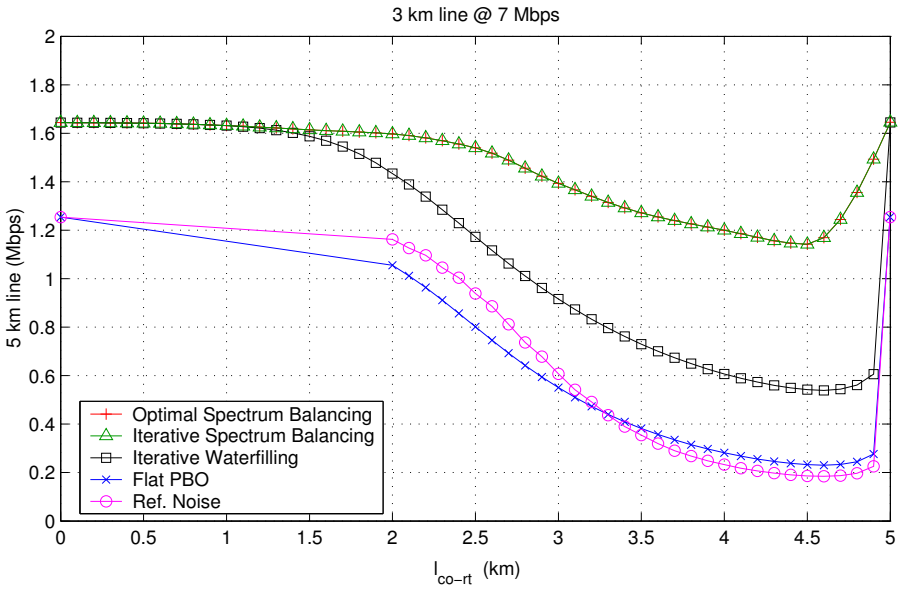

Figure 3.8: Data-rate on 5 km CO line vs. $l_{\mathrm{co-rt}}$

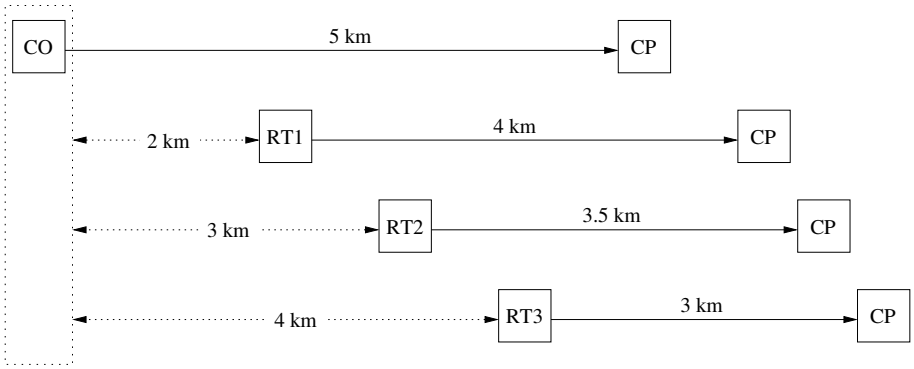

Figure 3.9: Downstream ADSL 4 User Scenario

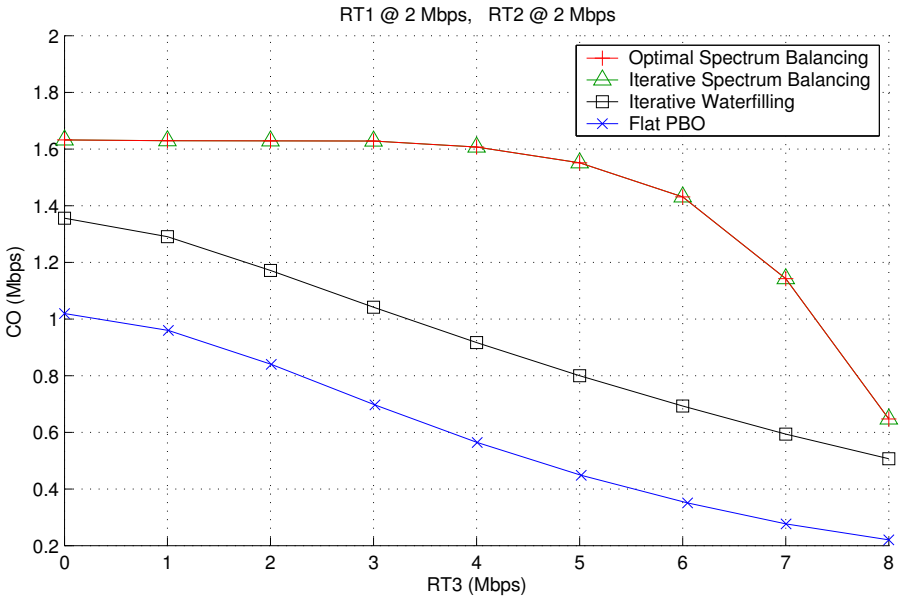

Figure 3.10: Rate Region - 4 User Scenario

The target rates on RT1 and RT2 have both been set to 2 Mbps. For a variety of different target rates on RT3, the CO attempted to maximize its own data-rate either by transmitting at full power in iterative waterfilling, or by setting its corresponding weight to unity in iterative spectrum balancing and optimal spectrum balancing. This produced the rate regions shown in Fig. 3.10.

The rate regions in Fig. 3.10 show the substantial gains that optimal spectrum balancing achieves over iterative waterfilling. For example, consider the case when a minimum service of 1 Mbps must be provided to the CO line. Fig. 3.10 shows that with iterative waterfilling the maximum achievable rate on RT3 is then 3.3 Mbps. Compare this with optimal spectrum balancing where the rate on RT3 can be increased to 7.3 Mbps whilst still maintaining 1 Mbps on the CO line. So the achievable rate on RT3 can be doubled through the use of optimal spectrum balancing. These results are summarized in Tab. 3.2.

The corresponding PSDs are shown in Fig. 3.11 for iterative waterfilling and Fig. 3.12 for optimal spectrum balancing. The PSDs from iterative spectrum balancing are not shown since they are nearly identical to those from optimal spectrum balancing. Note that with iterative waterfilling the PBO on the RTs is flat with frequency. Contrast this with iterative spectrum balancing where the PBO varies dramatically with frequency. Crosstalk coupling is minimal at low frequencies so with optimal spectrum balancing the RTs transmit at full

Method	RT3 (Mbps)	CO (Mbps)
Flat PBO	0.3	1.0
Iterative Waterfilling	3.3	1.0
Iterative Spectrum Balancing	7.3	1.0
Optimal Spectrum Balancing	7.3	1.0

Table 3.2: Rate Comparison - 4 User Scenario (RT1-2 @ 2 Mbps)

power on the lower tones. As frequency increases the RTs reduce their power to protect the CO. The level of PBO increases with the nearness of an RT's transmitter to the receiver of the CO line. At 430 kHz the CO line becomes inactive due to poor channel-SNR. Above this frequency the CO line no longer needs to be protected and the PSDs of the RTs increase abruptly. RT3 still does some PBO to protect RT1. At 750 kHz RT1 becomes inactive due to poor channel-SNR on its line. As a result the PSD on RT3 increases again.

It should be clear that optimal performance requires PBO that varies with frequency. Optimal spectrum balancing adapts the transmit spectra to match the crosstalk coupling strength and the type of active users on each particular tone. This leads to a large performance gain over iterative waterfilling, which can only implement frequency flat PBO.

Note that, as the iterative spectrum balancing and optimal spectrum balancing rate region coincide in Fig. 3.10, iterative spectrum balancing gives close to optimal performance in this scenario. After simulating iterative spectrum balancing in a broad range of scenarios, it appears to be near-optimal in general. A detailed study of why iterative spectrum balancing yields near-optimal performance is an important area for future work. We postulate that this is due to the hierarchal structure of crosstalk, by which we mean that far-end users do not cause substantial crosstalk to near-end users. For example, in this scenario the CO causes significant interference to no-one, and RT n only causes significant interference to the CO and RT m, $\forall m < n$. This appears to enable the iterative, user-by-user line-search of iterative spectrum balancing to converge to the globally optimal solution.

3.6.2 Near-far Problem in VDSL

Upstream VDSL transmission is simulated with 4×600m lines and 4×1200m lines as depicted in Fig. 3.13. Each modem has a maximum transmit power of 11.5 dBm available. The usual PSD constraint is not applied in the optimal spectrum balancing or iterative waterfilling algorithms. A spectral mask is applied to the flat PBO, reference noise and reference PSD methods and is set at -60 dBm/Hz[7]. Alien crosstalk is incorporated into the background noise using ETSI model A. FDD bandplan 998 is used with the frequency bands

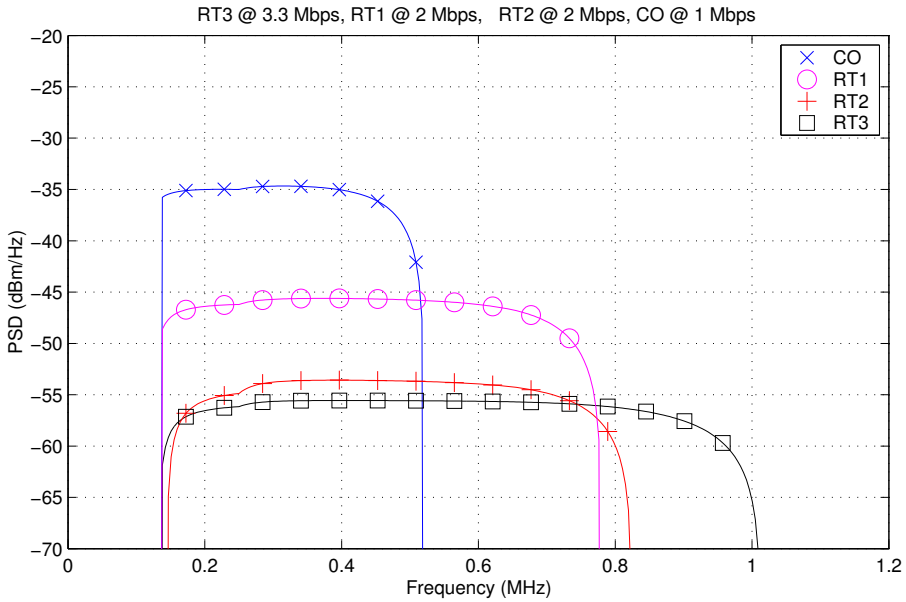

Figure 3.11: Iterative Waterfilling PSDs

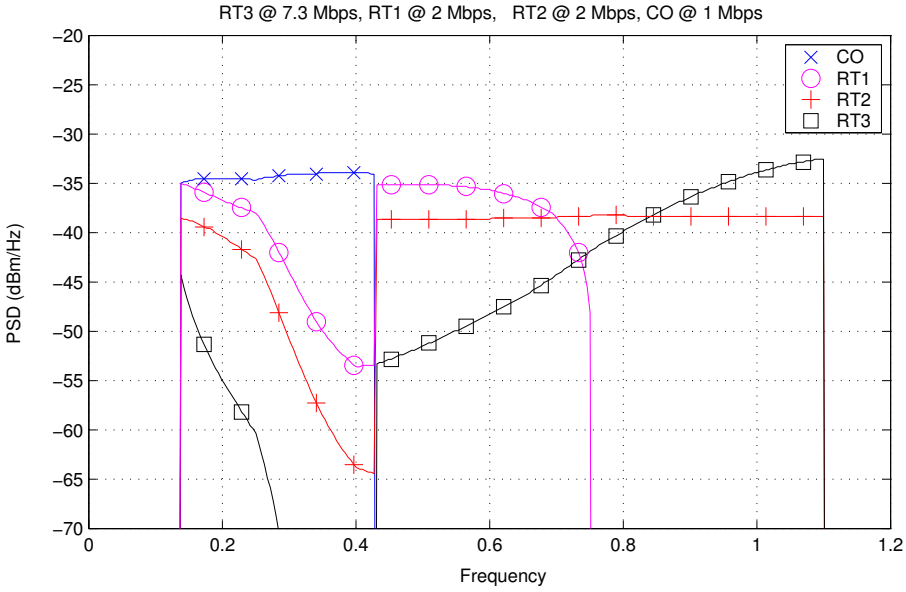

Figure 3.12: Optimal Spectrum Balancing PSDs

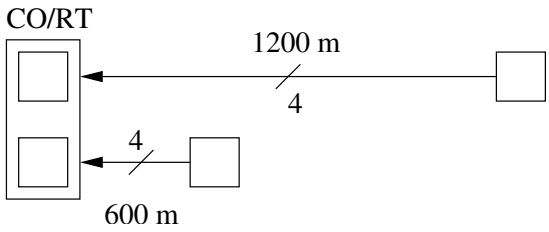

Figure 3.13: Upstream VDSL Scenario

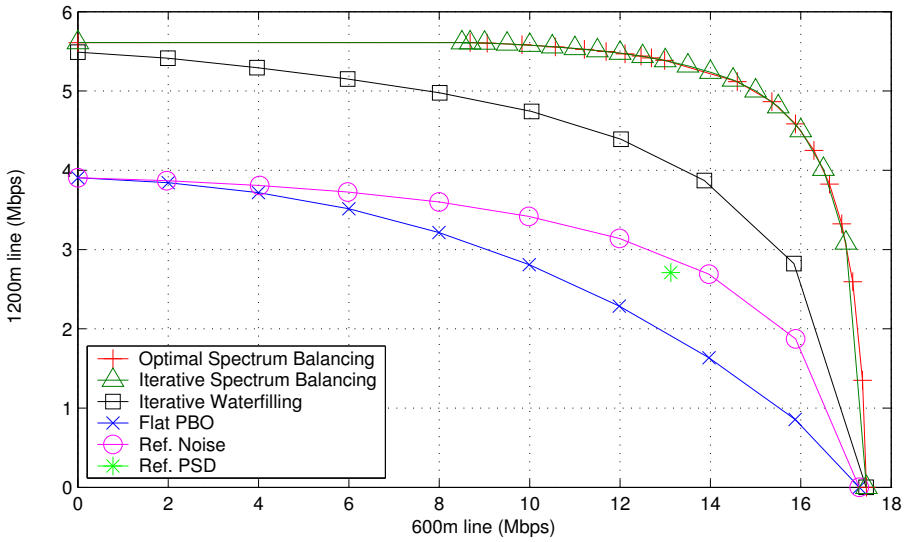

Figure 3.14: Rate Regions in Upstream VDSL

corresponding to amateur radio frequencies notched off. For more details see [7]. In the reference noise method the reference length was set to 1200 m and the reference noise to ETSI model A.

Fig. 3.14 shows the rate regions corresponding to various spectrum management algorithms. Included are flat PBO, iterative waterfilling, the reference noise method and the reference PSD method, which is currently adopted in VDSL standards[9][7].

The PSDs corresponding to a 5 Mbps service on the 1200 m lines are depicted in Fig. 3.15 and Fig. 3.16. Under the 998 FDD bandplan there are two separate upstream bands: 3.75 - 5.2 MHz and 8.5 - 12 MHz. CP modems may not use frequencies between these bands since they are reserved for downstream

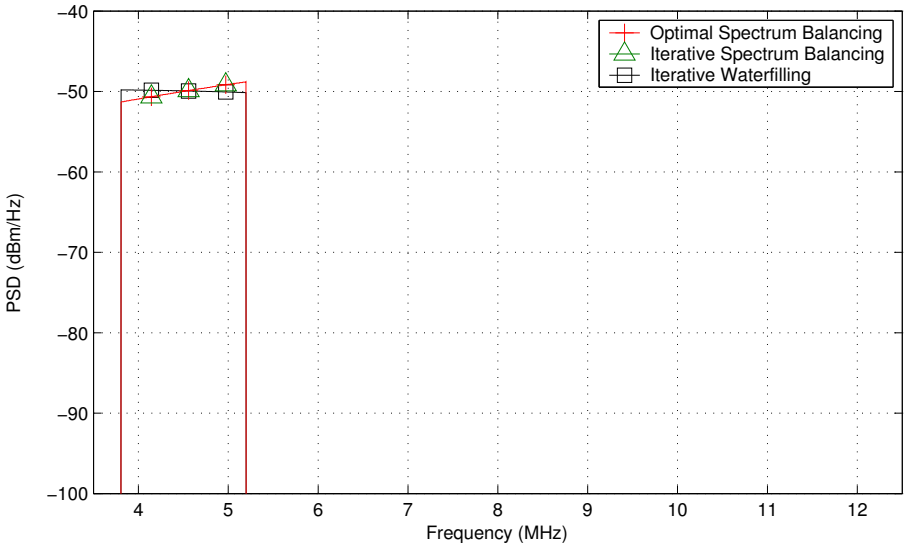

Figure 3.15: PSD on 1200m lines in Upstream VDSL (1200m line @ 5 Mbps)

transmission by the CO modems.

The optimal PSD of the 600 m lines in the first upstream band is quite flat with some power back-off (PBO) applied. In the second upstream band PBO is not required since the 1200 m lines are not active. There the optimal PSD on the 600 m lines increases with frequency as the crosstalk coupling between the four different 600 m lines rises.

The 600 m lines see a relatively flat channel up to 10 MHz. So with iterative waterfilling they transmit with almost the same PSD of -80 dBm/Hz in the first band and in the second band below 10 MHz. The crosstalk coupling from the 1200 m lines into the 600 m lines is minimal due to the 600 m of attenuation the signals from the 1200 m lines experience before coupling begins. As such the PSD of the 1200 m lines does not affect the PSD adopted by the 600 m lines. As with optimal spectrum balancing, iterative waterfilling requires a large amount of PBO in the first band to ensure the desired service rate for the 1200 m lines. Unfortunately the transmit PSD has now also been decreased in the second transmission band. This is unnecessary since the 1200 m lines are inactive in the second band. The rate of the 600 m lines is decreased without benefiting the 1200 m lines and this leads to inferior performance.

As shown in Tab. 3.3, using optimal spectrum balancing instead of iterative waterfilling allows the data-rate on the 600 m lines to be increased from 7.7 Mbps to 15 Mbps whilst still maintaining a 5 Mbps service on the 1200 m

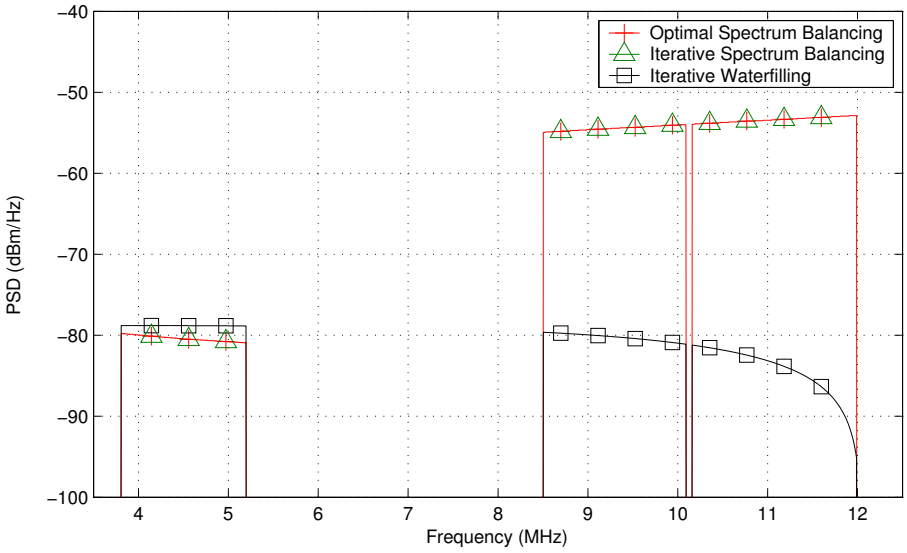

Figure 3.16: PSD on 600m lines in Upstream VDSL (1200m Line @ 5 Mbps)

lines. Again the data-rate is approximately doubled through the use of optimal spectrum balancing.

Iterative spectrum balancing yields identical spectra and rate regions to optimal spectrum balancing, resulting in optimal performance.

Note the optimal rate regions for both the ADSL and VDSL scenarios are convex as was predicted in Section 3.3.5.

Table 3.3: Achievable Rates in Upstream VDSL

Scheme	1200 m Rate	600 m Rate
Ref. PSD	2.7 Mbps	13.1 Mbps
Flat PBO	3.9 Mbps	0.0 Mbps
Ref. Noise	3.9 Mbps	0.0 Mbps
Iterative Waterfilling	5.0 Mbps	7.7 Mbps
Iterative Spectrum Balancing	5.0 Mbps	15.0 Mbps
Optimal Spectrum Balancing	5.0 Mbps	15.0 Mbps

3.6.3 Discrete Bitloading

The same simulations are made with discrete bitloading, with each modem forced to adopt an integer bitloading value. The maximum bitloading $b_{\max}$ is set to 14. All other simulation parameters are the same. In the iterative waterfilling algorithm the Levin-Campello algorithm is used to ensure integer bitloadings on each tone[15].

In the ADSL scenario using optimal spectrum balancing, instead of iterative waterfilling, allows the data-rate on the RT distributed line to be increased from 3.1 Mbps to 7.3 Mbps, whilst still maintaining a 1 Mbps service on the CO distributed line.

In the VDSL scenario using optimal spectrum balancing, instead of iterative waterfilling, allows the data-rate on the 600 m lines to be increased from 3.4 Mbps to 13 Mbps, whilst still maintaining a 5 Mbps service on the 1200 m lines.

Iterative spectrum balancing yields identical spectra and rate regions to optimal spectrum balancing, resulting in optimal performance.

3.7 Summary

This chapter presented a centralized algorithm for optimal spectrum balancing in DSL. The algorithm calculates the spectra for the modems within a network to achieve optimal performance, thereby operating on the rate region boundary. The algorithm can operate under a combination of total power and spectral mask constraints, and can use either continuous or discrete bitloading.

Through the use of a dual decomposition, the inner loop of the proposed algorithm solves the spectrum management problem independently on each tone. The result is a computationally tractable and efficient algorithm. Simulations show that the proposed algorithm yields significant gains over existing spectrum management techniques, typically doubling the achievable data-rate.

Whilst this chapter has focused on the problem of spectrum management in DSL, the algorithm is also applicable to any communication system where inter-user interference is a problem. Optimal spectrum balancing could also be applied to broadband cable networks, high-speed Ethernets or fixed wireless links.

A patent has been filed on this material[32]. Optimal spectrum balancing is now part of the draft ANSI standard on Dynamic Spectrum Management[8].

Part II

Multi-user Signal Coordination

Overview

Crosstalk is a major problem in modern DSL systems such as ADSL and VDSL. Crosstalk can be mitigated through the coordination of DSL modems, which can be done either on a spectral or signal level. Spectral coordination was discussed in Part I of this thesis. Signal coordination is discussed in this part.

To facilitate signal coordination either the transmitting or receiving modems must be co-located. So, unlike spectral coordination, signal coordination cannot be applied in mixed CO/RT scenarios. Furthermore, signal coordination has a higher run-time complexity than conventional DSL modems, which increases production cost.

Nevertheless signal coordination gives superior performance, allowing DSL to be delivered at higher speeds and to a larger number of customers. The increased revenue compensates for higher production costs and will make signal coordination an important technology in the medium term.

In the US, signal coordination is used between co-located CO receivers. Reception is done in a joint fashion; the signals received on each line are combined to cancel crosstalk whilst preserving the signal of interest.

Chapter 4 discusses crosstalk canceler design. Existing techniques are based on decision feedback between the different users within the binder. To prevent error propagation decoding must be done before decisions are fed back, which leads to a high computational complexity and latency. To address this problem, a simple linear canceler is presented based on the well known ZF criterion. This technique has a low complexity and latency. It is shown that, due to a special property of upstream DSL channels, this design operates close to the theoretical channel capacity. A low complexity algorithm is proposed for spectra optimization when crosstalk cancellation is employed.

In the downstream, signal coordination is used between co-located CO transmitters. Transmission is done in a joint fashion; predistortion is introduced into the signal of each user prior to transmission. This predistortion is chosen such that it annihilates with the crosstalk introduced in the channel. As a

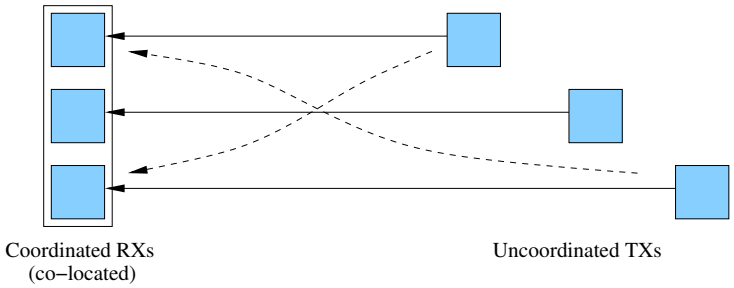

Figure 3.17: Co-located Receivers

result the customer premises (CP) modems receive a crosstalk free signal.

This technique, known as crosstalk precoding, is discussed in Chapter 5. Existing precoder designs lead either to poor performance or require the replacement of CP modems. Millions of CP modems are currently in use, owned and operated by a multitude of customers. Replacing these modems presents a huge legacy issue. To address this problem a simple linear precoder is presented based on a channel diagonalizing criterion. The precoder has a low complexity and works with existing CP modems. It is shown that, due to a special property of downstream DSL channels, this design operates close to the theoretical channel capacity. A low complexity algorithm is proposed for spectra optimization when crosstalk precoding is employed.

As a by-product, the work in Chapters 4 and 5 produced as set of bounds on the determinants and inverses of diagonally dominant matrices. These are listed in Appendix B.

Despite the low complexity of the techniques presented in Chapters 4 and 5, signal coordination still requires a much higher complexity than is available in existing DSL modems. Crosstalk cancellation and precoding have a complexity that scales quadratically with the number of lines within a binder. For typical binders, which contain anywhere from 20 to 100 lines, these techniques are outside the scope of present day implementation and may remain so for several years. Chapter 6 addresses this problem through a technique known as partial cancellation.

It is well known that the majority of crosstalk experienced on a line comes from the 3 to 4 surrounding pairs in the binder. Furthermore, since crosstalk coupling varies dramatically with frequency, the worst effects of crosstalk are limited to a small selection of tones. Partial cancelers exploit these facts to achieve the majority of the performance of full cancellation at a fraction of the complexity. Whilst the idea of partial cancellation has been discussed in literature, no work has specifically focused on partial canceler design.

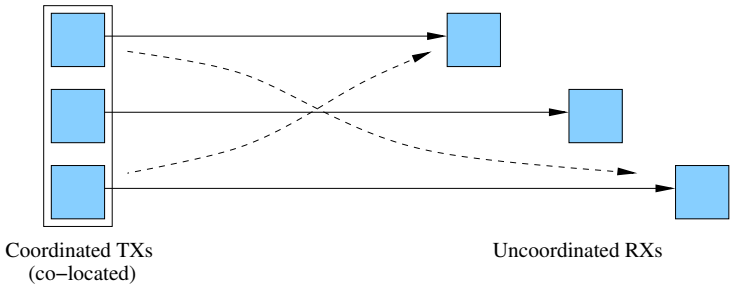

Figure 3.18: Co-located Transmitters

Chapter 6 investigates partial canceler and precoder design, which is in essence a problem of resource allocation. Given a limited amount of available run-time complexity, a modem must distribute this across lines and tones such that the data-rate is maximized. Chapter 6 presents the optimal algorithm for partial canceler design and several simpler, sub-optimal algorithms. These algorithms are shown to achieve 90% of the data-rate of full cancellation at less than 30% of the complexity.

The techniques developed in this part assume knowledge of the direct and crosstalk channels in the network. This requires the use of multi-user channel identification techniques. When CO modems are co-located this is straightforward to implement in practice and requires only that the training sequences of the modems are mutually orthogonal. For a detailed study of multi-user channel identification in a DSL context see [112, 11, 57].

This material in Chapter 4 has been published as [34, 22, 28, 23] and submitted for publication as [18]. This material in Chapter 5 has been published as [17, 29], submitted for publication as [19] and submitted to standardization as [33]. The material in Chapter 6 has been published as [27, 25, 24, 26] and has been patented by Alcatel[30, 31].

Chapter 4

Receiver Coordination

4.1 Introduction

In conventional DSL modems the transmitted symbol is estimated based on the received signal on the corresponding line[1]

$$\hat{x}_k^n = f_{\text{su-rx}}\left(y_k^n\right),$$

where $f_{\text{su-rx}}(\cdot)$ denotes the single-user receiver operation. Crosstalk is treated as background noise; it decreases the receiver-side SNR and limits the achievable bitloading to

$$b_k^n = \log_2\left(1 + \frac{1}{\Gamma}\frac{\left|h_k^{n,n}\right|^2 s_k^n}{\sum_{m\neq n}\left|h_k^{n,m}\right|^2 s_k^m + \sigma_k}\right).$$

Part I of this thesis investigated ways of intelligently setting the transmit spectra s_k^n of the modems in an attempt to minimize the effects of crosstalk. These techniques were referred to as spectral coordination since they optimize the spectra of the modems in the network.

In upstream transmission the receiving modems are often co-located in a common *central office* (CO). This makes possible a second, higher level of coordination, whereby the modems coordinate not just their spectra, but their signals as well. Essentially a DSL modem now estimates the transmitted signal based on the received signals on every line

$$\hat{x}_k^n = f_{\text{mu-rx}}\left(\mathbf{y}_k\right), \tag{4.1}$$

[1]The work in this chapter was done in close collaboration with Dr. George Ginis, Texas Instruments, San Jose, CA.

where $f_{\mathrm{mu-rx}}(\cdot)$ denotes the multi-user receiver operation and

$$\mathbf{y}_k \triangleq [y_k^1, \ldots, y_k^N]^T.$$

Such receiver coordination makes it possible to filter out the crosstalk on each line. Afterwards each modem sees only the signal of interest and the filtered background noise. As this chapter will show, since the DSL channel is well-conditioned, in the spatially white noise case it is possible to achieve a data-rate very close to the single-user bound

$$b_k^n = \log_2\left(1 + \frac{1}{\Gamma}\frac{\|\mathbf{h}_k^n\|^2 s_k^n}{\sigma_k}\right),$$

where the nth column of $\mathbf{H}_k$ is defined $\mathbf{h}_k^n \triangleq [\mathbf{H}_k]_{\mathrm{col}\,n}$.

Several crosstalk canceler designs have been proposed. A decision feedback structure was shown to achieve close to the theoretical channel capacity[59] and is described in more detail in Section 4.4. Unfortunately this structure suffers from error propagation. To minimize the effects of error propagation each user's data-stream must be decoded before decisions are fed back. This leads to a high computational complexity and a latency that grows with the number of users in the binder. Binders can contain hundreds of lines. As a result, this design is inapplicable in real-time applications such as voice over IP or video conferencing.

A simpler crosstalk canceler design is the linear *zero-forcing* (ZF) canceler[92], which is described in Section 4.5. This design has a low complexity, no latency and does not suffer from error propagation. Furthermore since it is based on a ZF criterion it removes all crosstalk. Despite these advantages it is well known that ZF criteria can lead to severe noise enhancement in ill-conditioned channels.

Section 4.5 analyzes the performance of the linear ZF canceler in a DSL environment. It is shown that, due to the well conditioned structure of the DSL channel matrix, ZF designs lead to negligible noise enhancement. Section 4.5 derives bounds to show that the linear ZF canceler operates close to the single-user bound. These bounds allow performance of the linear ZF canceler to be predicted without the need for explicit knowledge of the crosstalk channels, which simplifies service provisioning significantly. These bounds are a major contribution of this chapter.

Alternative cancellation techniques have also been proposed that use turbo coding principles to facilitate cancellation[41, 113, 49]. Other techniques exploit the cyclostationarity of crosstalk[74, 10]. The advantage of these methods is that they do not require signal coordination, and can instead be applied independently on each modem. Unfortunately these techniques are extremely complex and give poor performance when more than one crosstalker exists.

Other techniques use joint linear processing at both the transmit and receive side of the link[71, 70, 90]. This requires co-location of both CO and *customer premises* (CP) modems, which is typically not the case since different customers are usually situated at different locations. Furthermore, it has been shown that the theoretical channel capacity is achievable with receiver-side coordination only, so using coordination on both ends of the link does not improve performance[105].

The rest of this chapter is organized as follows. The system model for a network of DSL modems transmitting to a single CO is given in Section 4.2. A property of the upstream DSL channel, known as *column-wise diagonal dominance* (CWDD), is explored. As described in Section 4.3, from an information theoretical perspective, the upstream DSL channel is a *multi-access channel* (MAC). This allows the single-user bound from Section 2.2.3 to be applied to bound the capacity of the channel. Section 4.4 describes the multi-user *decision feedback canceler* (DFC) and the problems it has with error propagation, high complexity and latency.

To address the problems of the DFC, Section 4.5 describes a much simpler linear design that has a low complexity, no latency and is free from error propagation. Section 4.5 uses the CWDD property to formulate a lower bound on the performance of the linear canceler. This bound shows that the linear canceler operates close to the single-user bound. Section 4.6 describes power loading algorithms for use with the linear canceler. This can be seen as the combination of spectral and signal coordination for the upstream DSL channel. Existing power loading algorithms for the MAC are extremely complex, having a polynomial complexity in the number of lines and tones. Application of the linear canceler decouples the power allocation problem between lines, and this simplifies power allocation significantly. The PSD for each line can then be found through a low-complexity waterfilling procedure. Section 4.7 compares the performance of the different cancelers.

4.2 System Model and CWDD

For crosstalk cancellation to be applied, the receiving modems must be co-located at a common CO. This makes it straightforward to synchronize the modems, and transmission can be modeled independently on each tone, as described in Chapter 2,

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k.$$

Typically the noise is spatially white, and we make this assumption here

$$\mathcal{E}\left\{\mathbf{z}_k \mathbf{z}_k^H\right\} = \sigma_k \mathbf{I}_N. \tag{4.2}$$

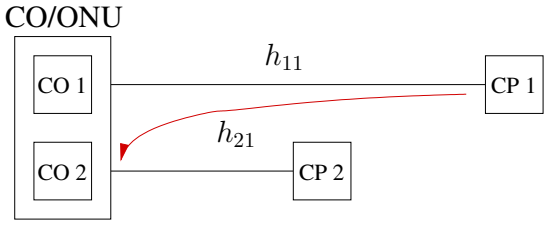

Figure 4.1: Column-wise Diagonal Dominance $|h_{11}| \gg |h_{21}|$

Since the receiving modems are co-located, the crosstalk signal transmitted from a disturber into a victim must propagate through the full length of the disturber's line. This is depicted in Fig. 4.1, where CP 1 is the disturber and CO 2 is the victim. The shielding between twisted pairs increases the attenuation. As a result, the crosstalk channel matrix $\mathbf{H}_k$ is *column-wise diagonally dominant* (CWDD), since on each column of $\mathbf{H}_k$ the diagonal element has the largest magnitude

$$|h_k^{m,m}| \gg |h_k^{n,m}|\,,\ \forall m \neq n. \tag{4.3}$$

CWDD implies that the crosstalk channel $h_k^{n,m}$ from a disturber m into a victim n is always weaker than the direct channel of the disturber $h_k^{m,m}$. The degree of CWDD can be characterized with the parameter α_k

$$|h_k^{n,m}| \leq \alpha_k\,|h_k^{m,m}|\,,\ \forall m \neq n. \tag{4.4}$$

Note that crosstalk cancellation is based on joint reception. As such it requires the co-location of receiving modems. So in all channels where crosstalk cancellation can be applied, the CWDD property holds. CWDD has been verified through extensive measurement campaigns of real binders. In 99% of lines α_k is bounded

$$\alpha_k \leq K_{\text{xf}} f_k \sqrt{d_{\text{coupling}}},$$

where $K_{\text{xf}} = -22.5$ dB and f_k is the frequency on tone k in MHz[7]. Here d_{coupling} is the coupling length between the disturber and the victim in kilometers, as defined in (2.14). The coupling length can be upper bounded by the longest line length in the binder. Hence

$$\alpha_k \leq K_{\text{xf}} f_k \sqrt{l_{\max}}, \tag{4.5}$$

where $l_{\max}$ denotes the length of the longest line in the binder. To find a value for α_k that is independent of the particular binder configuration, $l_{\max}$ can be set to 1.2 km, which is the maximum deployment length for DSL. On typical lines α_k is less than -11.3 dB. The following sections show that CWDD ensures a well-conditioned crosstalk channel matrix. This results in the near-optimality of the linear ZF canceler.

