

KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT INGENIEURSWETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
AFDELING ESAT-SCD: SISTA/COSIC/DOCARCH
Kasteelpark Arenberg 10 – B-3001 Leuven

RESOURCE MANAGEMENT AND OPTIMIZATION IN MULTI-USER DSL SYSTEMS

Promotor:
Prof. dr. ir. M. Moonen

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de ingenieurswetenschappen

door

Paschalis TSIAFLAKIS

December 2009



KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT INGENIEURSWETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
AFDELING ESAT-SCD: SISTA/COSIC/DOCARCH
Kasteelpark Arenberg 10, B-3001 Leuven

RESOURCE MANAGEMENT AND OPTIMIZATION IN MULTI-USER DSL SYSTEMS

Jury:
Prof. em. dr. ir. Y. Willems, voorzitter
Prof. dr. ir. M. Moonen, promotor
Prof. dr. A. Bultheel
Prof. dr. M. Diehl
Prof. dr. ir. L. Vandendorpe (U.C.L.)
Prof. dr. ir. J. Vandewalle

Proefschrift voorgedragen
tot het behalen van de
graad van Doctor in de
Ingenieurswetenschappen

door

Paschalis TSIAFLAKIS

December 2009

© 2009 Katholieke Universiteit Leuven, Groep Wetenschap & Technologie, Arenberg Doctoraatsschool, W. de Croylaan 6, 3001 Heverlee, België

Alle rechten voorbehouden. Niets uit deze uitgave mag vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

ISBN 978-94-6018-153-5
D/2009/7515/135

Voorwoord

Na een opleiding als industrieel en burgerlijk ingenieur, belandde ik in oktober 2004 opeens in een kantoortje van het departement Elektrotechniek, met slechts één grote verantwoordelijkheid voor de volgende vijf jaar, namelijk het afleggen van een doctoraat. Ik had slechts een vaag vermoeden van wat doctoreren juist inhield, maar het stelde me gerust dat ik maar één verantwoordelijkheid had. Bovendien hield dat in dat ik lekker met wiskunde mocht spelen en ik werd er zelfs nog voor betaald ook. Het was te mooi om waar te zijn. Maar al gauw begon ik in te zien, waarom ik ervoor betaald werd. Het was niet zo eenvoudig en de oplossingen van de vraagstukken kon je niet terugvinden achteraan in het handboek. Gelukkig besefte ik snel dat ik er niet alleen voor stond, want naast mijn zijde stonden vele mensen die er waren voor mij, in zware en in goede tijden.

Nu ben ik er. Nu is het tijd om die mensen te bedanken.

Vooreerst zou ik mijn promotor Prof. Marc Moonen willen bedanken. Marcs onderzoeksdiscipline maar vooral ook zijn manier van lesgeven, presenteren en synthetiseren wisten me reeds te boeien tijdens mijn studiejaren, en waren dan ook de hoofdredenen om te kiezen voor een doctoraat bij hem, een kans die hij me bood ondanks mijn Limburgs accent. Onze talloze discussies in zijn bureau tijdens de 'friday meetings', in zijn wagen onderweg naar projectvergaderingen, of in het buitenland tijdens de zovele conferenties, waren voor mij een groot plezier, maar vooral ook een belangrijke leerschool die mij enorm geïnspireerd hebben. Marc, ik heb veel geleerd van je aanpak, je positivisme, en je visie op problemen. Bedankt voor de opleiding en de richting die je me gaf, voor de vrijheid die ik kreeg en je vertrouwen en appreciatie voor het verrichtte onderzoekswerk.

Verder wens ik de leden van mijn begeleidings- en examencommissie, Prof. Adhemar Bultheel, Prof. Wim Dehaene, Prof. Moritz Diehl, Prof. Luc Vandendorpe, Prof. Joos Vandewalle en de voorzitter Prof. Yves Willems, te bedanken voor hun onmiddellijke bereidheid om deel uit te maken van de commissies, maar ook voor het nalezen van het proefschrift en de waardevolle suggesties en

aanvullingen.

I am also very grateful to Prof. Mung Chiang for hosting me at Princeton University. Thank you for giving me the opportunity to collaborate with such an excellent research team, and for all your interaction and interest in my research work. This has been a very rewarding experience for me.

Ik zou ook graag enkele professoren willen bedanken, waarmee ik het genoegen heb gehad te discussiëren over een aantal interessante en zeer uitdagende (NP-hard zoals wij dat noemen) optimalisatie-problemen. Bedankt Prof. Moritz Diehl, Prof. François Glineur, Prof. Ion Necoara, Prof. Johan Suykens en Prof. Lieven Vandenberghe, voor de discussies, voor jullie enthousiasme, en voor de interesse die jullie bij me hebben aangewakkerd voor het vakdomein van optimalisatietheorie.

To my colleagues Prof. Yung Yi and Prof. Chee Wei Tan, your ideas and our many discussions have been enlightening and inspiring to me. Thank you for the nice and productive time we had at Princeton University.

To my colleagues at Alcatel-Lucent: Dr. Mamoun Guenach, Dr. Jochen Maes, Dr. Michael Peeters, Jan Verlinden, and Dr. Geert Ysebaert, working with you has been a great pleasure.

To the (ex-)group members at the Katholieke Universiteit Leuven: Alexander, Amir, Ann, Beier, Bram, Bruno, Deepak, Geert R., Geert V.M., Geert L., Geert Y., Gert, Hilde, Imad, Jan, Kim, Koen, Ludovico, Matteo, Olivier, Pepe, Prabin, Raphael, Rodrigo, Romain, Sam, Simon, Sylwester, T.J., Toon and Vincent. Thank you for providing such an enjoyable workplace over the last five years, for all inspiring discussions, and also for all evenings in the city of Leuven. (I don't seem to remember all of them, but they were fun for sure.) Toon, a special thank to you for all your help during the last (writing) phase. To Toni, Niels and Ernesto, thank you for all the great coffee breaks.

Ook een speciale dank aan de personen die mij geholpen hebben bij alle administratie die er bij een doctoraat komt kijken. Bedankt Eliane, Evelyn, Ida, Ilse, John, Lut, Péla and Veerle.

Een speciale dank ook aan al mijn (niet-ESAT) vrienden, voor de interesse en vooral voor de ontspanning die gezorgd hebben voor de nodige afleiding. Hier gaan we op drinken.

Ik ben het F.W.O.-Vlaanderen erkentelijk voor de financiële ondersteuning gedurende de eerste vier jaar van mijn doctoraat.

Tenslotte kom ik bij het belangrijkste deel, het thuisfront:
Vader, moeder, ik ben jullie ongelooflijk dankbaar voor alles wat jullie voor mij gedaan hebben. Jullie hebben mij alle kansen geboden in het leven, jullie hebben altijd in mij geloofd, en jullie hebben mij altijd gesteund in alles, zelfs in heel zware tijden. Ik besef heel goed dat ik hier niet zou zijn zonder jullie. Ik heb nooit de kans gehad om jullie hiervoor te bedanken.... Bedankt vader, bedankt moeder! Jullie hebben mij opgevoed met goede waarden, waar respect, doorzetting en liefde centraal staan. Jullie hebben mij gemaakt zoals ik nu ben. Bedankt voor de warme plek die jullie mij altijd geboden hebben (in Genk, Hasselt of zelfs in Leuven) en voor de leuke momenten die we samen beleefd hebben. Bedankt dat ik iedere dag op jullie kan rekenen.

Venizelos, ook jou wil ik hierbij speciaal in het licht plaatsen, ook jij hebt bijgedragen tot de persoon die ik ben. Bedankt voor de leuke tijden die we gehad hebben en nog zullen beleven in de toekomst, bedankt dat je zo een toffe broer bent. Kristina, ook jij verdient een speciaal pluimpje, omdat je een toffe schoonzus en een heel goede moeder bent, maar vooral ook voor het temmen van mijn broer. Aristoteles, Omero, en Iason, jullie zijn mijn drie favoriete neefjes. Bedankt voor al die leuke spelletjes die we gespeeld hebben en voor al die leuke weekenduitstapjes die we samen gedaan hebben. Jullie speelvreugde gaven me steeds weer de kracht om door te gaan tot het einde.

Ik zou ook graag mijn tweede familie, de "schoon"familie, enorm willen bedanken. Agnes en Rob, bedankt voor jullie interesse, vriendschap en gezellig samenzijn. Bedankt voor de leuke momenten die variëren van het installeren van de elektriciteitsleiding in de kelder tot het samen supporteren op de top van de Mont Ventoux. Graag zou ik ook moemoe Lieke, en Peter en Mieke, willen bedanken voor de supergezellige familiefeesten en zondagnamiddagen waar ik altijd naar uitkijk, zeker als er een chocoladekoek op mij te wachten staat. Bedankt allemaal.

Tenslotte, mijn grootste dank gaat uit naar Elke. Elke, jouw levensenergie, enthousiasme, intelligentie, geduld en vertrouwen in mij, hebben op mij ingewerkt als een superdynamo die maar door blijft draaien en die mij elke dag weer zoveel levensenergie geeft. Bedankt voor alle steun en alle prachtige momenten die wij samen hebben doorgebracht en nog zullen doorbrengen in de toekomst. Bedankt dat je er altijd bent voor me. Bedankt dat jij jij bent.

Nu ben ik er. Nu is het tijd om die mensen te bedanken.

Paschalis Tsiaflakis
December 2009

Abstract

Digital subscriber line (DSL) technology is currently the most widely deployed broadband access technology and will continue to play an important role during the next decade. However, to cope with the bandwidth-intensive and heterogeneous quality-of-service (QoS) requirements of the many emerging broadband applications and services, it is essential to further improve DSL technology.

One of the major sources of performance degradation of current DSL systems is the electromagnetic coupling among different twisted pairs within the same cable bundle, which is also referred to as *crosstalk*. The presence of crosstalk transforms DSL systems into a very challenging multi-user multi-carrier interference environment, where different users, i.e. modems, can significantly impact each other's data transmission. This is also referred to as the DSL crosstalk problem.

Multi-user coordination techniques such as spectrum and signal coordination are recognized as a key approach for crosstalk mitigation. These techniques consist of coordinating the transmission of multiple users so as to prevent or even remove the impact of crosstalk, resulting in huge performance gains.

More specifically, spectrum coordination consists of coordinating the transmit spectrum of each modem so as to prevent the impact of crosstalk. Signal coordination consists of applying joint signal processing to a set of colocated transmitters and/or receivers so as to remove the crosstalk impact. These two types of coordination can also be combined together to obtain a joint spectrum and signal coordination approach.

The spectrum and/or signal coordination algorithms that were available at the start of this thesis were however quite limited in scope. They mainly focused on crosstalk mitigation so as to maximize physical layer data rates. Furthermore, these algorithms are either optimal but too complex, or low-complex but quite suboptimal.

The main topic of this thesis is the development of resource management and optimization algorithms for such spectrum and signal coordination. The main

focus is on the development of numerically robust and low-complexity algorithms that achieve near-optimal performance, even for large-scale DSL systems. Moreover, it is investigated how to extend the traditional data rate maximization approach towards a wider scope of objectives that are better aligned with the application and network QoS requirements.

The problem of optimal spectrum coordination is first tackled with the proposal of the branch-and-bound optimal spectrum balancing algorithm (BB-OSB), that succeeds in finding the same globally optimal solution as the existing optimal spectrum balancing (OSB) algorithm, but with a significantly reduced computational complexity. Then several distributed spectrum coordination algorithms (ASB2, CA-DSB, DSB, MS-DSB) are proposed with a different level of required centralized control. They can be seen as distributed water-filling type of procedures, that are steered by a centralized controller through message-passing. These algorithms require very-low complexity where simulation results demonstrate that they perform near-optimally, even for large-scale DSL scenarios. Finally, an improved dual decomposition based optimization approach is developed for spectrum coordination. It is shown how this novel approach further improves the convergence speed of existing state-of-the-art spectrum coordination algorithms by one order of magnitude, with the same computational complexity.

A second part of this thesis focuses on the extension of the traditional physical layer data rate maximization approach of spectrum coordination towards a much wider scope. An initial framework is proposed to study upper layer performance, such as throughput and delay, by combining upper layer scheduling and spectrum coordination together, where a number of polynomial-time algorithms are proposed. Then, within the emerging theme of Green Information Technology, a *Green DSL* framework is developed, that involves a range of energy-aware spectrum coordination problem formulations. An efficient solution approach is proposed to tackle the corresponding Green DSL problems. Using this Green DSL framework, it is demonstrated that huge energy savings can be obtained with only a minor degradation in data rate performance. Finally, it is shown that fairness can become an issue in applying Green DSL, where the concept of *fair greening* is introduced. It involves a number of approaches to incorporate different notions of fairness into the design objectives of spectrum coordination.

The last part of this thesis focuses on joint spectrum and signal coordination. A dual decomposition based algorithm is proposed for multi-user partial signal coordination with and without (ON/OFF) spectrum coordination. It is shown that this approach improves on existing algorithms in complexity as well as in performance. Finally, a scheme is proposed for joint spectrum and receiver signal coordination, which is shown to improve on existing schemes, especially for DSL scenarios with significant spatially correlated alien crosstalk.

Korte Inhoud

'Digital subscriber line' (DSL)-technologie is tegenwoordig de meest populaire technologie voor breedband-internettoegang en zal ook in het volgende decennium een belangrijke rol blijven spelen. Om echter tegemoet te komen aan de breedbandintensieve en heterogene kwaliteitsvereisten van toekomstige breedbandapplicaties en -diensten, is het van essentieel belang om de performantie van DSL-technologie verder te verbeteren.