4.3 Theoretical Capacity

We start with a bound on the theoretical capacity of the upstream DSL channel with coordinated receivers. This will prove useful in evaluating crosstalk canceler performance since it provides an upper bound on the achievable data-rate with any possible crosstalk cancellation scheme.

Theorem 4.1 *With spatially white background noise the capacity of user n with a fixed transmit spectrum s_k^n is upper bounded*

$$R_n \leq f_s \sum_k b_{k,\text{mac-dsl}}^n \tag{4.6}$$

where

$$b_{k,\text{mac-dsl}}^n \triangleq \log_2 \left(1 + \sigma_k^{-1} s_k^n \Gamma^{-1} \left|h_k^{n,n}\right|^2 \left[1 + (N-1)\alpha_k^2\right]\right).$$

Proof: CO modems are co-located and do reception in a joint fashion, so from an information theoretical perspective this is a multi-access channel. The single-user bound developed in Section 2.2.3 applies and limits the achievable bitloading of user n on tone k

$$b_k^n \leq \log_2 \left(1 + \sigma_k^{-1} s_k^n \Gamma^{-1} \left\|\mathbf{h}_k^n\right\|_2^2\right).$$

Here the SNR-gap to capacity Γ accounts for the sub-optimality of practical coding schemes. The CWDD property (4.4) leads to the bound

$$\begin{aligned} \left\|\mathbf{h}_k^n\right\|_2^2 &= \left|h_k^{n,n}\right|^2 + \sum_{m \neq n} \left|h_k^{m,n}\right|^2, \\ &\leq \left|h_k^{n,n}\right|^2 \left[1 + \alpha_k^2 (N-1)\right], \end{aligned}$$

which leads to (4.6). ■

Multi-user techniques are often used in wireless systems and lead to large increases in the *signal power* at the receiver. The observation is that if the path from transmit antenna n to receive antenna n is weak, then the path from transmit antenna n to receive antenna m might be strong. The result is a statistical averaging across spatial dimensions, an effect known as spatial diversity, which leads to large improvements in performance[54].

In DSL channels there is, unfortunately, no equivalent to spatial diversity. This can be seen in equation (4.6). Here the CWDD of $\mathbf{H}_k$ implies that very little increase can be made in the signal power through the use of multiple receivers. This is the case since the channel from transmitter n to receiver m is much

weaker than the direct channel from transmitter n to receiver n. Note that the benefit, although small, increases with the crosstalk channel strength α_k and the number of crosstalkers N.

Although spatial diversity is negligible, the use of co-ordinated reception is by no means fruitless. Instead of benefiting through spatial diversity, the primary benefit in DSL channels is crosstalk cancellation. That is, co-ordinated reception does not increase signal power in DSL, but instead decreases *interference power*.

4.4 Decision Feedback Canceler

Decision feedback equalizers are traditionally used for cancelling *inter-symbol interference* (ISI) in frequency selective channels. In a multi-user context the same principle can be applied to remove *inter-user interference*, otherwise known as crosstalk. In this case the decision feedback operates across users rather than time[59, 61, 60].

The structure of the *decision feedback canceler* (DFC) is now described. Consider the QR decomposition of the crosstalk channel matrix

$$\mathbf{H}_k \stackrel{\text{qr}}{=} \mathbf{Q}_k \mathbf{R}_k, \tag{4.7}$$

where $\mathbf{Q}_k$ is a unitary matrix and $\mathbf{R}_k$ is upper triangular. The DFC applies the linear feed-forward filter $\mathbf{Q}_k^H$ to the received vector to yield

$$\begin{aligned} \widetilde{\mathbf{y}}_k &= \mathbf{Q}_k^H \mathbf{y}_k, \\ &= \mathbf{R}_k \mathbf{x}_k + \widetilde{\mathbf{z}}_k, \end{aligned} \tag{4.8}$$

where the filtered noise $\widetilde{\mathbf{z}}_k \triangleq \mathbf{Q}_k^H \mathbf{z}_k$[59]. If the noise is spatially white (4.2) then filtering with the unitary matrix $\mathbf{Q}_k^H$ does not alter the noise statistics

$$\begin{aligned} \mathcal{E}\left\{\widetilde{\mathbf{z}}_k \widetilde{\mathbf{z}}_k^H\right\} &= \mathcal{E}\left\{\mathbf{Q}_k \mathbf{z}_k \mathbf{z}_k^H \mathbf{Q}_k^H\right\}, \\ &= \sigma_k \mathbf{I}_N \\ &= \mathcal{E}\left\{\mathbf{z}_k \mathbf{z}_k^H\right\}. \end{aligned}$$

If the noise is spatially coloured then a noise pre-whitening must be applied prior to the DFC, which leads to a more complex receiver structure[105]. However, in DSL the assumption of spatially white noise is often a valid one.

From (4.8) it is clear that the transmission channel has been transformed into an upper triangular channel $\mathbf{R}_k$. This channel is causal in the sense that there is an order in the crosstalk of the users. User N experiences crosstalk from no-one; user $N-1$ experiences crosstalk only from user N; user $N-2$ experiences crosstalk only from users N and $N-1$; and so on.

This causal structure admits the use of decision feedback to remove crosstalk. User N experiences no crosstalk. Hence the signal of user N can be detected, and the crosstalk it causes to the other components of $\mathbf{y}_k$ can be removed. At this point user $N-1$ can be detected free from crosstalk, and the crosstalk it causes to the remaining users can be removed. This procedure iterates until all users have been detected. The estimate for user n is thus formed

$$\hat{x}_k^n = \text{dec}\left[\frac{1}{r_k^{n,n}}\left(y_k^n - \sum_{m=n+1}^{N} r_k^{n,m}\hat{x}_k^m\right)\right],$$

where dec[.] denotes the decision operation and $r_k^{n,m} \triangleq [\mathbf{R}_k]_{n,m}$[59]. It is typically assumed that no decisions errors are made

$$\hat{x}_k^m = x_k^m, \forall m > n, \tag{4.9}$$

which leads to the following estimate for the symbol of user n

$$\hat{x}_k^n = \text{dec}\left[x_k^n + \frac{\widetilde{z}_k^n}{r_k^{n,n}}\right].$$

The data-rate of user n on tone k is then

$$b_{k,\text{dfc}}^n = \log_2(1 + \sigma_k^{-1}\Gamma^{-1}s_k^n\,|r_k^{n,n}|^2). \tag{4.10}$$

The CWDD property can be used to show that $|r_k^{n,n}| \simeq |h_k^{n,n}|$[59]. As a result, for small α_k, the DFC operates very close to the single-user bound

$$b_{k,\text{dfc}}^n \simeq b_{k,\text{mac-dsl}}^n,$$

So the DFC gives near-optimal performance. It should be noted, however, that this performance analysis is based on the assumption of error-free decisions (4.9). For this to be valid a perfect channel code must be used, which has infinite decoding complexity and delay[48].

In practice a sub-optimal code will be used, which can lead to decision errors, error propagation and poor performance. Furthermore, decoding of each user's codeword must be done before decisions are fed back. This increases complexity substantially and leads to a latency that grows with the number of lines in the binder.[2] Typical binders contain hundreds of lines, where the increase in latency due to the DFC can be substantial. As a result the DFC cannot support real-time applications such as voice over IP and video conferencing.

[2]In DSL systems the codewords are interleaved across the entire DMT block to add robustness against deep frequency nulls, which result from line properties such as bridged taps. Furthermore, the codeword may be interleaved across several DMT blocks to add robustness against impulse noise. This means that the codewords are already quite long, and the latency is typically at the limit required for most applications.

4.5 Near-optimal Linear Canceler

This section describes a simple linear canceler. Unlike the DFC, this structure has a low complexity, no latency and supports real-time applications. The structure is based on the *zero-forcing* (ZF) criterion, which leads to the following estimate of the transmitted vector

$$\begin{aligned} \widehat{\mathbf{x}}_k &= \mathbf{H}_k^{-1}\mathbf{y}_k, \\ &= \mathbf{x}_k + \mathbf{H}_k^{-1}\mathbf{z}_k. \end{aligned} \tag{4.11}$$

Each user experiences a crosstalk free channel, affected only by the filtered background noise.

It is well known that ZF designs lead to severe noise-enhancement when the channel matrix $\mathbf{H}_k$ is ill-conditioned. Fortunately CWDD ensures that the channel matrix is well-conditioned; so the linear ZF canceler leads to negligible noise enhancement and each user achieves a data-rate close to the single-user bound. To see this consider the *singular value decomposition* (SVD) of $\mathbf{H}_k$

$$\mathbf{H}_k \overset{\text{svd}}{=} \mathbf{U}_k \Lambda_k \mathbf{V}_k^H.$$

The CWDD of $\mathbf{H}_k$ ensures that its columns are approximately orthogonal. That is (4.3) implies that

$$\mathbf{h}_k^{m\,H}\,\mathbf{h}_k^n \simeq \begin{cases} \left|h_k^{n,n}\right|^2, & n = m; \\ 0, & n \neq m. \end{cases}$$

As a result

$$\begin{aligned} \mathbf{V}_k \Lambda_k^H \Lambda_k \mathbf{V}_k^H &= \mathbf{H}_k^H \mathbf{H}_k \\ &\simeq \operatorname{diag}\left\{ \left|h_k^{1,1}\right|^2, \ldots, \left|h_k^{N,N}\right|^2 \right\}, \end{aligned}$$

This implies that the right singular vectors can be closely approximated as $\mathbf{V}_k \simeq \mathbf{I}_N$. The linear ZF filter can then be approximated as

$$\begin{aligned} \mathbf{H}_k^{-1} &= \mathbf{V}_k \Lambda_k^{-1} \mathbf{U}_k^H, \\ &\simeq \Lambda_k^{-1} \mathbf{U}_k^H. \end{aligned}$$

Since $\mathbf{U}_k$ is unitary it causes no noise enhancement. Furthermore, Λ_k^{-1} is diagonal so it scales the noise and signal powers equally. So, due to the CWDD of $\mathbf{H}_k$, filtering the received signal with the matrix $\mathbf{H}_k^{-1}$ causes negligible noise enhancement. This allows the linear ZF canceler to achieve near-optimal performance, operating close to the single-user bound in DSL channels. This observation is made rigorous through the following theorem.

Theorem 4.2 *If* $A_{min}^{(m)} \geq \alpha_k m B_{max}^{(m)}$, $m = 1 \ldots N-1$; *then the data-rate achieved by the linear ZF canceler can be lower bounded*

$$R_n \geq f_s \sum_k b_{k,\mathrm{zf-bound}}^n,$$

where

$$b_{k,\mathrm{zf-bound}}^n \triangleq \log_2 \left(1 + \Gamma^{-1} \sigma_k^{-1} s_k^n \left|h_k^{n,n}\right|^2 f^{-1}(N, \alpha_k)\right), \tag{4.12}$$

$$f(N, \alpha_k) \triangleq \left(\frac{A_{\max}^{(N-1)}}{A_{\min}^{(N)}}\right)^2 + (N-1)\left(\frac{B_{\max}^{(N-1)}}{A_{\min}^{(N)}}\right)^2, \tag{4.13}$$

$$\begin{bmatrix} A_{\max}^{(m)} \\ B_{\max}^{(m)} \end{bmatrix} \triangleq \left(\prod_{i=1}^{m} \begin{bmatrix} 1 & (i-1)\alpha_k \\ \alpha_k & (i-1)\alpha_k \end{bmatrix}\right) \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \tag{4.14}$$

and

$$A_{\min}^{(m)} \triangleq 1 - \sum_{i=1}^{m} \alpha_k (i-1) B_{\max}^{(i-1)}. \tag{4.15}$$

Proof: Eq. (4.11) implies that, after application of the linear ZF canceler, the soft estimate of the transmitted symbol is

$$\widehat{x}_k^n = x_k^n + \left[\mathbf{H}_k^{-1}\right]_{\mathrm{row}\ n} \mathbf{z}_k.$$

Hence the post-cancellation signal power is s_k^n, the post cancellation interference power is zero and the post cancellation noise power is

$$\begin{aligned} \widetilde{\sigma}_{k,n} &\triangleq \mathcal{E}\left\{\left|\left[\mathbf{H}_k^{-1}\right]_{\mathrm{row}\ n} \mathbf{z}_k\right|^2\right\}, \\ &= \left\|\left[\mathbf{H}_k^{-1}\right]_{\mathrm{row}\ n}\right\|^2 \sigma_k, \end{aligned} \tag{4.16}$$

where (4.2) is applied in the second line. Hence the data-rate achieved by the linear ZF canceler is

$$b_{k,\mathrm{zf}}^n(s_k^n) = \log_2(1 + \Gamma^{-1} \widetilde{\sigma}_{k,n}^{-1} s_k^n). \tag{4.17}$$

Define the matrix $\mathbf{G}_k \triangleq [g_k^{n,m}]$, where $g_k^{n,m} \triangleq h_k^{n,m}/h_k^{m,m}$. Now

$$\mathbf{H}_k = \mathbf{G}_k \mathrm{diag}\{h_k^{1,1}, \ldots, h_k^{N,N}\},$$

hence

$$\mathbf{H}_k^{-1} = \mathrm{diag}\{h_k^{1,1}, \ldots, h_k^{N,N}\}^{-1} \mathbf{G}_k^{-1}, \tag{4.18}$$

and

$$\left[\mathbf{H}_k^{-1}\right]_{n,m} = \frac{1}{h_k^{n,n}} \left[\mathbf{G}_k^{-1}\right]_{n,m}. \tag{4.19}$$

Since the receivers are co-located at the CO, the US channel matrix is CWDD (4.4). This implies that $\mathbf{G}_k \in \mathbb{A}^{(N)}$, where $\mathbb{A}^{(N)}$ denotes the set of $N \times N$ diagonally dominant matrices, as defined in Appendix B. So Theorem B.2 can be applied to bound the elements of $\mathbf{G}_k^{-1}$. This implies

$$\left|\left[\mathbf{H}_k^{-1}\right]_{n,m}\right| \leq \begin{cases} |h_k^{n,n}|^{-1} A_{\max}^{(N-1)}/A_{\min}^{(N)}, & n = m; \\ |h_k^{n,n}|^{-1} B_{\max}^{(N-1)}/A_{\min}^{(N)}, & n \neq m; \end{cases}$$

where $A_{\max}^{(N)}$ and $B_{\max}^{(N)}$ are defined in (4.14) and $A_{\min}^{(N)}$ is defined in (B.4). Hence

$$\left\|\left[\mathbf{H}_k^{-1}\right]_{n,m}\right\|^2 \leq |h_k^{n,n}|^{-2} f(N, \alpha_k),$$

where $f(N, \alpha_k)$ is defined as in (4.13). Together with (4.16) this yields

$$\widetilde{\sigma}_{k,n} \leq \sigma_k |h_k^{n,n}|^{-2} f(N, \alpha_k).$$

Combining this with (4.17) leads to (4.12), which concludes the proof. ■

The function $f(N, \alpha_k)$ can be interpreted as an upper bound on the noise enhancement caused by the linear ZF canceler. In CWDD channels $f(N, \alpha_k) \simeq 1$. As a result each modem operates at a rate

$$b_{k,\mathrm{zf}}^n \simeq \log_2\left(1 + \Gamma^{-1}\sigma_k^{-1} s_k^n |h_k^{n,n}|^2\right).$$

So the linear ZF canceler completely removes crosstalk with negligible noise enhancement.

Note that the bound can be used to guarantee a data-rate without explicit knowledge of the crosstalk channels. This is because the bound only depends on the binder size, direct channel gain, and background noise power. Good models for these characteristics exist based on extensive measurement campaigns. Crosstalk channels on the other hand are poorly understood and actual channels can deviate significantly from the few empirical models that exist, see for example Fig. 2.5. This can make provisioning of services difficult. Using the bound (4.12) allows us to overcome this problem. The bound tells us that the crosstalk channel gain is not important as long as CWDD is observed. CWDD is a well understood and modeled phenomenon. As a result (4.12) allows provisioning to be done in a reliable and accurate fashion.

A note of explanation may be necessary at this point. It may seem that CWDD allows us to easily predict, or at least bound, the crosstalk power that a receiver experiences. This is *not* true. The crosstalk power that a receiver experiences depends on the magnitude of elements along a *row*, not *column*, of $\mathbf{H}_k$. This in turn depends on the configuration of the other lines within the binder, which varies dramatically from one scenario to another. For example, in the scenario

in Fig. 4.2, the crosstalk from the 150m line into the 1200m line is stronger than the direct signal on the 1200m line itself. So the crosstalk from the other lines into the 1200m line cannot be bounded without knowledge of the entire binder configuration. So, without the bound, knowledge of the entire binder is necessary to predict the performance of a single line. This makes provisioning of services extremely difficult. CWDD, on the other hand, applies to all lines when receivers are co-located. No knowledge of the actual binder configuration is necessary. Using (4.12) the performance of a line can be estimated using only locally available information about the line itself, such as its direct channel attenuation and background noise for which reliable models and statistical data exist.

The value for α_k from (4.5) is based on worst 1% case models. Hence for 99% of lines α_k will be smaller. So in 99% of lines a data-rate above the bound (4.12) is achieved. So the bound is a useful tool not just for theoretical analysis, but for provisioning of services as well.

Simulations in Section 4.7 use the bound together with (4.4) to show that the linear ZF canceler operates close to the single-user bound.

4.6 Spectra Optimization

Part I of this thesis discussed the coordination of DSL modems on a spectral level. Each modem generates its signal independently, however the transmit spectra are designed in a joint fashion to mitigate crosstalk.

When signal coordination is used crosstalk can be filtered at the receiver side. It is, however, still interesting to optimize the transmit spectra of each modem to achieve maximum performance. This can be viewed as the combination of signal and spectra coordination. Each transmitter is subject to a total power constraint

$$\sum_k s_k^n \leq P_n, \ \forall n. \tag{4.20}$$

As in Part I, the goal is to maximize the rate of user N, subject to target rate constraints on the other users in the network. Following the same development in Section 3.3.5, this can be reformulated as a weighted rate-sum optimization

$$\max_{\mathbf{s}_1,\ldots,\mathbf{s}_N} \sum_n w_n R_n \quad \text{s.t.} \quad \sum_k s_k^n \leq P_n, \ \forall n; \qquad s_k^n \geq 0, \ \forall n, k. \tag{4.21}$$

In contrast to Part I, R_n now represents the rate of modem n *with* crosstalk cancellation. The data-rate R_n is a function of the transmit PSDs $\mathbf{s}_1, \ldots, \mathbf{s}_N$,

and also depends on the type of crosstalk canceler used. If an optimal, decision-feedback based canceler is used, the objective function becomes convex[100]. Solving (4.21) then requires the solution of a KN-dimensional convex optimization. Although the cost function is convex, no closed form solution is known and numerical techniques must be used instead[100]. Conventional numerical optimization techniques, such as interior point methods, have a polynomial complexity in the dimensionality of the search space. In ADSL $K = 256$, whilst in VDSL $K = 4096$. The resulting search thus has an extremely high dimensionality, for which conventional optimization techniques are prohibitively complex.

A low complexity, iterative algorithm has been proposed for the special case where an unweighted rate-sum is maximized, that is $w_n = 1$ for all n[108]. Unfortunately, since this algorithm cannot optimize a weighted rate-sum, it cannot ensure that the target rates are achieved. These target rates are essential to ensure that each customer achieves their desired quality-of-service.

In this section a spectra coordination algorithm is developed for use with the ZF canceler. Since the ZF canceler removes all crosstalk, the spectrum coordination problem decouples into an independent power loading for each user. This reduces complexity considerably. Furthermore, Theorem 4.2 ensures that this approach operates close to the single-user bound.

4.6.1 Theoretical Capacity

We start by extending the single-user bound from Section 4.3 to DSL modems that may vary their transmit spectra under a total power constraint. The resulting upper bound is useful for evaluating crosstalk canceler performance with optimized spectra.

Theorem 4.3 *When the transmit PSD s_k^n is allowed to vary under a total power constraint (4.20), the capacity for user n can be upper bounded*

$$R_n \leq f_s \sum_k b_{k,\mathrm{mac-wf}}^n,$$

where the upper bound is defined

$$b_{k,\mathrm{mac-wf}}^n \triangleq \log_2 \left(1 + \sigma_k^{-1}\Gamma^{-1} s_{k,\mathrm{mac-wf}}^n \left|h_k^{n,n}\right|^2 \left[1 + \alpha_k^2(N-1)\right]\right), \qquad (4.22)$$

the single-user waterfilling PSD is defined

$$s_{k,\mathrm{mac-wf}}^n \triangleq \left[\frac{1}{\lambda_n} - \frac{\Gamma\sigma_k}{\left|h_k^{n,n}\right|^2 \left[1 + \alpha_k^2(N-1)\right]}\right]^+, \qquad (4.23)$$

and λ_n is chosen such that power constraint on line n is tight

$$\sum_k s_k^n = P_n. \tag{4.24}$$

This theorem is an intuitively simple extension of the single-user bound from Theorem 4.1. The proof, however, is not so straightforward and is now detailed. The following Lemma will prove useful in the proof.

Lemma 4.4 *If $g(\mathbf{x}) \geq f(\mathbf{x})$, $\forall \mathbf{x}$, then*

$$\max_{\mathbf{vx} \leq p} g(\mathbf{x}) \geq \max_{\mathbf{vx} \leq p} f(\mathbf{x}), \tag{4.25}$$

where $\mathbf{x}$ is the vector over which the optimization takes place, and the vector $\mathbf{v}$ and scalar p impose a linear constraint on $\mathbf{x}$.

Proof: Define

$$\mathbf{x}_f \triangleq \arg\max_{\mathbf{vx} \leq p} f(\mathbf{x}).$$

Since $g(\mathbf{x}) \geq f(\mathbf{x})$, $\forall \mathbf{x}$,

$$g(\mathbf{x}_f) \geq f(\mathbf{x}_f). \tag{4.26}$$

Now define

$$\mathbf{x}_g \triangleq \arg\max_{\mathbf{vx} \leq p} g(\mathbf{x}).$$

The optimality of $\mathbf{x}_g$ in $g(\mathbf{x})$, over the subspace defined by the constraints $\mathbf{vx} \leq p$, implies that

$$\begin{aligned} g(\mathbf{x}_g) &\geq g(\mathbf{x}_f), \\ &\geq f(\mathbf{x}_f), \end{aligned}$$

where (4.26) is applied in the second line. This implies (4.25). ∎

Corollary 4.5 *Limit the total power of the transmit PSD such that $\sum_k s_k^n \leq P_n$. Under this constraint*

$$\begin{aligned} &\max_{\mathbf{s}_n} \sum_k \log_2 \left(1 + \sigma_k^{-1} s_k^n \Gamma^{-1} \left\|\mathbf{h}_k^n\right\|^2\right) \\ &\quad \leq \max_{\mathbf{s}_n} \sum_k \log_2 \left(1 + \sigma_k^{-1} s_k^n \Gamma^{-1} \left|h_k^{n,n}\right|^2 \left[1 + \alpha_k^2 (N-1)\right]\right), \end{aligned} \tag{4.27}$$

where $\mathbf{s}_n \triangleq [s_1^n, \ldots, s_K^n]$.

Proof: Let $\mathbf{x} = \mathbf{s}_n$, $p = P_n$, $\mathbf{v} = 1_{1\times K}$,

$$f(\mathbf{s}_n) = \sum_k \log_2\left(1 + \sigma_k^{-1}\Gamma^{-1}s_k^n \left\|\mathbf{h}_k^n\right\|^2\right),$$

and

$$g(\mathbf{s}_n) = \sum_k \log_2\left(1 + \sigma_k^{-1}\Gamma^{-1}s_k^n \left|h_k^{n,n}\right|^2 \left[1 + \alpha_k^2(N-1)\right]\right).$$

CWDD (4.3) implies that

$$\left|h_k^{n,n}\right|^2 \left[1 + (N-1)\alpha_k^2\right] \geq \left\|\mathbf{h}_k^n\right\|^2,$$

hence $g(\mathbf{s}_n) \geq f(\mathbf{s}_n)$, $\forall \mathbf{s}_n$. Lemma 4.4 now implies (4.27), which completes the proof. ■

We now proceed with the proof to Theorem 4.3.

Proof: The single-user bound developed in Section 2.2.3 applies, limiting the achievable rate of user n on tone k

$$b_k^n \quad \leq \quad \log_2\left(1 + \sigma_k^{-1}\Gamma^{-1}s_k^n \left\|\mathbf{h}_k^n\right\|_2^2\right).$$

Hence

$$R_n \leq \max_{\sum_k s_k^n \leq P_n} f_s \sum_k \log_2\left(1 + \sigma_k^{-1}\Gamma^{-1}s_k^n \left\|\mathbf{h}_k^n\right\|_2^2\right).$$

Corollary 4.5 now implies

$$R_n \leq \max_{\sum_k s_k^n \leq P_n} f_s \sum_k \log_2\left(1 + \sigma_k^{-1}\Gamma^{-1}s_k^n \left|h_k^{n,n}\right|^2 \left[1 + \alpha_k^2(N-1)\right]\right).$$

In this optimization the objective function is concave, and the total power constraint forms a convex set. Hence the *Karush-Kuhn-Tucker* (KKT) conditions are sufficient for optimality. Examining the KKT conditions leads to (4.22), (4.23) and (4.24), which completes the proof. ■

4.6.2 Near-Optimal Linear Canceler

Transmit spectra optimization with the ZF canceler is now considered. The optimization problem is now

$$\begin{aligned} \mathbf{s}_1^{\text{zf}}, \ldots, \mathbf{s}_N^{\text{zf}} \quad &= \quad \arg\max_{\mathbf{s}_1,\ldots,\mathbf{s}_N} \sum_n \sum_k w_n b_{k,\text{zf}}^n; \\ &\text{s.t.} \quad \sum_k s_k^n \leq P_n, \ \forall n; \end{aligned}$$

$$s_k^n \geq 0, \ \forall n, k;$$

where the optimal solution $\mathbf{s}_n^{\text{zf}} \triangleq [s_{1,\text{zf}}^n, \ldots, s_{K,\text{zf}}^n]$. Here $b_{k,\text{zf}}^n$ denotes the rate achieved with the linear ZF canceler, which was defined in (4.17) as

$$b_{k,\text{zf}}^n \triangleq \log_2 \left(1 + s_k^n \widetilde{\sigma}_{k,n}^{-1} \Gamma^{-1}\right).$$

Observe that, when using the ZF canceler, the bitrate of each user depends only on its own transmit PSD. It is independent of the PSDs of the other users since all interference will be removed. The optimization problem is now decoupled between users, allowing the optimal power allocation to be found independently for each user. This also implies that a single PSD is optimal regardless of the choice of weights w_n.

Since the objective function is concave and the constraints form a convex set, the KKT conditions are sufficient for optimality. Examining these leads to the classic waterfilling equation

$$s_{k,\text{zf}}^n \quad = \quad \left[\frac{1}{\lambda_n} - \Gamma \widetilde{\sigma}_{k,n}\right]^+. \tag{4.28}$$

The *waterfilling level* λ_n must be chosen such that the total power constraint for user n is tight, that is $\sum_k s_{k,\text{zf}}^n = P_n$.

Conventional waterfilling algorithms can be applied to find the correct waterfilling level with $\mathcal{O}(K \log K)$ complexity. So the overall complexity of power allocation is $\mathcal{O}(NK \log K)$ with the linear ZF canceler[3]. This is a significant reduction when compared to existing power allocation algorithms for the multi-access channel, which have $\mathcal{O}(N^4 K \log K)$ complexity in the unweighted rate-sum case, and polynomial complexity in KN in the weighted rate-sum case[108].

Theorem 4.2 shows that, as a result of CWDD, the ZF canceler operates close to the single-user bound. So using the ZF canceler in combination with the power allocation (4.28) gives near-optimal performance. This is confirmed through simulation in the following section.

4.7 Performance

This section evaluates the performance of the ZF canceler in a binder of 8 VDSL lines. The line lengths range from 150m to 1200m in 150m increments,

[3]Note that this is a significant reduction in complexity when compared to the power allocation problem in Chapter 3. The basic reason is that power allocation in a MAC typically involves a convex optimization, where-as power allocation in an IC involves a non-convex optimization.

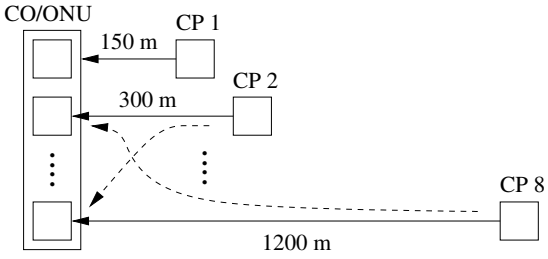

Figure 4.2: Upstream VDSL scenario

as shown in Fig. 4.2. For all simulations the line diameter is 0.5 mm (24-AWG). Direct and crosstalk channels are generated using the empirical models described in Section 2.1.2. The target symbol error probability is 10^{-7} or less. The coding gain is set to 3 dB and the noise margin is set to 6 dB. As per the VDSL standards the tone spacing Δ_f is set to 4.3125 kHz and DMT symbol rate f_s to 4 kHz[9][7]. The modems use 4096 tones, and the 998 FDD bandplan. Background noise is generated using ETSI noise model A[7]. Performance is compared with the DFC and the single-user bound.

4.7.1 Fixed Transmit Spectra

Current VDSL standards require that modems transmit under a spectral mask of -60 dBm/Hz[9][7]. This section evaluates the performance of the linear ZF canceler when all modems are operating at this mask.

Fig. 4.4 shows the data-rate achieved by each of the lines with the different crosstalk cancelers. The linear ZF canceler achieves substantial gains, typically 30 Mbps or more, over conventional systems with no cancellation. As can be seen the linear ZF canceler achieves near-optimal performance, operating close to the single-user bound. This is a direct result of the CWDD of $\mathbf{H}_k$, which ensures that the ZF canceler causes negligible noise enhancement. The noise enhancement caused by the ZF canceler on the 600m line is plotted for each tone in Fig. 4.3. As can be seen the noise enhancement is less than 0.16 dB, which has negligible effect on performance.

Fig. 4.5 shows the data-rate achieved by the linear ZF canceler as a percentage of the single-user bound. Performance does not drop below 99% of the single-user bound. The lower bound on the performance of the linear ZF canceler (4.12) is also included for comparison. As can be seen the bound is quite tight and guarantees that the linear ZF canceler will achieve at least 92% of the single-user bound.

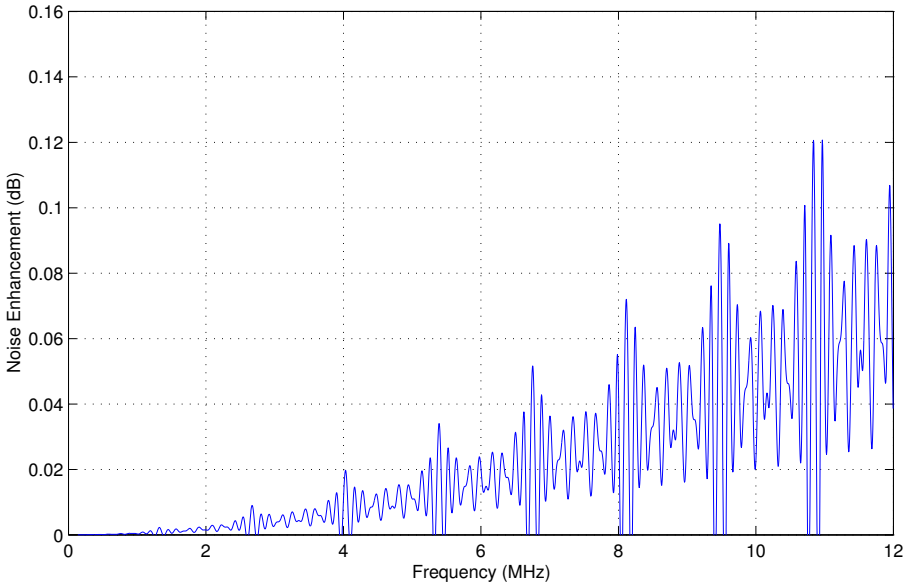

Figure 4.3: Noise enhancement of ZF Canceler on 600 m. line

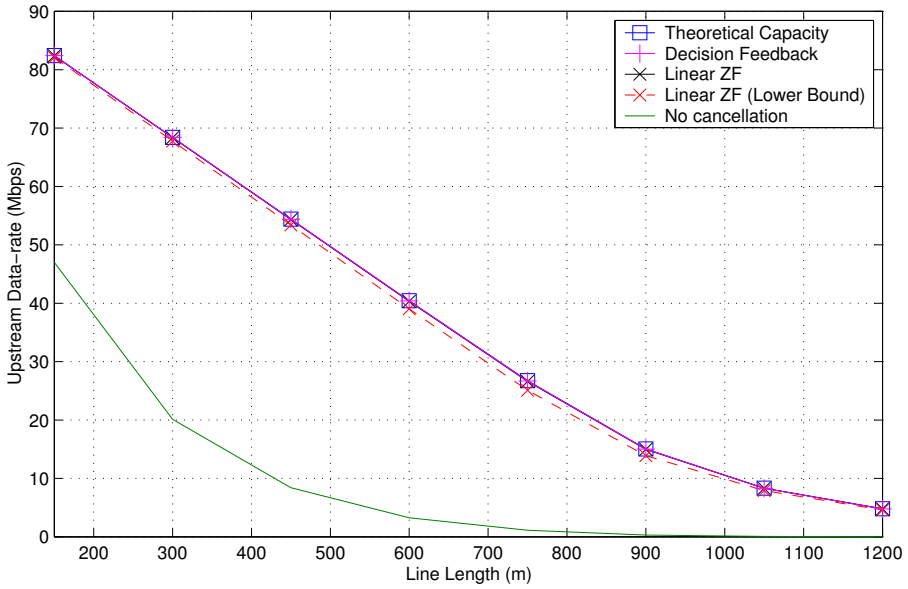

Figure 4.4: Upstream Data-rate with Different Cancelers

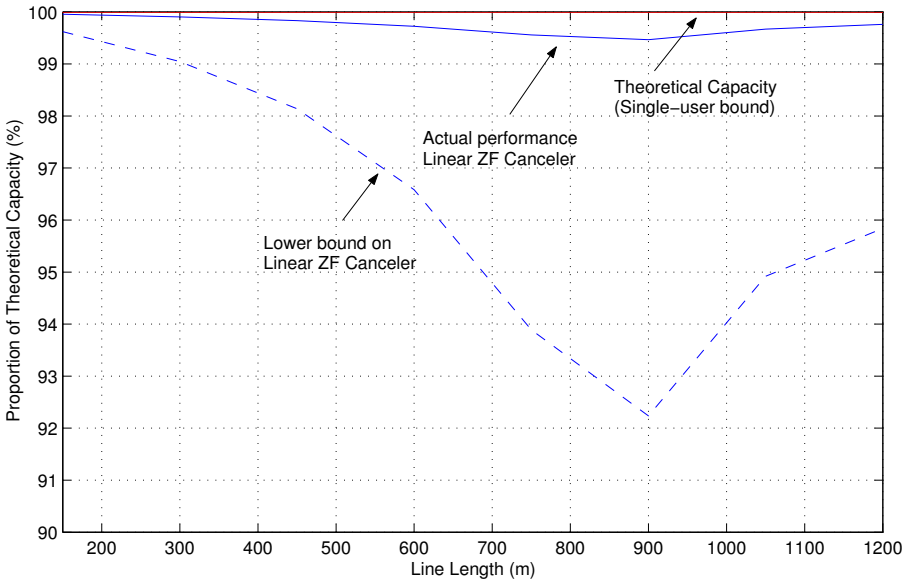

Figure 4.5: Proportion of Single-user Bound Achieved by ZF Canceler

4.7.2 Optimized Transmit Spectra

Whilst current VDSL standards require the use of spectral masks, there is growing interest in the use of adaptive transmit spectra[8]. This section investigates the performance of the linear ZF canceler with optimized spectra (4.28).