Eén van de grootste belemmeringen voor het verder verbeteren van DSL systemen, is de elektromagnetische koppeling tussen de verschillende getwiste koperdraden binnen éénzelfde kabelbundel, ook wel *overspraak* genoemd. Overspraak transformeert DSL systemen tot een meerdere-gebruiker interferentieomgeving, waarbij verschillende gebruikers, i.e. modems, elkaars datatransmissie significant kunnen beïnvloeden. Dit fenomeen wordt ook wel benoemd als het DSL overspraakprobleem.

Meerdere-gebruiker coördinatietechnieken zoals spectrum- en signaalcoördinatie zijn algemeen gekend als een efficiënte aanpak voor overspraakonderdrukking. Bij deze technieken wordt de transmissie van meerdere gebruikers zodanig gecoördineerd dat de overspraakimpact voorkomen of zelfs verwijderd wordt, resulterend in enorme performantiewinsten.

Meer specifiek wordt bij spectrumcoördinatie de destructieve impact van overspraak voorkomen door het zorgvuldig coördineren van de zendspectra van alle modems. Bij signaalcoördinatie wordt de overspraakimpact daarentegen actief verwijderd door een gemeenschappelijke signaalverwerking van de zend- en/of ontvangstsignalen. Deze twee types van coördinatie kunnen verder ook gecombineerd worden tot een gemeenschappelijke spectrum- en signaalcoördinatieaanpak.

De algoritmen voor spectrum- en/of signaalcoördinatie die beschikbaar waren bij de start van deze thesis waren echter beperkt in functionaliteit en performantie. Ze waren namelijk vooral gericht op overspraakonderdrukking met als doel de datasnelheden te verhogen. Bovendien waren ze ofwel optimaal maar

te complex, ofwel laag-complex maar zeer suboptimaal.

Het hoofdonderwerp van deze thesis is de ontwikkeling van middelenbeheer en -optimalisatie-algoritmen voor zulke spectrum- en/of signaalcoördinatie. Hierbij ligt de hoofdklemtoon op de ontwikkeling van numeriek robuuste en lage-complexiteitsalgoritmen met een quasi-optimale performantie en die toepasbaar zijn voor grootschalige DSL systemen. Bovendien wordt er bestudeerd hoe de traditionele aanpak van datasnelheidsmaximalisatie kan uitgebreid worden naar een breder gamma van ontwerpdoelen die beter afgestemd zijn op de applicatie- en netwerkkwaliteitsvereisten.

Het probleem van optimale spectrumcoördinatie wordt eerst aangepakt met het voorstel van het 'branch-and-bound optimal spectrum balancing' (BB-OSB) algoritme, dat erin slaagt om dezelfde globaal optimale oplossing te bekomen als het bestaande 'optimal spectrum balancing' (OSB) algoritme, maar dan met een significant gereduceerde computationele complexiteit. Daarnaast worden verschillende gedistribueerde spectrumcoördinatie algoritmen (ASB2, CA-DSB, DSB, MS-DSB) voorgesteld die gekenmerkt worden door een verschillende graad van vereiste centrale controle. Deze algoritmen beschikken over een gedistribueerde 'water-filling'-structuur die bovendien sturing toelaat vanuit een centrale controller door middel van boodschappen. Ze vereisen een zeer lage computationele complexiteit en simulatieresultaten tonen aan dat ze leiden tot quasi-optimale spectrumcoördinatie, zelfs voor grootschalige DSL systemen. Tenslotte wordt er een verbeterde 'duale decompositie'-gebaseerde optimalisatie-aanpak ontwikkeld voor spectrumcoördinatie. Er wordt aangetoond dat deze nieuwe aanpak toelaat de convergentiesnelheid van bestaande spectrumcoördinatie-algoritmen met één ordegrootte te verbeteren, zonder de complexiteit te verhogen.

Een tweede deel van deze thesis richt zich op de uitbreiding van de traditionele 'fysische laag'-datasnelheidsmaximalisatie-aanpak van spectrumcoördinatie naar een breder gamma van ontwerpdoelen. Eerst wordt een initieel kader voorgesteld dat toelaat om 'hogere laag'-performantiematen te bestuderen, zoals verwerkingscapaciteit ('throughput') en vertraging ('delay'). Dit wordt verkregen door het combineren van 'hogere laag'-scheduling en spectrumcoördinatie, waarvoor een aantal algoritmen voorgesteld worden die in polynomiale tijd kunnen worden uitgevoerd. Hierna wordt binnen de nieuwe problematiek van *groene informatietechnologie* een kader ontwikkeld voor *Groene DSL*, dat bestaat uit een reeks van energie-bewuste spectrumcoördinatie-probleemformuleringen. Hiervoor wordt verder ook een efficiënte oplossingsaanpak voorgesteld. Gebruikmakend van het Groene DSL kader, wordt er aangetoond dat enorme energiebesparingen mogelijk zijn met slechts een minieme opoffering van datasnelheidsperformantie. Tenslotte wordt er aangetoond dat *eerlijkheid (fairness)* een belangrijke kwestie wordt bij het toepassen van Groene DSL. Daartoe wordt het concept van '*fair greening*' geïntroduceerd, dat een aantal aanpakken omvat om eerlijkheid te integreren in de ontwerpdoelen van spectrumcoördinatie.

Het laatste deel van deze thesis richt zich op gemeenschappelijke spectrum- en signaalcoördinatie. Eerst wordt een 'duale decompositie'-gebaseerde aanpak voorgesteld voor partiële signaalcoördinatie met of zonder (AAN/UIT) spectrumcoördinatie. Er wordt aangetoond dat deze aanpak bestaande algoritmen zowel in complexiteit als in performantie verbetert. Tenslotte wordt een transmissieschema voorgesteld voor gemeenschappelijke spectrum- en ontvangstsignaalcoördinatie, dat bestaande algoritmen significant verbetert, vooral voor DSL scenario's met spatiaal gecorreleerde (externe) overspraakinvloeden.

Glossary

Mathematical Notation

$(\cdot)^T$	transpose operation
$(\cdot)^H$	matrix Hermitian transpose
$\|\cdot\|$	Euclidean norm
$\|x\|$	absolute value of (real or complex) scalar x
$[x]^+$	$\max(0, x)$
$[x]_0$	$\max(0, x)$
$[x]_b^a$	$\max(\min(x, a), b)$
x^c	complex conjugate of complex scalar x
$\mathbf{x} \geq \mathbf{y}$	componentwise inequality between vectors $\mathbf{x}$ and $\mathbf{y}$
$(\mathbf{x})_l$	lth element of vector $\mathbf{x}$
$[\mathbf{X}]_{n,m}$	element from nth row and mth column of matrix X
$\mathbf{X} \geq \mathbf{Y}$	componentwise inequality between matrices $\mathbf{X}$ and $\mathbf{Y}$
$\det(\mathbf{X})$	matrix determinant
$\frac{\partial f}{\partial x}$	partial derivative of f with respect to x
$\mathrm{diag}(\mathbf{x})$	diagonal matrix with vector x as diagonal
∇f	gradient of function f
$E(\cdot)$	statistical average operation
$\max(x, y)$	maximum of scalars x and y
$\min(x, y)$	minimum of scalars x and y
$\mathcal{O}(\cdot)$	order
$\mathbb{P}(\cdot)$	probability operator
$\rho(\mathbf{X})$	spectral radius of matrix $\mathbf{X}$
$\mathrm{rem}(x, a)$	remainder after dividing x by a

Fixed Symbols

b_k^n	bit loading of user n on tone k
$\mathbf{b}_k$	vector containing the bit loadings of all users on tone k
$\mathbf{b}^n$	vector containing the bit loadings of user n on all tones

$c_k^{n,m}$	partial crosstalk cancellation configuration for line n for the interference from line m on tone k
c_k	channel capacity on tone k
$d_k(\mathbf{s}_k)$	prox-function on tone k
$\mathcal{F}$	greening fairness index
f_s	DMT symbol rate
$g(\boldsymbol{\lambda})$	Lagrange dual function
$h_k^{n,n}$	direct channel frequency response of line n on tone k
$h_k^{n,m}$	FEXT crosstalk channel frequency response from line m into line n on tone k
$\mathbf{H}_k$	$N \times N$ crosstalk channel matrix on tone k
int_k^n	sum interference from users $m \neq n$ into the line of user n on tone k
I	identity matrix
k	tone index
K	number of tones available for transmission
$\mathcal{K}$	set of tones available for transmission $= \{1, \ldots, K\}$
L	Lagrangian
L_c	Lipschitz constant
L_k	decomposed Lagrangian for tone k
$L(t)$	Lyapunov function on time slot t
n	user index
N	number of interfering users
$\mathcal{N}$	set of interfering users $= \{1, \ldots, N\}$
P^n	sum total transmit power allocated to user n
$P^{n,\text{tot}}$	total transmit power budget available to user n
$\mathbf{P}^{\text{tot}}$	vector containing the total available transmit power budgets for all users
$Q^n(t)$	queue buffer length of user n at time slot t
R^n	achieved data rate of user n
$R^{n,\text{target}}$	minimum data rate constraint for user n
$\mathcal{R}$	achievable rate region
s_k	transmit power on tone k of a particular transmitter (in a single-user setting)
s_k^n	transmit power of transmitter n on tone k
$\mathbf{s}_k$	vector containing transmit powers of all transmitters on tone k
$\mathbf{s}^n$	transmit spectrum of transmitter n, i.e. vector containing transmit powers of transmitter n on all tones
$s_k^{n,*}$	optimal transmit power allocation of transmitter n on tone k
s_k^{mask}	spectral mask of transmit power on tone k, i.e. upper bound on s_k, for a particular transmitter in a single-user setting
$s_k^{n,\text{mask}}$	spectral mask of transmitter n on tone k, i.e. upper

	bound on s_k^n
$\mathcal{S}$	set of transmit power allocations over all transmitters and tones that satisfy spectral mask and per-user total power constraints as defined by DSL standards
w^n	relative weight for user n in weighted data rate sum maximization
$\mathbf{w}$	vector containing the relative weights w_n for all users n
x_k^n	symbol transmitted by transmitter n on tone k
$\mathbf{x}_k$	vector containing symbols of all transmitters transmitted on tone k
y_k^n	symbol received by receiver n on tone k
$\mathbf{y}_k$	vector containing symbols of all receivers received on tone k
z_k	noise signal on tone k received by a particular receiver (in a single-user setting)
z_k^n	noise signal received by receiver n on tone k
$\mathbf{z}_k$	vector containing the received noise signals of all receivers on tone k
Γ	SNR-gap
Δ_s	discrete accuracy on transmit power
λ_n	Lagrange multiplier corresponding to the constraint $P^n \leq P^{n,\text{tot}}$
$\boldsymbol{\lambda}$	vector containing the Lagrange multipliers λ_n of all users n
$\boldsymbol{\Lambda}$	throughput region
σ_k	noise power on tone k received by a particular user (in a single-user setting)
σ_k^n	noise power on tone k received by receiver n
$\boldsymbol{\sigma}_k$	vector containing the received noise powers of all receivers on tone k
ω_n	relative weight for user n in weighted data rate sum maximization
$\boldsymbol{\omega}$	vector containing the relative weights ω_n for all users

Acronyms and Abbreviations

ADSL	asymmetric digital subscriber line
AFE	analog front-end
ANSI	American National Standards Institute
ASB	autonomous spectrum balancing
AWG	American Wire Gauge
AWGN	additive white Gaussian noise
BB-OSB	branch-and-bound optimal spectrum balancing

BER	bit error ratio
BPL	Broadband over Powerline
CA-DSB	convex approximation distributed spectrum balancing
CO	central office
CP	customer premises
cWRS	constrained weighted rate sum maximization
DAC	digital-to-analog converter
DBS	Direct Broadcast Satellite
DFE	decision feedback equalizer
DMT	discrete multitone
DS	downstream
DSB	distributed spectrum balancing
DSL	digital subscriber line
DSLAM	digital subscriber line access multiplexer
DSM	dynamic spectrum management
e.g.	*exempli gratia*: for example
FEQ	frequency-domain equalizer
FEXT	far-end crosstalk
FFT	fast Fourier transform
FTTB	Fiber to the Building
FTTC	Fiber to the Cabinet
FTTH	Fiber to the Home
FTTN	Fiber to the Node
GP	geometric programming
HAN	home area networks
HDTV	high definition television
ICI	inter-carrier interference
ICT	information and communication technology
i.e.	*id est*: that is
IFFT	inverse fast Fourier transform
ISB	iterative spectrum balancing
ISDN	Integrated Services Digital Network
ISI	inter-symbol interference
ISP	internet service provider
IW	iterative water-filling
IWF	iterative water-filling
kbps	kilobits per second
kHz	kilohertz
KKT	Karush-Kuhn-Tucker
LAN	local area network
MAC	multiple access channel
MAC-ISB	multiple access channel iterative spectrum balancing
MAC-OSB	multiple access channel optimal spectrum balancing
Mbps	megabits per second
MHz	megahertz

MIMO	multiple input multiple output
MMSE-GDFE	minimum mean square error generalized decision feedback equalization
MS-DSB	multiple starting point distributed spectrum balancing
NEXT	near-end crosstalk
OLT	optical line terminal
OSB	optimal spectrum balancing
PBnB	prismatic branch-and-bound
PCC	partial crosstalk cancellation
PLC	power line communications
PSD	power spectral density
QAM	quadrature amplitude modulation
QoS	quality-of-service
RFI	radio frequency interference
RT	remote terminal
SC	spectrum coordination
SCALE	successive convex approximation for low-complexity
SMC	spectrum management center
SNR	signal-to-noise ratio
SINR	signal-to-interference-and-noise ratio
s.t.	subject to
UMTS	Universal Mobile Telecommunications Service
US	upstream
VDSL	very high speed digital subscriber line
VoD	video on demand
VoIP	voice over IP
WiFi	Wiress Fidelity
WiMAX	Worldwide Interoperability for Microwave Access
ZF	zero-forcing

Contents

Voorwoord **i**

Abstract **v**

Korte Inhoud **vii**

Glossary **xi**

Contents **xvii**

1 Introduction and Overview **1**

1.1 DSL Broadband Access 1

1.2 DSL Basics 6

1.2.1 DSL Environment 6

1.2.2 Discrete Multitone Modulation 7

1.2.3 Bit Loading 11

1.2.4 Power Loading 12

1.3 The Crosstalk Problem 13

1.3.1 Crosstalk Types 15

1.3.2 Crosstalk Channel Modeling 16

1.3.3 Crosstalk Scenarios 18

1.4 Thesis Problem Statement and Contributions 20
1.4.1 Problem Statement 20
1.4.2 Contributions 23
1.5 Dynamic Spectrum Management 26
1.6 State-of-the-Art Spectrum Coordination 29
1.6.1 System Model 32
1.6.2 Spectrum Coordination Problem Modeling 33
1.6.3 Spectrum Coordination Algorithms 37
1.7 State-of-the-Art Signal Coordination 48
1.8 Chapters and Publications Overview 53
Bibliography 54

2 Centralized Spectrum Coordination 63

2.1 Introduction 65
2.2 System Model 67
2.3 Optimal Spectrum Balancing 69
2.3.1 The Spectrum Management Problem 69
2.3.2 Dual Decomposition 70
2.3.3 Solving the Dual Problem 72
2.4 An Efficient Lagrange Multiplier Search Algorithm 74
2.4.1 Single-user Spectrum Management 74
2.4.2 Multi-user Spectrum Management 78
2.5 A Low Complexity Branch and Bound Approach to the Per-Tone Optimization Problem 83
2.6 Simulation Results 88
2.6.1 Symmetric Scenario 89
2.6.2 Asymmetric Scenario 90

2.7 Conclusion 95

Bibliography 97

3 Distributed Spectrum Coordination 99

3.1 Introduction 101

3.2 System Model 103

3.3 Spectrum Management Problem 104

3.4 A KKT Perspective on State-of-the-Art Distributed Spectrum Management Algorithms 105

3.4.1 KKT perspective on Iterative Water-filling 105

3.4.2 KKT Perspective on SCALE 108

3.4.3 KKT Perspective on ASB 111

3.5 Novel Distributed Spectrum Management Algorithms 113

3.5.1 Convex Approximation Distributed Spectrum Balancing (CA-DSB) 114

3.5.2 Distributed Spectrum Balancing (DSB) 117

3.5.3 Multiple Starting Point Distributed Spectrum Balancing (MS-DSB) 118

3.5.4 Per-Tone Penalty Water-filling Algorithms 122

3.5.5 Complexity Analysis 124

3.5.6 Convergence Analysis 125

3.6 Simulation Results 128

3.6.1 ADSL performances 129

3.6.2 VDSL performances 136

3.7 Conclusion 138

Bibliography 140

4 Geometric Programming Based Spectrum Coordination 143

4.1 Introduction 145

4.2 System Model 147
4.3 Spectrum Management Problems 148
4.4 Optimality Certificate for Polynomial Time Spectrum Management 149
4.4.1 Polynomial Time Spectrum Management 149
4.4.2 Optimality Analysis 152
4.4.3 Algorithm DSM-GP 154
4.5 Simulation Results 155
4.6 Conclusion 156
Bibliography 157

5 Cross-Layer Spectrum Coordination 159

5.1 Introduction 161
5.2 System Model and Performance Metrics 162
5.2.1 System Model 162
5.2.2 Performance Metrics 163
5.3 Throughput-Optimal Scheduling and DSM Algorithms 164
5.4 Polynomial Complexity Scheduling Achieving Throughput Optimality 166
5.4.1 δ-Randomized DSM Algorithms 166
5.4.2 Delay Performance of δ-Randomized DSM 168
5.4.3 Significant Delay Decrease with Small Extra Complexity 169
5.5 Simulation Results 171
5.6 Conclusion 172
Bibliography 173

6 Green DSL: Energy-Aware Spectrum Coordination 175

6.1 Introduction 177

6.2 System Model 178
6.3 Green DSL 179
6.3.1 Unifying Green DSL Formulation 179
6.3.2 General Systematic Solution Procedure 181
6.4 Performance Impact of Green DSL 183
6.5 Conclusion 188
Bibliography 189

7 General Green DSL: Fair and Energy-Aware Spectrum Coordination 191

7.1 Introduction 193
7.2 DSL System Model 195
7.3 Greening DSL 196
7.4 Fairness in Greening DSL 196
7.4.1 Fair Greening Formulations 196
7.4.2 Unified Formulation and Solution Algorithm 199
7.5 Simulation Results 201
7.6 Conclusion 204
Bibliography 204

8 Improved Dual Decomposition Based Spectrum Coordination 207

8.1 Introduction 209
8.2 System Model 211
8.3 Dynamic Spectrum Management 212
8.3.1 Dynamic Spectrum Management Problem 212
8.3.2 Dynamic Spectrum Management Algorithms 213
8.4 Improved Dual Decomposition 215

8.4.1 An Improved Dual decomposition Approach for Iterative Convex Approximation Based DSM Algorithms 218

8.4.2 An Improved Dual Decomposition Approach for Direct DSM Algorithms 226

8.5 An Interleaving Procedure for Recovering the Primal Solution from the Dual Solution 228

8.6 Simulation Results 230

8.6.1 Convergence Speed Up for Iterative Convex Approximation based DSM 231

8.6.2 Convergence Speed Up for Direct DSM 232

8.7 Conclusion 237

Bibliography 241

9 Cancellation Resource Optimization for Linear Partial Crosstalk Cancellation 245

9.1 Introduction 247

9.2 Dual Decomposition 248

9.2.1 System Model 248

9.2.2 Partial Crosstalk Cancellation Problem 249

9.2.3 Per-tone Search Complexity Reduction 252

9.3 Single-User Algorithms and Complexity Comparison 253

9.3.1 Single-User Resource Allocation Algorithm 254

9.3.2 Single-User Dual Decomposition Algorithm 256

9.4 Multi-User Algorithms and Complexity Comparison 260

9.4.1 Multi-user Resource Allocation Algorithm 260

9.4.2 Multi-user Dual Decomposition Algorithm 261

9.5 Simulation Results 265

9.6 Conclusion 266

Bibliography 269

10 Joint Spectrum and Cancellation Resource Optimization for Linear Partial Crosstalk Cancellation 271

10.1 Introduction 273

10.2 Joint Spectrum Management and Constrained Partial Crosstalk Cancellation 275

10.2.1 System Model 275

10.2.2 Problem Statement 277

10.2.3 Dual Decomposition 278

10.2.4 Complexity 281

10.3 Complexity Reduction 281

10.3.1 Partial Crosstalk Cancellation 282

10.3.2 Spectrum Management: ON/OFF Power Loading . . . 284

10.4 Simulation Results 286

10.5 Conclusion 298

Bibliography 300

11 Joint Spectrum and Receiver Signal Coordination 303

11.1 Introduction 305

11.2 System Model 306

11.3 Multiple Access Channel Optimal Spectrum Balancing 307

11.4 Multiple Access Channel Iterative Spectrum Balancing (MAC-ISB) 309

11.5 Simulation Results 310

11.6 Conclusion 312

Bibliography 312

12 Conclusions and Further Research 315

12.1 Summary and Conclusions 315

12.2 Further Research 320

Bibliography 322

Publication List **325**

Curriculum Vitae **329**

Chapter 1

Introduction and Overview

1.1 DSL Broadband Access

Over the last decade, the internet has evolved enormously and one has witnessed a true revolution of broadband internet usage. This is mainly driven by the many available and emerging broadband internet applications and services. For instance, applications like Facebook, YouTube, online gaming, peer-to-peer file sharing, voice over IP (VoIP), video-conferencing, video on demand (VoD) and high definition TV (HDTV) have become very popular, but also social services such as e-health, e-government, tele-working and tele-education are nowadays starting to gain a lot of importance.

The emergence of the many bandwidth intensive applications and services has resulted in relentlessly increasing demands for higher broadband bandwidth. This evolution has obviously also lead to increasing performance requirements for the underlying broadband technologies. For instance, even with the latest data compression techniques, HDTV requires in the order of 15 to 20 Mbps of downstream bandwidth [1]. As a result, the telecom industry is facing a major challenge: the provision of high quality-of-service (QoS) broadband connections to millions of customers worldwide to support existing and future broadband applications and services.

One crucial part of the broadband connection is the *broadband access network*. This part connects subscribers to their immediate service provider, which is then connected to the internet core network. The broadband access network has consistently been regarded as the main bottleneck in the broadband connection [2]. This is because the offered bandwidth of broadband access technologies significantly lags behind the bandwidth provided at the end-user side (with

high speed local area networks (LAN) or even home area networks (HAN)) and at the internet core network side. Therefore it is essential to improve broadband access technologies so as to close this bandwidth gap.

Broadband access technologies can generally be classified into two groups [1]: *fixed line technologies* and *wireless technologies*. Fixed line broadband technologies consist of a direct physical connection to the end-user, where the most popular examples are digital subscriber lines, cable modem technology, power line communications (PLC) and optical fiber connections (FTTx). Wireless broadband technology refers to technologies that use microwaves to provide a connection to the end-user. An overview of a number of popular broadband access technologies will now be given.

Digital subscriber line (DSL or xDSL) technology refers to a family of technologies for high speed digital communication over the wires, i.e. twisted pairs, of the local telephone network infrastructure. This infrastructure basically consists of multiple twisted pairs that are bundled into large cable bundles, and where many cable bundles are deployed from the telephone central office (CO) or a remote terminal (RT) towards the end-users. Each end-user has one (or two) dedicated twisted pairs to connect with the CO (or RT), where the latter has a high bandwidth connection to the internet core network.

The history of digital communication over the telephone twisted pairs goes back to the eighties. Telecom equipment manufacturers were then working hard to develop Integrated Services Digital Network (ISDN) technology, which was able to provide a digital connection of 128 kbit/s in both *upstream* (US) (towards the CO) and *downstream* (DS) (towards the end-user) directions. However, the main interest of telephone network operators (telcos) was to move to a much faster digital broadband connection over their local telephone infrastructure so as to deploy video services, such as VoD, to residential customers, in response to the cable TV companies that were starting to deploy telephone services in their coaxial cable infrastructure. In the early nineties, the American National Standards Institute (ANSI) put up a competition to establish a new standard for digital transmission over the local telephone twisted pairs. It was a technology spin-out of Professor J.M. Cioffi's research group at Stanford University that suggested the approach of "Discrete Multitone (DMT) modulation" that finally won the competition. DMT has become the basis today for most modern DSL standards in use. This first resulted in the proposal of asymmetric digital subscriber line (ADSL), with a transmission speed of approximately 6 Mbps downstream and 640 kbps upstream. The asymmetry of the service was justified by the asymmetric profile of video services, i.e. high bandwidth for video downstream and low bandwidth for upstream signaling. However as these data rates were still not sufficient for video services, the driving application shifted from VoD to internet data access. Soon improved follow-up DSL technologies were proposed and standardized: ADSL2 in 2002, ADSL2plus (ADSL2+) in 2003, very high speed DSL (VDSL) in 2004 and VDSL2 in 2006. With the

introduction of VDSL and VDSL2, the original goal of video services was revisited, as these technologies are able to deliver very high bit rates over large distances, e.g. 25 Mbps over 1 km with VDSL and even 100-200 Mbps over a distance of 300 m with VDSL2 [3]. This has recently enabled new services such as broadband triple play, which is a service consisting of voice, video and data traffic, all over the same twisted pair.

Cable modem technology is one of the biggest competitors of DSL. It uses the coaxial cable TV infrastructure to offer a broadband connection to the end-user. The coaxial cable is a medium that has a much larger useful bandwidth compared to the telephone twisted pair medium. However cable modem broadband access relies on a shared network architecture, i.e. the same cable line connects to many end-users. As a consequence cable modem service can slow down significantly when many people in the same neighborhood access their cable modem internet connection simultaneously.

Broadband over Powerline (BPL) is another broadband access technology for high speed data transmission over the existing power lines. Here, the end-users can simply plug in their modem into an ordinary wall outlet to receive broadband internet access. Although BPL already exists for several years, it has not been implemented in large-scale. This has several reasons, one of them being the fact that power lines are inherently a very noisy environment. BPL can be seen as an interesting option for rural areas where neither cable nor telephone infrastructure is available.