Fig. 4.6 shows the data-rates achieved on each line. The use of optimized spectra yields a gain of 5-8 Mbps. The benefit is more substantial on the longer lines, where a 5 Mbps gain can double the data-rate.

Fig. 4.6 shows that spectra coordination gives maximum benefit on long lines. This is to be expected since on long lines the direct channel gain decreases more rapidly with frequency. Note that the benefit of adaptive spectra, when crosstalk has already been cancelled, comes primarily from the modem loading power in the best parts of the channel, which are typically in the lower frequencies.

The single-user bound (4.22) is included for comparison. As can be seen the linear ZF canceler operates close to the single-user bound. Fig. 4.7 shows the data-rate achieved by the linear ZF canceler as a percentage of the single-user bound. Performance does not drop below 99% of the single-user bound. The lower bound on the data-rate of the ZF canceler is also included for comparison. The bound (4.12) guarantees that the ZF canceler and waterfilled PSD achieve

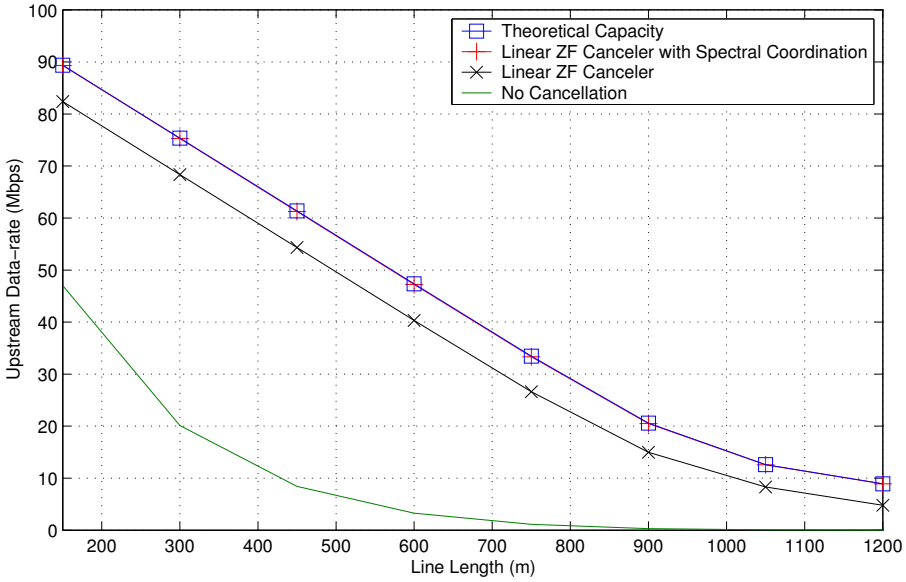

Figure 4.6: Upstream Data-rate with Coordinated Transmit Spectra

92% of the single-user bound.

4.8 Summary

This chapter investigated the design of crosstalk cancelers for upstream DSL. Existing designs, which are based on decision feedback, suffer from error propagation, high complexity and long latency. A linear ZF canceler is proposed, which has a low complexity and no latency.

An oft-cited problem with the ZF design is that it leads to severe noise enhancement in ill-conditioned channels. Fortunately DSL channels with co-located receivers are column-wise diagonal dominant. This ensures that the DSL channel is well conditioned, and that the noise enhancement caused by the ZF design is negligible.

An upper bound on the capacity of the multi-user DSL channel was derived. This single-user bound shows that spatial diversity in the DSL environment is negligible. Therefore the best outcome that a crosstalk canceler can achieve is the complete suppression of crosstalk without noise enhancement.

A lower bound on the performance of the linear ZF canceler was derived. This

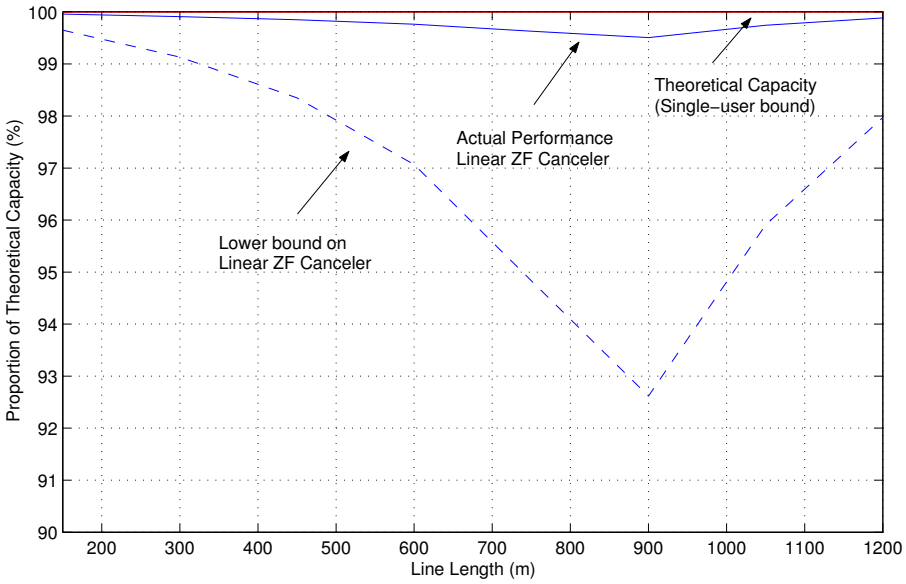

Figure 4.7: Proportion of Single-user Bound Achieved by ZF Canceler

bound depends only on the binder size, direct channel gain and background noise for which reliable models and statistical data exist. As a result the performance of the linear ZF canceler can be accurately predicted, which simplifies service provisioning considerably. This bound shows that the linear ZF canceler operates close to the single-user bound. So the linear ZF canceler is a low complexity, low latency design with guaranteed near-optimal performance.

The combination of spectral optimization and crosstalk cancellation was considered. The bounds were extended to DSL systems with optimized spectra. Spectra optimization in a multi-access channel in general involves a highly complex optimization. Since the linear ZF canceler decouples transmission on each line, the spectrum on each modem can be optimized independently, leading to a significant reduction in complexity.

Chapter 5

Transmitter Coordination

5.1 Introduction

The previous chapter investigated receiver coordination as a means of improving DSL performance. In upstream communication, the receiving modems are co-located at the *central office* (CO). Reception is done a joint fashion; the signals received on each line are combined to cancel crosstalk whilst preserving the signal of interest. The results were significant, with the data-rate on each line typically increasing by 30 Mbps or more[1].

In *downstream* (DS) communication the receiving modems reside within different *customer premises* (CP). The receiving modems are not co-located, so joint reception and crosstalk cancellation is impossible. Fortunately, in DS communication the transmitting modems are co-located at the CO. So joint *transmission* is possible. Predistortion is introduced into each signal before transmission. The predistortion is chosen such that it annihilates with the crosstalk introduced in the binder, a technique known as crosstalk precoding.

Define $\widetilde{x}_k^n$ as the symbol intended for receiver n on tone k, and x_k^n as the signal sent by transmitter n on tone k. In conventional DSL systems each modem transmits the symbol intended for the corresponding receiver

$$x_k^n = \widetilde{x}_k^n.$$

Recall that k denotes the tone index and lies in the range $1 \ldots K$. The signal

[1] The work in this chapter was done in close collaboration with Dr. George Ginis, Texas Instruments, San Jose, CA.

received by modem n on tone k is then

$$y_k^n = h_k^{n,n}\widetilde{x}_k^n + \sum_{m\neq n} h_k^{n,m}\widetilde{x}_k^m + z_k^n.$$

The first term is the signal of interest, the middle term the interference, and the third term the background noise. Since the receivers are not co-located they must treat crosstalk as background noise. This decreases the SNR and limits the data-rate to

$$b_k^n = \log_2\left(1 + \frac{1}{\Gamma}\frac{\left|h_k^{n,n}\right|^2 s_k^n}{\sum_{m\neq n}\left|h_k^{n,m}\right|^2 s_k^m + \sigma_k^n}\right).$$

In downstream communication the transmitting modems are often co-located at a common CO. Transmission can then be coordinated to mitigate crosstalk. Essentially, each CO modem transmits a mixture of the symbols intended for the different CP modems

$$x_k^n = \widetilde{x}_k^n + \overline{f}_{k,\mathrm{mu-tx}}^n\left(\widetilde{x}_k^1, \ldots, \widetilde{x}_k^N\right), \tag{5.1}$$

where $\overline{f}_{k,\mathrm{mu-tx}}^n(\cdot)$ denotes the multi-user precoding operation for transmitter n and N denotes the number of lines in the binder. The signal transmitted contains the symbol of interest $\widetilde{x}_k^n$ plus a predistortion term $\overline{f}_{k,\mathrm{mu-tx}}^n\left(\widetilde{x}_k^1, \ldots, \widetilde{x}_k^N\right)$. The signal received by modem n on tone k is then

$$y_k^n = h_k^{n,n}\widetilde{x}_k^n + \sum_{m\neq n} h_k^{n,m}\widetilde{x}_k^m + \sum_m h_k^{n,m}\overline{f}_{k,\mathrm{mu-tx}}^m + z_k^n.$$

The predistortion is chosen such that it annihilates with the crosstalk introduced in the binder

$$\sum_m h_k^{n,m}\overline{f}_{k,\mathrm{mu-tx}}^m = -\sum_{m\neq n} h_k^{n,m}\widetilde{x}_k^m, \ \forall n.$$

Hence

$$\begin{aligned} y_k^n &= h_k^{n,n}x_k^n + \sum_{m\neq n} h_k^{n,m}x_k^m + z_k^n, \\ &= h_k^{n,n}\widetilde{x}_k^n + \sum_{m\neq n} h_k^{n,m}\widetilde{x}_k^m + \sum_m h_k^{n,m}\overline{f}_{k,\mathrm{mu-tx}}^m + z_k^n, \\ &= h_k^{n,n}\widetilde{x}_k^n + z_k^n, \end{aligned}$$

and each modem receives only the signal of interest and background noise. It will be shown that, since the DSL channel is well conditioned, the achievable data-rate is very close to the single-user bound

$$b_{k,\mathrm{bc}}^n \triangleq \log_2\left(1 + (\sigma_k^n)^{-1}\Gamma^{-1}\left(\sum_m \left|h_k^{n,m}\right|\sqrt{s_k^m}\right)^2\right),$$

where the nth row of $\mathbf{H}_k$ is defined $\overline{\mathbf{h}}_k^n \triangleq [\mathbf{H}_k]_{\text{row } n}$.

Several crosstalk precoder designs have been proposed. The simplest is a linear structure based on the ZF criterion[92] and is described in more detail in Section 5.4. Unfortunately with this design all modems experience the channel of the weakest line in the binder. When the channels of the different lines vary significantly, due to varying line lengths or bridged taps, this design leads to poor performance.

A decision feedback structure, based on the *Tomlinson-Harashima precoder* (THP), was shown to achieve close to the single-user bound[59] and is described in more detail in Section 5.5. Unfortunately this structure requires a change of CP modems. Millions of CP modems are currently in use, owned and operated by a multitude of customers. Replacing these modems presents a huge legacy issue. Furthermore, CO modems and CP modems are typically manufactured by different hardware vendors, making joint designs difficult.

To address this problem, this chapter presents a novel linear precoder based on a channel diagonalizing criterion. This technique has a low complexity and does not require the replacement of CP modems. Section 5.6 analyzes the performance of the diagonalizing precoder in a DSL environment. It is shown that, due to the well conditioned structure of the DSL channel matrix, the diagonalizing design leads to negligible transmit power enhancement. Section 5.6 derives bounds which show that the diagonalizing precoder operates close to the single-user bound. These bounds allow performance of the DP to be predicted without the need for explicit knowledge of the crosstalk channels, which simplifies service provisioning significantly. The diagonalizing precoder and the associated bound are the major contributions of this chapter.

The rest of this chapter is organized as follows. The system model for a network of DSL modems transmitting from a single CO to a multitude of CPs is given in Section 5.2. A property of the downstream DSL channel, known as *row-wise diagonal dominance* (RWDD), is explored. As described in Section 5.3, from an information theoretical perspective, the downstream DSL channel is a *broadcast channel* (BC). This allows the single-user bound from Section 2.2.4 to be applied to bound the capacity of the channel. Section 5.4 describes the *zero-forcing precoder* (ZFP) and the problems it has with transmit power enhancement. Section 5.5 describes the multi-user THP, and shows that it requires replacement of CP modems.

To address these problems, Section 5.6 describes a much simpler linear design, the *diagonalizing precoder* (DP), which has a low complexity and works with existing CP modems. Section 5.6 uses the RWDD property to formulate a lower bound on the performance of the DP. This bound shows that the DP operates close to the single-user bound.

Section 5.7 describes power loading algorithms for use with the DP. This can be seen as the combination of spectral and signal coordination for the upstream DSL channel. Existing power loading algorithms for the BC are extremely complex, having a polynomial complexity in the number of lines and tones. Application of the DP decouples the power allocation problem between lines, and this simplifies power allocation significantly. The PSD for each line can then be found through a low-complexity waterfilling procedure. The performance of the different precoders is evaluated in Section 5.8.

5.2 System Model and RWDD

For crosstalk precoding to be applied, the transmitting modems must be co-located at a common CO. This makes it straightforward to synchronize the modems, and transmission can be modeled independently on each tone as described in Chapter 2

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k.$$

The transmit PSD on each line must obey a spectral mask constraint. That is

$$s_k^n \leq s_k^{\text{mask}}, \ \forall k, n. \tag{5.2}$$

Since the transmitting modems are co-located, the crosstalk signal transmitted from a disturber into a victim must propagate through the full length of the victim's line. This is depicted in Fig. 5.1, where CO2 is the disturber and CP1 is the victim. The shielding between twisted pairs increases the attenuation. As a result, the crosstalk channel matrix $\mathbf{H}_k$ is *row-wise diagonally dominant* (RWDD), since on each row of $\mathbf{H}_k$ the diagonal element has the largest magnitude

$$|h_k^{n,n}| \gg |h_k^{n,m}|, \ \forall m \neq n. \tag{5.3}$$

RWDD implies that the crosstalk channel $h_k^{n,m}$ from a disturber m into a victim n is always weaker than the direct channel of the victim $h_k^{n,n}$. Contrast this with the CWDD described in the previous chapter, where the crosstalk channel $h_k^{n,m}$ from a disturber into a victim is always weaker than the direct channel of the disturber $h_k^{m,m}$.

The degree of RWDD can be characterized with the parameter α_k

$$|h_k^{n,m}| \leq \alpha_k \, |h_k^{n,n}|, \ \forall m \neq n. \tag{5.4}$$

Note that crosstalk precoding is based on joint transmission. As such it requires the co-location of transmitting modems. So in all channels where crosstalk precoding can be applied, the RWDD property holds. RWDD has been verified through extensive measurement campaigns of real binders. In 99% of lines α_k is bounded

$$\alpha_k \leq K_{\text{xf}} f_k \sqrt{d_{\text{coupling}}},$$

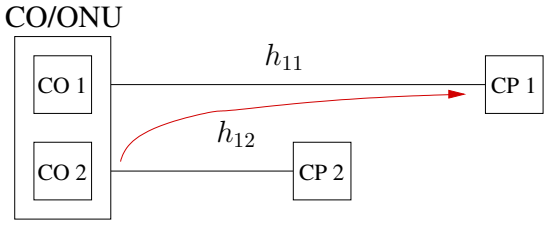

Figure 5.1: Row-wise Diagonal Dominance $|h_{11}| \gg |h_{12}|$

where $K_{\text{xf}} = -22.5$ dB and f_k is the frequency on tone k in MHz[7]. Here d_{coupling} is the coupling length between the disturber and the victim in kilometers, as defined in (2.14). The coupling length can be upper bounded by the longest line length in the binder. Hence

$$\alpha_k \leq K_{\text{xf}} f_k \sqrt{l_{\max}}, \tag{5.5}$$

where $l_{\max}$ denotes the length of the longest line in the binder. To find a value for α_k that is independent of the particular binder configuration, $l_{\max}$ can be set to 1.2 km, which is the maximum deployment length for DSL. On typical lines α_k is less than -11.3 dB. The following sections show that RWDD ensures a well-conditioned crosstalk channel matrix. This results in the near-optimality of the DP.

5.3 Theoretical Capacity

We start with a bound on the capacity of the downstream DSL channel with coordinated transmitters. This will prove useful in evaluating crosstalk precoder performance since it provides an upper bound on the achievable data-rate with any possible crosstalk precoding scheme.

Theorem 5.1 *The capacity for user n when all transmitters are subject to a spectral mask s_k^{mask} is upper bounded*

$$R_n \leq f_s \sum_k b_{k,\text{bc-dsl}}^n \tag{5.6}$$

where

$$b_{k,\text{bc-dsl}}^n \leq \log_2 \left(1 + (\sigma_k^n)^{-1} \Gamma^{-1} s_k^{\text{mask}} |h_k^{n,n}|^2 [1 + (N-1)\alpha_k]^2\right).$$

Proof : We follow similar lines to the proof of Theorem 4.1 and again make use of the single-user bound. First consider the channel capacity. CO modems are co-located and transmit jointly, so from an information theoretical perspective this is a broadcast channel. The single-user bound developed in Section 2.2.4 applies. From (2.21) the achievable bitloading of user n on tone k is limited

$$b_k^n \leq \log_2 \left(1 + (\sigma_k^n)^{-1} \Gamma^{-1} \left(\sum_m |h_k^{n,m}| \sqrt{s_k^m} \right)^2 \right). \tag{5.7}$$

Here the SNR-gap to capacity Γ is included to account for the sub-optimality of practical coding schemes. Combining this with (5.4), (5.7) and the spectral mask constraint (5.2) leads to (5.6), which completes the proof. ■

As in the previous chapter, the lack of spatial diversity in DSL can be seen in equation (5.6). Here the RWDD of $\mathbf{H}_k$ implies that very little increase can be made in the received signal power through the use of multiple transmitters. This is the case since the channel from transmitter n to receiver m is much weaker than the direct channel from transmitter m to receiver m. Note that the benefit, although small, increases with the crosstalk channel strength α_k and the number of crosstalkers N.

Although spatial diversity is negligible, the use of co-ordinated transmission is by no means fruitless. Instead of benefiting through spatial diversity, the primary benefit in DSL channels is crosstalk precoding. That is, co-ordinated transmission does not increase signal power in DSL channels, but instead decreases *interference power*.

5.4 Zero Forcing Precoder

The simplest precoder design is the *zero forcing precoder* (ZFP). Define the vector

$$\widetilde{\mathbf{x}}_k \triangleq [\widetilde{x}_k^1, \ldots, \widetilde{x}_k^N]^T,$$

which contains the symbols intended for each user on tone k. The ZFP multiplies the true symbols $\widetilde{\mathbf{x}}_k$ with a precoding matrix $\mathbf{P}_{k,\text{zf}}$ prior to transmission. The transmitted symbols are then

$$\mathbf{x}_k = \mathbf{P}_{k,\text{zf}} \widetilde{\mathbf{x}}_k.$$

The ZFP is based on a zero-forcing criteria, which leads to the following precoding matrix

$$\mathbf{P}_{k,\text{zf}} \triangleq \frac{1}{\beta_{k,\text{zf}}} \mathbf{H}_k^{-1},$$

where the scaling factor is defined

$$\beta_{k,\mathrm{zf}} \triangleq \max_n \left\| \left[\mathbf{H}_k^{-1}\right]_{\mathrm{row}\, n} \right\|. \tag{5.8}$$

The scaling factor $\beta_{k,\mathrm{zf}}$ ensures that compliance with the spectral masks (5.2) is maintained after precoding. Consider

$$\begin{aligned} x_k^n &= \left[\mathbf{P}_{k,\mathrm{zf}}\right]_{\mathrm{row}\, n} \widetilde{\mathbf{x}}_k, \\ &= \beta_{k,\mathrm{zf}}^{-1} \sum_m \left[\mathbf{H}_k^{-1}\right]_{n,m} \widetilde{x}_k^m. \end{aligned} \tag{5.9}$$

The power of the true symbols are set to obey the spectral mask

$$\widetilde{s}_k^n \leq s_k^{\mathrm{mask}},\ \forall k, n; \tag{5.10}$$

where $\widetilde{s}_k^n \triangleq \mathcal{E}\left\{|\widetilde{x}_k^n|^2\right\}$. Under this condition

$$\begin{aligned} s_k^n &= \beta_{k,\mathrm{zf}}^{-2} \mathcal{E}\left\{ \left| \sum_m \left[\mathbf{H}_k^{-1}\right]_{n,m} \widetilde{x}_k^m \right|^2 \right\} \\ &= \beta_{k,\mathrm{zf}}^{-2} \sum_m \left| \left[\mathbf{H}_k^{-1}\right]_{n,m} \right|^2 \widetilde{s}_k^n, \\ &\leq \beta_{k,\mathrm{zf}}^{-2} \left\| \left[\mathbf{H}_k^{-1}\right]_{\mathrm{row}\, n} \right\|^2 s_k^{\mathrm{mask}}, \\ &\leq s_k^{\mathrm{mask}}, \end{aligned}$$

where (5.8) is used in the last line. Hence the ZFP maintains compliance with the spectral mask constraints.

During transmission the predistortion introduced by the ZFP annihilates with the crosstalk. The received vector

$$\begin{aligned} \mathbf{y}_k &= \mathbf{H}_k \mathbf{P}_{k,\mathrm{zf}} \widetilde{\mathbf{x}}_k + \mathbf{z}_k, \\ &= \beta_{k,\mathrm{zf}}^{-1} \widetilde{\mathbf{x}}_k + \mathbf{z}_k, \end{aligned}$$

and each user experiences a crosstalk free channel. All users experience the same direct channel gain $\beta_{k,\mathrm{zf}}^{-1}$. Unfortunately, this causes all users to experience the worst channel in the binder. To see this, first note that with the ZFP each user achieves a data-rate of

$$\begin{aligned} b_{k,\mathrm{zfp}}^n &= \log_2\left(1 + (\sigma_k^n)^{-1} \Gamma^{-1} \widetilde{s}_k^n \beta_{k,\mathrm{zf}}^{-2}\right), \\ &\leq \log_2\left(1 + (\sigma_k^n)^{-1} \Gamma^{-1} s_k^{\mathrm{mask}} \beta_{k,\mathrm{zf}}^{-2}\right), \end{aligned}$$

when the modem transmits at the spectral mask. Now consider the upper bound on the theoretical capacity (5.6). Since $b_{k,\mathrm{zfp}}^n \leq b_{k,\mathrm{opt}}^n,\ \forall n$; this implies that

$$\beta_{k,\mathrm{zf}}^{-2} \leq \left|h_k^{n,n}\right|^2 \left[1 + (N-1)\alpha_k\right]^2,\ \forall n;$$

$$= \quad [1+(N-1)\alpha_k]^2 \min_n |h_k^{n,n}|^2 .$$

Hence

$$b_{k,\text{zfp}}^n \leq \log_2 \left(1 + (\sigma_k^n)^{-1} \Gamma^{-1} s_k^{\text{mask}} [1+(N-1)\alpha_k]^2 \min_m |h_k^{m,m}|^2\right).$$

Since $[1+(N-1)\alpha_k]^2 \simeq 1$, all users in the binder will experience a direct channel gain of $\min_m |h_k^{m,m}|$. When the line lengths vary significantly, or if one of the lines in the binder contains a bridged tap, the weakest channel in the binder is significantly weaker than the other channels. In this case the ZFP gives extremely sub-optimal performance. Consider, for example, a scenario with ten 300m lines and one 1200m line. With the ZFP all lines will experience the direct channel of the 1200m line. In many cases the ZFP leads to even worse performance than without crosstalk precoding as is shown in Section 5.8.

5.5 Tomlinson-Harashima Precoder

In the previous chapter it was shown that a decision feedback structure can be used to cancel crosstalk. The users were detected in an iterative fashion. After each user is detected, their interference is subtracted from the remaining undetected users. In downstream transmission a similar structure can be used for crosstalk precoding[59, 58]. This can be seen as the multi-user extension of the Tomlinson-Harashima precoder, which is commonly used for precoding against ISI in single-user channels[67, 91].

The structure of the *Tomlinson-Harashima Precoder* (THP) is now described. Consider the QR decomposition of the crosstalk channel matrix

$$\mathbf{H}_k^T \stackrel{\text{qr}}{=} \mathbf{Q}_k \mathbf{R}_k, \tag{5.11}$$

where $\mathbf{Q}_k$ is a unitary matrix and $\mathbf{R}_k$ is upper triangular. Here $(.)^T$ is used to denote the transpose operation. Now

$$\mathbf{H}_k = \mathbf{R}_k^T \mathbf{Q}_k^T.$$

Prior to transmission, the signal is pre-multiplied with $\mathbf{Q}_k$ such that

$$\mathbf{x}_k = \mathbf{Q}_k \acute{\mathbf{x}}_k, \tag{5.12}$$

where $\acute{\mathbf{x}}_k \triangleq [\acute{x}_k^1, \ldots, \acute{x}_k^N]^T$. The received vector is then

$$\begin{aligned} \mathbf{y}_k &= \mathbf{H}_k \mathbf{Q}_k \acute{\mathbf{x}}_k + \mathbf{z}_k, \\ &= \mathbf{R}_k^T \acute{\mathbf{x}}_k + \mathbf{z}_k. \end{aligned} \tag{5.13}$$

Since $\mathbf{Q}_k$ is unitary, compliance with the spectral masks (5.2) is maintained after the precoding operation. To see this, first note that the power of $\acute{x}_k^n$ is set to obey the spectral mask

$$\acute{s}_k^n \leq s_k^{\text{mask}}, \ \forall n;$$

where $\acute{s}_k^n \triangleq \mathcal{E}\left\{|\acute{x}_k^n|^2\right\}$. It is assumed that the signal $\acute{x}_k^n$ of each user is independent, which is approximately true[58]. Under this assumption

$$\begin{aligned} s_k^n &= \mathcal{E}\left\{\left|\sum_m [\mathbf{Q}_k]_{n,m}\, \acute{x}_k^m\right|^2\right\}, \\ &= \sum_m \left|[\mathbf{Q}_k]_{n,m}\right|^2 \acute{s}_k^n, \\ &\leq s_k^{\text{mask}}, \ \forall n; \end{aligned}$$

where (5.12) is used in the first line, and the unitarity of $\mathbf{Q}_k$ is used in the last line.

From (5.13) it is clear that the transmission channel has been transformed into a lower triangular channel $\mathbf{R}_k^T$. This channel is causal in the sense that there is an order in the crosstalk of the users. User 1 experiences crosstalk from no-one; user 2 experiences crosstalk only from user 1; user 3 experiences crosstalk only from users 1 and 2; and so on.

This causal structure admits the use of the Tomlinson-Harashima precoder to precompensate for the effects of crosstalk. User 1 experiences no crosstalk. Hence the signal of user 1 can be transmitted directly; that is

$$\acute{x}_k^1 = \widetilde{x}_k^1,$$

where $\widetilde{x}_k^n$ denotes the true symbol intended for user n. At this point the signal transmitted by user 1 is known. This allows the remaining users to predistort their signals, and annihilate the crosstalk introduced by user 1. User 2 then operates free from crosstalk. The signal transmitted by user 2 is known, which allows the remaining users to predistort their signals, and annihilate the crosstalk introduced by user 2. This procedure iterates until all users have predistorted their signals to annihilate all crosstalk introduced in the channel. Each user is then detected free from crosstalk. The symbol transmitted by user n on tone k after Tomlinson-Harashima precoding is thus

$$\acute{x}_k^n = \text{mod}_{s_k^{\text{mask}}}\left[\widetilde{x}_k^n - \sum_{m=1}^{n-1} \frac{r_k^{m,n}}{r_k^{n,n}} \acute{x}_k^m\right], \tag{5.14}$$

where the modulo operation is defined

$$\text{mod}_M[x] \triangleq x - \sqrt{M}\left\lfloor \frac{x + \sqrt{M}/2}{\sqrt{M}} \right\rfloor,$$

$\lfloor . \rfloor$ denotes the rounding-down operation[59], and $r_k^{n,m} \triangleq [\mathbf{R}_k]_{n,m}$. The modulo ensures that spectral mask compliance is maintained after precoding. That is, if $\widetilde{\mathbf{x}}_k$ obeys the spectral masks, then $\acute{\mathbf{x}}_k$ does as well.

At the receiver a second modulo operation is applied to estimate the transmitted symbol

$$\begin{aligned}
\hat{x}_k^n &= \mathrm{mod}_{s_k^{\mathrm{mask}}}\left[\frac{y_k^n}{r_k^{n,n}}\right], \\
&= \mathrm{mod}_{s_k^{\mathrm{mask}}}\left[\acute{x}_k^n + \sum_{m\neq n}\frac{r_k^{m,n}}{r_k^{n,n}}\acute{x}_k^m + \frac{z_k^n}{r_k^{n,n}}\right], \\
&= \mathrm{mod}_{s_k^{\mathrm{mask}}}\left[\widetilde{x}_k^n + \frac{z_k^n}{r_k^{n,n}}\right], \\
&= \widetilde{x}_k^n + \frac{\widetilde{z}_k^n}{r_k^{n,n}}.
\end{aligned}$$

The second line makes use of (5.13) and the property $\mathrm{mod}\left(a \times \mathrm{mod}(b)\right) = \mathrm{mod}\left(ab\right)$. The third line makes use of (5.14) and the property

$$\mathrm{mod}\left(\mathrm{mod}(a) + \mathrm{mod}(b)\right) = \mathrm{mod}\left(a + b\right).$$

The effective noise $\widetilde{z}_k^n$ in the fourth line is similar to the original noise z_k^n except that it exhibits a wrap-around effect on the edges of the QAM constellation. If the QAM constellation has many symbols, the wrap-around effect is rare and has negligible impact on performance[59]. Under this assumption the data-rate achieved by user n on tone k is

$$b_{k,\mathrm{th}}^n = \log_2(1 + (\sigma_k^n)^{-1}\,\Gamma^{-1}\widetilde{s}_k^n\,|r_k^{n,n}|^2).$$

The RWDD of the channel matrix can be used to show that $|r_k^{n,n}| \simeq |h_k^{n,n}|$[59]. As a result, for small α_k, the THP operates very close to the single-user bound

$$b_{k,\mathrm{th}}^n \simeq b_{k,\mathrm{bc-dsl}}^n.$$

Whilst the THP gives near-optimal performance, its major drawback is that it requires a modulo operation at the receiver. This requires a change of the CP modem design. Millions of CP modems are currently in use, owned and operated by a multitude of customers. Replacing these modems presents a huge legacy issue. Furthermore, CO modems and CP modems are typically manufactured by different hardware vendors, making joint designs difficult.

5.6 Near-optimal Linear Precoder

To address the problems of the THP, this section presents a simple linear precoder that works with existing CP modems. Like the THP, this precoder oper-

ates close to the single-user bound. The *diagonalizing precoder* (DP) multiplies the true symbols $\widetilde{\mathbf{x}}_k$ with a precoding matrix $\mathbf{P}_{k,\mathrm{dp}}$ prior to transmission. The transmitted symbols are

$$\mathbf{x}_k = \mathbf{P}_{k,\mathrm{dp}}\widetilde{\mathbf{x}}_k. \tag{5.15}$$

The DP is based on a channel diagonalizing criterion. After precoding, each user should see their own direct channel free from crosstalk. Contrast this with the ZFP where after precoding each user experiences a channel gain of unity, scaled by $\beta_{k,\mathrm{zf}}^{-1}$. The DP precoding matrix is defined

$$\mathbf{P}_{k,\mathrm{dp}} \triangleq \frac{1}{\beta_{k,\mathrm{dp}}}\mathbf{H}_k^{-1}\mathrm{diag}\left\{h_k^{1,1},\ldots,h_k^{N,N}\right\},$$

where the scaling factor is

$$\beta_{k,\mathrm{dp}} \triangleq \max_n \left\| \left[\mathbf{H}_k^{-1}\right]_{\mathrm{row}\ n} \mathrm{diag}\left\{h_k^{1,1},\ldots,h_k^{N,N}\right\}\right\|. \tag{5.16}$$

As with the ZF precoder, the scaling factor $\beta_{k,\mathrm{dp}}$ ensures that after precoding compliance with the spectral masks (5.2) is maintained. That is, if $\widetilde{\mathbf{x}}_k$ obeys the spectral masks, then $\mathbf{x}_k$ will as well.

During transmission the predistortion introduced by the DP annihilates with the crosstalk. The received vector is then

$$\begin{aligned}\mathbf{y}_k &= \mathbf{H}_k\mathbf{P}_{k,\mathrm{dp}}\widetilde{\mathbf{x}}_k + \mathbf{z}_k,\\ &= \beta_{k,\mathrm{dp}}^{-1}\mathrm{diag}\left\{h_k^{1,1},\ldots,h_k^{N,N}\right\}\widetilde{\mathbf{x}}_k + \mathbf{z}_k.\end{aligned} \tag{5.17}$$

Application of the DP diagonalizes the channel matrix. Each user now experiences their direct channel, scaled by $\beta_{k,\mathrm{dp}}$ and completely free from interference. RWDD in the crosstalk channel matrix implies that $\beta_{k,\mathrm{dp}} \simeq 1$. As a result, each user operates close to their single-user bound, and the DP is a near-optimal precoding structure. To see this consider the *singular value decomposition* (SVD) of $\mathbf{H}_k$

$$\mathbf{H}_k \overset{\mathrm{svd}}{=} \mathbf{U}_k\Lambda_k\mathbf{V}_k^H.$$

The RWDD of $\mathbf{H}_k$ (5.3) ensures that its rows are approximately orthogonal

$$\overline{\mathbf{h}}_k^{m\,H}\overline{\mathbf{h}}_k^{n} \simeq \begin{cases}|h_k^{n,n}|^2, & n=m;\\ 0, & n\neq m.\end{cases}$$

As a result

$$\begin{aligned}\mathbf{U}_k\Lambda_k\Lambda_k^H\mathbf{U}_k^H &= \mathbf{H}_k\mathbf{H}_k^H\\ &\simeq \mathrm{diag}\left\{\left|h_k^{1,1}\right|^2,\ldots,\left|h_k^{N,N}\right|^2\right\}.\end{aligned}$$

This implies that the left singular vectors can be closely approximated as $\mathbf{U}_k \simeq \mathbf{I}_N$. Furthermore, the singular values can be approximated

$$\begin{aligned} \left|[\Lambda_k]_{n,n}\right| &\simeq \left\|\overline{\mathbf{h}}_k^n\right\|, \\ &\simeq |h_k^{n,n}|. \end{aligned} \tag{5.18}$$

So

$$\mathbf{H}_k \simeq \Lambda_k \mathbf{V}_k^H. \tag{5.19}$$

This indicates that the channel can be approximately diagonalized by precoding with the matrix $\mathbf{V}_k$. Since $\mathbf{V}_k$ is unitary the precoding operation will maintain compliance with the spectral masks. The channel is approximately diagonalized, so each user experiences a channel that is almost crosstalk free.

Equation (5.19) leads to the following approximation for the precoding matrix

$$\begin{aligned} \mathbf{P}_{k,\mathrm{dp}} &\simeq \frac{1}{\beta_{k,\mathrm{dp}}} \mathbf{V_k} \mathbf{\Lambda_k^{-1}} \mathrm{diag}\left\{h_k^{1,1}, \ldots, h_k^{N,N}\right\}, \\ &\simeq \frac{1}{\beta_{k,\mathrm{dp}}} \mathbf{V}_k, \end{aligned}$$

where (5.18) is applied in the second line. The motivation behind the DP design is now clear. Since $\mathbf{P}_{k,\mathrm{dp}} \simeq \mathbf{V}_k$, the DP precoding matrix is close to unitary. Hence application of the DP causes negligible increase in the transmit power. Scaling with $\beta_{k,\mathrm{dp}}$ is not necessary to maintain compliance with the spectral masks, and $\beta_{k,\mathrm{dp}} \simeq 1$. As a result the DP achieves near-optimal performance, operating close to the single-user bound. This observation is made rigorous through the following theorem.

Theorem 5.2 *If $A_{min}^{(m)} \geq \alpha_k m B_{max}^{(m)}$, $m = 1 \ldots N-1$; the data-rate achieved by the DP can be lower bounded*

$$R_n \geq f_s \sum_k b_{k,\mathrm{dp-bound}}^n, \tag{5.20}$$

where

$$b_{k,\mathrm{dp-bound}}^n \triangleq \log_2\left(1 + \Gamma^{-1} \left(\sigma_k^n\right)^{-1} \widetilde{s}_k^n \left|h_k^{n,n}\right|^2 f^{-1}(N, \alpha_k)\right),$$

and $f(N, \alpha_k)$, $A_{min}^{(n)}$ and $B_{\max}^{(n)}$ are defined in (4.13), (4.15) and (4.14) respectively.

Proof: Equation (5.17) implies that after application of the DP the signal at receiver n is

$$y_k^n = \beta_{k,\mathrm{dp}}^{-1} h_k^{n,n} \widetilde{x}_k^n + z_k^n.$$

Hence the received signal power for user n on tone k is $\beta_{k,\mathrm{dp}}^{-2}\widetilde{s}_k^n\left|h_k^{n,n}\right|^2$, the received interference power is zero, and the received noise power is σ_k^n. So the data-rate achieved by the diagonalizing precoder is

$$b_{k,\mathrm{dp}}^n(\widetilde{s}_k^n) = \log_2\left(1+\Gamma^{-1}\left(\sigma_k^n\right)^{-1}\beta_{k,\mathrm{dp}}^{-2}\widetilde{s}_k^n\left|h_k^{n,n}\right|^2\right). \tag{5.21}$$

Define the matrix $\overline{\mathbf{G}}_k \triangleq [\overline{g}_k^{n,m}]$, where $\overline{g}_k^{n,m} \triangleq h_k^{n,m}/h_k^{n,n}$. Now

$$\mathbf{H}_k = \mathrm{diag}\left\{h_k^{1,1},\ldots,h_k^{N,N}\right\}\overline{\mathbf{G}}_k,$$

hence

$$\mathbf{H}_k^{-1}\mathrm{diag}\left\{h_k^{1,1},\ldots,h_k^{N,N}\right\} = \overline{\mathbf{G}}_k^{-1}. \tag{5.22}$$

Since the transmitters are co-located at the CO, the DS channel is RWDD (5.4). This implies that $\overline{\mathbf{G}}_k \in \mathbb{A}^{(N)}$, where $\mathbb{A}_n^{(N)}$ denotes the set of $N \times N$ diagonally dominant matrices, as defined in Appendix B. Theorem B.2 can be applied to bound the elements of $\overline{\mathbf{G}}_k^{-1}$ as follows

$$\left|\left[\overline{\mathbf{G}}_k^{-1}\right]_{n,m}\right| \leq \begin{cases} A_{\max}^{(N-1)}/A_{\min}^{(N)}, & n=m; \\ B_{\max}^{(N-1)}/A_{\min}^{(N)}, & n\neq m. \end{cases} \tag{5.23}$$

Combining (5.16) and (5.22) implies $\beta_{k,\mathrm{dp}}^2 \leq f(N,\alpha_{k,})$. Combining this with (5.21) leads to (5.20), which concludes the proof. ■

The function $f(N,\alpha_{k,n})$ can be interpreted as an upper bound on the scaling factor $\beta_{k,\mathrm{dp}}^2$. Recall that the scaling factor is the increase in transmit power that results from precoding with $\mathbf{H}_k^{-1}\mathrm{diag}\left\{h_k^{1,1},\ldots,h_k^{N,N}\right\}$. It is included to ensure that the DP does not increase the transmit power on any line. In RWDD channels $f(N,\alpha_k) \simeq 1$. As a result each modem operates at a rate

$$b_{k,\mathrm{zf}}^n \simeq \log_2\left(1+\Gamma^{-1}\sigma_k^{-1}\widetilde{s}_k^n\left|h_k^{n,n}\right|^2\right).$$

So the DP completely removes crosstalk with negligible increase in transmit power. A scaling factor close to unity can be chosen, and each user operates close to their single-user bound.

Note that the bound can be used to guarantee a data-rate without explicit knowledge of the crosstalk channels. This is because the bound only depends on the binder size, direct channel gain, and background noise power. Good models for these characteristics exist based on extensive measurement campaigns. Crosstalk channels on the other hand are poorly understood and actual channels can deviate significantly from the few empirical models that exist, see for example Fig. 2.5. This can make provisioning of services difficult.

Using the bound (5.20) allows us to overcome this problem. The bound tells us that the crosstalk channel gain is not important as long as RWDD is observed. RWDD is a well understood and modeled phenomenon. As a result (5.20) allows provisioning to be done in a reliable and accurate fashion.

The value for α_k from (5.5) is based on worst 1% case models. Hence for 99% of lines $\alpha_{k,n}$ will be smaller. So in 99% of lines a data-rate above the bound (5.20) is achieved. So the bound is a useful tool not just for theoretical analysis, but for provisioning of services as well.

Simulations in Section 5.8 use the bound together with (5.4) to show that the DP operates close to the single-user bound.

5.7 Spectra Optimization

Part I of this thesis discussed the coordination of DSL modems on a spectral level. Each modem generates its signal independently, however the transmit spectra are designed in a joint fashion to mitigate crosstalk.

When signal coordination is used crosstalk can be precoded at the transmitter side. It is, however, still interesting to optimize the transmit spectra of each modem to maximize performance. This can be viewed as the combination of signal and spectra coordination. Each transmitter is subject to a total power constraint

$$\sum_k s_k^n \leq P_n, \ \forall n. \tag{5.24}$$

As in Part I, the objective is to maximize the rate of user N, subject to target rate constraints on the other users in the network. Following the same development in Section 3.3.5, this can be reformulated as a weighted rate-sum optimization

$$\max_{\mathbf{s}_1,\ldots,\mathbf{s}_2} \sum_n w_n R_n \quad \text{s.t.} \quad \sum_k s_k^n \leq P_n, \ \forall n; \qquad s_k^n \geq 0, \ \forall n,k. \tag{5.25}$$

In contrast to Part I, R_n now represents the rate of modem n *with* crosstalk precoding. The data-rate R_n is a function of the transmit PSDs $\mathbf{s}_1, \ldots, \mathbf{s}_N$, and also depends on the type of crosstalk precoder used. If an optimal precoder is used, the objective function becomes convex[99]. Solving (5.25) then requires the solution of a KN-dimensional convex optimization. Although the cost function is convex, no closed form solution is known and numerical techniques must be used instead[99]. Conventional numerical optimization techniques, such as interior point methods, have a polynomial complexity in the dimensionality of

the search space. In ADSL $K = 256$, whilst in VDSL $K = 4096$. The resulting search thus has a high dimensionality, for which conventional optimization techniques are prohibitively complex.

A low complexity, iterative algorithm has been proposed for the special case where an unweighted rate-sum is maximized, that is $w_n = 1$ for all n[106]. Unfortunately, since this algorithm cannot optimize a weighted rate-sum, it cannot ensure that the target rates are achieved. These target rates are essential to ensure that each customer achieves their desired quality-of-service.

In this section a spectra coordination algorithm is developed for use with the DP. Since the DP removes all crosstalk, the spectrum coordination problem decouples into an independent power loading problem for each user. This reduces complexity considerably. Furthermore, Theorem 5.2 ensures that this approach operates close to the single-user bound.

5.7.1 Theoretical Capacity

We start by extending the single-user bound from Section 5.3 to DSL modems that may vary their transmit spectra under a total power constraint. The resulting upper bound is useful for evaluating crosstalk precoder performance with optimized spectra.

Theorem 5.3 *When the transmit PSD s_k^n is allowed to vary under a total power constraint (5.24), the capacity for user n can be upper bounded*

$$\begin{aligned} R_n \quad &\leq \quad \max_{\mathbf{s}_1,\ldots,\mathbf{s}_N} f_s \sum_k b_{k,\mathrm{bc-wf}}^n; \\ &\text{s.t.} \quad \sum_k s_k^m \leq P_m, \ \forall m; \\ &\qquad\quad s_k^m \geq 0, \ \forall m,k; \end{aligned} \tag{5.26}$$

where

$$b_{k,\mathrm{bc-wf}}^n \triangleq \log_2 \left(1 + (\sigma_k^n)^{-1} \Gamma^{-1} \left| h_k^{n,n} \right|^2 \left(\sqrt{s_k^n} + \alpha_k \sum_{m \neq n} \sqrt{s_k^m} \right)^2 \right).$$

Proof: Combining the single-user bound for broadcast channels (2.21) and RWDD (5.3) yields

$$b_k^n \leq \log_2 \left(1 + (\sigma_k^n)^{-1} \Gamma^{-1} \left| h_k^{n,n} \right|^2 \left(\sqrt{s_k^n} + \alpha_k \sum_{m \neq n} \sqrt{s_k^m} \right)^2 \right).$$

Including the total power constraints (5.24) leads to (5.26). ■

5.7.2 Near-optimal Linear Precoder

Transmit spectra optimization with the DP is now considered. The optimization problem is now

$$\begin{aligned} \widetilde{\mathbf{s}}_1^{\mathrm{dp}}, \ldots, \widetilde{\mathbf{s}}_N^{\mathrm{dp}} &= \arg \max_{\widetilde{\mathbf{s}}_1, \ldots, \widetilde{\mathbf{s}}_N} \sum_n \sum_k w_n b_{k,\mathrm{dp}}^n(\widetilde{s}_k^n); \\ \text{s.t.} \quad & \sum_k s_k^n \leq P_n, \ \forall n; \\ & s_k^n \geq 0, \ \forall n, k; \end{aligned} \tag{5.27}$$

where the optimal solution $\widetilde{\mathbf{s}}_n^{\mathrm{dp}} \triangleq [\widetilde{s}_1^{n,\mathrm{dp}}, \ldots, \widetilde{s}_K^{n,\mathrm{dp}}]$ and $\widetilde{\mathbf{s}}_n \triangleq [\widetilde{s}_1^n, \ldots, \widetilde{s}_K^n]$. Here $b_{k,\mathrm{dp}}^n$ denotes the bitloading achieved with the DP, which was defined in (5.21) as

$$b_{k,\mathrm{dp}}^n = \log_2 \left(1 + (\sigma_k^n)^{-1} \Gamma^{-1} \beta_{k,\mathrm{dp}}^{-2} \left|h_k^{n,n}\right|^2 \widetilde{s}_k^n\right), \tag{5.28}$$

Since the transmit spectra are being optimized spectral masks do not apply. The system instead operates under a total power constraint (5.24). As a result, the scaling factor $\beta_{k,\mathrm{dp}}$ can be discarded by setting

$$\beta_{k,\mathrm{dp}} = 1, \ \forall k. \tag{5.29}$$

From (5.15), the signal sent by transmitter n on tone k is

$$x_k^n = \sum_m p_{k,\mathrm{dp}}^{n,m} \widetilde{x}_k^m,$$

where the elements of the precoding matrix are defined $p_{k,\mathrm{dp}}^{n,m} \triangleq [\mathbf{P}_{k,\mathrm{dp}}]_{n,m}$. Hence the power on line n is

$$s_k^n = \sum_m \left|p_{k,\mathrm{dp}}^{n,m}\right|^2 \widetilde{s}_k^m. \tag{5.30}$$

Combining (5.28), (5.29) and (5.30), the original optimization problem (5.27) becomes

$$\begin{aligned} \widetilde{\mathbf{s}}_1^{\mathrm{dp}}, \ldots, \widetilde{\mathbf{s}}_N^{\mathrm{dp}} &= \arg \max_{\widetilde{\mathbf{s}}_1, \ldots, \widetilde{\mathbf{s}}_N} \sum_n \sum_k w_n \log_2 \left(1 + \sigma_k^{-1} \Gamma^{-1} \left|h_k^{n,n}\right|^2 \widetilde{s}_k^n\right); \\ \text{s.t.} \quad & \sum_k \sum_m \left|p_{k,\mathrm{dp}}^{n,m}\right|^2 \widetilde{s}_k^m \leq P_n, \ \forall n; \\ & \widetilde{s}_k^n \geq 0, \ \forall n, k. \end{aligned}$$

Observe that, when using the DP, the bitrate of each user depends only on its own transmit PSD. It is independent of the PSDs of the other users since all interference will be pre-filtered. Unfortunately, the optimization is still coupled between users. This is the case since the total power constraint on each modem

Algorithm 5.1 Optimal Power Allocation with the DP

repeat

for each n: $\widetilde{s}_k^n = \left[w_n \left(\sum_m \lambda_m p_{k,\mathrm{dp}}^{m,n} \right)^{-1} - \Gamma \sigma_k^n \left| h_k^{n,n} \right|^{-2} \right]^+, \ \forall k$

for each n: $\lambda_n = \left[\lambda_n + \epsilon \left(\sum_k \sum_m p_{\mathrm{dp},k}^{n,m} \widetilde{s}_k^m - P_n \right) \right]^+$

until convergence

must be satisfied after the precoding operation. As a result the PSD sent by a particular user $\widetilde{s}_k^n$ is not equal to the transmit PSD of the corresponding modem s_k^n. These PSDs are coupled through the precoding matrix $\mathbf{P}_{k,\mathrm{dp}}$ and as a result the optimization must be done jointly across all users. Nevertheless, it is still possible to optimize the transmit PSDs efficiently through the use of a dual objective.

First note that the objective function is concave and the constraints form a convex set. As a result the KKT conditions are sufficient for optimality. Examining these leads to

$$\widetilde{s}_{k,\mathrm{dp}}^n = \left[\frac{w_n}{\sum_m \lambda_m p_{k,\mathrm{dp}}^{m,n}} - \frac{\Gamma \sigma_k^n}{|h_k^{n,n}|^2} \right]^+ .$$

The power allocation of each user is coupled through the Lagrangian multipliers $\lambda_1, \ldots, \lambda_N$. The Lagrangian multipliers must be chosen such that for each line n the power constraint is tight

$$\sum_k \sum_m p_{k,\mathrm{dp}}^{n,m} \widetilde{s}_{k,\mathrm{dp}}^m = P_n,$$

or the corresponding Lagrangian multiplier λ_n is zero. An efficient solution can be found with Alg. 5.1, which has polynomial complexity in N and linear complexity in K[2]. This is a significant reduction compared to existing power allocation algorithms for the broadcast channel, which have polynomial complexity in KN[106].

Theorem 5.2 shows that, as a result of RWDD, the DP operates close to the single-user bound. So using the DP in combination with Alg. 5.1 gives near-optimal performance. This is confirmed through simulation in the following section.

[2]Note that this is a significant reduction in complexity when compared to the power allocation problem in Chapter 3. The basic reason is that power allocation in a BC typically involves a convex optimization, where-as power allocation in an IC involves a non-convex optimization.

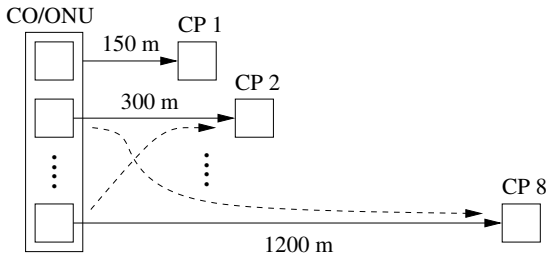

Figure 5.2: Downstream VDSL scenario

5.8 Performance

This section evaluates the performance of the DP in a binder of 8 VDSL lines. The line lengths range from 150m to 1200m in 150m increments as shown in Fig. 5.2. For all simulations the line diameter is 0.5 mm (24-AWG). Direct and crosstalk channel transfer functions are generated using the empirical models described in Section 2.1.2. The target symbol error probability is 10^{-7} or less. The coding gain is set to 3 dB and the noise margin to 6 dB. As per the VDSL standards the tone spacing Δ_f is set to 4.3125 kHz and DMT symbol rate f_s to 4 kHz[9][7]. The modems use 4096 tones, and the 998 FDD bandplan. Background noise is generated using ETSI noise model A[7]. Performance is compared with the ZFP, THP and the single-user bound.

5.8.1 Fixed Transmit Spectra

Current VDSL standards require that modems transmit under a spectral mask of -60 dBm/Hz[9][7]. This section evaluates the performance of the DP when all modems are operating at this mask.

Fig. 5.4 shows the data-rate achieved on each of the lines with the different crosstalk precoding schemes. As predicted, the ZFP gives quite poor performance, with all lines forced to operate at the rate of the weakest line in the binder, which in this case is the 1200m line. In fact, for all of the lines shorter than 1200m, the ZF precoder results in worse performance than with no precoding at all.

The DP avoids the problems of the ZFP, and achieves substantial gains, typically 30 Mbps or more, over conventional systems with no cancellation. As can be seen in Fig. 5.4 the DP achieves near-optimal performance, operating close to the single-user bound. This is a direct result of the RWDD of $\mathbf{H}_k$, which ensures that the scaling parameter $\beta_{k,\mathrm{dp}}$ is always close to unity. The scaling parameters of the ZFP $\beta_{k,\mathrm{zf}}$ and the DP $\beta_{k,\mathrm{dp}}$ are plotted for each tone in Fig.

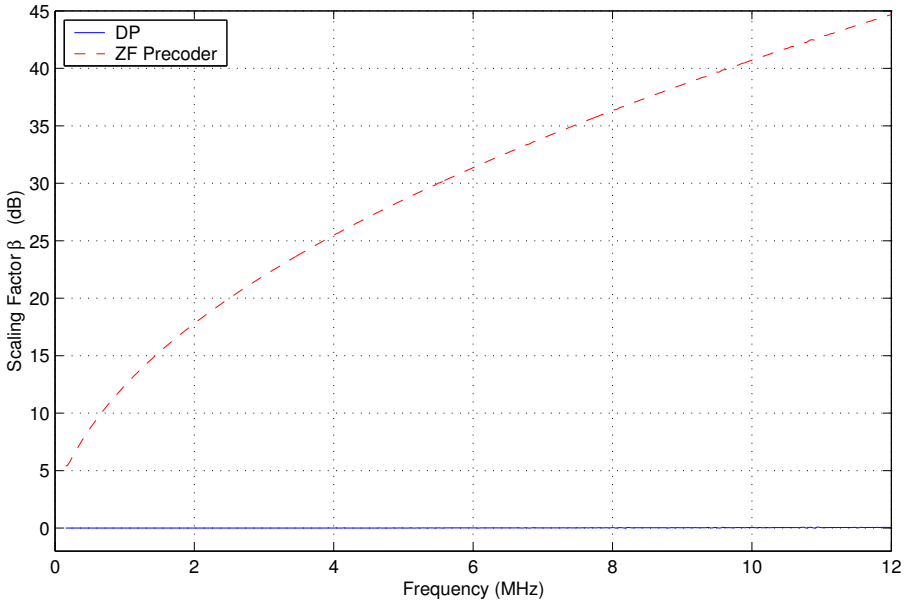

Figure 5.3: Scaling Factor β_k

5.3. The scaling parameter in the ZFP is quite large, which results in poor performance on the shorter lines. On the other hand the scaling parameter in the DP is typical close to unity, which has negligible effect on performance.

Fig. 5.5 shows the data-rate of the DP as a percentage of the single-user bound. Performance does not drop below 99% of the single-user bound. The lower bound on the performance of the DP (5.20) is also included for comparison. As can be seen the bound is quite tight and guarantees that the DP will achieve at least 97% of the single-user bound.

5.8.2 Optimized Transmit Spectra

Whilst current VDSL standards require the use of spectral masks, there is growing interest in the use of adaptive transmit spectra[8]. This section investigates the performance of the DP with spectra optimized according to Alg. 5.1.

The data-rates achieved on each line are shown in Fig. 5.6. The use of optimized spectra yields a gain of 5-8 Mbps. The benefit is more substantial on the longer lines, where a 5 Mbps gain can double the achievable bitrate.

Fig. 5.6 shows that crosstalk precoding gives maximum benefit on short lines,

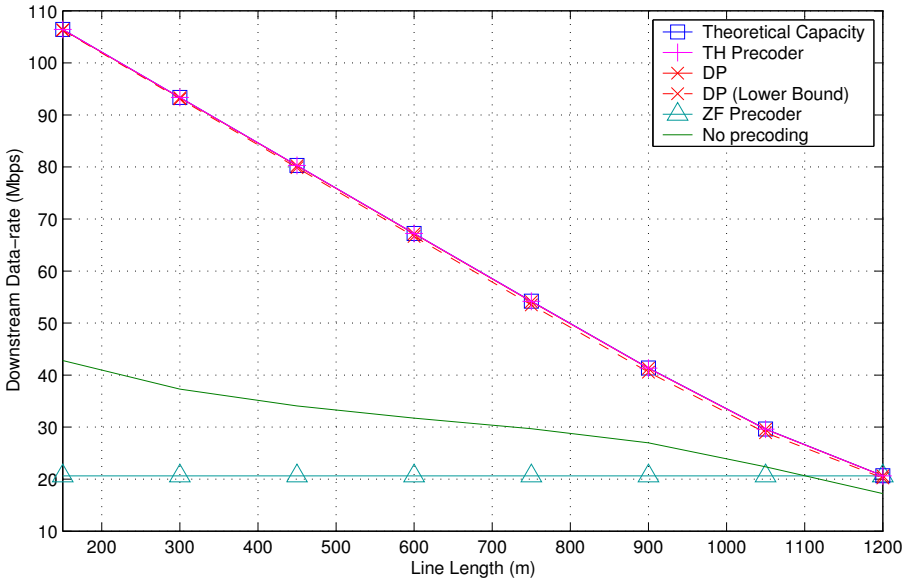

Figure 5.4: Downstream Data-rate of Different Precoders

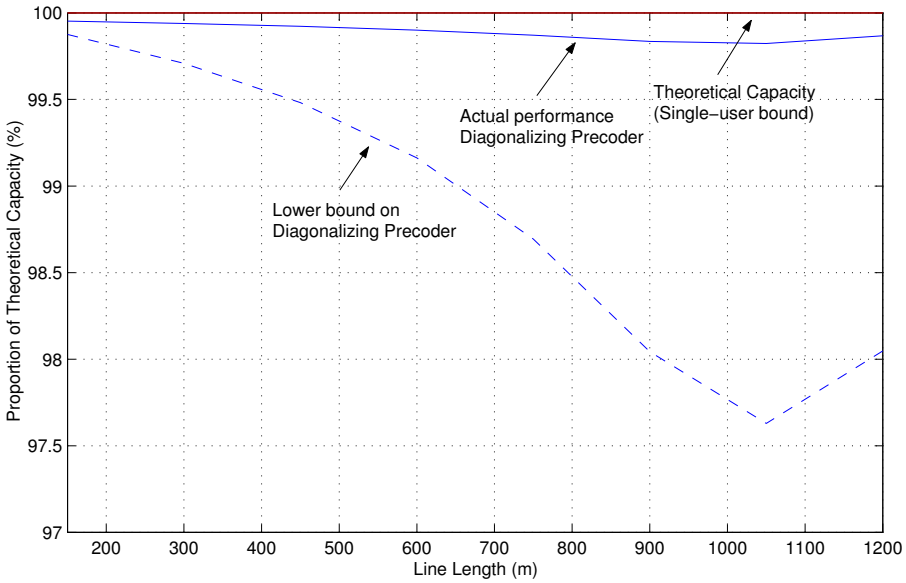

Figure 5.5: Proportion of Theoretical Capacity Achieved by DP

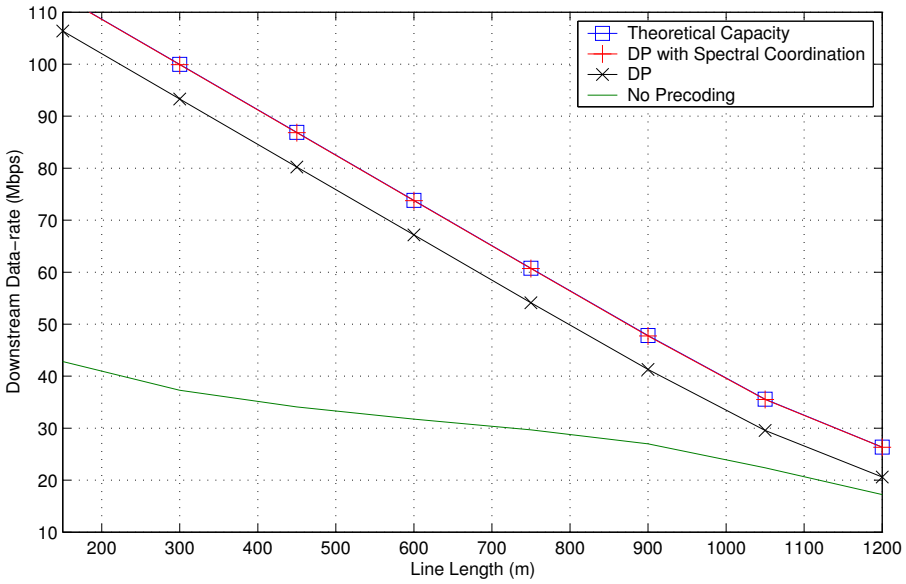

Figure 5.6: Downstream Data-rate Achieved with Coordinated Transmit Spectra

whilst spectra coordination gives maximum benefit on long lines. This is to be expected since short lines are crosstalk limited, whilst long lines are noise limited.

The single-user bound (5.26) is included for comparison. As can be seen the DP operates close to the single-user bound. Fig. 5.7 shows the data-rate achieved by the DP as a percentage of the single-user bound. Performance does not drop below 99% of the single-user bound. The lower bound on the performance of the DP is also included for comparison. The bound (5.20) guarantees that the DP with optimized spectra from Alg. 5.1 achieves 97% of the single-user bound.

5.9 Summary

The previous chapter investigated crosstalk cancellation for upstream DSL, a technique based on receiver coordination. In downstream transmission, receivers are not co-located, and crosstalk cancellation is impossible. Instead crosstalk must be mitigated through transmitter coordination, a technique known as crosstalk precoding.

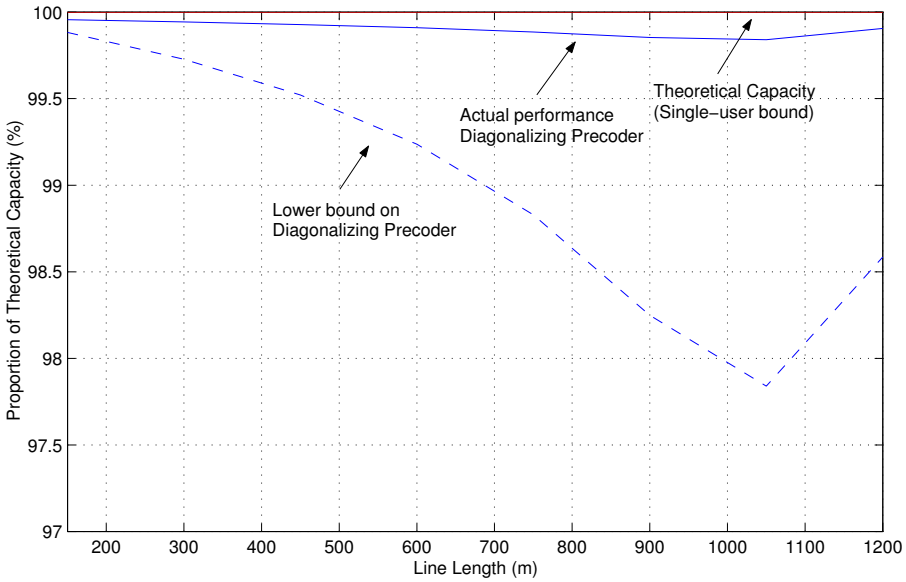

Figure 5.7: Proportion of Theoretical Capacity Achieved by ZF Canceler

This chapter investigated the design of crosstalk precoders for DSL. Existing designs suffer from poor performance or require the replacement of customer premises modems, which presents a huge legacy issue.

A novel linear precoder based on a channel diagonalizing criterion was proposed. This design has a low complexity and works with existing customer premises modems.

Any linear crosstalk precoder must include a scaling factor to ensure that spectral masks are maintained after precoding. In some cases this scaling factor decreases performance by forcing certain modems to operate below their transmit mask. Fortunately DSL channels with co-located transmitters are row-wise diagonal dominant. This ensures that the scaling factor of the proposed precoder is close to unity, and has negligible impact on performance.

An upper bound on the capacity of the multi-user DSL channel was derived. This single-user bound shows that spatial diversity in the DSL environment is negligible. Therefore the best outcome that a crosstalk precoder can achieve is the complete pre-filtering of crosstalk without increasing the transmit power.

A lower bound on the data-rate of the *diagonalizing precoder* (DP) was derived. This bound depends only on the binder size, direct channel gain and background noise for which reliable models and statistical data exist. As a

result the performance of the DP can be accurately predicted, which simplifies service provisioning considerably. This bound shows that the DP operates close to the single-user bound. So the DP is a low complexity design that works with existing customer premises modems, and has guaranteed near-optimal performance.

The combination of spectral optimization and crosstalk precoding was considered. The bounds were extended to DSL systems with optimized spectra. Spectra optimization in a broadcast channel generally involves a highly complex optimization. Since the DP decouples transmission on each line, the spectrum on each modem can be optimized through a dual decomposition, leading to a significant reduction in complexity.

Chapter 6

Partial Coordination

6.1 Introduction

Crosstalk is *the* dominant source of performance degradation in modern DSL systems. In upstream transmission, *receivers* (RX) can be coordinated on a signal level since they are co-located at the *central office* (CO). This allows crosstalk cancellation to be employed. The received signals on different modems are combined such that crosstalk is cancelled and only the signal of interest remains. Chapter 4 showed that a simple linear canceler, based on the ZF criterion, achieves near-optimal performance and leads to spectacular performance gains[1].

In downstream transmission, *transmitters* (TX) can be coordinated on a signal level since they are co-located at the CO. This allows crosstalk precoding to be employed. The symbols to be transmitted are predistorted, such that the predistortion and crosstalk annihilate during transmission. Chapter 5 showed that a simple linear precoder, based on the channel diagonalizing criterion, achieves near-optimal performance and leads to spectacular performance gains.

Whilst the benefits of crosstalk cancellation and precoding are large, the complexity is extremely high, even with the simple linear schemes described above. For example, in a bundle with 20 users transmitting on 4096 tones and operating at a block rate of 4000 blocks/s the complexity of linear crosstalk cancellation exceeds 6.5 billion multiplications/s. This is outside the scope of current implementation and may remain so for several years. Other techniques such as soft-interference cancellation and non-linear crosstalk cancellation add even

[1]The work in this chapter was done in close collaboration with Dr. George Ginis, Texas Instruments, San Jose, CA.

more complexity.

To address this problem, this chapter develops the concepts of partial cancellation and precoding. It is well known that the majority of the crosstalk experienced by a modem comes from only a few dominant crosstalkers within the binder. Furthermore, since crosstalk coupling varies dramatically with frequency, the worst effects are limited to a small subset of tones. By exploiting these properties, partial cancelers and precoders achieve the majority of the performance of full crosstalk cancellation at a fraction of the run-time complexity.

The rest of this chapter is organised as follows: The system model is described in Sec. 6.2. Sec. 6.3 discusses the selectivity of crosstalk, and shows that the majority of crosstalk experienced by a line comes from only a few crosstalkers within the binder and occurs over only a small selection of tones. This crosstalk selectivity is exploited by partial cancelers and precoders to reduce run-time complexity. Sec. 6.4 describes the design of partial cancelers and shows that, in order to achieve maximum reduction in run-time complexity, both the space and frequency selectivity of crosstalk must be exploited. Partial precoder design is discussed in Sec. 6.5. The performance of the different partial cancelers and precoders is evaluated in Sec. 6.7.

6.2 System Model

Since either transmitting or receiving modems are co-located, it is straightforward to synchronize the modems. Transmission can then be modeled independently on each tone as described in Chapter 2

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k. \tag{6.1}$$

In this chapter the number of lines in the binder is defined as $N+1$ to simplify notation. So the matrix $\mathbf{H}_k$ is of size $(N+1) \times (N+1)$, whilst the vectors $\mathbf{y}_k$, $\mathbf{x}_k$ and $\mathbf{z}_k$ are of length $N+1$. Typically the noise is spatially white, and we make this assumption here

$$\mathcal{E}\left\{\mathbf{z}_k \mathbf{z}_k^H\right\} = \sigma_k \mathbf{I}_N. \tag{6.2}$$

6.3 Crosstalk Selectivity

Fig. 6.2 shows a selection of crosstalk channels from a set of measurements of a 24 AWG cable. As can be seen the severity of crosstalk varies significantly with both frequency and space. We make two observations:

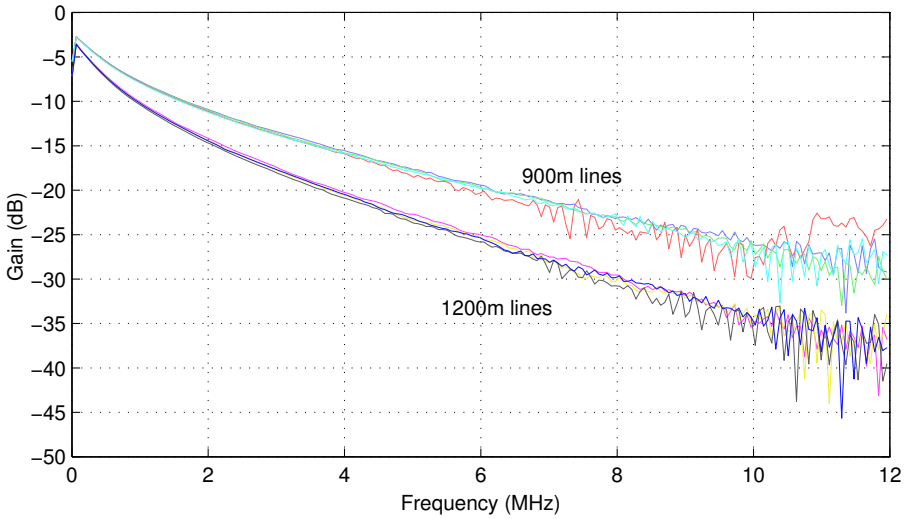

Figure 6.1: Direct channels from a measured binder

First, since electromagnetic coupling follows a distanced squared law, the majority of the crosstalk that a line experiences comes from the 4 or 5 surrounding lines within the binder. We refer to this as the *space-selectivity* of crosstalk. This is illustrated in Fig. 6.3. The near-far effect also gives rise to space-selectivity; in upstream transmission, modems that are located closer to the CO will cause more crosstalk than those located further away.

To illustrate the space-selectivity of crosstalk we calculated the proportion of crosstalk caused by the i largest crosstalkers into user n on tone k. All users had identical transmit PSDs, so crosstalker m on tone k was larger than crosstalker m' at tone k if $|h_k^{n,m}| > |h_k^{n,m'}|$. The result was averaged across all tones k and every line n within the binder and is plotted in Fig. 6.4. As can be seen close to 80% of the crosstalk energy is caused by the 3 largest crosstalkers.

Second, the crosstalk coupling varies significantly with frequency. Electromagnetic coupling increases with frequency and reflections within the binder cause nulls in the transfer function. We refer to this as the *frequency selectivity* of crosstalk.To illustrate this we evaluated the proportion of crosstalk contained within the i strongest tones between TX n and RX m. Tone k is said to be stronger than tone k' if $|h_k^{n,m}| > |h_{k'}^{n,m}|$. The result was averaged across all TXs n and RXs m and is plotted in Fig 6.5. As can be seen almost 80% of the crosstalk energy is contained within half of the tones.