Optical fiber, also called fiber, is a fourth transmission medium that can be used for broadband access. This medium is characterized by a much lower attenuation and interference, and so also a much larger transmission bandwidth that outperforms all other broadband access media. The ultimate goal is to move towards a Fiber to the Home (FTTH) broadband access network, where all end-users have an individual high speed physical fiber connection. However, installing such a fiber access network is difficult and time-consuming, and most importantly very expensive. The current approach is to gradually extend the existing fiber network towards the end-users and then cover the last mile using other broadband access technologies such as DSL or BPL. The term FTTx refers to a set of technologies that bring fiber a step closer to the subscriber. For instance, Fiber to the Node (FTTN), also called Fiber to the Cabinet (FTTCab), brings the fiber to a cabinet serving a neighborhood. Customers are then connected to this cabinet using for instance DSL transmission over existing twisted pair infrastructure. Another example is Fiber to the Building (FTTB), which brings the fiber in close proximity of a building, from where individual physical connections (twisted pairs, power lines, etc) start towards the end-users in the building.

Wireless broadband access is a wireless alternative for providing broadband access to the end-user. Over the last decade, several wireless solutions have

been proposed. WiFi (Wireless Fidelity) is a popular wireless technology that is used for in-home or in-office wireless local area networks (WLAN) with data rates of up to 54 Mbps, shared over all users. WiFi *hotspots* are access areas (e.g. in restaurants, airports and universities) where WiFi internet connectivity is provided, and are becoming quite popular nowadays. WiMAX (Worldwide Interoperability for Microwave Access) is another wireless technology that is designed to deliver broadband connectivity of up to 11 Mbps over a larger distance, i.e. multiple kilometers depending on line-of-sight (LOS) connectivity. Satellite broadband technologies also exist. For instance, Direct Broadcast Satellite (DBS) provides two-way high-speed data transmission over a geostationary satellite. However, these systems generally suffer from large delays, which is unacceptable for many broadband services. Another wireless option is high speed cellular access. For instance, UMTS (Universal Mobile Telecommunications Service), also known as 3G, provides a broadband connection of up to 2 Mbps. Although some of the wireless technologies look very promising, it should be mentioned that the wireless channel is a much more challenging medium compared to a fixed physical channel (twisted pair, cable, fiber). More specifically, wireless channels have a larger attenuation per unit distance, and they are generally more time varying, interference-limited and shared. Furthermore, one important issue is the scarcity of the radio spectrum which limits the maximum bandwidth assigned to a wireless technology. For these reasons, it is unlikely that a wireless technology will ever act as the sole technology covering all broadband access, but it should rather be seen as a complementary technology that will eventually coexist with fixed line broadband solutions.

Nowadays, DSL technology remains the most popular broadband access technology. The main reason of its popularity is the fact that it is much more cost effective than many other broadband access technologies, such as cable modem, fiber, or satellite. This is because it reuses the millions of existing telephone twisted pairs and transforms them into high quality broadband connections, where only the end-user's modem, also referred to as customer premises equipment (CPE), and the CO telecommunication equipment needs to be upgraded.

In a recent survey [4] by Point Topic Ltd, a marketing group that provides broadband usage statistics every quarter, it was reported that in the first quarter of 2009, the distribution of technologies for all worldwide broadband access connections consisted of 64.6% (or 277.35 million) DSL subscribers, 20.6% (or 88.73 million) cable modem subscribers, 12.4% (or 53.4 million) fiber subscribers, and 2.3% (or 9.73 million) subscribers using other technologies.

In Figure 1.1, the evolution over time is shown for the three most popular broadband access technologies. It shows that over the last three years the number of DSL subscribers has been consistently increasing with a larger slope than the number of cable modem and fiber subscribers. Furthermore in [5], it is predicted that the number of DSL subscribers will even raise up to 331 million in 2012.

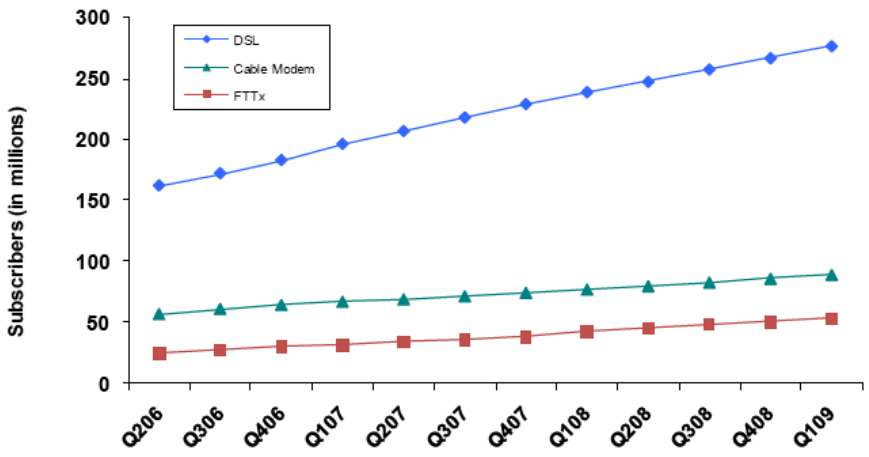


Figure 1.1: Trends of the three most popular broadband access technologies over the different quarters of the last three years. This figure was published in [4] by Point Topic Ltd and is reprinted by permission.

It is expected that the number of fiber subscribers will significantly increase over time. However this will require extensive and expensive infrastructure updates and it will take many years to move towards fully fiber-fed access networks. In the meantime access infrastructure will evolve to hybrid fiber solutions (FTTN, FTTB) and also here DSL will play an important role in covering the last-mile connectivity to the end-user.

Considering these numbers and trends, one can expect that DSL broadband access will continue to play an important role during the next decade. However, it is imperative that DSL technology is improved so as to stay competitive with other broadband access technologies and so as to cope with the many stringent, increasing, and heterogeneous QoS requirements, in particular:

- application specific QoS: large data rates, small delay performance, link robustness, symmetry, etc.
- network specific QoS: large reach, high stability, low power consumption, low signalling overhead, fairness, etc.

This thesis focuses on *resource management and optimization* in multi-user DSL systems so as to improve DSL performance and to ensure the necessary QoS requirements for future DSL broadband access.

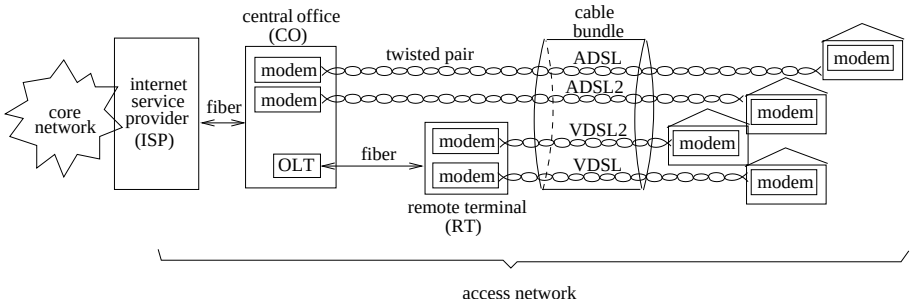


Figure 1.2: DSL broadband access scenario

1.2 DSL Basics

1.2.1 DSL Environment

The telephone network infrastructure consists of multiple twisted pairs that are bundled into cable bundles, and where multiple cable bundles start from telephone central offices or fiber-fed remote terminals towards more than one billion households worldwide. In Figure 1.2, a small DSL broadband access scenario is depicted. It consists of a central office, that is fiber-connected to the internet service provider (ISP), which is then connected to the core network. Two twisted pairs start from the CO towards two far users, at a maximum distance of 5.5 km, offering ADSL and ADSL2 service. There is a remote terminal that is fiber-connected to the central office via the optical line terminal (OLT). As the RT can be brought closer to the customers, it is possible to offer DSL services, such as VDSL and VDSL2, that can deliver much higher data rates over small distances (300m-900m). The cable bundle consists in total of four twisted pairs. For each twisted pair physical connection, a pair of DSL modems, one at each side, is required to set up the high-speed data transmission.

The telephone access network was originally designed to offer telephone services, with a twisted pair bandwidth usage between 0 and 4 kHz. To deliver broadband data rates, current DSL technologies transmit over the same twisted pairs up to much larger frequencies, i.e. even up to 30 MHz for VDSL2 technology. This results however in a DSL environment with many channel impairments that have to be dealt with, e.g.:

- The channel attenuation of a twisted pair strongly depends on the length of the line and also on the frequency, i.e. at higher frequencies much larger signal attenuations occur. The DSL transmission channel can thus be viewed as a very *frequency selective* channel. This results in *inter-symbol interference* (ISI), which basically means that consecutively transmitted

symbols interfere which each other at the receiver.
- The large bandwidth usage results in spectral overlap with amateur radio bands (ham radio, above 1.8 MHz), that can cause significant narrowband *radio frequency interference* (RFI) to the DSL transmission.
- Devices, such as in-house dimmers, refrigerators and switching of AC motors can cause significant *impulse noise*, which consists of energy spikes with random amplitude and spectral content.
- Electromagnetic coupling or *crosstalk* between multiple twisted pairs within the same cable bundle results in huge interference degradation, especially at the higher frequency range used by DSL systems.

In order to enable high quality data transmission, DSL transceivers have to be designed carefully so as to be robust against these channel impairments. One advantage is that the DSL channel is typically slowly time-varying, unlike wireless channels, as channel variations are mainly caused by changing temperature and weather conditions [6]. This allows for offline (and online) optimization (and adaptation) of resources and parameter settings so as to configure the DSL transceiver towards a better and more robust performance under the specific line conditions.

1.2.2 Discrete Multitone Modulation

DSL transceivers can be divided into two main groups: (i) those that use single-carrier transmission, and (ii) those that use multi-carrier transmission. Single-carrier DSL transceivers use single-carrier modulation to transmit a data stream over the frequency selective DSL channel. Most modern DSL modems in use are however based on multi-carrier transmission. The concept of multi-carrier transmission is to transmit multiple independent data streams, in parallel, over multiple independent carriers, also called subchannels or tones. More specifically, a set of narrow-band *quadrature amplitude modulation* (QAM) signals are transmitted over consecutive independent tones. Each of the individual per-tone QAM signals corresponds to a low bit rate transmission over a small bandwidth. The sum of all these per-tone bit rates adds up to a large total bit rate transmission over the corresponding twisted pair.

In the multi-carrier setting, the frequency selective DSL channel can be viewed as being approximated by consecutive independent tones with a flat frequency response, as shown in Figure 1.3. This decomposition over tones has a number of advantages. For instance, as the channel frequency response is almost flat within a tone, each of the per-tone QAM transmissions is free of ISI. In contrast, single-carrier transmission sees a non-flat channel frequency response and hence it significantly suffers from ISI. Such single-carrier systems require the usage of complex equalization in order to tackle the ISI problem. The concept of equalization is to add processing to the receiver or to the transmitter so as to flatten the frequency selectivity of the transmission.

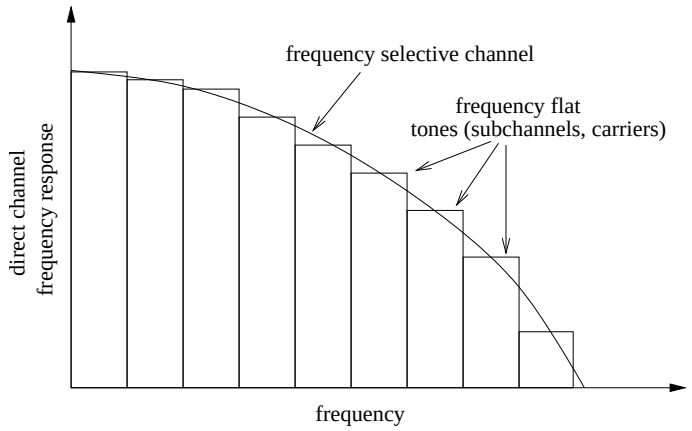


Figure 1.3: Approximation of the frequency selective channel with multi-carrier tones with flat channel frequency responses

Discrete multitone (DMT) modulation is the all-digital multi-carrier implementation that is used by modern DSL systems. It is based on cheap Fast Fourier Transform (FFT) operations at the transmitter and at the receiver, which results in a very low-complexity implementation.

In Figure 1.4, the *DMT transmitter* is depicted. At the input, there is an incoming bit stream, which can be encoded (e.g., using a Reed-Solomon (RS) channel encoder) and interleaved. From this bit stream, a block of B bits is serial-to-parallel converted where each of the parallel branches are allocated a number of bits, which is also referred to as *bit loading*. The exact number of bits in each branch depends on the quality of the channel and the level of noise, as will be explained in more detail in Section 1.2.3 on bit loading. The sets of bits are then converted into their corresponding QAM constellation with a certain scaling, which corresponds to power loading, as will be explained in more detail in Section 1.2.4. Furthermore, there is a mirroring operation, which corresponds to using a flipped copy of the upper half QAM symbols as the lower half symbols, combined with a complex conjugate operation, i.e. $x_n^c(b)$ is the complex conjugate of $x_n(b)$. The set of obtained symbols is fed into an Inverse Fast Fourier Transform (IFFT) operation, resulting in a set of real valued symbols $\{x_l^t(b), l = 1, \ldots, L\}$, because of the imposed complex conjugate symmetry at the input of the IFFT. Finally, a cyclic prefix of length C is added, which consists of a copy of the C last samples $\{x_l^t(b), l = L - C + 1, \ldots, L\}$ to the front. The result is serialized and transformed into an analog signal that is suitable for transmission over the twisted pair, via a number of digital filters, a digital-to-analog converter (DAC), and analog front-end (AFE) processing.