So the effects of crosstalk vary considerably with both space and frequency. Furthermore, the majority of its effects are contained within a small selection

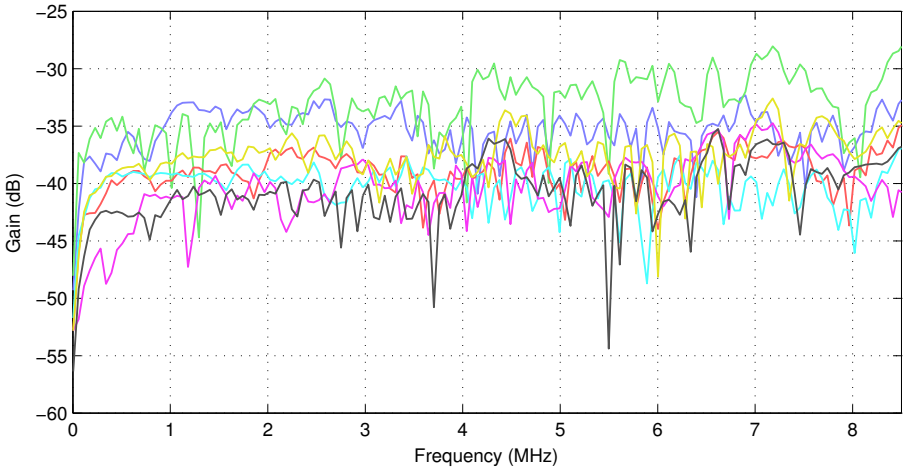

Figure 6.2: Crosstalk channels from a measured binder

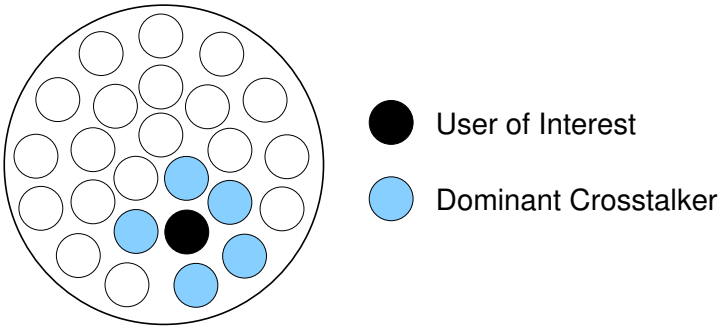

Figure 6.3: Geometry of a 25-pair Bundle

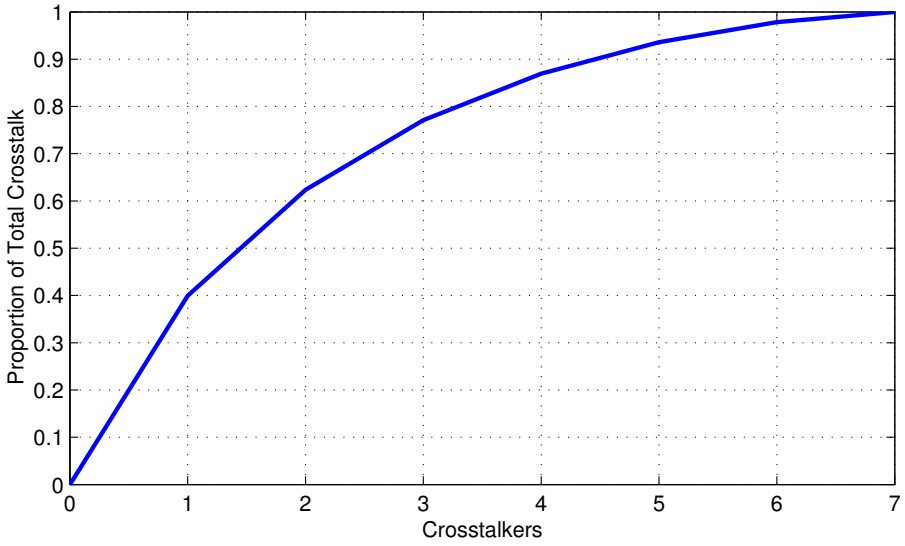

Figure 6.4: Proportion of Crosstalk caused by i largest crosstalkers

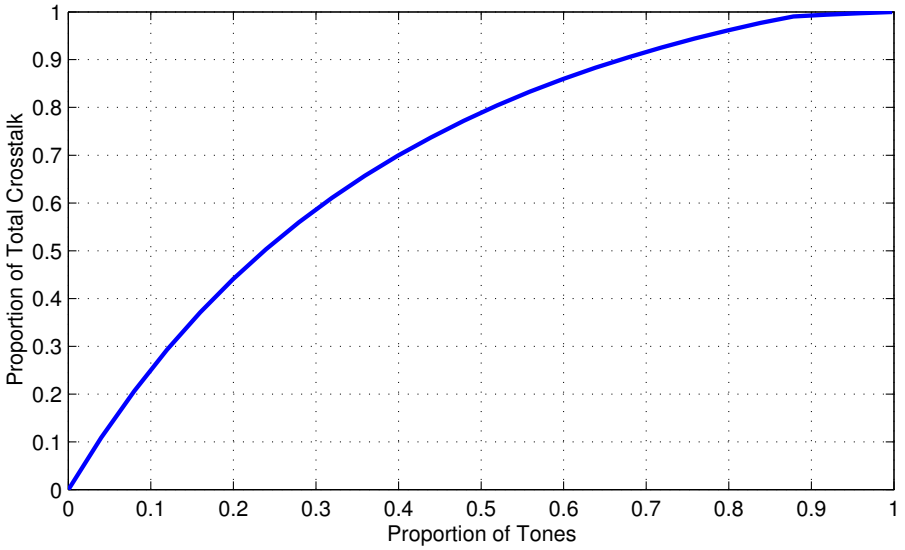

Figure 6.5: Proportion of Crosstalk contained within i worst tones

of crosstalkers and tones. These observations suggest that the majority of the performance of full crosstalk cancellation can be achieved by limiting cancellation to the largest crosstalkers, a technique known as partial cancellation. Some tones will see more significant crosstalkers than others. An efficient design should scale between conventional single-user detection (SUD) and full crosstalk cancellation on a tone-by-tone basis. On each tone the design should choose the appropriate degree of crosstalk cancellation based on the severity of crosstalk experienced. By only canceling the largest crosstalkers and by varying the degree of crosstalk cancellation on each tone, partial cancellation can approach the performance of full cancellation at a fraction of the run-time complexity.

6.4 Partial Receiver Coordination

6.4.1 Principle

This section investigates the design of partial cancelers, which are used in upstream transmission where the RXs are co-located at the CO. The estimate of the transmitted symbols is formed through a linear combination of the received signals

$$\hat{\mathbf{x}}_k = \mathbf{W}_k \mathbf{y}_k.$$

In the detection of user n, the RX observes the direct line of user n to recover the signal, and $r_{k,n}$ additional lines to enable crosstalk cancellation. The parameter $r_{k,n}$ varies both with the tone k and user n to match the severity of crosstalk experienced by each user on each tone. Note that $r_{k,n} = N$ corresponds to full cancellation, whilst $r_{k,n} = 0$ corresponds to no cancellation. Define the set of extra observation lines

$$\mathbb{M}_k^n \triangleq \{m_{k,n}(1), \ldots, m_{k,n}(r_{k,n})\}.$$

In a partial canceler, the estimate of the transmitted symbol is formed using a linear combination of the received signals on the *observation lines only.* The received signals on the other lines are not used in estimating x_k^n. Hence $\mathbf{W}_k$ has a sparse structure

$$[\mathbf{W}_k]_{n,m} = 0, \ \forall m \notin \{n, \mathbb{M}_k^n\}. \tag{6.3}$$

A ZF design is chosen for the partial canceler, since it was shown in Sec. 4.5 to achieve near-optimal performance. Under the ZF criterion, the partial cancellation filter removes all crosstalk from crosstalkers in the set $\mathbb{M}_k^n$. Hence

$$[\mathbf{W}_k \mathbf{H}_k]_{n,m} = \begin{cases} 1, & m = n; \\ 0, & m \in \mathbb{M}_k^n. \end{cases} \tag{6.4}$$

Note that, due to the sparseness of $\mathbf{W}_k$, crosstalk cancellation for user n on tone k now requires only $r_{k,n}$ multiplications/DMT block. Recall that $r_{k,n}$ denotes the number of crosstalkers cancelled when detecting user n on tone k. Compare this with the N multiplications required for full cancellation. This technique has many similarities to hybrid selection/combining from the wireless field[64, 65, 63, 66], where selection is also used between receive antennas to reduce complexity and the number of analog front-ends (AFE) required.

6.4.2 Design

The design of the partial cancellation filter $\mathbf{W}_k$ is now described. In the detection of user n on tone k the signals received on lines $\{n, \mathbb{M}_k^n\}$ are used. Define the vector containing the corresponding received symbols

$$\overline{\mathbf{y}}_k^{\mathrm{rx}\,n} \triangleq \left[y_k^n,\, y_k^{m_{k,n}(1)} \cdots\, y_k^{m_{k,n}(r_{k,n})}\right]^T,$$

the vector containing the corresponding transmitted symbols

$$\overline{\mathbf{x}}_k^{\mathrm{rx}\,n} \triangleq \left[x_k^n,\, x_k^{m_{k,n}(1)} \cdots\, x_k^{m_{k,n}(r_{k,n})}\right]^T,$$

and the vector containing the corresponding noise

$$\overline{\mathbf{z}}_k^{\mathrm{rx}\,n} \triangleq \left[z_k^n,\, z_k^{m_{k,n}(1)} \cdots\, z_k^{m_{k,n}(r_{k,n})}\right]^T.$$

Define the corresponding partial channel matrix $\overline{\mathbf{H}}_k^{\mathrm{rx}\,n}$

$$\overline{\mathbf{H}}_k^{\mathrm{rx}\,n} \triangleq \begin{bmatrix} h_k^{n,n} & [\mathbf{H}_k]_{\mathrm{row}\,n,\,\mathrm{cols}\,\mathbb{M}_k^n} \\ [\mathbf{H}_k]_{\mathrm{rows}\,\mathbb{M}_k^n,\,\mathrm{col}\,n} & [\mathbf{H}_k]_{\mathrm{rows}\,\mathbb{M}_k^n,\,\mathrm{cols}\,\mathbb{M}_k^n} \end{bmatrix}, \tag{6.5}$$

where $[\mathbf{A}]_{\mathrm{rows}\,\mathbb{X},\,\mathrm{cols}\,\mathbb{Y}}$ denotes the sub-matrix formed from the rows $\mathbb{X}$ and columns $\mathbb{Y}$ of matrix $\mathbf{A}$. Define the set of lines not observed in the detection of user n on tone k

$$\begin{aligned} \underline{\mathbb{M}}_k^n &\triangleq \{1, \ldots, n-1,\ n+1, \ldots, N+1\} \setminus \mathbb{M}_k^n, \\ &= \left\{\underline{m}_{k,n}(1), \ldots, \underline{m}_{k,n}(N - r_{k,n})\right\}, \end{aligned}$$

where $\mathbb{A} \setminus \mathbb{B}$ denotes the elements contained in set $\mathbb{A}$ and not in set $\mathbb{B}$. Define the vector $\underline{\mathbf{x}}_k^n$ containing the corresponding transmitted symbols

$$\underline{\mathbf{x}}_k^{\mathrm{rx}\,n} \triangleq \left[x_k^{\underline{m}_{k,n}(1)} \cdots\, x_k^{\underline{m}_{k,n}(N-r_{k,n})}\right]^T,$$

and the vector containing the corresponding noise

$$\underline{\mathbf{z}}_k^{\mathrm{rx}\,n} \triangleq \left[z_k^{\underline{m}_{k,n}(1)} \cdots\, z_k^{\underline{m}_{k,n}(N-r_{k,n})}\right]^T,$$

Define the corresponding partial channel matrix

$$\underline{\mathbf{H}}_k^{\mathrm{rx}\,n} \triangleq \begin{bmatrix} [\mathbf{H}_k]_{\mathrm{row}\,n,\,\mathrm{cols}\,\underline{\mathbb{M}}_k^n} \\ [\mathbf{H}_k]_{\mathrm{rows}\,\mathbb{M}_k^n,\,\mathrm{cols}\,\underline{\mathbb{M}}_k^n} \end{bmatrix}.$$

Consider the reduced system model, which only contains the signals observed in the detection of user n at tone k

$$\overline{\mathbf{y}}_k^{\mathrm{rx}\,n} = \overline{\mathbf{H}}_k^{\mathrm{rx}\,n}\overline{\mathbf{x}}_k^{\mathrm{rx}\,n} + \underline{\mathbf{H}}_k^{\mathrm{rx}\,n}\underline{\mathbf{x}}_k^{\mathrm{rx}\,n} + \overline{\mathbf{z}}_k^{\mathrm{rx}\,n}. \tag{6.6}$$

Define the $1 \times (r_{k,n}+1)$ vector that contains the non-zero elements of $\mathbf{W}_k$ used in the detection of user n as

$$\overline{\mathbf{w}}_k^n \triangleq \left[w_k^{n,n},\, w_k^{n,m_{k,n}(1)} \cdots w_k^{n,m_{k,n}(r_{k,n})}\right],$$

where $w_k^{n,m} \triangleq [\mathbf{W}_k]_{n,m}$. The estimate of user n's symbol is formed

$$\begin{aligned} \hat{x}_k^n &= [\mathbf{W}_k]_{\mathrm{row}\,n}\,\mathbf{y}_k, \\ &= \overline{\mathbf{w}}_k^n\overline{\mathbf{y}}_k^{\mathrm{rx}\,n}, \end{aligned} \tag{6.7}$$

where (6.3) is used in the second line. Under the ZF criterion (6.4), the partial cancellation filter removes all crosstalk from crosstalkers in the set $\mathbb{M}_k^n$. The ZF criterion for user n is equivalent to

$$\overline{\mathbf{w}}_k^n\overline{\mathbf{H}}_k^{\mathrm{rx}\,n} = \overline{\mathbf{e}}_{\mathrm{rx}\,1}^H,$$

where $\overline{\mathbf{e}}_{\mathrm{rx}\,n}$ is the nth column of the $(r_{k,n}+1) \times (r_{k,n}+1)$ identity matrix. Hence

$$\overline{\mathbf{w}}_k^n \triangleq \overline{\mathbf{e}}_{\mathrm{rx}\,1}^H \left(\overline{\mathbf{H}}_k^{\mathrm{rx}\,n}\right)^{-1}. \tag{6.8}$$

6.4.3 Achievable Rate

The rate achieved with the partial canceler is now analysed. From (6.6), (6.7) and (6.8), the estimate of the transmitted symbol is

$$\widehat{x}_k^n = x_k^n + \overline{\mathbf{w}}_k^n\underline{\mathbf{H}}_k^{\mathrm{rx}\,n}\underline{\mathbf{x}}_k^{\mathrm{rx}\,n} + \overline{\mathbf{w}}_k^{\mathrm{rx}\,n}\overline{\mathbf{z}}_k^{\mathrm{rx}\,n}. \tag{6.9}$$

The first term is the transmitted signal whilst the second and third terms are the residual crosstalk and filtered noise respectively. Define

$$\overline{\mathbf{G}}_k^{\mathrm{rx}\,n} = \overline{\mathbf{H}}_k^{\mathrm{rx}\,n}\mathrm{diag}\{h_k^{n,n}, h_k^{m_{k,n}(1),m_{k,n}(1)}, \ldots, h_k^{m_{k,n}(r_{k,n}),m_{k,n}(r_{k,n})}\}^{-1}. \tag{6.10}$$

Since the RXs are co-located at the CO, $\overline{\mathbf{H}}_k^{\mathrm{rx}\,n}$ is CWDD (4.4). This implies that $\overline{\mathbf{G}}_k^{\mathrm{rx}\,n} \in \mathbb{A}^{(N)}$, where $\mathbb{A}_n^{(N)}$ denotes the set of $N \times N$ diagonally dominant matrices, as defined in Appendix B. Now

$$\overline{\mathbf{H}}_k^{\mathrm{rx}\,n} = \overline{\mathbf{G}}_k^{\mathrm{rx}\,n}\mathrm{diag}\{h_k^{n,n}, h_k^{m_{k,n}(1),m_{k,n}(1)}, \ldots, h_k^{m_{k,n}(r_{k,n}),m_{k,n}(r_{k,n})}\}. \tag{6.11}$$

Combining this with (6.8) yields

$$\overline{\mathbf{w}}_k^n = (h_k^{n,n})^{-1} \left[\left(\overline{\mathbf{G}}_k^{\mathrm{rx}\,n} \right)^{-1} \right]_{\mathrm{row}\,1}. \tag{6.12}$$

Since $\overline{\mathbf{G}}_k^{\mathrm{rx}\,n} \in \mathbb{A}^{(N)}$, theorem B.2 can now be applied to bound the elements of $\overline{\mathbf{G}}_k^{\mathrm{rx}\,n}$. This implies

$$|[\overline{\mathbf{w}}_k^n]_i| \leq \begin{cases} |h_k^{n,n}|^{-1} A_{\max}^{(N-1)} / A_{\min}^{(N)}, & i = 1; \\ |h_k^{n,n}|^{-1} B_{\max}^{(N-1)} / A_{\min}^{(N)}, & \text{otherwise.} \end{cases}$$

CWDD implies that $B_{\max}^{(N-1)} \ll 1$, $A_{\max}^{(N-1)} \simeq 1$ and $A_{\min}^{(N)} \simeq 1$. Hence

$$|[\overline{\mathbf{w}}_k^n]_i| \simeq \begin{cases} |h_k^{n,n}|^{-1}, & i = 1; \\ 0, & \text{otherwise.} \end{cases}$$

From (6.9) the post-cancellation signal power is s_k^n and the post cancellation noise power is

$$\begin{aligned} \mathcal{E}\left\{ |\overline{\mathbf{w}}_k^n \mathbf{z}_k|^2 \right\} &= \|\overline{\mathbf{w}}_k^n\|_2^2 \, \sigma_k, \\ &\simeq |h_k^{n,n}|^{-2} \, \sigma_k, \end{aligned}$$

where (6.2) is applied in the first line. The residual interference

$$\begin{aligned} \overline{\mathbf{w}}_k^n \underline{\mathbf{H}}_k^{\mathrm{rx}\,n} \underline{\mathbf{x}}_k^{\mathrm{rx}\,n} &\simeq |h_k^{n,n}|^{-1} [\underline{\mathbf{H}}_k^{\mathrm{rx}\,n}]_{\mathrm{row}\,1} \underline{\mathbf{x}}_k^{\mathrm{rx}\,n}, \\ &= |h_k^{n,n}|^{-1} \sum_{m \in \underline{\mathbb{M}}_k^n} h_k^{n,m} x_k^m. \end{aligned}$$

The power of the residual interference is thus

$$\mathcal{E}\left\{ |\overline{\mathbf{w}}_k^{\mathrm{rx}\,n} \underline{\mathbf{H}}_k^{\mathrm{rx}\,n} \underline{\mathbf{x}}_k^{\mathrm{rx}\,n}|^2 \right\} \simeq |h_k^{n,n}|^{-2} \sum_{m \in \underline{\mathbb{M}}_k^n} |h_k^{n,m}|^2 \, s_k^m.$$

The *signal to interference plus noise ratio* (SINR) at the input of the decision device is thus

$$\mathrm{SINR}_k^n \simeq \frac{|h_k^{n,n}|^2 \, s_k^n}{\sum_{m \in \underline{\mathbb{M}}_k^n} |h_k^{n,m}|^2 \, s_k^m + \sigma_k}, \tag{6.13}$$

and the achievable data-rate for user n on tone k is

$$b_k^n = \log_2 \left(1 + \Gamma^{-1} \mathrm{SINR}_k^n \right).$$

There are two interesting observations to make a this point. First, as is expected, the partial canceler completely removes crosstalk from interferers in the set $\mathbb{M}_k^n$. More surprisingly however, the partial canceler does not change the statistics of the crosstalk from the interferers outside the set $\mathbb{M}_k^n$. It also does not change the statistics of the noise. So *the CWDD of* $\mathbf{H}_k$ *ensures that the ZF partial canceler does not enhance the power of the crosstalk from interferers outside* $\mathbb{M}_k^n$ *nor of the noise.*

6.5 Partial Transmitter Coordination

6.5.1 Principle

This section investigates the design of partial precoders, which are used in downstream transmission where the TXs are co-located at the CO. The precoded signal is formed through a linear combination of the symbols intended for each user

$$\mathbf{x}_k = \mathbf{P}_k \widetilde{\mathbf{x}}_k, \tag{6.14}$$

where $\widetilde{\mathbf{x}}_k$ denotes the vector of symbols intended for each RX. Partial precoder design is slightly different to partial canceler design as described in the previous section. Partial cancelers are designed in a row-wise fashion. RX n chooses which crosstalkers it would like to eliminate from its received signal, and then sets its crosstalk cancellation filter $[\mathbf{W}_k]_{\text{row } n}$ accordingly. Partial precoders must be designed in a column-wise fashion. TX m chooses which RXs it would like to protect from its own signal, and then sets its crosstalk precoding filter $[\mathbf{P}_k]_{\text{col } m}$ accordingly. Unfortunately it is very difficult to find a formulation that allows TX m to choose which RXs to protect, and the resulting optimization becomes extremely complex.

As an alternative, we chose a design procedure that begins with the RX, as in partial canceler design. Based on the potential benefit of removing crosstalk on a tone, each RX decides which TXs it would like to be protected from through precoding. Define the set of TXs that RX n would like to be protected from as

$$\mathbb{M}_k^n \triangleq \{m_{k,n}(1), \ldots, m_{k,n}(r_{k,n})\}.$$

Now, define the RXs that would like to be protected from TX m as

$$\mathbb{N}_k^m \triangleq \{n : m \in \mathbb{M}_k^n\} = \{n_{k,m}(1), \ldots, n_{k,m}(t_{k,m})\}, \tag{6.15}$$

where $t_{k,m}$ denotes the cardinality of $\mathbb{N}_k^m$ and is, in fact, the number of receivers who would like crosstalker m precoded out of their received signal on tone k. In a partial precoder, the signals on lines $\mathbb{N}_k^m$ are precoded against crosstalk from TX m. The signals on the other lines are not precoded against TX m. Hence $\mathbf{P}_k$ has a sparse structure

$$[\mathbf{P}_k]_{n,m} = 0, \ \forall n \notin \{m, \mathbb{N}_k^m\}. \tag{6.16}$$

A DP design is chosen for the partial precoder, since it was shown in Sec. 5.6 to achieve near-optimal performance. Under the DP criterion, the partial precoding filter prevents crosstalk from being caused to any of the receivers in the set $\mathbb{N}_k^m$. Hence

$$[\mathbf{H}_k \mathbf{P}_k]_{n,m} = \begin{cases} \beta_{k,\text{partial}}^{-1} h_k^{m,m}, & n = m; \\ 0, & n \in \mathbb{N}_k^m, \end{cases} \tag{6.17}$$

where the scaling factor $\beta_{k,\text{partial}}$ is added to ensure that spectral masks are maintained after precoding

$$\begin{aligned} \beta_{k,\text{partial}} &= \min \beta_{k,\text{partial}}, \\ &\text{s.t.} \quad \max_n \|[\mathbf{P}_k]_{\text{row } n}\| \leq 1. \end{aligned} \tag{6.18}$$

Note that, due to the sparseness of $\mathbf{P}_k$, crosstalk precoding for TX m on tone k now requires only $t_{k,m}$ multiplications/DMT block in contrast to the N multiplications required for full precoding.

6.5.2 Design

The design of the partial precoding filter $\mathbf{P}_k$ is now described. When precoding for crosstalk from TX m, the signals of TXs $\{m, \mathbb{N}_k^m\}$ are modified. Define the vector containing the corresponding received symbols

$$\overline{\mathbf{y}}_k^{\text{tx}\,m} \triangleq \left[y_k^m,\, y_k^{n_{k,m}(1)} \cdots y_k^{n_{k,m}(t_{k,m})}\right]^T,$$

the vector containing the corresponding transmitted symbols

$$\overline{\mathbf{x}}_k^{\text{tx}\,n} \triangleq \left[x_k^m,\, x_k^{n_{k,m}(1)} \cdots x_k^{n_{k,m}(t_{k,m})}\right]^T,$$

and the vector containing the corresponding noise

$$\overline{\mathbf{z}}_k^{\text{tx}\,m} \triangleq \left[z_k^m,\, z_k^{n_{k,m}(1)} \cdots z_k^{n_{k,m}(t_{k,m})}\right]^T.$$

Define the corresponding partial channel matrix $\overline{\mathbf{H}}_k^{\text{tx}\,m}$

$$\overline{\mathbf{H}}_k^{\text{tx}\,m} \triangleq \begin{bmatrix} h_k^{m,m} & [\mathbf{H}_k]_{\text{row } m,\, \text{cols}\, \mathbb{N}_k^m} \\ [\mathbf{H}_k]_{\text{rows}\, \mathbb{N}_k^m,\, \text{col } m} & [\mathbf{H}_k]_{\text{rows}\, \mathbb{N}_k^m,\, \text{cols}\, \mathbb{N}_k^m} \end{bmatrix}, \tag{6.19}$$

Define the set of lines whose signals are not modified during the precoding of user m on tone k

$$\begin{aligned} \underline{\mathbb{N}}_k^m &\triangleq \{1, \ldots, m-1,\ m+1, \ldots, N+1\} \setminus \mathbb{N}_k^m, \\ &= \left\{\underline{n}_{k,m}(1), \ldots, \underline{n}_{k,m}(N - t_{k,m})\right\}, \end{aligned}$$

Define the vector $\underline{\mathbf{x}}_k^{\text{tx}\,m}$ containing the corresponding transmitted symbols

$$\underline{\mathbf{x}}_k^{\text{tx}\,m} \triangleq \left[x_k^{\underline{n}_{k,m}(1)} \cdots x_k^{\underline{n}_{k,m}(N-t_{k,m})}\right]^T,$$

and the vector containing the corresponding noise

$$\underline{\mathbf{z}}_k^{\text{tx}\,m} \triangleq \left[z_k^{\underline{n}_{k,m}(1)} \cdots z_k^{\underline{n}_{k,m}(N-t_{k,m})}\right]^T,$$

Define the corresponding partial channel matrix

$$\underline{\mathbf{H}}_k^{\mathrm{tx}\,m} \triangleq \begin{bmatrix} [\mathbf{H}_k]_{\mathrm{row}\,m,\,\mathrm{cols}\,\underline{\mathbb{N}}_k^m} \\ [\mathbf{H}_k]_{\mathrm{rows}\,\mathbb{N}_k^m,\,\mathrm{cols}\,\underline{\mathbb{N}}_k^m} \end{bmatrix}.$$

Consider the reduced system model, which only contains the signals used in the precoding of TX m on tone k

$$\overline{\mathbf{y}}_k^{\mathrm{tx}\,m} = \overline{\mathbf{H}}_k^{\mathrm{tx}\,m}\overline{\mathbf{x}}_k^{\mathrm{tx}\,m} + \underline{\mathbf{H}}_k^{\mathrm{tx}\,m}\underline{\mathbf{x}}_k^{\mathrm{tx}\,m} + \overline{\mathbf{z}}_k^{\mathrm{tx}\,m}. \tag{6.20}$$

Define the $(t_{k,m}+1)\times 1$ vector that contains the non-zero elements of $\mathbf{P}_k$ used in the precoding of TX m as

$$\overline{\mathbf{p}}_k^m \triangleq \left[p_k^{m,m},\, p_k^{n_{k,m}(1),m}\ \cdots\ p_k^{n_{k,m}(t_{k,m}),m}\right]^T,$$

where $p_k^{n,m} \triangleq [\mathbf{P}_k]_{n,m}$. Hence

$$[\mathbf{P}_k]_{n,m} = \begin{cases} [\overline{\mathbf{p}}_k^m]_1, & n = m; \\ [\overline{\mathbf{p}}_k^m]_{i+1}, & n = n_{k,m}(i); \\ 0, & \text{otherwise.} \end{cases} \tag{6.21}$$

Under the DP criterion (6.17), the partial precoder protects all RXs in the set $\mathbb{N}_k^m$ from the crosstalk of TX m. The DP criterion for TX m is thus

$$\overline{\mathbf{H}}_k^{\mathrm{tx}\,m}\overline{\mathbf{p}}_k^m = \beta_{k,\mathrm{partial}}^{-1} h_k^{m,m}\overline{\mathbf{e}}_{\mathrm{tx}\,1},$$

where $\overline{\mathbf{e}}_{\mathrm{tx}\,m}$ is the mth column of the $(t_{k,m}+1)\times(t_{k,m}+1)$ identity matrix. Hence

$$\overline{\mathbf{p}}_k^m \triangleq \beta_{k,\mathrm{partial}}^{-1} h_k^{m,m}\left(\overline{\mathbf{H}}_k^{\mathrm{tx}\,m}\right)^{-1}\overline{\mathbf{e}}_{\mathrm{tx}\,1}, \tag{6.22}$$

6.5.3 Achievable Rate

The rate achieved with the partial precoder is now analysed. Using (6.1), (6.17), (6.15) and (6.14) the signal at RX n is

$$\begin{aligned} y_k^n &= [\mathbf{H}_k]_{\mathrm{row}\,n}\mathbf{P}_k\widetilde{\mathbf{x}}_k + z_k^n, \\ &= \beta_{k,\mathrm{partial}}^{-1} h_k^{n,n}\widetilde{x}_k^n + \sum_{m\in\underline{\mathbb{M}}_k^n} h_k^{n,m}\sum_i p_k^{m,i}\widetilde{x}_k^i + z_k^n. \end{aligned} \tag{6.23}$$

The first term is the transmitted signal whilst the second and third terms are the residual crosstalk and noise respectively. Define

$$\overline{\mathbf{G}}_k^{\mathrm{tx}\,m} = \mathrm{diag}\{h_k^{m,m}, h_k^{n_{k,m}(1),n_{k,m}(1)}, \ldots, h_k^{n_{k,m}(t_{k,m}),n_{k,m}(t_{k,m})}\}^{-1}\overline{\mathbf{H}}_k^{\mathrm{tx}\,m}. \tag{6.24}$$

Since the TXs are co-located at the CO, $\overline{\mathbf{H}}_k^{\mathrm{tx}\,m}$ is RWDD (4.4). This implies that $\overline{\mathbf{G}}_k^{\mathrm{tx}\,m} \in \mathbb{A}^{(N)}$, where $\mathbb{A}_n^{(N)}$ denotes the set of $N \times N$ diagonally dominant matrices, as defined in Appendix B. Now

$$\overline{\mathbf{H}}_k^{\mathrm{tx}\,m} = \mathrm{diag}\{h_k^{m,m}, h_k^{n_{k,m}(1),n_{k,m}(1)}, \ldots, h_k^{n_{k,m}(t_{k,m}),n_{k,m}(t_{k,m})}\}\overline{\mathbf{G}}_k^{\mathrm{tx}\,m}. \quad (6.25)$$

Combining this with (6.22) yields

$$\overline{\mathbf{p}}_k^m = \beta_{k,\mathrm{partial}}^{-1}\left[\left(\overline{\mathbf{G}}_k^{\mathrm{tx}\,m}\right)^{-1}\right]_{\mathrm{col}\,1}. \quad (6.26)$$

Since $\overline{\mathbf{G}}_k^{\mathrm{tx}\,n} \in \mathbb{A}^{(N)}$, theorem B.2 can now be applied to bound the elements of $\overline{\mathbf{G}}_k^{\mathrm{tx}\,n}$. This implies

$$|[\overline{\mathbf{p}}_k^m]_i| \leq \begin{cases} \beta_{k,\mathrm{partial}}^{-1} A_{\max}^{(N-1)}/A_{\min}^{(N)}, & i = 1; \\ \beta_{k,\mathrm{partial}}^{-1} B_{\max}^{(N-1)}/A_{\min}^{(N)}, & \text{otherwise.} \end{cases}$$

RWDD implies that $B_{\max}^{(N-1)} \ll 1$, $A_{\max}^{(N-1)} \simeq 1$ and $A_{\min}^{(N)} \simeq 1$. Hence

$$|[\overline{\mathbf{p}}_k^m]_i| \simeq \begin{cases} \beta_{k,\mathrm{partial}}^{-1}, & i = 1; \\ 0, & \text{otherwise.} \end{cases}$$

Hence from (6.21)

$$\left|[\mathbf{P}_k]_{n,m}\right| \simeq \begin{cases} \beta_{k,\mathrm{partial}}^{-1}, & n = m; \\ 0, & \text{otherwise.} \end{cases} \quad (6.27)$$

From (6.18) and (6.27) it is clear that setting $\beta_{k,\mathrm{partial}}$ to unity will approximately maintain the spectral masks. Combining this with (6.23) and (6.27) leads to the following approximation for the received signal

$$y_k^n \simeq h_k^{n,n}\widetilde{x}_k^n + \sum_{m \in \underline{\mathbb{M}}_k^n} h_k^{n,m}\widetilde{x}_k^m + z_k^n.$$

The power of the residual interference is

$$\begin{aligned} \mathcal{E}\left\{\left|\sum_{m \in \underline{\mathbb{M}}_k^n} h_k^{n,m}\widetilde{x}_k^m\right|^2\right\} &= \sum_{m \in \underline{\mathbb{M}}_k^n} |h_k^{n,m}|^2\, \mathcal{E}\left\{|\widetilde{x}_k^m|^2\right\}, \\ &\simeq \sum_{m \in \underline{\mathbb{M}}_k^n} |h_k^{n,m}|^2\, s_k^m, \end{aligned}$$

where (6.14) and (6.27) are applied in the second line. The SINR at the input of the decision device is thus

$$\mathrm{SINR}_k^n \simeq \frac{|h_k^{n,n}|^2\, s_k^n}{\sum_{m \in \underline{\mathbb{M}}_k^n} |h_k^{n,m}|^2\, s_k^m + \sigma_k}, \quad (6.28)$$

and the achievable data-rate for user n on tone k is

$$b_k^n = \log_2\left(1 + \Gamma^{-1}\mathrm{SINR}_k^n\right).$$

There are two interesting observations to make a this point. First, as is expected, the partial precoder completely removes crosstalk from interferers in the set $\mathbb{M}_k^n$. More surprisingly however, the partial precoder does not significantly change the statistics on any of the lines. That is, $\mathcal{E}\left\{\mathbf{x}_k\mathbf{x}_k^H\right\} \simeq \mathcal{E}\left\{\widetilde{\mathbf{x}}_k\widetilde{\mathbf{x}}_k^H\right\}$. So *the RWDD of* $\mathbf{H}_k$ *ensures that the DP partial precoder does not increase the power of, or introduce correlation in to, the signals on the lines.*

6.6 Complexity Distribution

The previous sections described the procedure for designing partial cancelers and precoders for a particular tone with a given complexity $r_{k,n}$. Typically, the crosstalk canceler or precoder must be designed according to a given complexity budget for all tones $k = 1 \ldots K$. This section investigates the distribution of complexity across frequency.

Let the available complexity of the crosstalk canceler be cK multiplications / DMT-block / user, such that

$$\sum_k |\mathbb{M}_k^n| \leq cK, \ \forall n,$$

where $|\mathbb{A}|$ denotes the cardinality of set $\mathbb{A}$. This corresponds to c times the complexity of the conventional frequency domain equalizer, which is implemented in existing DSL modems. The problem is now to distribute the available complexity across frequency such that the net data-rate is maximized

$$\max_{\left\{\mathbb{M}_k^n\right\}_{k=1,\ldots,K}} \sum_k b_k^n, \quad \text{s.t.} \sum_k |\mathbb{M}_k^n| \leq cK. \tag{6.29}$$

This section investigates algorithms for solving this problem. The optimal algorithm is described, together with several simpler heuristic methods. All algorithms exploit space-selectivity, frequency-selectivity or a combination of both in order to reduce run-time complexity. The following sections describe complexity distribution in partial cancelers, however the algorithms are also directly applicable to partial precoders.

6.6.1 Line Selection

The majority of the crosstalk experienced by a modem comes from only a few dominant crosstalkers within the binder. We refer to this as the space-selectivity of crosstalk, and it can be exploited to reduce the run-time complexity of crosstalk cancellation. In practice, this corresponds to the partial

Algorithm 6.1 Line Selection Only

$$r_{k,n} = c,\ \forall n, k.$$

canceler processing only a subset of $\mathbb{M}_k^n$ of the signals on the CO lines when detecting user n. This section first investigates the optimal choice for the subset $\mathbb{M}_k^n$. The problem is to maximize the data-rate given a limited amount of run-time complexity,

$$\max_{\mathbb{M}_k^n} b_k^n \text{ s.t. } |\mathbb{M}_k^n| \leq r_{k,n}, \tag{6.30}$$

where b_k^n is the rate of user n on tone k. From (6.13) and (6.28), maximizing SINR_k^n, and thus data-rate b_k^n, corresponds to minimizing the amount of crosstalk in the set $\underline{\mathbb{M}}_k^n$. So the data-rate is maximized by setting $\mathbb{M}_k^n$ to contain the largest crosstalkers of user n on tone k. Define the indices of the crosstalkers of user n on tone k sorted in order of crosstalk strength

$$\begin{array}{ll} \text{order} & \{q_{k,n}(1),\ \ldots,\ q_{k,n}(N)\}\,; \\ \text{s.t.} & \left|h_k^{n,q_{k,n}(i)}\right|^2 s_k^{q_{k,n}(i)} \geq \left|h_k^{n,q_{k,n}(i+1)}\right|^2 s_k^{q_{k,n}(i+1)},\ \forall i; \\ & q_{k,n}(i) \neq n,\ \forall i. \end{array}$$

Remark 6.1 Optimal Receive Line Selection

In CWDD US channels and RWDD DS channels, the set $\mathbb{M}_k^n$ which maximizes the data-rate of user n on tone k subject to a complexity constraint of $r_{k,n}$ multiplications/DMT-block (see optimisation in (6.30)) is

$$\mathbb{M}_k^n = \{q_{k,n}(1),\ \ldots,\ q_{k,n}(r_{k,n})\}\,.$$

Proof: Follows from examination of (6.13) for US and (6.28) for DS. ■

This suggests the simple partial canceler design in Alg. 6.1. Accordingly, the partial canceler simply cancels the c largest crosstalkers on each tone.

This algorithm is only capable of exploiting space-selectivity to reduce run-time complexity. The degree of partial cancellation is equal on all tones, so the algorithm cannot exploit the frequency-selectivity of the crosstalk channel. As will be shown, this leads to poor performance compared to algorithms that exploit both space and frequency-selectivity. The advantage of this algorithm is its simplicity. It requires only $\mathcal{O}(KN)$ multiplications and K length-N sorts to initialize the partial canceler for each user. Here we define initialization complexity as the complexity of determining $\mathbb{M}_k^n$, $\forall k$. Initialization complexity does not include actual calculation of the crosstalk cancellation parameters $\overline{\mathbf{w}}_k^n$ for each tone. This requires $\mathcal{O}(\sum_k (r_{k,n}+1)^3)$ multiplications for user n regardless of the partial cancellation algorithm employed.

The initialization complexities, in terms of multiplication and logarithm operations per user, of the different partial cancellation algorithms are listed in Tab. 6.1. The required number of sort operations are listed in Tab. 6.2. All algorithms have equal run-time complexity. The initialization complexity for partial precoders is identical to that of the partial cancelers.

6.6.2 Tone Selection

The previous section described an algorithm that only exploits the space selectivity of crosstalk, i.e. the fact that crosstalk varies significantly between different lines. Crosstalk coupling also varies significantly with frequency, a property that can also be exploited to reduce run-time complexity.

In low frequencies crosstalk coupling is minimal so crosstalk cancellation yields minimal gains. In higher frequencies crosstalk coupling is severe. However, in high frequencies the direct channel attenuation is so great that the modems can only achieve a low bitloading, even in the absence of crosstalk. This limits the gains of crosstalk cancellation. Thus the largest gains from crosstalk cancellation are experienced in the intermediate frequencies, and this is where most of the available run-time complexity should allocated.

Consider a crosstalk canceler that only operates on a subset of tones $\mathbb{K}_n$ when detecting user n. This section investigates the optimal choice for the subset $\mathbb{K}_n$. The problem is to maximize the data-rate given a limited amount of run-time complexity,

$$\max_{\mathbb{K}_n} \sum_{k\in\mathbb{K}_n} b_k^n(N) + \sum_{k\notin\mathbb{K}_n} b_k^n(0) \text{ s.t. } |\mathbb{K}_n| \leq cK/N, \tag{6.31}$$

where $b_k^n(r_{k,n})$ is defined as the rate achieved by user n on tone k when the $r_{k,n}$ largest crosstalkers are cancelled,

$$b_k^n(r_{k,n}) \triangleq \log_2\left(1 + \frac{1}{\Gamma}\frac{\left|h_k^{n,n}\right|^2 s_k^n}{\sum_{i=r_{k,n}+1}^{N}\left|h_k^{n,q_{k,n}(i)}\right|^2 s_k^{q_{k,n}(i)} + \sigma_k}\right). \tag{6.32}$$

Define the gain of full crosstalk cancellation, $r_{k,n} = N$, as

$$g_{k,n} \triangleq b_k^n(N) - b_k^n(0),$$

and define the tone indices ordered by this as[2]

$$\begin{array}{ll} \text{order} & \{k_n(1), \ldots, k_n(K)\}; \\ \text{s.t.} & g_{k_n(i),n} \geq g_{k_n(i+1),n},\ \forall i. \end{array}$$

[2]Note that $g_{k,n}$ can be efficiently calculated on a logarithmic scale by dividing the arguments of the logarithms in $b_k^n(N)$ and $b_k^n(0)$.

Algorithm 6.2 Tone Selection Only

$$r_{k,n} = \begin{cases} N, & k \in \{k_n(1), \ldots, k_n(cK/N)\}\,; \\ 0, & \text{otherwise.} \end{cases}$$

Remark 6.2 Optimal Tone Selection

The set $\mathbb{K}_k^n$ which maximizes the data-rate of user n subject to a complexity constraint of cK multiplications/DMT-block (see optimisation in (6.31)) is

$$\mathbb{K}_n = \{k_n(1), \ldots, k_n(cK/N)\}\,.$$

This suggests the simple partial canceler design in Alg. 6.2. This algorithm employs full crosstalk cancellation, $r_{k,n} = N$, on the cK/N tones with the largest gain and no cancellation on all other tones. This leads to a run-time complexity of cK multiplications/DMT-block/user.

Note that in this algorithm $r_{k,n}$ is restricted to the values 0 and N; it is not possible to only cancel the largest crosstalkers and the space-selectivity of the crosstalk channel is ignored. The initialization complexity of this algorithm is $\mathcal{O}(KN)$ multiplications and one sort of size K, per user.

6.6.3 Simple Joint Tone-Line Selection

So far, sections 6.1 and 6.2 have described partial cancellation algorithms that exploit only one form of selectivity in the crosstalk channel. To achieve the maximum reduction in run-time complexity it is necessary to exploit both space and frequency-selectivity together. An efficient algorithm should adapt the degree of crosstalk cancellation on each tone to match the potential gain. In practice this means that $r_{k,n}$ must be allowed to take on values other than 0 or N. It must also be allowed to take on different values as the tone k varies.

Sec. 6.4.3 showed that observing the direct line of a crosstalker allows the partial canceler to remove the crosstalk from that line completely. Hence line selection is equivalent to choosing the set of crosstalkers to be cancelled. When combined with tone selection, the partial canceler design problem is essentially a choice of which crosstalker-tone pairs to cancel in the detection of a given user. Similarly, in a crosstalk precoder, the design problem is to chose which crosstalker-tone pairs to precode against, when protecting a given user.

The rate improvement from canceling a particular crosstalker on a particular tone is dependent on the other crosstalkers that have already been canceled on that tone; there is an inherent coupling in crosstalker selection and this complicates the line selection problem substantially. The algorithm described

Algorithm 6.3 Simple Tone-Line Selection

$$\mathbb{M}_k^n = \{m \,:\, (m,k) \in \{d_n(1), \ldots, d_n(cK)\}\}\,.$$

in this section removes this coupling by ignoring the effect of other crosstalkers in the system. This greatly simplifies crosstalker-tone pair selection with only a small performance penalty, as is demonstrated in Sec. 6.7.

Define the gain of cancelling crosstalker m on tone k in the detection of user n, and in the absence of all other crosstalkers[3]

$$\overline{g}_{k,n}(m) \triangleq \log\left(1 + \frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^n}{\sigma_k}\right) - \log\left(1 + \frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^n}{|h_k^{n,m}|^2 s_k^m + \sigma_k}\right).$$

Define crosstalker-tone pair $d_n(i) \triangleq (m_n(i), k_n(i))$ and its corresponding gain $\overline{g}_n\left(d_n(i)\right) \triangleq \overline{g}_{k_n(i),n}\left(m_n(i)\right)$. Now define the indicies of the crosstalker-tone pairs ordered by gain

$$\begin{array}{ll} \text{order} & \{d_n(1), \ldots, d_n(KN)\}\,; \\ \text{s.t.} & \overline{g}_n\left(d_n(i)\right) \geq \overline{g}_n\left(d_n(i+1)\right),\ \forall i. \end{array}$$

This suggests the simplified joint tone-line selection algorithm in Alg. 6.3. In the detection of user n, the partial canceler observes the signal at line m on tone k if tone-line pair (m,k) has a large gain

$$(m,k) \in \{d_n(1), \ldots, d_n(cK)\}\,.$$

This leads to a run-time complexity of cK multiplications/DMT-block/user. The benefit of this algorithm is its low complexity. Pair selection for one user has a complexity of $\mathcal{O}(KN)$ multiplications and one sort of size KN. Furthermore, this algorithm exploits both the space and frequency-selectivity of the crosstalk channel, which allows it to cancel the largest crosstalkers on the tones where they do the most harm. In Sec. 6.7 it is shown that this algorithm leads to near-optimal performance.

6.6.4 Optimum Joint Tone-Line Selection

It is interesting to evaluate the sub-optimality of the different algorithms with an upper bound achieved by the truly optimal partial cancellation algorithm. The partial canceler design problem is formulated in (6.29). Remark 6.1 implies that the optimal set of extra lines, $\mathbb{M}_k^n$, to observe when detecting user n on tone k corresponds to the largest crosstalkers of user n on tone k. This allows us

[3]Note that $\overline{g}_{k,n}(m)$ can be efficiently calculated on a logarithmic scale by dividing the arguments of each log function.

to determine $\mathbb{M}_k^n$ in a straight-forward way once $r_{k,n}$ has been determined. The problem is now one of *resource allocation*. Given cK multiplications/user, the optimal algorithm must distribute these across tones such that the data-rate is maximized

$$\max_{\{r_{k,n}\}_{k=1,\ldots,K}} \sum_k b_k^n, \quad \text{s.t.} \sum_k r_{k,n} \leq cK.$$

An exhaustive search could require the evaluation of up to N^K different allocations. In VDSL $K = 4096$, for which an exhaustive search is numerically intractable.

Due to the structure of the problem a greedy algorithm, listed as Alg. 6.4, can be applied to iteratively find the optimal allocation for some values of c. The greedy algorithm is optimal, efficient and has a tractable complexity.

The algorithm cannot find a solution for any arbitrary value of c, nevertheless the range of values of c for which the algorithm can find a solution are so closely spaced that this is not a practical problem. Define the value of canceling r crosstalkers on tone k as

$$v_{k,n}(r) = \frac{b_k^n(r) - b_k^n(0)}{r}.$$

Recall that $b_k^n(r)$ is the rate achieved by user n on tone k when the r largest crosstalkers are canceled and is evaluated using (6.32). Value $v_{k,n}(r)$ is the increase in data-rate (benefit) divided by the increase in run-time complexity (cost). It measures the increase in data-rate per multiplication for user n when r multiplications are spent on tone k. The algorithm begins by initializing $v_{k,n}(r)$ for all values of r and k. It then proceeds as follows

1. Find choice of tone k and number of cancelled crosstalkers r with largest value $v_{k,n}(r)$. Store this as (k_s, r_s).
2. Set the number of lines to be observed on tone k_s to r_s, and
$$\mathbb{M}_{k_s}^n = \{q_{k_s,n}(1), \ldots, q_{k_s,n}(r_s)\}.$$
3. Set the value of canceling r_s or fewer crosstalkers on tone k_s to zero. This prevents re-selection of previously selected pairs.
4. Update the value of canceling $r_s + 1$ or more crosstalkers on tone k_s. The data-rate increase and cost should be relative to the currently selected number of crosstalkers.

The algorithm iterates through steps 1-4 until the allocated complexity exceeds cK. This yields an upper bound on partial cancellation performance for a given complexity. Since the algorithm allocates at most N multiplications in each

Algorithm 6.4 Optimal Line-Tone Selection

init $v_{k,n}(r) = \left(b_k^n(r) - b_k^n(0)\right)/r, \quad \forall\, k,\, r > 0;$
repeat
$(k_s, r_s) = \arg\max_{(k,r)} v_{k,n}(r);$
$r_{k_s,n} = r_s;$
$v_{k_s,n}(r) = 0, \quad \forall r \leq r_s;$
$v_{k_s,n}(r) = \left(b_{k_s}^n(r) - b_{k_s}^n(r_s)\right)/\left(r - r_s\right), \quad \forall r > r_s;$
while $\sum_k r_{k,n} < cK$.

iteration, the total allocated complexity will be at the most $cK + N$. In VDSL $K = 4096$, and typically $cK \gg N$. Hence the difference between the desired run-time complexity and that of the solution provided by the algorithm is minimal. So the upper bound is tight.

Like Alg. 6.3, this algorithm exploits both the space and frequency-selectivity of the crosstalk channel to reduce run-time complexity. The algorithm generates a resource allocation at the end of each iteration that is optimal in the sense that of all resource allocations of equal run-time complexity the allocation generated by this algorithm achieves the highest data-rate. Unfortunately, this algorithm is considerably more complex than Alg. 6.3. Pair selection for a single user requires $\mathcal{O}(KN^2)$ multiplications and $\mathcal{O}(KN)$ logarithm operations. It is hard to define the exact sorting complexity since it varies significantly with the scenario. Sorting complexity is typically much higher than any of the other algorithms and can require up to KN sort operations that can each have sizes as large as KN.

6.6.5 Complexity Distribution between Users

So far we have limited the run-time complexity of detecting each user to cK such that

$$\sum_k |\mathbb{M}_k^n| \leq cK,\ \forall n.$$

If crosstalk cancellation of all lines in a binder is integrated into a single processing module at the CO, then multiplications can be shared between users. That is, the true constraint is on the total complexity of crosstalk cancellation for *all users*

$$\sum_{n=1}^{N+1} \sum_k |\mathbb{M}_k^n| \leq cK\,(N+1)\,.$$

The available complexity can be distributed based on the target rates of each user. Denote the number of multiplications/DMT-block allocated to user n as

Table 6.1: Initialization Complexity of Different Algorithms (per user)

Scheme	Initialization Complexity	
	Multiplications	Logs
Line Selection Only	KN	0
Tone Selection Only	$K(N+5)$	0
Simple Joint Selection	$3K(N+1)$	0
Optimal Joint Selection	$K(0.5N^2+2.5N+3)$	$K(N+1)$

Table 6.2: Initialization Complexity of Different Algorithms (per user)

Scheme	Sort Operations		
	Sort Size N	Sort Size K	Sort Size KN
Line Selection Only	K	0	0
Tone Selection Only	0	1	0
Simple Joint Selection	0	0	1
Optimal Joint Selection	0	0	KN

κ_n. Then

$$\kappa_n = \mu_n cK\left(N+1\right),\ \forall n; \quad \text{s.t.} \sum_n \mu_n = 1.$$

Here μ_n is the parameter that determines the proportion of computing resources allocated to user n. This allows partial cancellation to be viewed as a resource allocation not only across tones, but across users as well. Given a fixed number of multiplications, these must be divided between users based on the target rates of each user. In a similar fashion to the multi-user power allocation from Chapter 3, a rate region can be defined as the set of achievable rate-tuples under a given complexity constraint. This allows an operator to visualise the different trade-offs that can be achieved between the rates of different users inside a binder.

Limiting crosstalk cancellation on each tone to the users who benefit the most leads to further reductions in run-time complexity with minimal performance loss. This is demonstrated in Sec. 6.7.1.

6.7 Performance

This section compares the performance of the partial cancellation algorithms described in the previous section. Performance is compared over a range of sce-

narios in crosstalk channels that exhibit both space and frequency-selectivity. As will be shown, the ability to exploit both space and frequency-selectivity is essential for achieving low run-time complexity in all scenarios.

In all simulations the line diameter is 0.5 mm (24-AWG). The target symbol error probability is 10^{-7} or less. The coding gain and noise margin are set to 3 dB and 6 dB respectively. The maximum bitloading is not constrained. As per the VDSL standards the tone spacing Δ_f and DMT symbol rate f_s are set to 4.3125 kHz and 4 kHz respectively[9][7]. The modems use a transmit mask of -60 dBm/Hz according to FDD bandplan 998[9]. Background noise was ETSI Type A[7].

6.7.1 Upstream

This section describes the performance of the different algorithms in upstream transmission. We use semi-empirical transfer functions from the ETSI VDSL standards[7]. Note that in these channel models each user has identical crosstalk channels to all crosstalkers of equal line length. That is, the variation of crosstalk channel attenuation with the separation between lines within the binder is not modeled. However, when a binder consists of lines of varying length the model does capture the near-far effect. All users will see the modems located closest to the CO (near-end) as the largest sources of crosstalk. On the other hand when a binder consists of lines of equal length all users will see equal crosstalk from all other users. There will then be no space-selectivity in the crosstalk channel model.

In reality we would expect more space-selectivity than is contained within these channel models. Hence we can expect the reduction in run-time complexity to be even larger than that shown here. The number of lines in the binder is always 8, so $N = 7$.

Equidistant Lines

In the first scenario the binder contains 8×1000m lines. Since the lines are of equal length the crosstalk channels exhibit frequency-selectivity only; no space-selectivity is present. Shown in Fig. 6.6 are the rates achieved by each of the algorithms versus run-time complexity. Complexity is shown as a percentage relative to full crosstalk cancellation ($c = N$).

Alg. 6.1 can only exploit space-selectivity. There is no space-selectivity in this scenario so this algorithm gives extremely poor performance. Worst of all, we actually see a non-convex rate vs. run-time complexity curve. So partial cancellation gives worse performance than time-sharing. In other words, a system could do full crosstalk cancellation for a fraction of the time, and no cancellation for the rest, and this would lead to better performance than Alg.

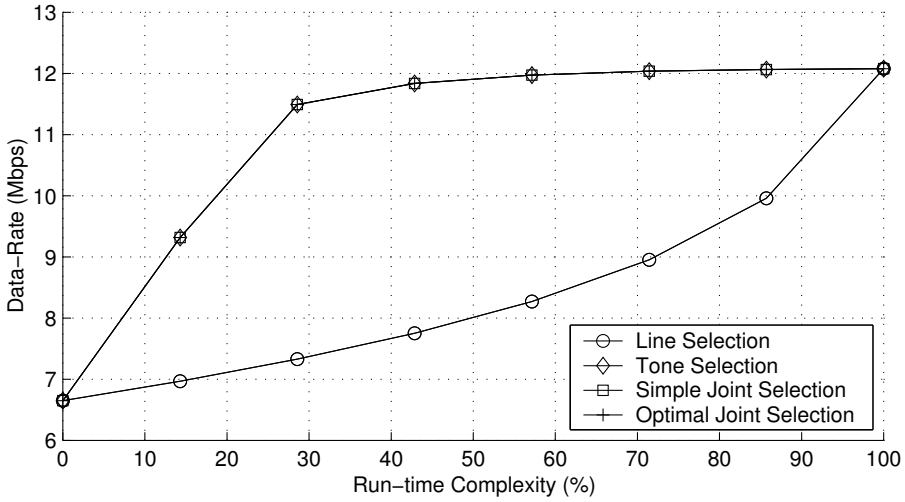

Figure 6.6: Data-rate vs. Run-time Complexity (Equidistant Lines)

6.1 with the same run-time complexity. The reason for this is as follows. As the number of crosstalkers canceled $p_{k,n}$ increases, the increase in signal-to-interference ratio (SIR) grows rapidly.

This is illustrated through the following example. Consider a binder with seven crosstalkers. Let the crosstalkers have identical crosstalk channels χ_k^n to user n as is the case in our simulation. Cancelling the first crosstalker causes the SIR to increase from

$$\frac{1}{7}|h_k^{n,n}|^2|\chi_k^n|^{-2} \longrightarrow \frac{1}{6}|h_k^{n,n}|^2|\chi_k^n|^{-2}.$$

Cancelling the sixth crosstalker gives a much larger SIR increase from

$$\frac{1}{2}|h_k^{n,n}|^2|\chi_k^n|^{-2} \longrightarrow |h_k^{n,n}|^2|\chi_k^n|^{-2}.$$

In general cancelling the pth crosstalker leads to an SIR increase by a factor of

$$(N-p+1)(N-p)^{-1}.$$

So the increase in SIR grows with rapidly with p as $p \to N$. Recall that $b_k^n = \log(1+SINR_k^n) \simeq SINR_k^n$ for low $SINR_k^n$. So when crosstalkers have equal strength and the SINR is low, data-rate gain will grow rapidly with the number of crosstalkers cancelled p. This is why cancelling N crosstalkers typically gives greater than N times the data-rate gain of canceling one crosstalker. This leads to the non-convex rate-complexity curve of Fig. 6.6.

When the channel exhibits space-selectivity the first crosstalker causes much more interference than the second and so on. This effect counter-acts the rapid growth of SIR with p. As a result the best trade-off between performance and complexity usually occurs somewhere between no cancellation and full cancellation.

Alg. 6.2 cannot exploit space-selectivity. In this scenario this is not a problem since all crosstalkers have equal strength. Alg. 6.2 can implement a form of frequency-sharing. This is analogous to the time-sharing just discussed and allows this algorithm to cancel e.g. 6 crosstalkers on half of the tones instead of 3 crosstalkers on all of the tones. For this reason Alg. 6.2 will always give a convex rate vs. complexity curve. Comparing the performance of Alg. 6.2 to the optimal algorithm shows that Alg. 6.2 achieves near-optimal performance in this scenario.

Alg. 6.3 also gives near-optimal performance. Note that with 29% of the complexity of full crosstalk cancellation the partial canceler achieves 89% of the performance gains.

Near-Far Scenario

In this scenario the binder contains 4×300m lines and 4×1200m lines. The lines suffer the near-far effect and all users see the 300m lines as the dominant source of crosstalk. This space-selectivity assists the partial cancellation algorithms in reducing run-time complexity. Frequency-selectivity is present in this scenario and is most pronounced on far-end lines. Near-end lines have relatively flat channels and benefit less from algorithms which exploit frequency-selectivity alone.

Fig. 6.7 shows the rates of the 300m near-end lines vs. complexity for each of the algorithms. Fig. 6.8 shows the same for the 1200m far-end lines.

Alg. 6.1 cannot exploit frequency-selectivity. On near-end lines frequency-selectivity is minimal and reasonable performance is still achieved. Again we see a non-convex rate-complexity curve however above 43% complexity Alg. 6.1 gives near-optimal performance. On far-end users frequency-selectivity is pronounced and Alg. 6.1 gives poor performance.

Alg. 6.2 cannot exploit space-selectivity and on near-end users this leads to poor performance; the rates achieved are identical to time sharing. On far-end users frequency-selectivity is pronounced and this algorithm still achieves reasonable performance despite its inability to exploit space-selectivity.

Alg. 6.3 exploits both space and frequency-selectivity. As a result it achieves near-optimal performance for both near and far-end users. With 43% complexity this algorithm achieves 99% of the performance gains on near-end users. On

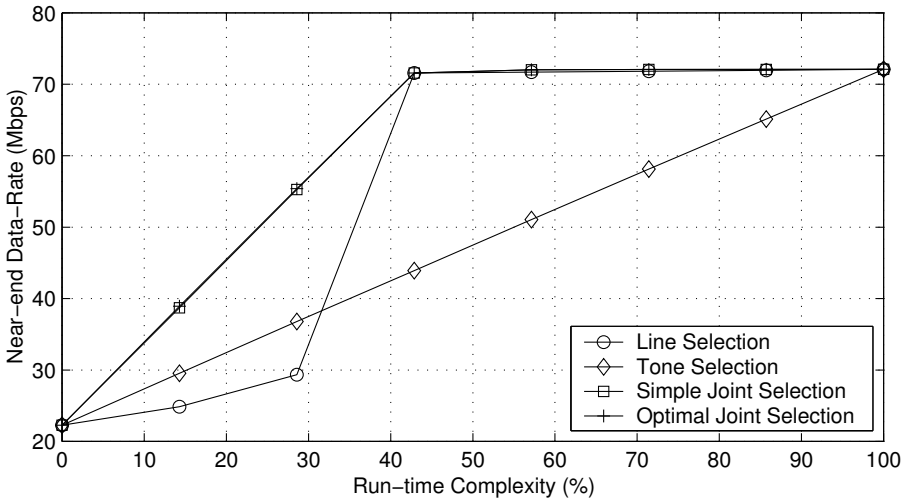

Figure 6.7: Near-end Data-rate vs. Run-time Complexity

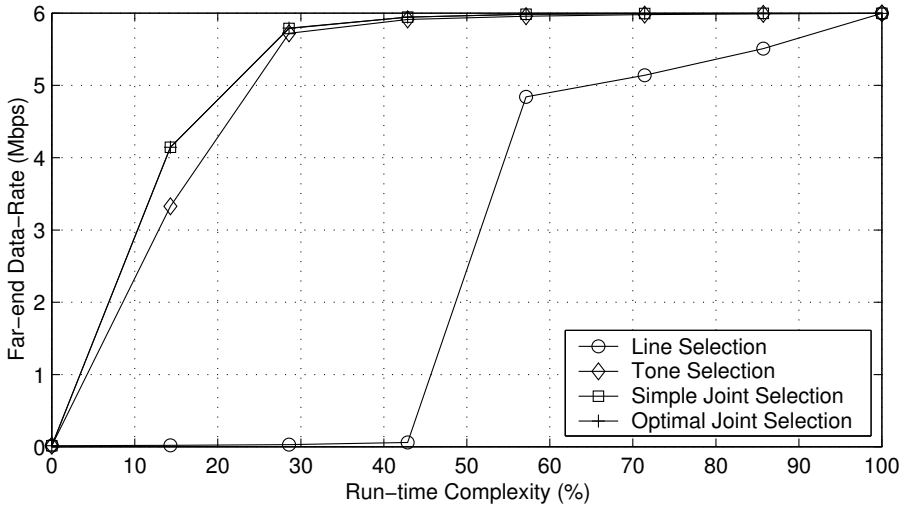

Figure 6.8: Far-end Data-rate vs. Run-time Complexity

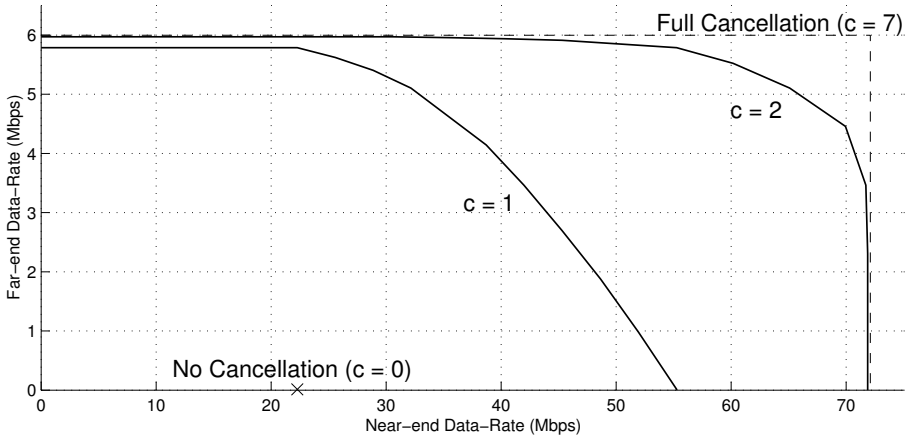

Figure 6.9: Achievable Rate Regions vs. Complexity (Simple Joint Selection)

far-end users 29% complexity achieves 97% of the performance gains.

Complexity Distribution between Users

The distribution of run-time complexity between users, as described in Sec. 6.6.5, is now discussed. Fig. 6.9 shows the rate regions under varying complexities c using Alg. 6.3. The rate region was constructed by dividing multiplications between the two classes of near-end and far-end users. Users of one class receive an equal number of multiplications; $2\mu_{\text{near}}cK$ and $2\mu_{\text{far}}cK$ multiplications/DMT-block for the near-end and far-end users respectively. By varying the parameter μ_{far} the boundary of the rate region is traced. Note that $\mu_{\text{near}} = 1-\mu_{\text{far}}$. Fig. 6.9 shows that with 29% of the run-time complexity of full cancellation ($c = 2$) the partial canceler achieves the majority of the operating points within the rate region.

In Fig. 6.10 the rate regions of the different partial cancellation algorithms are compared for $c = 2$. Note the considerably larger rate region which is achieved by exploiting both space and frequency-selectivity in Alg. 6.3 and Alg. 6.4.

It is clear that the performance of algorithms that exploit only one type of selectivity, such as Alg. 6.1 and Alg. 6.2, varies considerably with the scenario. By exploiting both space and frequency-selectivity Alg. 6.3 consistently gave near-optimal performance. This algorithm also has a significantly lower complexity than the optimal algorithm, Alg. 6.4. So Alg. 6.4 is a computationally simple approach for partial crosstalk canceler design that leads to low initialization complexity, low run-time complexity and near-optimal performance.

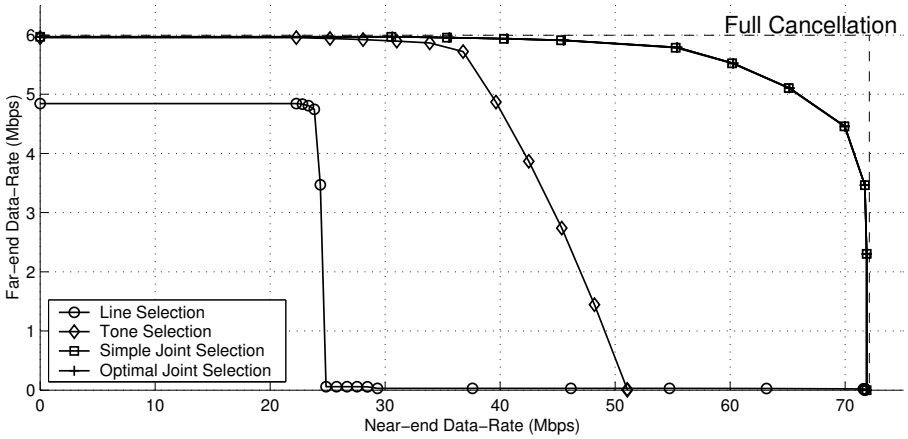

Figure 6.10: Achievable Rate Regions of Different Algorithms (29% complexity)

6.7.2 Downstream

This section describes the performance of the different algorithms in downstream transmission. Again it will be made clear that the ability to exploit both space and frequency-selectivity is essential for achieving the lowest possible run-time complexity. The simulations use a set of measured crosstalk channels from a 0.5 mm (24 AWG) 8-pair cable. The first 4 pairs are 900 m long, the last 4 are 1200 m The direct channels are depicted in Fig. 6.1 and the crosstalk channels in Fig. 6.2. Other simulation parameters are the same as in the upstream scenario.

The distribution of run-time complexity between users, as described in Sec. 6.6.5, is evaluated. Fig. 6.11 shows the rate regions for varying complexities c using Alg. 6.3. Fig. 6.9 indicates that with 30% of the run-time complexity ($c = 2.4$) of full precoding, the partial precoder achieves the majority of the operating points in the rate region.

In Fig. 6.10 the rate regions of the different algorithms are compared with 20% complexity ($c = 1.6$). Note the considerably larger rate region that is achieved by exploiting both space and frequency-selectivity in Alg. 6.3.

To illustrate the potential gains, consider the case when the 1200 m lines have a target rate of 20 Mbps. Tab. 6.3 then shows the rates that can be achieved on the 900 m lines. The allocation of complexity between users is shown, along with the rate gain as a proportion of full cancellation. By definition the rate gain of no cancellation is 0%, and the rate gain of full cancellation is 100%.

Essentially the algorithms allocate just enough complexity to the 1200 m lines

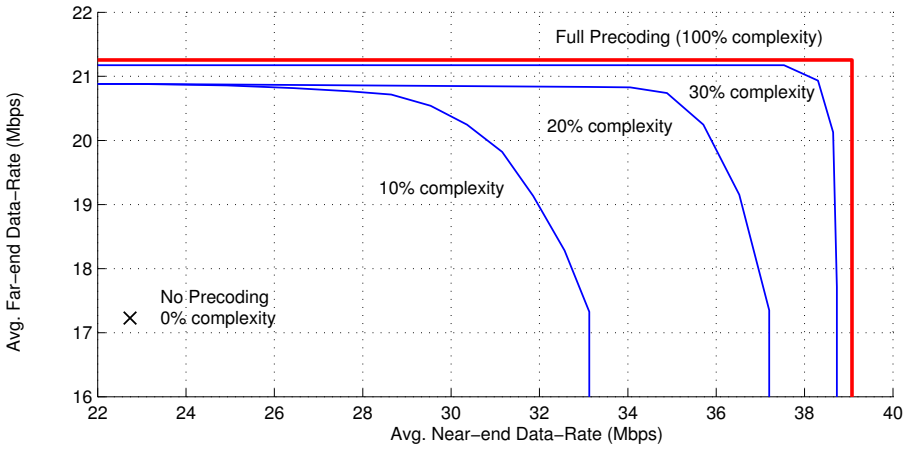

Figure 6.11: Achievable rate regions vs. complexity (Simple joint selection)

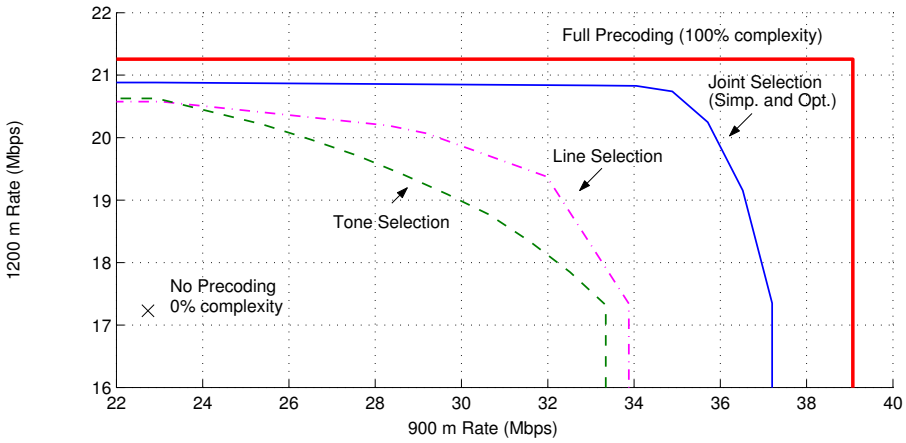

Figure 6.12: Achievable rate regions of different algorithms (20% complexity)

Table 6.3: Achievable data-rates with different algorithms

Precoding Technique	μ_{far}	Rate Gain (%) 900 m	Rate Gain (%) 1200 m	Rate (Mbps) 900 m	Rate (Mbps) 1200 m
None	-	0	0	22.7	17.2
Tone Selection	0.8	23	70	26.4	20.0
Line Selection	0.6	41	70	29.5	20.0
Simple Joint Selection	0.2	80	70	35.9	20.0
Optimal Joint Selection	0.2	80	70	35.9	20.0
Full	-	100	100	39.1	21.2

Table 6.4: Complexity distribution with different algorithms

Precoding Technique	μ_{far}	Complexity (%) Total	Complexity (%) 900 m	Complexity (%) 1200 m
None	-	0	0	0
Tone Selection	0.8	20	8	32
Line Selection	0.6	20	16	24
Simple Joint Selection	0.2	20	32	8
Optimal Joint Selection	0.2	20	32	8
Full	-	100	100	100

such that their target rate is achieved. This corresponds to finding the smallest possible μ_{far}, that still achieves 20 Mbps. Once this is done, any left over complexity is allocated to the 900 m lines. The better a partial precoding algorithm is, the smaller the value of μ_{far} it will be able to reach whilst still achieving the 1200 m target rate.