The *DMT receiver* is depicted in Figure 1.5 and basically performs the in-

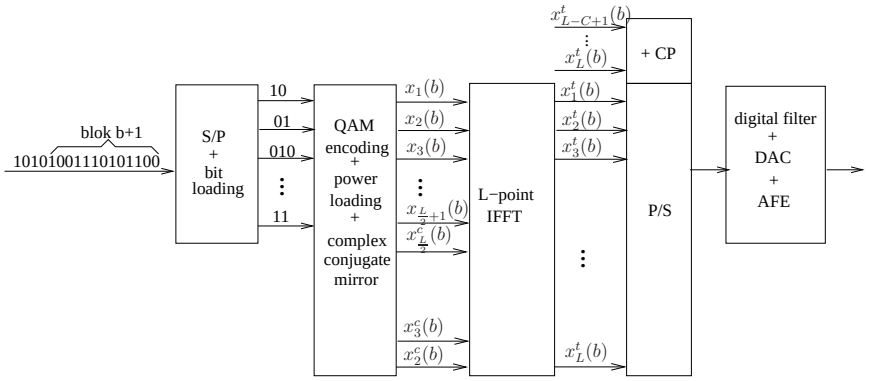


Figure 1.4: Block scheme of DMT transmitter

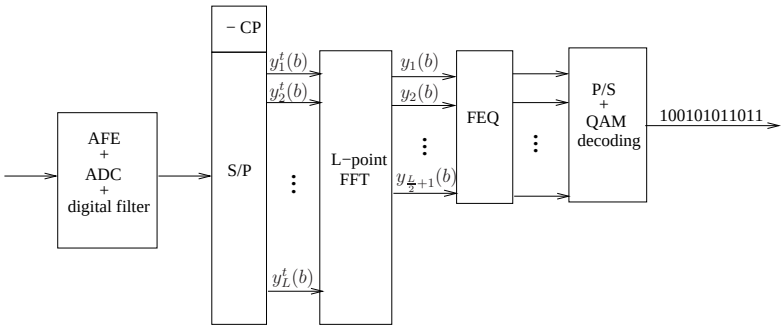


Figure 1.5: Block scheme of DMT receiver

verse operations of the DMT transmitter. The received analog signal is first transformed into a digital signal, via the analog front-end, the analog-to-digital (ADC) converter, and digital filtering. The cyclic prefix is removed and after a serial-to-parallel conversion, this is fed into an FFT operation. The first half of the obtained samples are equalized by a frequency domain equalizer (FEQ), which corresponds to a per-sample scaling operation. Finally a QAM decoding is performed so as to obtain the corresponding bits and this is serialized to the output, which can be followed by a channel decoder and de-interleaver.

It can be shown that when the cyclic prefix length C is longer than the channel impulse response, the DMT transmission results in a frequency domain block-by-block transmission that is fully decoupled over the tones. Formally this is denoted as follows:

$$y_l(b) = h_l(b)x_l(b) + z_l(b), \qquad l = 1, \ldots, \frac{L}{2}, \tag{1.1}$$

where b denotes the specific time block, $x_l(b)$ and $y_l(b)$ denote the frequency domain symbols transmitted and received over tone l during block b, respectively, $z_l(b)$ denotes the received noise on tone l during block b, and $h_l(b)$ denotes the direct channel frequency response for tone l during block b. Note that for each time block b, B bits are transmitted over the channel, where B depends on the sum of all bit allocations (see also Section 1.2.3 on bit loading). For convenience, we will disregard the block index b from (1.1), and also replace index l by index $k \in \mathcal{K}$, with $\mathcal{K}$ indicating the set of available tones for transmission. This notation is useful for the remaining of the text as $\mathcal{K}$ can include any subset of all available tones, e.g. for upstream or downstream transmission only. This results in the following expression for DMT transmission:

$$y_k = h_k x_k + z_k, \qquad k \in \mathcal{K} = \{1, \ldots, K\}. \tag{1.2}$$

Expression (1.2) basically states that the QAM symbol x_k at the input of the DMT transmitter's IFFT, is transmitted over tone h_k, without any interference from the other symbols $x_m, m \neq k$, and with a noise component corresponding to the received noise on tone k. So DMT transmission achieves a full decoupling over tones. This is a very important property of DMT transmission. It simplifies the design and optimization of the DSL transceiver so as to tackle the many channel impairments, as it decouples the design problem into multiple small per-tone problems that require much less complexity. For instance, DMT transceivers remove the need for complex equalization by decoupling the problem over tones. Comparing a DMT-based VDSL transceiver with $K = 4096$ and $C = 320$ with a single-carrier DSL transceiver employing decision feedback equalization (DFE) to combat ISI, it can be shown that the DMT transceiver requires 12 times less complexity [7, 8].

1.2.3 Bit Loading

One important component of the DMT transmission scheme is bit loading, which basically determines the number of bits that is allocated to each of the tones. A well-considered bit loading allows to improve the spectral efficiency of DMT transceivers. Intuitively, a good tone, i.e. a tone with a high signal-to-noise ratio (SNR), can carry more bits than a low SNR tone for the same probability of error.

The maximum number of bits that can be transmitted over a channel with probability of error that tends to zero, is called the channel capacity. Assuming additive white Gaussian noise (AWGN), the channel capacity for a tone k can be expressed as follows

$$c_k = \log_2(1 + \mathrm{SNR}_k),$$

where SNR_k is the received signal-to-noise ratio for tone k and is defined as

$$\mathrm{SNR}_k \triangleq \frac{|h_k|^2 s_k}{\sigma_k}, \tag{1.3}$$

and where the transmit power, also called the power spectral density (PSD), on tone k is denoted as $s_k \triangleq E(|x_k|^2)$, and the received noise power, also called the noise power spectral density, on tone k is denoted as $\sigma_k \triangleq E(|z_k|^2)$, with $E(.)$ denoting the statistical average operation.

In practice, the probability of error cannot be zero, and one can not achieve the channel capacity. For practical systems, an expression for the achievable bit rate under a given target probability of error can be obtained using the SNR gap concept [9]. More specifically, for DMT transmission, the achievable bit rate for tone k is given by

$$b_k = \log_2\left(1 + \frac{\mathrm{SNR}_k}{\Gamma}\right), \tag{1.4}$$

where Γ denotes the SNR-gap, which is a constant that is defined by the target probability of error, the desired noise margin, the modulation type and the coding gain of the used coding scheme (e.g. Reed-Solomon coding).

In a DMT system, the receiver can measure the received SNR on all tones, and feedback this information to the transmitter. The transmitter can then perform bit loading by determining the number of bits that is allocated to each of the tones using (1.4), which also fixes the QAM constellation sizes.

The total achievable data rate R can then be obtained by summing the allocated bits over all tones and can be expressed as follows

$$R = f_s \sum_{k \in \mathcal{K}} b_k,$$

where f_s denotes the DMT symbol rate, i.e. the number of QAM symbols that is transmitted over one tone during one second.

1.2.4 Power Loading

Another important component of the DMT transmission scheme is *power loading*. Power loading consists of allocating transmit powers to the independent per-tone QAM transmissions, which basically comes down to a scaling of the corresponding QAM constellations. Allocating more transmit power s_k to a particular tone k, increases its SNR and hence the maximum number of bits b_k that can be transmitted over that tone, as given by (1.4). So power and bit loading strongly depend on each other.

Power loading is subject to a number of constraints. Transmit powers typically have to satisfy a set of predefined spectral masks, which are per-tone upper bounds on the allocated transmit powers, so as to ensure spectral compatibility with other DSL services. The spectral mask constraints are given as follows

$$s_k \leq s_k^{\text{mask}}, \qquad k \in \mathcal{K}. \tag{1.5}$$

Furthermore, the total transmit power, i.e. the sum of all allocated per-tone transmit powers, is subject to a modem total power constraint P^{tot}, that arises from limitations on the AFE, as follows

$$\sum_{k \in \mathcal{K}} s_k \leq P^{\text{tot}}. \tag{1.6}$$

Optimizing the power loading subject to constraints (1.5) and (1.6) can significantly improve the spectral efficiency of the DMT transmission. Here, one can aim at two main objectives [9]: (i) maximize the total data rate, also referred to as rate-adaptive power loading , or (ii) minimize the total allocated transmit power subject to a minimum required data rate, also referred to as power or margin-adaptive power loading. The rate-adaptive power loading optimization problem is given as follows:

$$\begin{aligned} \{s_k^*, k \in \mathcal{K}\}: \quad & \underset{\{s_k, k \in \mathcal{K}\}}{\text{argmax}} \quad R \\ & \text{s.t.} \qquad \sum_{k \in \mathcal{K}} s_k \leq P^{\text{tot}} \\ & \qquad\quad 0 \leq s_k \leq s_k^{\text{mask}}, \qquad k \in \mathcal{K}. \end{aligned} \tag{1.7}$$

The objective function is concave and differentiable, and the constraints are affine. Thus, problem (1.7) is a convex problem, and its Karush-Kuhn-Tucker (KKT) conditions are sufficient for optimality. The optimal power loading

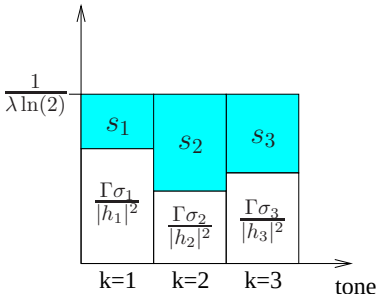


Figure 1.6: Water-filling power loading

$\{s_k^*, k \in \mathcal{K}\}$ can be found by solving the system obtained by the Karush-Kuhn-Tucker (KKT) conditions, which can be summarized as follows

$$s_k = \Big[\frac{1}{\lambda \ln(2)} - \frac{\Gamma\sigma_k}{|h_k|^2}\Big]_0^{s_k^{\text{mask}}}, \qquad k \in \mathcal{K}, \tag{1.8}$$

$$\lambda\big(\sum_{k\in\mathcal{K}} s_k - P^{\text{tot}}\big) = 0, \tag{1.9}$$

where $[.]_a^b$ denotes the operation $\max(\min(.,a),b)$, and where λ is an extra variable which is called the Lagrange multiplier. The optimal value for the transmit power s_k for given λ is given in closed form by (1.8). The optimal value for λ, i.e. that satisfies (1.9), can be easily found by using a bisection search since $\sum_{k\in\mathcal{K}} s_k$ is monotonically decreasing with increasing λ. Once the optimal transmit powers $s_k^*, k \in \mathcal{K}$, are found, one can obtain the optimal bit loading using (1.4) and (1.3) with $s_k = s_k^*$. Alternative efficient power and bit loading implementations for solving (1.7) are given in [9].

The optimization procedure for (1.7) is sometimes also referred to as water-filling [9]. Disregarding the bounds, (1.8) can be reformulated as follows

$$\frac{1}{\lambda \ln(2)} = \frac{\Gamma\sigma_k}{|h_k|^2} + s_k, \qquad k \in \mathcal{K}. \tag{1.10}$$

This is visualized in Figure 1.6, where the left-side term of (1.10) can be regarded as the overall water-level, which depends on λ, the first right-side term is the floor level in each tone k, and the transmit powers s_k correspond to the depth in each tone k.

1.3 The Crosstalk Problem

A twisted pair consists of two copper wires (the forward and return conductor) that are twisted together. The twist reduces the electromagnetic radiation

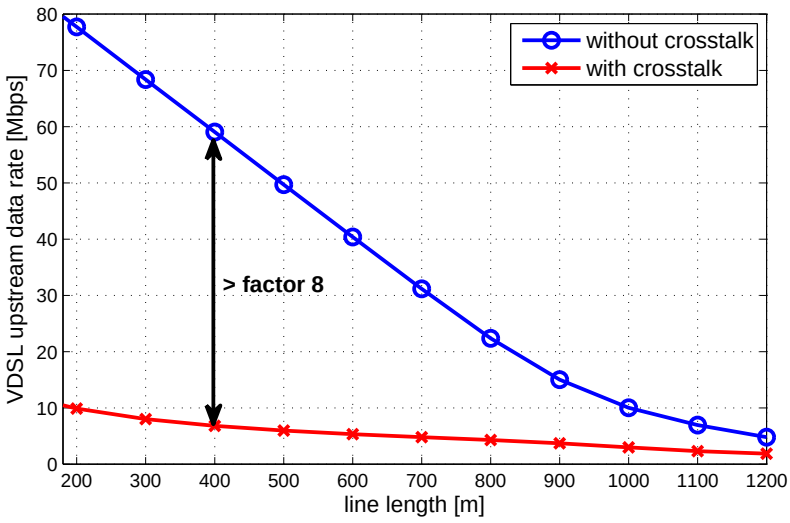


Figure 1.7: Impact of crosstalk on the data rate of an upstream VDSL service for different line lengths. Cable bundle consists of 40 equal-length twisted pairs.

towards the environment, and also the electromagnetic interference from the environment [10]. This works well in a frequency range between 0 and 4 kHz, which corresponds to the bandwidth for voice traffic.

To enable broadband data rates, new DSL technologies have systematically extended the frequency bandwidth usage, even up to 30 MHz for VDSL2. At these high frequencies the twisted pairs start to act as antennas, which results in electromagnetic coupling among the different twisted pairs within the same cable bundle. This type of interference is also referred to as *crosstalk*.

Crosstalk transforms the cable bundle transmission into a multi-user interference environment where transmission over one twisted pair can have a huge impact on the neighboring twisted pairs. The level of crosstalk depends on many factors. It varies both with line length and frequency. It is worst among adjacent lines inside the cable bundle. It does not change much over time, and it is symmetrical in that the same crosstalk is observed between two ends when measured in either direction [11].