With tone selection $\mu_{\text{far}} = 0.8$ is required achieve the target rate on the 1200 m lines. This allocates 80% of the available complexity to the 1200 m lines. With the remaining 20% the rate on the 900 m lines can be increased to 26.4 Mbps, which corresponds to 23% of the achievable rate gain.

Line selection gives better performance. Less complexity needs to be allocated to the 1200 m lines, and they achieve their target rate with $\mu_{\text{far}} = 0.6$. This leaves 40% of the available complexity to the 1200 m lines, allowing them to achieve 41% of the potential gains.

Joint selection gives a much higher performance than either line or tone selection alone. The 1200 m line target rate is achieved with only $\mu_{\text{far}} = 0.2$ and the 900 m lines can increase their rates to 35.9 Mbps, which is 80% of the achievable gain. This underscores the importance of exploiting both space and frequency selectivity when designing a partial precoder.

So using joint selection 70% of the achievable gains on the 1200 m lines, and 80% of the achievable gains on the 900 m lines can be achieved. This is done with only 20% of the run-time complexity of full precoding.

Hence Alg. 6.4 is a computationally simple approach for partial crosstalk precoder design that leads to low initialization complexity, low run-time complexity and near-optimal performance.

6.8 Summary

Crosstalk is *the* dominant source of performance degradation in modern DSL systems. Several crosstalk cancellation and precoding schemes have been proposed to address this. Whilst these schemes lead to large performance gains, they have high run-time complexities, typically beyond the scope of implementation for current systems.

Crosstalk channels in DSL are space and frequency selective. That is, the majority of crosstalk comes from a few users and its effects are limited to a subset of tones. Partial coordination exploits this by limiting operation to the tones and lines where it gives maximum benefit. As a result these schemes achieve the majority of the gains of full coordination at a fraction of the complexity.

This chapter presented several partial coordination algorithms. *Line Selection*

removes only the largest crosstalkers of each user. This allows it to exploit the space selectivity of crosstalk, however since the number of cancelled crosstalkers is the same on each tone, frequency selectivity cannot be exploited. *Tone Selection* uses full coordination on the tones that benefit the most, however since this is an 'all or nothing' approach, it cannot exploit space selectivity by only removing the dominant crosstalkers. *Joint Line-Tone Selection* gives the best performance, limiting coordination to the lines and tones that benefit most.

In upstream transmission, the *Simple Joint Line-Tone Selection* algorithm achieves 90% of the performance gains of full cancellation with only 29% of the run-time complexity. In downstream transmission, this algorithm achieves 80% of the performance gains of full precoding with 20% of the run-time complexity.

The distribution of run-time complexity between users was considered. This allows complexity to be allocated to the users who benefit most, leading to further reductions in complexity. In a similar fashion to the multi-user power allocation from Chapter 3, this led to the development of rate regions. The difference here is that computing power rather than transmit power is allocated between users.

This chapter has only considered applications to VDSL. Nevertheless, the techniques presented here can also be applied to other wireline systems such as ADSL and *Ethernet-in-the-first-mile* (EFM).

Chapter 7

Conclusions

The goal of this thesis was to develop practical multi-user techniques for mitigating crosstalk in DSL. This chapter summarizes the key results of the thesis and suggests topics for future work.

Multi-user techniques are based on the coordination of different users in a network. This can be done on either a spectral or signal level.

Part I of this thesis investigated multi-user spectra coordination. With spectral coordination the transmit spectra of the modems within a network are limited in some way to minimize the negative effects of crosstalk. Each modem must achieve a trade-off between maximizing its own data-rate and minimizing the crosstalk it causes to other modems within the network. The goal is to achieve a fair trade-off between the rates of the different users.

Chapter 3 investigated the design of optimal transmit spectra for a network of interfering DSLs. This problem was previously considered intractable since it requires the solution of a high-dimensional, non-convex optimization. Chapter 3 showed how the application of a dual-decomposition can solve the optimization in an efficient, tractable way. The resulting algorithm, which we name *optimal spectrum balancing*, achieves significant gains over existing spectra coordination algorithms, typically doubling or tripling the achievable data-rate. The *optimal spectrum management* algorithm is now part of the draft ANSI standard on Dynamic Spectrum Management[8].

Part II of this thesis investigated multi-user signal coordination. In a DSL network, the line-side transceivers are often co-located at the CO. This allows modems to be co-ordinated on a signal level.

In the upstream signal coordination is used between co-located CO receivers.

Reception is done in a joint fashion; the signals received on each line are combined to cancel crosstalk whilst preserving the signal of interest.

Chapter 4 discussed crosstalk canceler design. Existing techniques are based on decision feedback between the different users. To prevent error propagation decoding must be done before decisions are fed back, which leads to a high computational complexity and latency. To address this problem a simple linear canceler was presented based on the well known ZF criterion. This technique has a low complexity and latency, and operates close to the theoretical multi-user channel capacity.

In the downstream signal coordination is used between co-located CO transmitters. Transmission is done in a joint fashion; some predistortion is introduced into the signal intended for each user prior to transmission. This predistortion is chosen such that it annihilates with the crosstalk introduced in the channel. As a result the CP modems receive a crosstalk free signal.

This technique, known as crosstalk precoding, was discussed in Chapter 5. Existing precoder designs lead to poor performance or require the replacement of CP modems. Millions of CP modems are currently in use, owned and operated by a multitude of customers. Replacing these modems presents a huge legacy issue. To address this problem a simple linear precoder was presented based on a channel diagonalizing criterion. This design has a low complexity and works with existing CP modems and operates close to the theoretical multi-user channel capacity.

As a by-product, the work in Chapter 4 and 5 produced a set of bounds on the determinants and inverses of diagonally dominant matrices, which are listed in Appendix B.

Despite the low complexity of the techniques presented in Chapter 4 and 5, signal coordination still requires a much larger run-time complexity than existing DSL modems. Crosstalk cancellation and precoding have a complexity that scales quadratically with the number of lines within a binder. For typical binders, which contain anywhere from 20 to 100 lines, these techniques are outside the scope of present day implementation. Chapter 6 addressed this problem through a technique known as partial cancellation.

It is well known that the majority of crosstalk experienced on a line comes from the 3 to 4 surrounding pairs in the binder. Furthermore, since crosstalk coupling varies dramatically with frequency, the worst effects are limited to a small selection of tones. Partial cancelers exploit this to achieve the majority of the performance of full cancellation at a fraction of the complexity. Whilst the idea of partial cancellation has been discussed in literature, to date no work has specifically focused on partial canceler design.

Chapter 6 investigated partial canceler and precoder design, which is essentially

a resource allocation problem. Given a limited amount of available run-time complexity, a modem must distribute this across lines and tones such that the data-rate is maximized. Chapter 6 presented the optimal algorithm for partial canceler design and several simpler, sub-optimal algorithms. These algorithms were shown to achieve 90% of the data-rate of full cancellation at less than 30% of the complexity.

Multi-user techniques address the three key challenges of speed, reach and symmetry facing DSL systems today. This thesis developed practical multi-user techniques for mitigating crosstalk in DSL. The techniques proposed have low complexity, low latency, and are compatible with existing customer premises equipment. In addition to being practical, the techniques were also shown to yield near-optimal performance, operating close to the theoretical multi-user channel capacity. As DSL continues to roll-out, multi-user techniques such as those proposed in this thesis will prove essential in maintaining an edge over competing broadband technologies.

Further Research

This thesis is only the first step in developing practical, multi-user techniques for DSL. Many interesting questions remain unanswered, and form the basis for future research. Some interesting, unexplored areas are listed here.

Autonomous Spectra Coordination

The spectra coordination algorithm developed in Chapter 3, optimal spectrum balancing, is a centralized algorithm requiring a spectrum management center (SMC) for direct implementation. In practice such a SMC may not exist, or may operate under limited knowledge of the network characteristics; for example the crosstalk channels are unknown in current DSL networks.

For this reason, autonomous algorithms are preferred from a practical standpoint. Autonomous algorithms operate independently on each modem, and only make use of information that can be measured at the modem, such as the line SNR. Autonomous algorithms minimize the overhead required for communication with the SMC, and increase the speed at which modems can adapt to changing line conditions.

In future work it will be interesting to develop autonomous DSM algorithms based on the insight gained from optimal spectrum balancing. The goal is to find a simple, autonomous algorithm which yields near-optimal performance in a broad range of scenarios. Early work in this area looks promising[20].

Combining Spectra Coordination with Bandplan Design

The spectra coordination algorithms investigated in Chapter 3 were constrained to operate under a fixed bandplan. This bandplan determines whether a tone is to be used for upstream or downstream transmission.

Using a fixed bandplan fixes the symmetry of service, which is the ratio between upstream and downstream rates. In many scenarios it is advantageous for an operator to vary the symmetry of service based on each customer's requirements, and there is already some discussion in standardization of allowing variable band-plans in fibre-to-the-basement deployments[81].

Jointly designing the transmit spectra and bandplan is a complex and typically intractable problem. A similar approach to optimal spectrum balancing could be used to find an efficient solution[21].

Combining Spectral and Partial Coordination

Chapter 6 discussed the design of partial crosstalk cancelers and precoders under the assumption of a flat transmit PSD. Performance could be improved by optimizing the transmit spectrum of each line. The combination of spectra coordination with partial cancellation is a difficult problem since the choice of transmit spectra is inherently linked to the choice of which crosstalkers to cancel.

Early work suggests that a dual decomposition could be used, in a similar fashion to optimal spectrum balancing, to reduce the complexity of joint transmit spectra and partial canceler design[110]. Significant work must still be done to turn this idea into an efficient and practical algorithm.

Adaptive Partial Cancelers and Precoders

The partial cancelers and precoders discussed in Chapter 6 rely on full channel knowledge. In practice this is straight-forward to obtain using MIMO channel identification techniques[112, 11, 57]. Nevertheless, it is still interesting to develop adaptive methods for partial cancellation, since these will lead to lower complexity, and faster adaption to changing line conditions.

The design of adaptive filters with sparsity constraints is a well known problem in the research community, on which a large body of work exists[68, 51]. These techniques could be applied in a DSL context. The goal is to develop a partial canceler that can adapt its filter coefficients as new modems come online and as the channel changes, whilst still maintaining sparsity. Early work in this area looks promising[69].

Fundamental Work

The optimal spectrum balancing algorithm provides the optimal solution to the spectra coordination problem. At the same time, it shows that dual decomposition can be applied to non-convex problems. This is a novel application of the dual decomposition, and suggests a more general algorithm that can be used to solve a whole class of non-convex problems.

It is interesting to generalize the optimal spectrum balancing algorithm, and clearly define the class of problems that it can solve. Early work suggests that this approach can efficiently solve any non-convex problem, provided that the objective can be decomposed, and that there is a degree of smoothness in the parameters on different tones[110]. Under these assumptions, a frequency-sharing argument can be used to ensure a duality-gap of zero in the final solution and optimality[110]. This work could significantly broaden the application of optimal spectrum balancing and is an important area for future research.

Another fundamental problem is the capacity of the interference channel. In Chapter 3 it was assumed that each modem treats crosstalk as background noise as this is the approach currently adopted in DSL modems. However, more advanced interference cancellation techniques can be employed, even when signal coordination is not possible. The capacity region for the interference channel is a long-standing open problem in information theory, and an important area for future work.

Appendices

A Optimality of Optimal Spectrum Balancing

The optimal spectrum balancing algorithm is listed as Alg. A.1. If discrete bitloading is employed then the maximization in the function *optimize_s* is limited to the PSD combinations corresponding to valid bitloading combinations, as calculated by (3.10).

The algorithm operates as follows. It is necessary to search through both λ_1 and λ_2 to find values which place sufficient importance on the total power constraint terms within the Lagrangian (3.12). Variation of w makes it possible to map out the optimal, achievable points on the convex hull of the rate region.

The algorithm contains three loops, an outer loop that varies w, an intermediate loop that searches for λ_1 and an inner loop that searches for λ_2. Bisection is used in each search.

When searching for λ_n, it is first necessary to find a value of λ_n which ensures that the power constraint of user n is satisfied. This value is stored in $\lambda_n^{\max}$. Note that a larger λ_n places more emphasis on the power constraint of user n in the Lagrangian. As a result, using a larger λ_n will result in a lower total power for user n.

Once $\lambda_n^{\max}$ is found the algorithm proceeds to bisection. Note that after the algorithm has completed, for each user either $\sum_k s_k^n = P_n$ or the corresponding Lagrangian multiplier is driven to zero, $\lambda_n = 0$. Thus the Lagrangian and the original objective become equivalent. More rigorously,

Theorem A.1 *For each w Alg. A.1 returns a PSD combination that is optimal for the spectrum management problem (3.2). That is, for some R_1^{target}*

$$\begin{aligned} \mathbf{s}_1, \mathbf{s}_2 \quad &= \quad \arg\max_{\mathbf{s}_1, \mathbf{s}_2} R_2, \\ &\text{s.t.} \quad R_1 \geq R_1^{\text{target}}, \end{aligned} \tag{A.1}$$

$$\sum_k s_k^n \leq P_n, \ \forall n,$$
$$0 \leq s_k^n \leq s_k^{n,\max}, \forall n, k.$$

Here R_1^{target} is in fact the rate of user 1 at convergence of the algorithm. Varying w from 0 to 1 allows all optimal operating points that lie on the convex hull of the rate region to be found. If the rate region is convex then all optimal operating points are found.

Proof : To prove the optimality of Alg. A.1 as stated in Theorem A.1, it is first shown that the algorithm converges. It will then be shown that at convergence maximizing the Lagrangian is equivalent to maximizing the weighted rate-sum (3.11). This implies the optimality of the PSDs generated by the algorithm[1].

To prove the convergence of Alg. A.1, the convergence of a related routine is first examined. This routine finds the optimal value for λ_n, thereby ensuring that the total power constraint on user n (3.3) is satisfied. At this value of λ_n the routine finds the optimal PSD for user n. As will be shown in Corollary A.5 and A.6, the algorithms *optimize_*λ_1 and *optimize_*λ_2 can be seen as special cases of this routine for specific values of the optimisation function $f(\mathbf{s}_n)$. Lemma A.3 proves that this routine converges. This in turn implies the convergence of *optimize_*λ_1 and *optimize_*λ_2.

First define the objective function

$$G(\mathbf{s}_n, \lambda_n) \triangleq f(\mathbf{s}_n) + \lambda_n(P_n - \sum_k s_k^n). \tag{A.2}$$

Denote the optimal power allocation for a given λ_n

$$\mathbf{s}_n(\lambda_n) \triangleq \arg\max_{\mathbf{s}_n} G(\mathbf{s}_n, \lambda_n),$$

with $s_k^n(\lambda_n) \triangleq [\mathbf{s}_n(\lambda_n)]_k$. The routine for user n is listed as Alg. A.2.

[1] This derivation resulted from close collaboration with Prof. Wei Yu, University of Toronto, Canada.

Algorithm A.1 Optimal Spectrum Balancing

Main Function

for $w = 0 \ldots 1$
  $\mathbf{s}_1, \mathbf{s}_2 = \text{optimize\_}\lambda_1(w)$
end

Function $\mathbf{s}_1, \mathbf{s}_2 =$ ***optimize***$\_\lambda_1(w)$

$\lambda_1^{\max} = 1$, $\lambda_1^{\min} = 0$
while $\sum_k s_k^1 > P_1$
  $\lambda_1^{\max} = 2\lambda_1^{\max}$
  $\mathbf{s}_1, \mathbf{s}_2 = \text{optimize\_}\lambda_2(w, \lambda_1^{\max})$
end
repeat
  $\lambda_1 = (\lambda_1^{\max} + \lambda_1^{\min})/2$
  $\mathbf{s}_1, \mathbf{s}_2 = \text{optimize\_}\lambda_2(w, \lambda_1)$
  if $\sum_k s_k^1 > P_1$, then $\lambda_1^{\min} = \lambda_1$, else $\lambda_1^{\max} = \lambda_1$
until convergence

Function $\mathbf{s}_1, \mathbf{s}_2 =$ ***optimize***$\_\lambda_2(w,\lambda_1)$

$\lambda_2^{\max} = 1$, $\lambda_2^{\min} = 0$
while $\sum_k s_k^2 > P_2$
  $\lambda_2^{\max} = 2\lambda_2^{\max}$
  $\mathbf{s}_1, \mathbf{s}_2 = \text{optimize\_}s(w, \lambda_1, \lambda_2^{\max})$
end
repeat
  $\lambda_2 = (\lambda_2^{\max} + \lambda_2^{\min})/2$
  $\mathbf{s}_1, \mathbf{s}_2 = \text{optimize\_}s(w, \lambda_1, \lambda_2)$
  if $\sum_k s_k^2 > P_2$, then $\lambda_2^{\min} = \lambda_2$, else $\lambda_2^{\max} = \lambda_2$
until convergence

Function $\mathbf{s}_1, \mathbf{s}_2 =$ ***optimize***$\_s(w, \lambda_1, \lambda_2)$

for $k = 1 \ldots K$
  $s_k^1, s_k^2 = \arg\max_{s_k^1, s_k^2} L_k(s_k^1, s_k^2, w, \lambda_1, \lambda_2)$
    s.t. $0 \le s_k^n \le s_k^{n,\max}$, $\forall n$
  (solve by 2-D exhaustive search)
end

Algorithm A.2 Routine for user n

$\lambda_n^{\max} = 1$, $\lambda_n^{\min} = 0$;
while $\sum_k s_k^n > P_n$
  $\lambda_n^{\max} = 2\lambda_n^{\max}$;
  $\mathbf{s}_n = \arg\max_{\mathbf{s}_n} f(\mathbf{s}_n) + \lambda_n^{\max}(P_n - \sum_k s_k^n)$;
end
repeat
  $\lambda_n = (\lambda_n^{\max} + \lambda_n^{\min})/2$;
  $\mathbf{s}_n = \arg\max_{\mathbf{s}_n} f(\mathbf{s}_n) + \lambda_n(P_n - \sum_k s_k^n)$;
  if $\sum_k s_k^n > P_n$, then $\lambda_n^{\min} = \lambda_n$, else $\lambda_n^{\max} = \lambda_n$;
until convergence

The following Lemma is used to prove the convergence of this routine.

Lemma A.2 *Fix n. $\sum_k s_k^n(\lambda_n)$ is monotonic decreasing in λ_n. Furthermore* $\lim_{\lambda_n \to \infty} \sum_k s_k^n(\lambda_n) = 0$.

Proof: Consider two Lagrangian multipliers λ_n^a and λ_n^b and their corresponding optimal PSDs $\mathbf{s}_n^a \triangleq \mathbf{s}_n(\lambda_n^a)$ and $\mathbf{s}_n^b \triangleq \mathbf{s}_n(\lambda_n^b)$. Denote the elements of these PSDs as $s_k^{n,a}$ and $s_k^{n,b}$ respectively. Let

$$\lambda_n^b \geq \lambda_n^a. \tag{A.3}$$

Define

$$\begin{aligned} A &\triangleq f(\mathbf{s}_n^b) + \lambda_n^b(P_n - \sum_k s_k^{n,b}), \\ B &\triangleq f(\mathbf{s}_n^a) + \lambda_n^b(P_n - \sum_k s_k^{n,a}), \\ C &\triangleq f(\mathbf{s}_n^a) + \lambda_n^a(P_n - \sum_k s_k^{n,a}), \\ D &\triangleq f(\mathbf{s}_n^b) + \lambda_n^a(P_n - \sum_k s_k^{n,b}). \end{aligned}$$

Now $G(\mathbf{s}_n^b, \lambda_n^b) \geq G(\mathbf{s}_n^a, \lambda_n^b)$ by the optimality of $\mathbf{s}_n^b$ in $G(\mathbf{s}_n, \lambda_n^b)$. Hence $A \geq B$. Similarly the optimality of $\mathbf{s}_n^a$ in $G(\mathbf{s}_n, \lambda_n^a)$ implies $C \geq D$. Consider 3 cases:

In the first case let $P_n - \sum_k s_k^{n,a} \geq 0$. Combining this with (A.3) implies $B \geq C$. Now $A \geq B \geq C \geq D$ implies $A - D \geq B - C$. Hence

$$(\lambda_n^b - \lambda_n^a)(P_n - \sum_k s_k^{n,b}) \quad \geq \quad (\lambda_n^b - \lambda_n^a)(P_n - \sum_k s_k^{n,a}),$$

which implies

$$\sum_k s_k^{n,a} \geq \sum_k s_k^{n,b}. \tag{A.4}$$

In the second case let $P_n - \sum_k s_k^{n,b} \leq 0$. Combining this with (A.3) implies $D \geq A$. Now $C \geq D \geq A \geq B$ implies $C - B \geq D - A$. Hence

$$(\lambda_n^a - \lambda_n^b)(P_n - \sum_k s_k^{n,a}) \geq (\lambda_n^a - \lambda_n^b)(P_n - \sum_k s_k^{n,b}),$$

which again implies (A.4).

In the third case let $P_n - \sum_k s_k^{n,a} < 0$ and $P_n - \sum_k s_k^{n,b} > 0$. This implies $P_n - \sum_k s_k^{n,b} > P_n - \sum_k s_k^{n,a}$ and again leads to (A.4).

So in all cases a larger λ_n leads to a smaller $\sum_k s_k^n$. This implies that $\sum_k s_k^n$ is monotonic decreasing in λ_n.

The second part of the Lemma is now proved. From (A.2) it can be shown that for large λ_n, $G(\mathbf{s}_n, \lambda_n) \simeq \lambda_n(P_n - \sum_k s_k^n)$ with the approximation becoming exact as $\lambda_n \to \infty$. Hence $\lim_{\lambda_n \to \infty} \mathbf{s}_n(\lambda_n) = \arg\max_{\mathbf{s}_n} \lambda_n(P_n - \sum_k s_k^n) = \mathbf{0}_{1\times K}$ where $\mathbf{0}_{1\times K}$ is the length K vector with all elements equal to zero. ■

Lemma A.3 Routine for user n *converges. At convergence*

$$\mathbf{s}_n = \arg\max_{\mathbf{s}_n} f(\mathbf{s}_n) \text{ s.t. } \sum_k s_k^n \leq P_n.$$

Proof : The routine consists of two stages: a preamble that determines $\lambda_n^{\max}$ and the actual routine itself.

The preamble clearly converges since from Lemma A.2, $\sum_k s_k^n(\lambda_n) \to 0$ as $\lambda_n \to \infty$.

The convergence of the main part of *routine for user n* can be shown as follows: $\lambda_n^{\max} - \lambda_n^{\min}$ decreases by half in each iteration. Thus, λ_n converges to a fixed value. Let's now consider two cases, depending on whether $\sum_k s_k^n(\lambda_n^{\min}) > P_n$ or not.

Suppose that $\sum_k s_k^n(\lambda_n^{\min}) > P_n$ at $\lambda_n^{\min} = 0$, then since the preamble ensures that $\sum_k s_k^n(\lambda_n^{\max}) \leq P_n$, throughout the algorithm it is always the case that $\sum_k s_k^n(\lambda_n^{\min}) > P_n$ and $\sum_k s_k^n(\lambda_n^{\max}) \leq P_n$. Since $\lambda_n^{\max} \geq \lambda_n \geq \lambda_n^{\min}$, $\lambda_n^{\min}$ and $\lambda_n^{\max}$ converge to a fixed value, and since $\sum_k s_k^n(\lambda_n)$ is monotonic in λ_n, this implies that $\sum_k s_k^n(\lambda_n)$ must converge to P_n. On the other hand, suppose that $\sum_k s_k^n(\lambda_n^{\min}) \leq P_n$ at $\lambda_n^{\min} = 0$. Then, λ_n will converge to 0.

Hence the algorithm will converge and at convergence either $\sum_k s_k^n(\lambda_n) = P_n$ or $\lambda_n = 0$. So at convergence

$$G(\mathbf{s}_n, \lambda_n) = f(\mathbf{s}_n).$$

In the routine

$$\begin{aligned} \mathbf{s}_n &= \arg\max_{\mathbf{s}_n} G(\mathbf{s}_n, \lambda_n), \\ &= \arg\max_{\mathbf{s}_n} f(\mathbf{s}_n) \text{ s.t. } \sum_k s_k^n \leq P_n. \end{aligned}$$

To see this, clearly $\mathbf{s}_n$ satisfies the constraint. Further, if there is some other feasible $\mathbf{s}'_n$ that does better than $\mathbf{s}_n$ for the objective function $f(\mathbf{s}_n)$, then $\mathbf{s}'_n$ should do better than $\mathbf{s}_n$ for the objective $G(\mathbf{s}_n, \lambda_n)$ also. This is contradicted by the optimality of $\mathbf{s}_n$ in $G(\mathbf{s}_n, \lambda_n)$. Hence $\mathbf{s}_n$ must be optimal in $f(\mathbf{s}_n)$. ■

Lemma A.4 *The function* optimize_s *(from Alg. 3.2) yields PSDs* $\mathbf{s}_1$ *and* $\mathbf{s}_2$ *which maximize the Lagrangian.*

Proof : From function *optimize_s*

$$s_k^1, s_k^2 = \arg\max_{s_k^1, s_k^2} L_k(w, \lambda_1, \lambda_2, s_k^1, s_k^2).$$

Since $L = \sum_k L_k + \lambda_1 P_1 + \lambda_2 P_2$, and since optimisation of the Lagrangian is unconstrained (recall that the constraints are incorporated into the objective function and need not be explicitly enforced) this implies

$$\mathbf{s}_1, \mathbf{s}_2 = \arg\max_{\mathbf{s}_1, \mathbf{s}_2} L(w, \lambda_1, \lambda_2, \mathbf{s}_1, \mathbf{s}_2).$$

■

Define the rates of user 1 and user 2 with the PSDs $\mathbf{s}_1$ and $\mathbf{s}_2$ as $R_1(\mathbf{s}_1, \mathbf{s}_2)$ and $R_2(\mathbf{s}_1, \mathbf{s}_2)$ respectively.

Corollary A.5 *The function* optimize_λ_2 *(from Alg. A.1) converges. At convergence*

$$\begin{aligned} \mathbf{s}_2 &= \arg\max_{\mathbf{s}_1, \mathbf{s}_2} wR_1(\mathbf{s}_1, \mathbf{s}_2) + (1-w)\, R_2(\mathbf{s}_1, \mathbf{s}_2) + \lambda_1(P_1 - \sum_k s_k^1); \\ &\text{s.t.} \quad \sum_k s_k^2 \leq P_2. \end{aligned} \tag{A.5}$$

Proof: Let $n = 2$ and $f(\mathbf{s}_2) \triangleq \max_{\mathbf{s}_1} wR_1(\mathbf{s}_1, \mathbf{s}_2) + (1-w)R_2(\mathbf{s}_1, \mathbf{s}_2) + \lambda_1(P_1 - \sum_k s_k^1)$. Lemma A.4 implies that *optimize_*λ_1 and Alg. A.2 are equivalent. Hence Lemma A.3 implies *optimize_*λ_2 converges, and that at convergence (A.5) is satisfied. ■

Corollary A.6 *The function* optimize_λ_1 *(from Alg. A.1) converges. At convergence*

$$\begin{aligned} \mathbf{s}_1 \quad &= \quad \arg\max_{\mathbf{s}_1, \mathbf{s}_2} wR_1(\mathbf{s}_1, \mathbf{s}_2) + (1-w)R_2(\mathbf{s}_1, \mathbf{s}_2); \\ &\text{s.t.} \quad \sum_k s_k^1 \leq P_1, \ \sum_k s_k^2 \leq P_2. \end{aligned} \tag{A.6}$$

Proof: Let $n = 1$ and

$$f(\mathbf{s}_1) = \max_{\mathbf{s_2}} wR_1(\mathbf{s}_1, \mathbf{s}_2) + (1-w)R_2(\mathbf{s}_1, \mathbf{s}_2) \text{ s.t. } \sum_k s_k^2 \leq P_2.$$

Then Lemma A.3 and Corollary A.5 imply *optimize_*λ_1 converges, and that at convergence (A.6) is satisfied. ■

From Theorem 3.1, for any particular w, there exists some R_1^{target} for which the weighted rate-sum optimization (A.6) is equivalent to the original spectrum management problem (3.14). Hence for any particular w the weighted rate-sum optimization leads to an optimal operating point.

Corollary A.6 implies that for each value of w in Alg. 3.2, the PSD combination returned by the algorithm maximizes a weighted-rate sum. Hence the PSD combination is also an optimal solution to the spectrum management problem (3.14). Furthermore, Theorem 3.2 states that by varying w from 0 to 1 it is possible to map out all achievable operating points on the boundary of the convex hull of the rate region.

This appendix has only explicitly proved optimality for the 2-user case. The proof for the N-user case follows inductively from what is given here. This concludes the proof of Theorem A.1. ■

It should be made clear that even when the rate region is non-convex, the PSD combinations returned by the optimal spectrum balancing algorithm are optimal, resulting in an operating point on the boundary of the rate region. The convexity of the rate region affects only the ability of the proposed algorithm to explore all optimal operating points. It does *not* affect the optimality of the points found by the algorithm.

Note that the cost function on each tone L_k is still non-convex. Hence the optimization of L_k must be solved through exhaustive search, which has an

exponential complexity in N. The important observation is that since the optimization on each tone is solved independently the algorithm has a linear, rather than exponential, complexity in K. This results in a computationally tractable algorithm.

If the function *optimize_s* finds multiple PSD combinations that yield the same value for the Lagrangian L_k, then all PSD combinations are stored. This ensures that if multiple points in the rate region are optimal for the same weight w then each of these points is discovered.

B Bounds on Diagonally Dominant Matrices

Define the set $\mathbb{A}^{(N)}$ of $N \times N$ matrices, such that for any $\mathbf{A}^{(N)} \in \mathbb{A}^{(N)}$

$$\begin{aligned} \left| \left[\mathbf{A}^{(N)} \right]_{n,n} \right| &= 1; \\ \left| \left[\mathbf{A}^{(N)} \right]_{n,m} \right| &\leq \alpha_k, \ \forall n \neq m. \end{aligned}$$

Define the set $\mathbb{B}^{(N)}$ of $N \times N$ matrices, such that for any $\mathbf{B}^{(N)} \in \mathbb{B}^{(N)}$

$$\begin{aligned} \left| \left[\mathbf{B}^{(N)} \right]_{n,n} \right| &= 1, \ \forall n < N; \\ \left| \left[\mathbf{B}^{(N)} \right]_{N,N} \right| &\leq \alpha_k; \\ \left| \left[\mathbf{B}^{(N)} \right]_{n,m} \right| &\leq \alpha_k, \ \forall n \neq m. \end{aligned}$$

Theorem B.1 *Consider any* $\mathbf{A}^{(N)} \in \mathbb{A}^{(N)}$ *and* $\mathbf{B}^{(N)} \in \mathbb{B}^{(N)}$. *The magnitude of the determinants of* $\mathbf{A}^{(N)}$ *and* $\mathbf{B}^{(N)}$ *can be bounded as follows*

$$\left| \det(\mathbf{A}^{(N)}) \right| \leq A_{\max}^{(N)}, \tag{B.1}$$

$$\left| \det(\mathbf{B}^{(N)}) \right| \leq B_{\max}^{(N)}, \tag{B.2}$$

where

$$\begin{bmatrix} A_{\max}^{(m)} \\ B_{\max}^{(m)} \end{bmatrix} \triangleq \left(\prod_{i=1}^{m} \begin{bmatrix} 1 & (i-1)\alpha_k \\ \alpha_k & (i-1)\alpha_k \end{bmatrix} \right) \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

and

$$A_{\min}^{(m)} \triangleq 1 - \sum_{i=1}^{m} \alpha_k (i-1) B_{\max}^{(i-1)}.$$

Furthermore, if

$$A_{min}^{(m)} \geq \alpha_k n B_{\max}^{(m)}, \ \forall m < N; \tag{B.3}$$

then the following bound also holds

$$\left| \det(\mathbf{A}^{(N)}) \right| \geq A_{\min}^{(N)}. \tag{B.4}$$

Note that $|.|$ *denotes the absolute value operator, whilst* $\det(.)$ *denotes the determinant operator.*

Proof: The proof is based on induction. Begin by assuming that the bounds (B.1), (B.2) and (B.4) hold for any $N \times N$ matrices of the form $\mathbf{A}^{(N)}$ and $\mathbf{B}^{(N)}$ for some specific value of N. Now consider any matrix $\mathbf{A}^{(N+1)} \in \mathbb{A}^{(N+1)}$. Decompose $\mathbf{A}^{(N+1)}$ as

$$\mathbf{A}^{(N+1)} = \begin{bmatrix} & & & a_{1,N+1} \\ & \mathbf{A}^{(N)} & & \vdots \\ & & & a_{N,N+1} \\ a_{N+1,1} & \cdots & a_{N+1,N} & 1 \end{bmatrix},$$

where $a_{n,m} \triangleq \left[\mathbf{A}^{(N+1)}\right]_{n,m}$ and $\mathbf{A}^{(N)}$ is the submatrix containing the first N rows and columns of $\mathbf{A}^{(N+1)}$. By expanding the determinant along the last row of $\mathbf{A}^{(N+1)}$ it can be seen that

$$\begin{aligned} &\left|\det(\mathbf{A}^{(N+1)})\right| \\ &= \left|\det(\mathbf{A}^{(N)})\right. \\ &+ \textstyle\sum_{m=1}^{N}(-1)^{N+1-m} a_{N+1,m} \det\left(\begin{bmatrix} \overline{\mathbf{A}}_m^{(N)} & \mathbf{a}_{N+1} \end{bmatrix}\right)\Big|, \\ &\leq \left|\det(\mathbf{A}^{(N)})\right| + \textstyle\sum_{m=1}^{N} \alpha_k \left|\det\left(\begin{bmatrix} \overline{\mathbf{A}}_m^{(N)} & \mathbf{a}_{N+1} \end{bmatrix}\right)\right|, \end{aligned} \tag{B.5}$$

where $\overline{\mathbf{A}}_m^{(N)}$ is the sub-matrix formed by removing column m from $\mathbf{A}^{(N)}$ and $\mathbf{a}_{N+1} \triangleq [a_{1,N+1} \dots a_{N,N+1}]^T$. The second line exploits the fact that row permutation does not affect the magnitude of a determinant. Define the permutation matrix

$$\Pi_m \triangleq [\mathbf{e}_1 \cdots \mathbf{e}_{m-1}\ \mathbf{e}_{m+1} \dots \mathbf{e}_N\ \mathbf{e}_m],$$

where $\mathbf{e}_m$ is defined as the mth column of the $N \times N$ identity matrix. Note that $\Pi_m^T[\overline{\mathbf{A}}_m^{(N)}\ \mathbf{a}_{N+1}] \in \mathbb{B}^{(N)}$. Using the fact that row permutations have no effect on the magnitude of a determinant, together with (B.1), (B.2) and (B.5) now yields

$$\left|\det(\mathbf{A}^{(N+1)})\right| \leq A_{\max}^{(N)} + \alpha_k N B_{\max}^{(N)},$$

hence

$$A_{\max}^{(N+1)} = A_{\max}^{(N)} + \alpha_k N B_{\max}^{(N)}. \tag{B.6}$$

Now consider any matrix $\mathbf{B}^{(N+1)} \in \mathbb{B}^{(N+1)}$. Decompose $\mathbf{B}^{(N+1)}$ as

$$\mathbf{B}^{(N+1)} = \begin{bmatrix} & & & b_{1,N+1} \\ & \mathbf{C}^{(N)} & & \vdots \\ & & & b_{N,N+1} \\ b_{N+1,1} & \cdots & b_{N+1,N} & b_{N+1,N+1} \end{bmatrix},$$

where $b_{n,m} \triangleq \left[\mathbf{B}^{(N+1)}\right]_{n,m}$ and $\mathbf{C}^{(N)}$ is the submatrix containing the first N rows and columns of $\mathbf{B}^{(N+1)}$. By expanding the determinant along the last row

of $\mathbf{B}^{(N+1)}$ it can be seen that

$$\begin{aligned}&\left|\det(\mathbf{B}^{(N+1)})\right| \\ &= \left|b_{N+1,N+1}\det(\mathbf{C}^{(N)})\right. \\ &\left.+\sum_{m=1}^{N}(-1)^{N+1-m}b_{N+1,m}\det\left(\left[\begin{array}{cc}\overline{\mathbf{C}}_m^{(N)} & \mathbf{b}_{N+1}\end{array}\right]\right)\right|,\end{aligned} \tag{B.7}$$

where $\overline{\mathbf{C}}_m^{(N)}$ is the sub-matrix formed by removing column m from $\mathbf{C}^{(N)}$ and $\mathbf{b}_{N+1} \triangleq [b_{1,N+1}\dots b_{N,N+1}]^T$. Note that $\mathbf{C}^{(N)} \in \mathbb{A}^{(N)}$ and

$$\Pi_m^T\left[\begin{array}{cc}\overline{\mathbf{C}}_m^{(N)} & \mathbf{b}_{N+1}\end{array}\right] \in \mathbb{B}^{(N)}.$$