One of the impacts of crosstalk is a degradation in the maximum data rates that can be achieved over the twisted pairs. In fact, in modern DSL systems crosstalk is the dominant source of performance degradation. For illustration, we have simulated the achievable data rate of a VDSL upstream service for different line lengths, inside a cable bundle of 40 equal-length VDSL lines. In Figure 1.7, the upper curve and the lower curve indicate the case without and

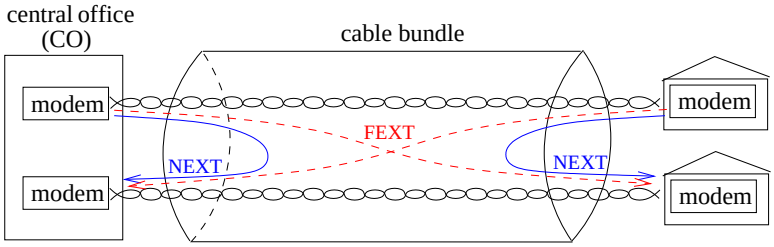


Figure 1.8: Near-end crosstalk (NEXT) and far-end crosstalk (FEXT) in a DSL cable bundle

with interference, respectively. The presence of crosstalk significantly impacts the DSL system in that it causes a huge decrease in data rate performance. For instance, for a loop length of 400 m there is a reduction in data rate performance from 60 Mbps to only 7 Mbps, i.e. an 8-fold reduction.

A thorough understanding of crosstalk is necessary to evaluate the impact on the DSL transmission, but also to develop and design solutions so as to tackle the crosstalk problem. In the next sections the types of crosstalk, crosstalk channel models and challenging crosstalk scenarios will be discussed in more detail.

1.3.1 Crosstalk Types

A receiver's crosstalk interference can have multiple origins, i.e. transmitters whose signal couples into the receiver's line. Generally, one can distinguish two types of crosstalk, which are also depicted in Figure 1.8:

- Near-end crosstalk (NEXT) refers to the crosstalk between transmitting modems and receiving modems at the same end of the cable bundle: more specifically, from the downstream transmission into the upstream reception of the modems at the CO, or from the upstream transmission into the downstream reception of the modems at the end-user side.
- Far-end crosstalk (FEXT) refers to the crosstalk between transmitters and receivers at the opposite ends of the cable bundle: more specifically, from the upstream (downstream) transmission of a modem at one side of the cable bundle into the upstream (downstream) reception of a modem at the other side of the cable bundle.

The level of NEXT depends on the frequency ($\sim f^{3/2}$) and is substantially independent of cable length and attenuation. FEXT depends on the frequency ($\sim f^2$) but is also proportional to the cable length [11]. NEXT is typically much more severe than FEXT. This is because NEXT couples directly from a transmitter to a receiver at the same side of the cable bundle, whereas FEXT

undergoes a large attenuation by travelling from one side of the cable bundle to the other.

NEXT coupling is typically avoided by the use of frequency-division duplex (FDD) or the so-called Zipper technique [12], which basically means that a different set of independent tones are allocated for upstream and downstream transmission. Most practical DSL systems use NEXT-avoiding mechanisms, which means that they only suffer from FEXT. In this thesis we will therefore only consider NEXT-free DMT systems, and focus on solutions to combat the FEXT crosstalk problem.

1.3.2 Crosstalk Channel Modeling

The crosstalk from one twisted pair (disturber pair) into another twisted pair (victim pair) is typically modeled by a *crosstalk channel frequency response*. It describes how the *transmit spectrum*, i.e. the transmit power over all tones, of the disturber pair signal is received (as NEXT or FEXT) by the victim pair. Assuming that the tone width is small, which is a valid assumption for DMT-based DSL systems, the crosstalk impact between two lines and for a particular tone k, can be modeled as follows

$$s_k^v = |h_k^{v,d}|^2 s_k^d, \tag{1.11}$$

where s_k^d (s_k^v) is the transmit (receive) power on tone k of the signal transmitted (received) at the disturber (victim) pair, and $|h_k^{v,d}|^2$ is the squared magnitude of the crosstalk channel frequency response from the disturber pair d into the victim pair v. Note that d and v refer to two interfering twisted pairs of a cable bundle. For a cable bundle of N twisted pairs, all crosstalk channel frequency responses on a particular tone k can be modeled by an $N \times N$ matrix as follows

$$\mathbf{H}_k \in \mathbb{C}^{N \times N} \text{ , with } h_k^{n,m} \triangleq [\mathbf{H}_k]_{n,m} \text{ , and } n, m \in \mathcal{N} = \{1, \ldots, N\}. \tag{1.12}$$

The diagonal elements of $\mathbf{H}_k$, i.e. $h_k^{n,n}, n \in \mathcal{N}$, represent the *direct channel frequency response* of line n, whereas the off-diagonal elements, i.e. $h_k^{n,m}, m \neq n$, represent the crosstalk channel frequency responses from line m into line n.

The direct and crosstalk channel frequency responses depend upon the characteristics, the structure, and the environment of the individual twisted pairs, and can vary a lot from pair to pair and from cable bundle to cable bundle. It is not feasible to perform laboratory or field tests to represent all likely circumstances that will be encountered in practice. Computer simulation is a much easier and faster alternative, but resides on accurate models of the direct and crosstalk channel frequency responses. In DSL standards [13][14], practical models have been proposed for different types of twisted pair (26-AWG, 24-AWG, 22-AWG, 19-AWG), and will be highlighted in the next two sections.

Table 1.1: Parameters of primary constants (1.13) for 26-AWG twisted pairs

r_{0c} [Ω/km]	286.176
a_c	0.14769620
l_0 [μH/km]	675.369
l_∞ [μH/km]	488.952
b	0.929
f_m [kHz]	806.339
$c_\infty[nF/km]$	49
g_0 [nMho/km]	43
g_e	0.70

Direct Channel Frequency Response

To model the direct channel frequency response of a particular type of twisted pair, one typically uses a two-port transmission line model. This model depends upon a number of primary constants, i.e., the resistance R [Ω/km], the inductance L [H/km], the capacitance C [F/km], and the conductance G [Mho/km]. These primary constants are defined as follows

$$\begin{aligned} R_k &= \sqrt[4]{r_{0c}^4 + a_c f_k^2}, \\ L_k &= \Big(l_0 + l_\infty (f_k/f_m)^b\Big)\Big(1 + (f_k/f_m)^b\Big)^{-1}, \\ C_k &= c_\infty, \\ G_k &= g_0 (f_k)^{g_e}, \end{aligned} \tag{1.13}$$

where $f_k = k\Delta_f$ refers to the frequency (in Hz) corresponding to tone k, Δ_f refers to the width of each tone, and where the values of the parameters are provided in DSL standards [13] for different types of twisted pairs. For instance, the parameters (with units) for a 26-AWG twisted pair are summarized in Table 1.1.

The source impedance is denoted as Z_S, and the load impedance as Z_L. An accurate estimate of the direct channel frequency response on a particular tone k and for a particular line (line n of a cable bundle containing N lines) with loop length d (in km), can then be given as follows

$$h_k^{n,n} = \frac{Z_L + Z_S}{Z_L \cosh(\gamma_k d) + Z_{0,k} \sinh(\gamma_k d) + \frac{Z_S Z_L}{Z_{0,k}} \sinh(\gamma_k d) + Z_S \cosh(\gamma_k d)}, \tag{1.14}$$

where the characteristic impedance of the transmission line Z_0 at tone k is

$$Z_{0,k} = \sqrt{\frac{R_k + j2\pi f_k L_k}{G + j2\pi f_k C_k}},$$

and the propagation constant γ_k per unit length at tone k is

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

FEXT Crosstalk Channel Frequency Response

Practical crosstalk channel frequency responses vary from pair to pair and from cable bundle to cable bundle, which leads to a certain degree of unpredictability. Extensive measurement campaigns have resulted in large databases with real practical measurements on the level of crosstalk for different parameters and settings. Based on these databases, a "1% worst-case model" is constructed, which corresponds to a crosstalk level that is larger than the crosstalk level of 99% of all real practical crosstalk measurements. The "1% worst-case FEXT channel crosstalk frequency response" from the transmitter of line m into the opposite receiver of line n for each tone k is given as follows

$$h_k^{n,m} = K_{xf}(f_k/f_0)\sqrt{l^{n,m}}|h_k^{n,n}(d^{n,m})|, \qquad m \neq n, \tag{1.15}$$

where $K_{xf} = 10^{-45/20}$ [14], $f_0 = 1$ MHz, and $l^{n,m}$ is the coupling path, which is the length (in km) of the cable over which the victim receiver of line n and the far-end disturbing transmitter of line m have a common cable path. The last factor in (1.15) is the direct channel gain for a channel of length $d^{n,m}$, where $d^{n,m}$ is the full length (in km) that is affecting the disturbing signal, i.e from the transmitter of line m to the receiver of line n.

The 1% worst-case model can be used to perform simulations to identify the 1% worst-case degradation for a particular DSL scenario, and also to design solutions for these worst-case scenarios. However this model cannot be used to assess the possible gains for a particular crosstalk mitigation technology as the gains would only be applicable to 1% of the lines, and therefore constitute a vast overestimation. Therefore full statistical models have been proposed in [15][16] that describe the offset from the 1% worst-case model. The usage of these models could provide a more realistic idea of the impact of crosstalk mitigation technologies for practical DSL scenarios. For further details on these models, we refer to [15] and [16].

1.3.3 Crosstalk Scenarios

The impact of crosstalk interference depends on the considered DSL scenario, i.e. the particular configuration of interfering twisted pairs that offer a DSL service (e.g., ADSL, VDSL) to a group of end-users. For some DSL scenarios, the presence of crosstalk only leads to minor or moderate performance degradation. However, there also exist DSL scenarios, namely those that suffer from the *near-far problem*, for which crosstalk can even lead to a complete loss of DSL service. The near-far problem occurs when a weak signal of interest (from

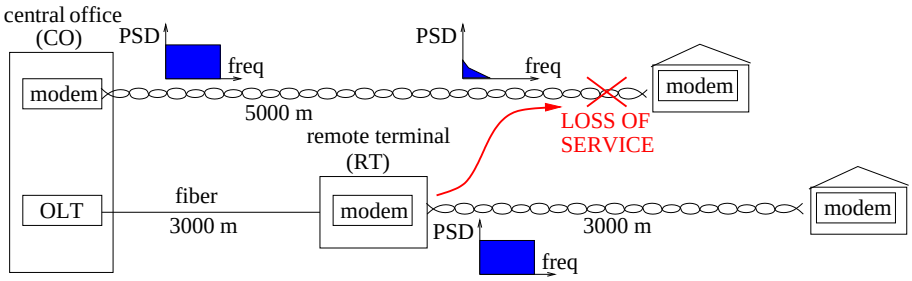


Figure 1.9: Near-far downstream ADSL scenario

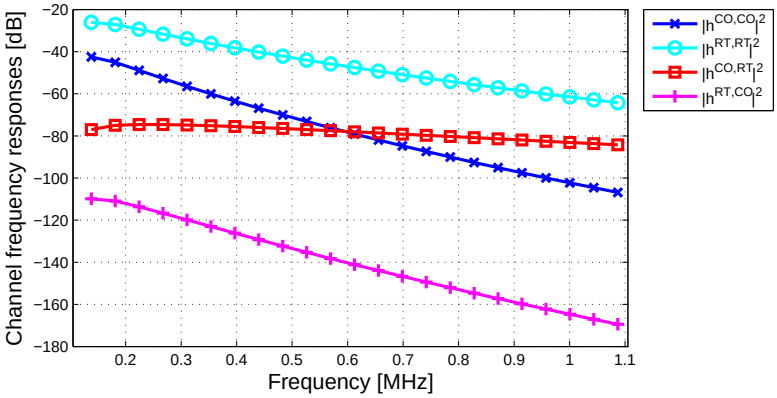


Figure 1.10: Squared magnitude of direct and crosstalk channel frequency responses for near-far downstream ADSL scenario of Figure 1.9

a far-end user) has to be detected under the strong interference of a near-end disturber.

Near-far scenarios are quite common for DSL systems. For instance, in Figure 1.9 a mixed CO-RT ADSL downstream scenario is shown, which consists of a 5 km twisted pair connection between the CO and a first end-user (the CO end-user), and a 3 km twisted pair connection between the RT and a second end-user (the RT end-user). The signal transmitted by the CO is attenuated by the twisted pair. As a result the CO-end-user receives a weak signal from the CO, that is distorted with a strong crosstalk signal from the near-end RT transmitter.

For illustration, the squared magnitudes of the channel frequency responses are shown in Figure 1.10 for the near-far ADSL downstream scenario of Figure 1.9. Here, $h^{\mathrm{CO,CO}}$ and $h^{\mathrm{RT,RT}}$ are the direct channel frequency responses of the CO and RT twisted pairs, respectively. The RT twisted pair has a smaller

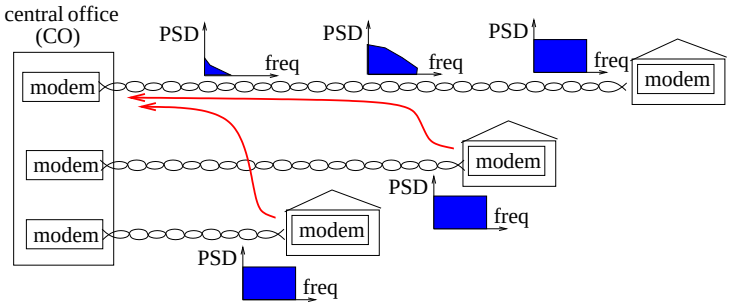


Figure 1.11: Near-far upstream VDSL scenario

attenuation as it corresponds to a shorter twisted pair length. Furthermore, $h^{\mathrm{CO,RT}}$ ($h^{\mathrm{RT,CO}}$) corresponds to the crosstalk channel frequency response of the downstream RT (CO) transmitter into the downstream CO (RT) receiver. The crosstalk interference from the RT into the CO is very large and even exceeds the direct channel gain of the CO twisted pair at frequencies above 600 kHz.

Another important type of near-far scenarios is the upstream VDSL scenario, where a particular example is shown in Figure 1.11. It consists of three CO-connected VDSL services with different loop lengths. The upstream transmission of the shorter lines acts as strong crosstalk interference to the VDSL upstream service over the longest line, and this can result in a significant performance degradation.

By mitigating the crosstalk problem, the performance of current interference-limited DSL systems can be significantly improved, especially for DSL scenarios that suffer from the near-far problem. The focus in this thesis is on the development of multi-user resource management and optimization procedures for crosstalk mitigation so as to improve the performance of DSL systems.

1.4 Thesis Problem Statement and Contributions

1.4.1 Problem Statement

Multi-User Resource Management and Optimization

The ubiquitous presence of twisted pair cabling offers the potential for telcos to deliver broadband services to a large number of end-users worldwide, with only a minimum of infrastructure investment. Over a single twisted pair, data rates of up to 500 Mbps can be realized [17]. However, the data rates of practical

DSL systems in use are much smaller, which indicates that the twisted pair infrastructure is being utilized much below its potential capacity.

One of the dominant reasons for this degraded performance is the presence of crosstalk interference amongst the twisted pairs within the same cable bundle. Crosstalk increases the noise floor for a particular DSL transmission and so decreases the achievable data rate for that particular line. It also decreases the reach of the DSL service, and increases the probability of instability of the system [18]. Crosstalk management, i.e. crosstalk mitigation and cancellation, is important so as to improve these issues, and so as to squeeze out the full potential of the twisted pair infrastructure.

Multi-user coordination techniques are recognized as a key approach for crosstalk management. Here, the basic concept is to coordinate the DSL transmission across multiple twisted pairs so as to mitigate the destructive impact of crosstalk amongst them.

Multi-user coordination techniques are based on an extended transmission model that explicitly incorporates the impact of crosstalk. For DMT-based DSL systems and under the standard assumptions (sufficiently long cyclic prefix, perfect synchronization, FDD), the full cable bundle transmission with N twisted pairs can be modeled independently for each tone k as follows

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k, \qquad k \in \mathcal{K}, \tag{1.16}$$

where $\mathbf{H}_k$ is as defined in (1.12). $\mathbf{y}_k = [y_k^1, \ldots, y_k^N]^T$ contains the received symbols on tone k, where y_k^n denotes the signal received on line n at tone k. $\mathbf{x}_k$ and $\mathbf{z}_k$ have similar structures. $\mathbf{x}_k$ denotes the vector of transmitted symbols on tone k. $\mathbf{z}_k$ denotes the vector of additive noise received on tone k and contains thermal noise, impulse noise, RFI, etc. Note that the set of tones $\mathcal{K}$ can refer to the set of upstream or downstream tones for upstream and downstream transmission, respectively.

Formula (1.16) can be seen as a multi-user extension of the single-user transmission (1.2), where the individual crosstalk channels on tone k are modeled with the off-diagonal components of $\mathbf{H}_k$. The signal received on line n and on tone k is then given as follows

$$\begin{aligned} y_k^n &= h_k^{n,n} x_k^n + \mathrm{int}_k^n + z_k^n, \\ \text{with } \mathrm{int}_k^n &= \sum_{m \neq n} h_k^{n,m} x_k^m, \end{aligned} \tag{1.17}$$

and consists of three terms: the attenuated signal of interest, the crosstalk interference int_k^n, and the received noise z_k^n. By jointly coordinating the transmission over all users, one can balance the impact of the crosstalk interference.

Multi-user coordination techniques can generally be divided into two main

types: *spectrum coordination*, and *signal coordination*. In this thesis the focus is on multi-user resource management and optimization for such spectrum and signal coordination, which will be explained in the next two sections.

Spectrum Coordination

Spectrum coordination consists of allocating the available transmit power resources across the users and tones so as to mitigate the impact of crosstalk interference. This basically comes down to solving a non-convex optimization problem where the objective and the constraints are functions of the transmit powers, and model the QoS requirements of the cable bundle DSL transmission.

The optimization algorithms for spectrum coordination that were available at the start of this thesis were quite limited in scope. They mainly focused on crosstalk mitigation so as to maximize physical layer data rates. Furthermore, these algorithms are too complex for practical use or too badly performing, i.e. the allocated transmit spectra correspond to a suboptimal data rate performance.

More specifically, the *optimal spectrum balancing* (OSB) algorithm was proposed in [19, 20], and succeeds in finding the optimal transmit power allocation, but requires a huge computational complexity. Another available algorithm was the *iterative spectrum balancing* (ISB) algorithm [21, 22] that significantly reduces the computational complexity compared to OSB, but performs suboptimally, especially for large-scale near-far scenarios. The *iterative water-filling* (IWF) algorithm was proposed in [23] and requires a very low computational complexity, but at a cost of a quite suboptimal performance.

To give a more concrete idea of the computational complexity, the simulation times required to execute OSB, ISB and IWF in MATLAB for a 4-user ADSL downstream scenario are 5 days, 1-2 minutes and less than 1 second, respectively. Practical cable bundles typically consist of 20-100 lines for which OSB is too complex to be applied.

A first goal of this thesis is the design of low-complexity and numerically robust spectrum coordination algorithms that significantly decrease the computational complexity of OSB while performing near-optimally, even for large-scale scenarios.

OSB and ISB are centralized spectrum coordination algorithms, which means that they are executed in a centralized location, with full channel knowledge. IWF is a distributed algorithm where each modem autonomously determines its transmit spectrum based on local measures. Here, each modem *selfishly* optimizes its own performance, which leads to a bad global network performance. A second goal in this thesis is the design of *distributed spectrum coordination* algorithms where each of the modems autonomously chooses its transmit spec-

trum based on local measures and in addition a centralized controller (e.g. the SMC) steers the modems towards a better social network performance. This type of approach is interesting as it combines local procedures, that can react fast on a changing channel environment, with a centralized control that steers the individual procedures through message-passing so as to obtain a better global network behaviour.

A third goal is to extend the limited physical layer data rate maximization approach to a much wider scope that includes objectives that are more relevant and are better aligned with the application and network QoS requirements.

Signal Coordination

Signal coordination consists of applying joint signal processing to a group of colocated transmitters and/or receivers so as to remove the crosstalk impact amongst the corresponding twisted pairs.

Many signal coordination techniques have been proposed in literature, consisting of feedback-based, linear and partial linear schemes. A more detailed explanation of these techniques is deferred to Section 1.7 on state-of-the-art signal coordination.

Joint signal and spectrum coordination combines the allocation of transmit spectra with the application of signal coordination, and further improves on the performance of applying signal coordination alone. However the corresponding procedures for resource management and optimization are very complex.

One important goal in this thesis is the design of low-complexity resource management and optimization algorithms for joint signal and spectrum coordination.

1.4.2 Contributions

An overview of this thesis is now given. The main contributions are summarized and the references to the corresponding publications are given.

In the remainder of Chapter 1, an overview is given of current state-of-the-art multi-user coordination techniques. First in Section 1.5, the area of dynamic spectrum management (DSM) [24, 25] is explained, which refers to a set of practical multi-user coordination approaches so as to improve DSL transmission, and that includes spectrum and signal coordination techniques. Then in Section 1.6 and 1.7, an overview is given of state-of-the-art spectrum and signal coordination techniques.

In Chapter 2 a centralized spectrum coordination algorithm is presented that has the same globally optimal performance as OSB, but significantly reduces the computational complexity of that required by OSB. The algorithm will

be referred to as *branch-and-bound optimal spectrum balancing* (BB-OSB). It consists of a dual decomposition approach that decomposes the problem into a convex master problem and many decoupled non-convex slave problems. For the master problem, a subgradient update approach with improved stepsize tuning is proposed that is based on the relation between the dual and primal variables. For each of the non-convex decomposed slave problems, a low-complexity branch-and-bound procedure is proposed. Whereas the OSB algorithm, as proposed in [19], requires a week of MATLAB simulation time to find the optimal solution for a 4-user ADSL downstream scenario, the BB-OSB algorithm succeeds in finding the same solution in less than 10 minutes. The contents of Chapter 2 have been published as [26, 27, 28].

In Chapter 3 a number of distributed spectrum coordination algorithms are presented with a different level of required message-passing. A first algorithm *autonomous spectrum balancing 2* (ASB2) is a fully autonomous algorithm. It is based on a similar *virtual reference line protection* concept as ASB [29], but consists of a low-complexity fixed point update approach that reduces the computational complexity compared to ASB when multiple virtual reference lines are used. A second and third algorithm are *convex approximation distributed spectrum balancing* (CA-DSB) and *distributed spectrum balancing* (DSB) that are distributed algorithms which can be implemented in a distributed fashion with the usage of message-passing. CA-DSB and DSB are based on an iterative convex approximation approach. They are very low-complexity algorithms that succeed in finding a locally optimal solution to the nonconvex spectrum coordination problem. A fourth algorithm is *multiple starting point distributed spectrum balancing* (MS-DSB), which is a semi-centralized spectrum coordination algorithm. This algorithm consists of a per-tone multiple starting point extension of the DSB algorithm. MS-DSB can be seen as a combination of multiple parallel fast local searches that start from carefully chosen starting points. Extensive simulation results show that the proposed algorithms succeed in efficiently tackling the spectrum coordination problems, e.g. a 4-user ADSL downstream scenario is tackled in less than 1 second. The MS-DSB algorithm even succeeds in finding the optimal transmit power allocations for practical large-scale scenarios. The contents of Chapter 3 have been published as [30, 31, 32].

In Chapter 4 it is shown that a particular non-convex spectrum coordination problem can be reformulated as a convex geometric programming problem under certain conditions on the channel environment and on the transmit power resource constraints. This result provides the first optimality certificate for a finite number of tones, and it is applicable to synchronous as well as asynchronous DSL transmission. There are two key innovations in the analysis: the use of geometric programming and M-matrix theory. Furthermore, the optimality analysis leads to a novel spectrum coordination algorithm, referred to as *DSM-GP*, with simulations demonstrating the usefulness of the optimality

certificate to realistic DSL scenarios. The contents of Chapter 4 have been published as [33].

In Chapter 5 an initial framework is proposed to study upper layer performance by looking at scheduling and spectrum coordination together. This contribution extends the physical layer data rate maximization approach to optimization of upper layer performance metrics such as throughput and delay, that may be much more important to improve user satisfaction. It is shown how optimal scheduling can be combined with optimal spectrum coordination, where algorithms are proposed that require only polynomial complexity. The contents of Chapter 5 have been published as [34].

In Chapter 6 traditional physical layer spectrum coordination, which focuses on data rate maximization, is extended towards a much wider energy-efficient scope. This leads to a unifying mathematical representation of energy-aware spectrum coordination problems which is referred to as '*Green DSL*'. A general solution procedure is furthermore proposed to tackle the corresponding spectrum coordination problems. Using this 'Green DSL' approach, the energy-rate trade-off is quantified for different practical settings. It is demonstrated that large transmit energy savings can be achieved with only a minor degradation in data rate performance. This is due to not just the logarithmic dependence of data rate on the signal-to-interference-and-noise ratio but also to the characteristics of crosstalk in DSL systems. The contents of Chapter 6 have been published as [35] and submitted as [36].

Chapter 7 introduces the concept of *fair greening*. It is first shown that straightforward application of Green DSL can sometimes lead to very unfair data rate and power allocations. Then four formulations of fair greening are proposed to tackle this unfairness problem. These formulations are unified into one general representation, and a unified algorithm, referred to as *General-Green-DSL*, is proposed to solve it. Simulations are then shown which quantify the intuitions on fairness for the four proposed fair greening formulations, as they offer a range of choices between maintaining a high sum data rate and enforcing various definitions of fairness. The contents of Chapter 7 have been published as [37].

Chapter 8 proposes an improved dual decomposition approach for spectrum coordination. Many spectrum coordination algorithms are based on a standard subgradient-based dual decomposition approach. In practice however, this approach is often found to lead to slow convergence or even no convergence at all, one of the reasons being the very difficult tuning of the stepsize parameters. In this chapter, a novel improved dual decomposition approach is presented that is inspired by recent advances in mathematical programming. It uses a smoothing technique for the Lagrangian combined with an optimal gradient based scheme for updating the Lagrange multipliers. The stepsize parameters are furthermore selected optimally removing the need for a tuning strategy.

It will be shown how this improved approach improves on the convergence of current state-of-the-art spectrum coordination algorithms based on iterative convex approximations (e.g. CA-DSB[31], SCALE[38]) with one order of magnitude. Furthermore the improved dual decomposition approach is applied to other DSM algorithms (OSB, ISB, ASB, ...) and further improvements are proposed so as to obtain fast and robust spectrum coordination algorithms. The effectiveness of the improved dual decomposition approach is demonstrated for several realistic multi-user DSL scenarios. The contents of Chapter 8 have been published as [39] and is accepted with mandatory revisions as [40].