Using the fact that row permutations have no effect on the magnitude of a determinant, together with (B.1), (B.2), and (B.7) now yields

$$\left|\det(\mathbf{B}^{(N+1)})\right| \leq \alpha_k A_{\max}^{(N)} + \alpha_k N B_{\max}^{(N)},$$

hence

$$B_{\max}^{(N+1)} = \alpha_k A_{\max}^{(N)} + \alpha_k N B_{\max}^{(N)}. \tag{B.8}$$

Combining (B.6) and (B.8) in matrix form yields

$$\left[\begin{array}{c}A_{\max}^{(N+1)} \\ B_{\max}^{(N+1)}\end{array}\right] = \left[\begin{array}{cc}1 & \alpha_k N \\ \alpha_k & \alpha_k N\end{array}\right]\left[\begin{array}{c}A_{\max}^{(N)} \\ B_{\max}^{(N)}\end{array}\right]. \tag{B.9}$$

We now proceed with the inductive proof. First note that $\left|\mathbf{A}^{(1)}\right| = 1$ and $\left|\mathbf{B}^{(1)}\right| \leq \alpha_k$, so (B.1) and (B.2) hold for $N = 1$. Hence through induction, (B.9) implies that (B.1) and (B.2) must hold for all N. This concludes the proof for the upper bounds (B.1) and (B.2). We now turn our attention to the lower bound (B.4). We assume that (B.4) holds for some specific value of N. Hence

$$\left|\det(\mathbf{A}^{(N)})\right| \geq A_{\min}^{(N)}.$$

Eq. (B.3) now implies

$$\begin{aligned}\left|\det(\mathbf{A}^{(N)})\right| &\geq \alpha_k N B_{\max}^{(N)}, \\ &\geq \left|\sum_{m=1}^{N}(-1)^{N+1-m}a_{N+1,m}\det([\overline{\mathbf{A}}_m^{(N)}\,\mathbf{a}_{N+1}])\right|,\end{aligned}$$

where (B.2) is used in the second line. Combining this with (B.5) and (B.2) now implies

$$\left|\det(\mathbf{A}^{(N+1)})\right| \geq \left|\det(\mathbf{A}^{(N)})\right| - \alpha_k N B_{\max}^{(N)},$$

which together with (B.4) yields

$$A_{\min}^{(N+1)} = A_{\min}^{(N)} - \alpha_k N B_{\max}^{(N)}. \tag{B.10}$$

First note that $\left|\mathbf{A}^{(1)}\right| = 1$ and $A^{(1)}_{\min} = 1$, so (B.4) holds for $N = 1$. Hence through induction, (B.10) implies that (B.4) holds for all N. This concludes the proof for the lower bound (B.4). ■

Theorem B.2 *If* $\mathbf{G} \in \mathbb{A}^{(N)}$ *and* $A^{(n)}_{min} \leq \alpha_k n B^{(n)}_{\max}$, $\forall n < N$; *then the magnitude of the elements of* $\mathbf{G}^{-1}$ *can be bounded*

$$\left|\left[\mathbf{G}^{-1}\right]_{n,m}\right| \leq \begin{cases} A^{(N-1)}_{max}/A^{(N)}_{min}, & n = m; \\ B^{(N-1)}_{max}/A^{(N)}_{min}, & n \neq m. \end{cases} \tag{B.11}$$

Proof : By definition of the matrix inverse

$$\left|\left[\mathbf{G}^{-1}\right]_{n,m}\right| = \left|\det(\overline{\mathbf{G}}^{m,n})\right| / \left|\det(\mathbf{G})\right|, \tag{B.12}$$

where $\overline{\mathbf{G}}^{m,n}$ is the sub-matrix formed by removing row m and column n from $\mathbf{G}$. Now $\mathbf{G} \in \mathbb{A}^{(N)}$ so from theorem B.1

$$\left|\det(\mathbf{G})\right| \geq A^{(N)}_{\min}. \tag{B.13}$$

If $m = n$ then $\overline{\mathbf{G}}^{m,n} \in \mathbb{A}^{(N-1)}$ and from theorem B.1

$$\left|\det(\overline{\mathbf{G}}^{m,m})\right| \leq A^{(N-1)}_{\max}, \forall m. \tag{B.14}$$

If $m \neq n$ then $\Pi^T_n \overline{\mathbf{G}}^{m,n} \Pi_m \in \mathbb{B}^{(N-1)}$ and from theorem B.1

$$\left|\det(\overline{\mathbf{G}}^{m,n})\right| = \left|\det(\Pi^T_n \overline{\mathbf{G}}^{m,n} \Pi_m)\right| \leq B^{(N-1)}_{\max}, \forall m \neq n. \tag{B.15}$$

Combining (B.12), (B.13), (B.14) and (B.15) yields (B.11), which concludes the proof. ■

Bibliography

[1] *Asymmetrical digital subscriber line (ADSL) transceivers*, ITU Std. G.992.1, 1999.

[2] *Physical Layer Parameters and Specification for 1000Mb/s Operation Over 4-pair of Category 5 Balanced Copper Cabling, Type 1000BASE-T*, IEEE Std. 802.3ab, 1999.

[3] *Asymmetrical digital subscriber line transceivers 2 (ADSL2)*, ITU Std. G.992.2, 2002.

[4] *Asymmetrical Digital Subscriber Line (ADSL) transceivers - Extended bandwidth ADSL2 (ADSL2+)*, ITU Std. G.992.5, 2003.

[5] *Physical layer management for Digital Subscriber Line (DSL) transcievers*, ITU Std. G.997.1, 2003.

[6] *Spectrum Management for Loop Transmission Systems*, ANSI Std. T1.417, Issue 2, 2003.

[7] *Very high speed Digital Subscriber Line (VDSL); Functional Requirements*, ETSI Std. TS 101 270-1, Rev. V.1.3.1, 2003.

[8] *Dynamic Spectrum Management*, ANSI Draft Std. T1E1.4/2003-018, Rev. 15, 2004.

[9] *Very-high bit-rate Digital Subscriber Lines (VDSL) Metallic Interface*, ANSI Std. T1.424, 2004.

[10] M. Abdulrahman and D. Falconer, "Cyclostationary crosstalk suppression by decision feedback equalization on digital subscriber loops," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 640–649, Apr 1992.

[11] C. Aldana and J. Cioffi, "Channel tracking for multiple input, single output systems using EM algorithm," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Helsinki, Finland, June 2001, pp. 586–590.

[12] E. Baccarelli, A. Fasano, and M. Biagi, "Novel Efficient Bit-Loading Algorithms for Peak-Energy-Limited ADSL-Type Multicarrier Systems," *IEEE Trans. Signal Processing*, vol. 50, pp. 1237–1247, May 2002.

[13] J. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5–14, May 1990.

[14] T. Bostoen, J. Verlinden, R. Cendrillon, and M. Moonen, "DSM in Practice: Iterative Waterfilling Implemented on ADSL Modems," in *Proc. IEEE Int. Conf. on Acoust., Speech and Sig. Processing (ICASSP)*, Montreal, May 2004.

[15] J. Campello, "Optimal discrete bit loading for multicarrier modulation systems," in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, 1998, p. 193.

[16] A. Carleial, "Interference channels," *IEEE Trans. Inform. Theory*, vol. 24, pp. 60–70, Jan 1978.

[17] R. Cendrillon, G. Ginis, and M. Moonen, "Improved Linear Crosstalk Precompensation for Downstream VDSL," in *Proc. IEEE Int. Conf. on Acoust., Speech and Sig. Processing (ICASSP)*, vol. 4, Montreal, May 2004, pp. 1053–1056.

[18] R. Cendrillon, G. Ginis, M. Moonen, and E. Van den Bogaert, "A Near-Optimal Linear Crosstalk Canceler for DSL," submitted to *IEEE Trans. Commun.*, 2004.

[19] R. Cendrillon, G. Ginis, M. Moonen, and J. Verlinden, "A Near-Optimal Linear Crosstalk Precoder for DSL," submitted to *IEEE Trans. Signal Processing*, 2004.

[20] R. Cendrillon and M. Moonen, "Iterative Spectrum Balancing for Digital Subscriber Lines," submitted to *IEEE Int. Conf. on Commun. (ICC)* 2005.

[21] R. Cendrillon and M. Moonen, "Optimal Bandplan Design for Digital Subscriber Lines," to be submitted to *IEEE Trans. Commun.*

[22] R. Cendrillon and M. Moonen, "Efficient Equalizers for Single and Multicarrier Environments with Known Symbol Padding," in *Proc. Int. Symposium on Signal Processing and its Applications (ISSPA)*, Kuala Lampur, August 2001, pp. 607–610.

[23] R. Cendrillon, M. Moonen, T. Bostoen, and G. Ginis, "The Linear Zero-Forcing Crosstalk Canceller is Near-optimal in DSL Channels," accepted for publication in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)* 2004.

[24] R. Cendrillon, M. Moonen, G. Ginis, K. Van Acker, T. Bostoen, and P. Vandaele, "Partial Crosstalk Cancellation Exploiting Line and Tone Selection in Upstream VDSL," in *Proc. Baiona Workshop on Signal Processing in Communications*, Spain, September 2003, pp. 229–234.

[25] R. Cendrillon, M. Moonen, G. Ginis, K. Van Acker, T. Bostoen, and P. Vandaele, "Partial Crosstalk Cancellation for Upstream VDSL," *EURASIP Journal on Applied Signal Processing*, vol. 10, pp. 1433–1448, Oct. 2004.

[26] R. Cendrillon, M. Moonen, G. Ginis, K. Van Acker, T. Bostoen, and P. Vandaele, "Partial Crosstalk Precompensation for Downstream VDSL," *Signal Processing*, vol. 84, pp. 2005–2019, Nov. 2004.

[27] R. Cendrillon, M. Moonen, D. Gore, and A. Paulraj, "Low Complexity Crosstalk Cancellation through Line Selection in Upstream VDSL," in *Proc. IEEE Int. Conf. on Acoust., Speech and Sig. Processing (ICASSP)*, Hong Kong, April 2003, pp. 692–695.

[28] R. Cendrillon, M. Moonen, and R. Suciu, "Simplified Power Allocation for the DSL Multi-access Channel through Column-wise Diagonal Dominance," in *Proc. Int. Symposium on Image and Signal Processing and Analysis (ISPA)*, Rome, September 2003, pp. 634–638.

[29] R. Cendrillon, M. Moonen, R. Suciu, and G. Ginis, "Simplified Power Allocation and TX/RX Structure for MIMO-DSL," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, San Francisco, December 2003.

[30] R. Cendrillon, M. Moonen, K. Van Acker, P. Vandaele, and T. Bostoen, "Method for Crosstalk Cancellation," Applicant: Alcatel Bell, European Patent EP1 109 329, 2002.

[31] R. Cendrillon, M. Moonen, K. Van Acker, P. Vandaele, and T. Bostoen, "Method to Provide Crosstalk Cancellation," Applicant: Alcatel Bell, United States Patent US5 887 032, 2003.

[32] R. Cendrillon, M. Moonen, K. Van Acker, J. Verlinden, T. Bostoen, and E. Van den Bogaert, "Power Control Method for Remotely Deployed Communication Service," Applicant: Alcatel Bell, International Patent Pending, 2003.

[33] R. Cendrillon, M. Moonen, J. Verlinden, and T. Bostoen, *A Linear Crosstalk Precoder that Achieves the Multi-user DSL Channel Capacity*, ANSI Std. Contrib., to be presented at T1E1.4 meeting, Dec., 2004.

[34] R. Cendrillon, O. Rousseaux, M. Moonen, E. Van den Bogaert, and J. Verlinden, "Waterfilling in MIMO Systems with Power Constraints on Each Transmitter," in *Proc. Symposium on Information Theory in the Benelux*, Netherlands, May 2003, pp. 219–226.

[35] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, *On the Optimality of Iterative Waterfilling in DSL*, ANSI Std. Contrib. T1E1.4/2003-325, 2004.

[36] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, *Optimal Spectrum Management*, ANSI Std. Contrib. T1E1.4/2003-365, Rev. 1, 2004.

[37] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, *Optimal Spectrum Management*, ANSI Std. Contrib. T1E1.4/2003-459, Rev. 2, 2004.

[38] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, *Proof of the Optimality of Optimal Spectrum Management*, ANSI Std. Contrib. T1E1.4/2003-365, 2004.

[39] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal Multi-user Spectrum Balancing for Digital Subscriber Lines," accepted for publication (subject to minor revision) in *IEEE Trans. Commun.*, 2004.

[40] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal Multi-user Spectrum Management for Digital Subscriber Lines," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Paris, June 2004, pp. 1–5.

[41] K. Cheong, W. Choi, and J. Cioffi, "Multiuser Soft Interference Canceler via Iterative Decoding for DSL Applications," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 363–371, Feb. 2002.

[42] S. Cherry, "The wireless last mile," *IEEE Spectr.*, vol. 40, pp. 18–22, Sept. 2003.

[43] G. Cherubini, "Optimum upstream power back-off and multiuser detection for VDSL," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, 2001, pp. 375–380.

[44] S. Chung, *DSL Handbook*. Auerbach, 2004, ch. 12 - Dynamic Spectrum Management.

[45] S. Chung and J. Cioffi, "Rate and power control in a two-user multicarrier channel with no coordination: the optimal scheme vs. a sub-optimal method," *IEEE Trans. Commun.*, vol. 51, pp. 1768–1772, Nov 2003.

[46] S. Chung and J. Cioffi, "The capacity region of frequency-selective Gaussian interference channels under strong interference," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, vol. 4, Paris, May 2003, pp. 11–15.

[47] J. Cioffi, *DSL Advances*. Prentice Hall, 2002, ch. 11 - Dynamic Spectrum Management.

[48] T. Cover and J. Thomas, *Elements of Information Theory.* Wiley, 1991.

[49] H. Dai and V. Poor, "Turbo multiuser detection for coded DMT VDSL systems," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 351–362, Feb. 2002.

[50] DSL Forum, "DSL Maintains Global Broadband Dominance: Wins North American Broadband Growth Race," Sept. 2004.

[51] I. Duff, A. Erisman, and J. Reid, *Direct Methods for Sparse Matrices.* Oxford Science Publications, 1986.

[52] J. L. Fang, "Modeling and characterization of copper access systems," Ph.D. dissertation, Stanford University, 2002.

[53] G. Forney and M. Eyuboglu, "Combined Equalization and Coding Using Precoding," *IEEE Commun. Mag.*, vol. 29, pp. 25–34, Dec. 1991.

[54] G. Foschini and M. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[55] N. Frigo, P. Iannone, and C. Reichmann, "A View of Fiber to the Home Economics," *IEEE Commun. Mag.*, vol. 42, pp. 516–523, Aug. 2004.

[56] M. Gagnaire, "An overview of broad-band access technologies," *Proc. IEEE*, vol. 85, pp. 1958–1972, Dec. 1997.

[57] S. Galli, C. Valenti, and K. Kerpez, "A frequency-domain approach to crosstalk identification in xDSL systems," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1497–1506, Aug 2001.

[58] G. Ginis and J. Cioffi, "A Multi-user Precoding Scheme achieving Crosstalk Cancellation with Application to DSL Systems," in *Proc. 34th Asilomar Conference*, Pacific Grove, CA, Oct. 2000, pp. 1627–1631.

[59] G. Ginis and J. Cioffi, "Vectored Transmission for Digital Subscriber Line Systems," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 1085–1104, June 2002.

[60] G. Ginis and J. Cioffi, "Vectored-DMT: a FEXT canceling modulation scheme for coordinating users," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, vol. 1, June 2001, pp. 305–309.

[61] G. Ginis, "Multi-line coordinated communication for broadband access networks," Ph.D. dissertation, Stanford University, 2002.

[62] A. Goldsmith and M. Effros, "The capacity region of broadcast channels with intersymbol interference and colored Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 47, pp. 211–219, Jan. 2001.

[63] D. Gore, R. Nabar, and A. Paulraj, "Selecting an optimal set of transmit antennas for a low rank matrix channel," in *Proc. IEEE Int. Conf. on Acoust., Speech and Sig. Processing (ICASSP)*, 2000, pp. 2785–2788.

[64] D. Gore and A. Paulraj, "Space-time block coding with optimal antenna selection," in *Proc. IEEE Int. Conf. on Acoust., Speech and Sig. Processing (ICASSP)*, 2001, pp. 2441–2444.

[65] D. Gore and A. Paulraj, "Statistical MIMO antenna sub-set selection with space-time coding," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, 2002, pp. 641–645.

[66] A. Gorokhov, "Antenna Selection Algorithms for MEA Transmission Systems," in *Proc. IEEE Int. Conf. on Acoust., Speech and Sig. Processing (ICASSP)*, 2002, pp. 2857–2860.

[67] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. 20, pp. 774–780, Aug 1972.

[68] G. Harikumar, C. Couvreur, and Y. Bresler, "Fast optimal and suboptimal algorithms for sparse solutions to linear inverse problems," in *Proc. Int. Conf. on Acoustics, Speech and Sig. Processing*, 1998, pp. 1877–1880.

[69] J. Homer, R. Cendrillon, and M. Moonen, "Adaptive Partial Crosstalk Cancellation for Digital Subscriber Lines," submitted to *EURASIP Journal on Applied Signal Processing*, 2004.

[70] M. Honig, P. Crespo, and K. Steiglitz, "Suppression of near- and far-end crosstalk by linear pre- and post-filtering," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 614–629, Apr 1992.

[71] M. Honig, K. Steiglitz, and B. Gopinath, "Multichannel signal processing for data communications in the presence of crosstalk," *IEEE Trans. Commun.*, vol. 38, pp. 551–558, Apr 1990.

[72] ITU, "Broadband - Opening up to the future," *Policy and Strategy Trends*, pp. 1–4, April 2003.

[73] K. Jacobsen, "Methods of upstream power backoff on very high-speed digital subscriber lines," *IEEE Commun. Mag.*, pp. 210–216, Mar. 2001.

[74] V. Joshi and D. Falconer, "Sequence estimation techniques for digital subscriber loop transmission with crosstalk interference," *IEEE Trans. Commun.*, vol. 38, pp. 1367–1374, Sep 1990.

[75] K. Kerpez, "Near-end crosstalk is almost Gaussian," *IEEE Trans. Commun.*, vol. 41, pp. 670–672, Jan 1993.

[76] I. Koffman and V. Roman, "Broadband wireless access solutions based on OFDM access in IEEE 802.16," vol. 40, pp. 96–103, 2002.

[77] Z. Mingliu and R. Wolff, "Crossing the digital divide: cost-effective broadband wireless access for rural and remote areas," vol. 42, pp. 99–105, Feb. 2004.

[78] T. Monath, N. Kristian, P. Cadro, D. Katsianis, and D. Varoutas, "Economics of Fixed Broadband Access Network Strategies," *IEEE Commun. Mag.*, vol. 41, pp. 132–139, Sept. 2003.

[79] V. Oksman, *Optimization of the PSD_REF for Upstream Power Back-off in VDSL*, ANSI Std. Contrib. T1E1.4/2001-102, 2001.

[80] A. Peled and A. Ruiz, "Frequency domain data transmission using reduced computational complexity algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, Denver, April 1980, pp. 964–967.

[81] A. Redfern, *A Bandplan Framework for VDSL2*, ANSI Std. Contrib. T1E1.4/2003-503, 2004.

[82] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inform. Theory*, vol. 27, pp. 786–788, Nov 1981.

[83] S. Schelstraete and V. Oksman, *Proposal for the North-American reference PSD*, ANSI Std. Contrib. T1E1.4/2001-157, 2001.

[84] B. Schrick and M. Riezenman, "Wireless broadband in a box," *IEEE Spectr.*, vol. 39, pp. 38–43, June 2002.

[85] A. Sendonaris, V. Veeravalli, and A. Aazhang, "Joint signaling strategies for approaching the capacity of twisted-pair channels," *IEEE Trans. Commun.*, vol. 46, pp. 673–685, May 1998.

[86] C. Shannon, "Two-way Communication Channels," in *Proc. Fourth Berkeley Symposium on Probability and Statistics*, vol. 1, 1961, pp. 611–644.

[87] F. Sjoberg, S. Wilson, and M. Isaksson, "Power back-off in the upstream of VDSL," in *Proc. Radiovetenskaplig Koferens (RVK'99)*, vol. 1, Karlskrona, Sweden, June 1999, pp. 436–440.

[88] K. Song, S. Chung, G. Ginis, and J. Cioffi, "Dynamic spectrum management for next-generation DSL systems," *IEEE Commun. Mag.*, vol. 40, pp. 101–109, Oct 2002.

[89] T. Starr, J. Cioffi, and P. Silverman, *Understanding Digital Subscriber Line Technology.* Prentice Hall, 1999.

[90] G. Taubock and W. Henkel, "MIMO Systems in the Subscriber-Line Network," in *Proc. of the 5th Int. OFDM-Workshop*, 2000, pp. 18.1–18.3.

[91] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *IEEE Electron. Lett.*, vol. 7, pp. 138–139, Mar 1971.

[92] M. Tsatsanis, *DSL Handbook*. Auerbach, 2004, ch. 13 - Multi-line DSL System Architectures.

[93] D. Tse and S. Hanly, "Multiaccess Fading Channels-Part I: Polymatroid Structure, Optimal Resources Allocation and Throughput Capacities," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2796–2815, Nov. 1998.

[94] E. Van den Bogaert, T. Bostoen, R. Zeroual, B. Van Wauwe, F. Van de Schueren, R. Cendrillon, and M. Moonen, "Ensuring The Robustness Against Non-Stationary Noise Of ADSL Transceivers When Applying DSM," in *Proc. Int. Conf. Commun. and Computer Networks*, Marbella, Spain, September 2004.

[95] E. Van den Bogaert, K. Van Acker, J. Verlinden, R. Cendrillon, and M. Moonen, "Waterfilled VDSL Echo Limitation for Rate-Reach Performance Improvement," submitted to *IEEE Int. Conf. on Commun. (ICC)*, 2005.

[96] M. Varanasi and T. Guess, "Optimum Decision Feedback Multiuser Equalization with Successive Decoding Achieves the Total Capacity of the Gaussian Multiple-Access Channel," in *Proc. of Asilomar Conf. on Signals, Syst. and Comput.*, 1997, pp. 1405–1409.

[97] J. Verlinden, E. Van den Bogaert, T. Bostoen, R. Cendrillon, and M. Moonen, "Protecting the robustness of ADSL and VDSL DMT modems when applying DSM," in *Proc. Int. Zurich Seminar on Commun. (IZS)*, 2004.

[98] J. Verlinden, E. Van den Bogaert, T. Bostoen, F. Zanier, M. Luise, R. Cendrillon, and M. Moonen, "Expected Rate-Reach Improvements of Dynamic Spectrum Management when applied to Real-Life Networks," in *Proc. Int. Conf. Commun. and Computer Networks*, Marbella, Spain, September 2004.

[99] P. Viswanath and D. Tse, "Sum capacity of the multiple antenna broadcast channel," in *Proc. Int. Symp. Inform. Theory (ISIT)*, 2002.

[100] P. Viswanath, D. Tse, and V. Anantharam, "Asymptotically Optimal Water-Filling in Vector Multiple-Access Channels," *IEEE Trans. Inform. Theory*, vol. 47, pp. 241–267, Jan. 2001.

[101] H. Weingarten, Y. Steinberg, and S. Shamai, "The Capacity Region of the Gaussian MIMO Broadcast channel," submitted to *IEEE Trans. Inform. Theory*, 2004.

[102] S. Weinstein and P. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun. Technol.*, vol. 19, pp. 626–634, Oct 1971.

[103] L. Xiao, M. Johansson, and S. Boyd, "Simultaneous routing and resource allocation via dual decomposition," *IEEE Trans. Commun.*, vol. 52, pp. 1136–1144, July 2004.

[104] W. Yu, "Spatial Multiplex in Downlink Multiuser Multiple-Antenna Wireless Environments," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, San Francisco, Dec. 2003.

[105] W. Yu and J. Cioffi, "Multiuser Detection in Vector Multiple Access Channels using Generalized Decision Feedback Equalization," in *Proc. Int. Conf. on Signal Processing, World Computer Congress*, 2000.

[106] W. Yu and J. Cioffi, "Sum Capacity of Gaussian Vector Broadcast Channels," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1875–1892, September 2004.

[107] W. Yu, G. Ginis, and J. Cioffi, "Distributed Multiuser Power Control for Digital Subscriber Lines," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 1105–1115, June 2002.

[108] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Iterative Water-filling for Gaussian Vector Multiple Access Channels," *IEEE Trans. Inform. Theory*, vol. 50, pp. 145–151, Jan 2004.

[109] W. Yu, "Competition and Cooperation in Multi-user Communication Environment," Ph.D. dissertation, Stanford University, 2002.

[110] W. Yu, "A Dual Decomposition Approach to the Sum Power Gaussian Vector Multiple Access Channel Sum Capacity Problem," in *Proc. 37th Annual Conf. on Information Sciences and Systems (CISS)*, March 2003.

[111] W. Yu, R. Lui, and R. Cendrillon, "Dual Optimization Methods for Multiuser OFDM Systems," accepted for publication in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)* 2004.

[112] C. Zeng, C. Aldana, A. Salvekar, and J. Cioffi, "Crosstalk identification in xDSL systems," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1488–1496, Aug 2001.

[113] C. Zeng and J. Cioffi, "Near-end Crosstalk Mitigation in ADSL Systems," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 949–958, June 2002.

List of Publications

Journal Papers

- R. Cendrillon, G. Ginis, M. Moonen, E. Van den Bogaert, "A Near-capacity Achieving Linear Crosstalk Precoder for DSL," submitted to *IEEE Trans. Signal Processing*, 2004.

- R. Cendrillon, G. Ginis, M. Moonen, E. Van den Bogaert, "A Near-capacity Achieving Linear Crosstalk Canceler for DSL," submitted to *IEEE Trans. on Commun.*, 2004.

- R. Cendrillon, W. Yu, M. Moonen, "Optimal Multi-user Spectrum Management for Digital Subscriber Lines," accepted for publication (subject to minor revision) in *IEEE Trans. on Commun.*, 2004.

- R. Cendrillon, M. Moonen, G. Ginis, K. Van Acker, T. Bostoen, P. Vandaele, "Partial Crosstalk Precompensation for Downstream VDSL," in *Signal Processing*, vol. 84, pp. 2005-2019, Nov. 2004.

- R. Cendrillon, M. Moonen, G. Ginis, K. Van Acker, T. Bostoen, P. Vandaele, "Partial Crosstalk Cancellation for Upstream VDSL," in *EURASIP Journal on Applied Signal Processing*, vol. 10, pp. 1433-1448, Sep. 2004.

Journal Papers in Preparation

- R. Cendrillon, M. Moonen, "Optimal Bandplan Design for Digital Subscriber Lines," to be submitted to *IEEE Trans. Signal Processing*, Dec. 2004.

- J. Homer, R. Cendrillon, M. Moonen, "Adaptive Partial Crosstalk Cancellation for Digital Subscriber Lines," to be submitted to *EURASIP Journal on Applied Signal Processing*, Dec. 2004.

- E. Van den Bogaert, R. Cendrillon, M. Moonen, "Ensuring The Robustness Against Non-Stationary Noise Of ADSL Transceivers When Applying DSM," to be submitted to *EURASIP Journal on Applied Signal Processing*, Dec. 2004.

Conference Papers

- R. Cendrillon, M. Moonen, "Iterative Spectrum Balancing for Digital Subscriber Lines," submitted to *IEEE Int. Conf. on Commun. (ICC)*, 2005.
- E. Van den Bogaert, K. Van Acker, J. Verlinden, R. Cendrillon, M. Moonen, "Waterfilled VDSL Echo Limitation for Rate-Reach Performance Improvement," submitted to *IEEE Int. Conf. on Commun. (ICC)*, 2005.
- R. Cendrillon, M. Moonen, T. Bostoen, G. Ginis, "The Linear Zero-Forcing Crosstalk Canceler is Near-optimal in DSL Channels," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dallas, Dec. 2004.
- W. Yu, R. Lui, R. Cendrillon, "Dual Optimization Methods for Multiuser OFDM Systems," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dallas, Dec. 2004.
- E. Van den Bogaert, T. Bostoen, R. Zeroual, B. Van Wauwe, and F. Van der Schueren, R. Cendrillon, M. Moonen, "Ensuring The Robustness Against Non-Stationary Noise Of ADSL Transceivers When Applying DSM," in *Proc. International Conference on Communication and Computer Networks* , Marbella, Spain, Sep. 2004.
- R. Cendrillon, G. Ginis, M. Moonen, J. Verlinden, T. Bostoen, "Improved Linear Crosstalk Precompensation for DSL," in *Proc. Int. Conf. on Acoust., Speech and Sig. Processing (ICASSP)*, May 2004.
- T. Bostoen, J. Verlinden, R. Cendrillon, M. Moonen, "DSM in Practice: Iterative Waterfilling Implemented on ADSL Modems," in *Proc. Int. Conf. on Acoust., Speech and Sig. Processing (ICASSP)*, May 2004.
- J. Verlinden, E. Van den Bogaert, T. Bostoen, R. Cendrillon, M. Moonen, "Protecting the robustness of ADSL and VDSL DMT modems when applying DSM," in *Proc. International Zurich Seminar on Communications (IZS)*, 2004.
- R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, T. Bostoen, "Optimal Multi-user Spectrum Management for Digital Subscriber Lines," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Paris, June 2004.

- R. Cendrillon, M. Moonen, R. Suciu, G. Ginis, "Simplified Power Allocation and TX/RX Structure for MIMO-DSL," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, San Francisco, Dec. 2003.

- R. Cendrillon, M. Moonen, G. Ginis, K. Van Acker, T. Bostoen, P. Vandaele, "Partial Crosstalk Cancellation Exploiting Line and Tone Selection in Upstream VDSL," in *Proc. of Sixth Baiona Workshop on Signal Processing in Communications*, Spain, Sep. 2003.

- R. Cendrillon, M. Moonen, R. Suciu, "Simplified Power Allocation for the DSL Multi-access Channel through Column-wise Diagonal Dominance," in *Proc. International Symposium on Image and Signal Processing and Analysis (ISPA)*, Rome, Sep. 2003.

- R. Cendrillon, O. Rousseaux, M. Moonen, E. Van den Bogaert, J. Verlinden, "Waterfilling in MIMO Systems with Power Constraints on Each Transmitter," in *Proc. Symposium on Information Theory in the Benelux*, Netherlands, May 2003.

- R. Cendrillon, M. Moonen, D. Gore, A. Paulraj, "Low Complexity Crosstalk Cancellation with Line Selection in Upstream VDSL," in *Proc. Int. Conf. on Acoust., Speech and Sig. Processing (ICASSP)*, Hong Kong, April 2003.

- R. Cendrillon, M. Moonen, "Efficient Equalizers for Single and Multicarrier Environments with Known Symbol Padding," in *Proc. of International Symposium on Signal Processing and its Applications (ISSPA)*, Kuala-Lampur, Malaysia, Aug. 2001.

Guest Editorials

- Special Issue on "Advanced Signal Processing Techniques for Digital Subscriber Lines," *EURASIP Journal on Applied Signal Processing,* to be published Dec. 2005.

Standard Contributions

- R. Cendrillon, M. Moonen, J. Verlinden, T. Bostoen, J. Cioffi "A Linear Crosstalk Precoder that Achieves the Multi-user DSL Channel Capacity," to be presented at ANSI T1E1.4 Working Group (DSL Access) Meeting, San Antonio, Dec. 2004.
- R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, T. Bostoen, "Proof of the Optimality of Optimal Spectrum Management," ANSI T1E1.4 Working Group (DSL Access) Meeting, contrib. 2004-460, North Carolina, May 2004.
- R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, T. Bostoen, "Optimal Spectrum Management," ANSI T1E1.4 Working Group (DSL Access) Meeting, contrib. 2004-459, North Carolina, May 2004; and contrib. 2003-365, Vancouver, Canada, Feb. 2004.[1]
- R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, T. Bostoen, "On the Optimality of Iterative Waterfilling in DSL," ANSI T1E1.4 Working Group (DSL Access) Meeting, contrib. 2003-325, San Diego, Dec. 2003.

Patents

- R. Cendrillon, M. Moonen, K. Van Acker, J. Verlinden, T. Bostoen, E. Van den Bogaert, Applicant: Alcatel Bell, 2004.[2]
- R. Cendrillon, M. Moonen, K. Van Acker, J. Verlinden, T. Bostoen, E. Van den Bogaert, "Power Control Method for Remotely Deployed Communication Service," Applicant: Alcatel Bell, Patent Pending, 2003.
- R. Cendrillon, M. Moonen, K. Van Acker, P. Vandaele, T. Bostoen, "Method for Crosstalk Cancellation," Applicant: Alcatel Bell, United States Patent US5887032, 2003; European Patent EP1109329, 2002.

[1] Now part of the draft ANSI standard on Dynamic Spectrum Management[8].
[2] Intellectual property issues prevent further details from being released at this point.

Curriculum Vitae

Raphael Cendrillon was born in Melbourne, Australia in 1978. He received the Electrical Engineering degree (Honours First Class) from the University of Queensland, Australia in 1999. Since 2000 he has been pursuing a Ph.D. at the Electrical Engineering Department (ESAT) of the Katholieke Universiteit Leuven, Belgium. In Spring 2002 he was a visiting scholar at the Information Systems Laboratory, Stanford University with Prof. John Cioffi. His research focuses on the application of information theory to multi-user communication problems in xDSL.

Mr. Cendrillon is sponsored by several Belgian government grants and industry sponsored projects. His work in xDSL is done in close collaboration with Alcatel Research and Innovation, Belgium, for which he was awarded the Alcatel Bell Scientific Prize, together with Prof. Marc Moonen, in 2004. He was also the recipient of an IEEE Travel Grant in 2003 and the K.U. Leuven Bursary for Advanced Foreign Scholars in 2004.