Chapter 9 focuses on resource allocation for partial signal coordination, where an optimal dual decomposition based algorithm is proposed. The complexity of the algorithm is shown to be more favourable compared to that of the resource allocation algorithm as proposed in [41]. Furthermore an extension is made to the multi-user case. The contents of Chapter 9 have been published as [42, 43].

Chapter 10 focuses on the problem of joint spectrum coordination and partial signal coordination. An algorithm is proposed that is based on a dual decomposition approach. The spectrum coordination part is limited to a type of ON/OFF spectrum coordination so as to reduce the computational complexity. The resulting algorithm is compared to existing algorithms that consist of independent spectrum coordination and partial signal coordination. The contents of Chapter 10 have been published as [44, 45, 46].

Chapter 11 focuses on joint spectrum and receiver signal coordination. More specifically, a transmission scheme is presented, referred to as MAC-OSB, which focuses on the weighted rate sum capacity and consists of spectrum coordination in combination with the generalized decision feedback equalizer. The proposed scheme incorporates per-user total power constraints, spectral mask constraints, and discrete bit or power loading constraints. Furthermore a low-complexity scheme, referred to as MAC-ISB, is presented which performs similar to MAC-OSB. Simulations are shown which demonstrate the large performance gains over existing schemes, especially for scenarios with significant alien crosstalk. The contents of Chapter 11 have been published as [47].

1.5 Dynamic Spectrum Management

Dynamic spectrum management (DSM) refers to a set of practical techniques to optimize DSL transmission for a cable bundle of twisted pairs. The techniques are defined for different levels of coordination and consist of optimizing the parameters and resources of the DSL transceivers by taking the crosstalk interference of the particular DSL scenario into account.

In Figure 1.12, an example of a possible DSL DSM infrastructure is shown. It

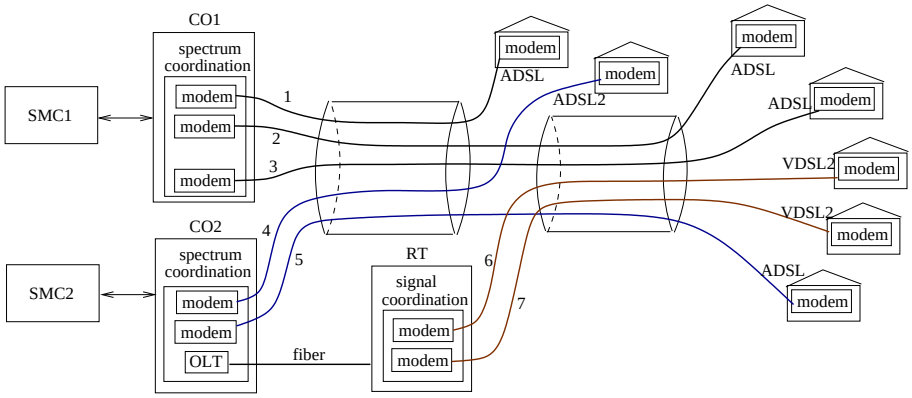


Figure 1.12: DSL DSM environment

consists of two central offices (CO1 and CO2) and one fiber-fed remote terminal (RT) that offer different types of DSL services (ADSL(2(+)), VDSL(2)) to the end-users. Furthermore there are two spectrum management centers (SMCs). The SMC enables the application of DSM. It can be seen as a set of computer workstations that collects data from the modems, e.g., line margins, transmit spectra, bits/tone tables, channel and noise characteristics, etc. After a processing step of the obtained data, the SMC then returns recommendations that can steer the modems towards a particular operating point through the configuration of DSL parameters (e.g. transmit spectra, data rates, forward-error-correction choices, allowed maximum margins, etc). Such recommendations can change daily or even more often [48].

DSM techniques can be classified into four levels based on the level of coordination among the interfering DSL modems [18]:

- **DSM level-0** corresponds to no DSM coordination. The SMC does not receive any information from the DSL modems under consideration.
- **DSM level-1** corresponds to the management of each twisted pair transmission, based on the reported characteristics of this cable. In this case, the SMC independently optimizes the resources for each single-line transmission.
- **DSM level-2** corresponds to the joint coordination of the DSL parameters (transmit spectra, data rates, etc) of multiple DSL lines by using the full channel knowledge (direct as well as crosstalk channel frequency responses). Multi-user spectrum coordination can be seen as a DSM level-2 technique.
- **DSM level-3** corresponds to signal coordination, i.e. signal processing (crosstalk cancellation and/or precoding) is performed on a set of received and/or transmitted symbols so as to remove crosstalk. It is sometimes

also referred to as *Vectored DSL*. This level of coordination can only be applied to transmitters and/or receivers that are colocated.

The attainable coordination, and so also the DSM level that can be applied, strongly depends on the situation, i.e. the particular DSL scenario, government regulation, etc. For instance, for the DSL infrastructure in Figure 1.12, there are only a number of coordinations possible. DSM level-2 and level-3 can be applied on lines 1, 2 and 3, where the crosstalk from lines 4, 5, 6 and 7 will act as noise sources to these lines. DSM level-2 can be applied to lines 4, 5, 6 and 7, and DSM level-3 to lines 4 and 5. At the RT, DSM level-2 as well as level-3 can be applied to lines 6 and 7.

The different levels of DSM techniques can differ significantly in complexity and in the required communication overhead. Generally, a higher level of coordination, corresponds to more complexity and communication overhead, but also to a better system performance.

The data rate performance of a full cable bundle transmission is not expressed by a single performance measure, but it is usually expressed using the concept of a *rate region*, also known as the *capacity region*. The rate region for a cable bundle is defined as the set of all possible data rate combinations for the corresponding DSL services that can be achieved for that particular channel and under the considered resource constraints (spectral mask and per-user total power constraints). An example of a two-user rate region is shown in Figure 1.13. All points within the region $\mathcal{R}$ can be achieved. The points on the boundary of the rate region correspond to the optimal operating points in terms of data rates. There is a clear trade-off between the achievable data rates of the DSL services due to the presence of crosstalk. Increasing the achievable data rate for a particular user, reduces the achievable data rate that can be achieved for the other users. If no crosstalk was present, the rate region would correspond to a square region.

The level of coordination determines the size of the rate region. Generally the rate region increases as more coordination is possible. For instance, the usage of signal coordination (DSM level-3) results in a larger rate region than for spectrum coordination (DSM level-2). The relation between the level of coordination, the computational complexity, and the performance is an important trade-off in the field of DSM.

In the next sections, we will focus on spectrum coordination and signal coordination, where we will give an overview of state-of-the-art solutions and developments, and their corresponding trade-offs in terms of computational complexity and performance.

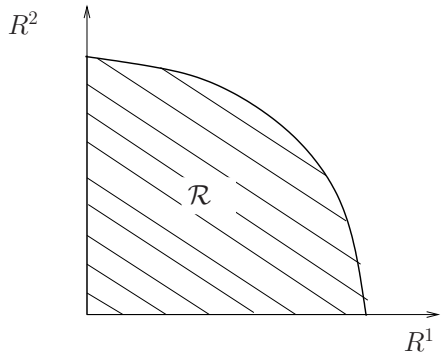


Figure 1.13: Rate region for a two-user cable bundle

1.6 State-of-the-Art Spectrum Coordination

The total amount of crosstalk interference that a particular modem n undergoes, depends on the transmit powers of the interfering modems m ($\neq n$). This can also be seen from (1.17). Increasing the transmit power for a modem generally results in an increased data rate over the corresponding twisted pair, but also impacts the performance of the other twisted pair transmissions due to crosstalk interference. The goal of spectrum coordination, which is also referred to as *spectrum management* or *spectrum optimization*, is to manage and optimize the allocation of the transmit spectra of all modems so as to mitigate the impact of crosstalk and so as to obtain a good trade-off for all single-line DSL transmissions.

In Figure 1.14, a VDSL scenario is shown that consists of three CO lines and three RT lines, which all interfere with each other inside the cable bundle. In the downstream direction, this corresponds to a near-far problem, as discussed in Section 1.3.3. In this case, spectrum coordination comes down to determining the transmit powers s_k^n across all users ($n \in \{1, 2, \ldots, 6\}$) for all tones ($k \in \mathcal{K} = \{1, 2, \ldots, 1608\}$) so as to improve the data rates. The underlined data rates are obtained when no spectrum coordination is applied, i.e. transmission at standardized spectral mask values, whereas the encircled data rates are the achievable data rates obtained with optimal spectrum coordination, i.e. an optimal allocation of the transmit spectra. So by optimizing the transmit powers, the data rates of the three CO lines are improved by a factor of 2, with even slightly increased data rates for the three RT lines. For illustration, the optimal transmit spectra that maximize the data rates, are shown in Figure 1.15.

Spectrum coordination is a very attractive solution as it requires only a minor infrastructure investment and it allows to significantly improve the performance by simply configuring the transmit spectrum settings of the modems. Note that

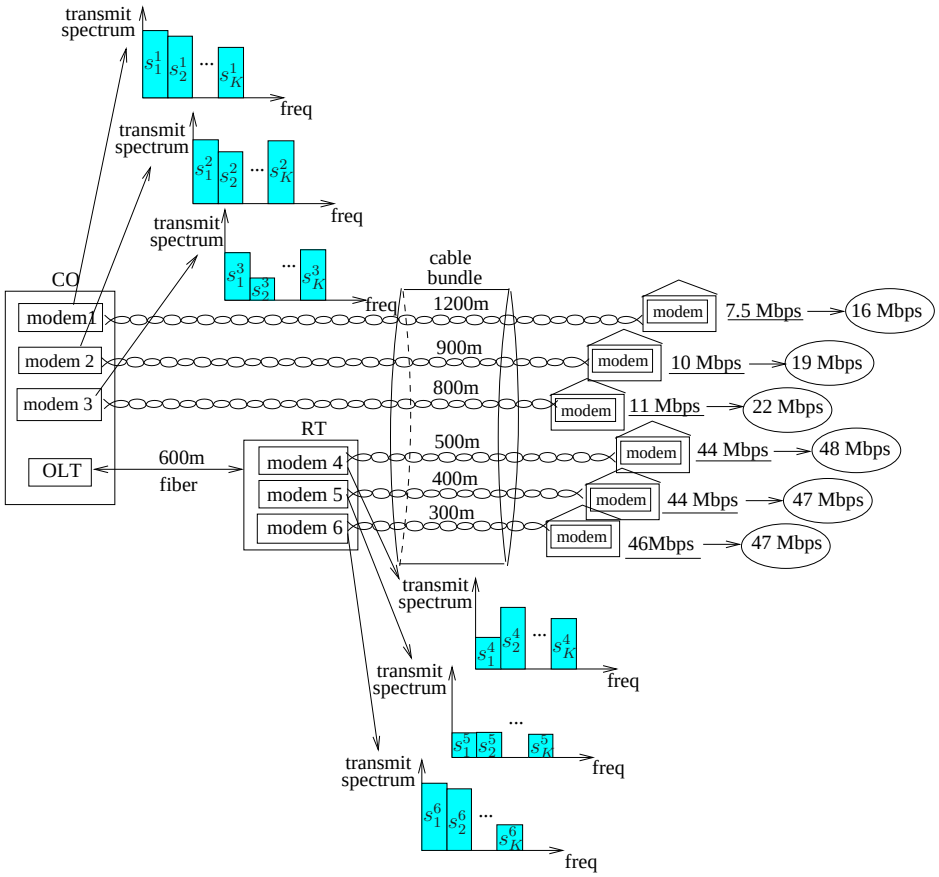


Figure 1.14: Mixed CO-RT VDSL downstream scenario with corresponding data rates for (i) transmit powers at standardized spectral masks (underlined), and (ii) optimal transmit powers (encircled)

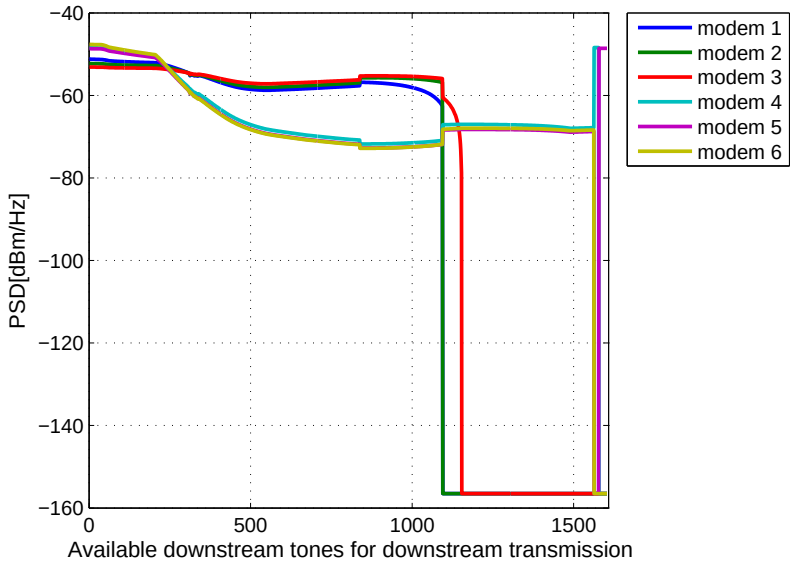


Figure 1.15: Optimal spectrum coordination, i.e. the optimal allocation of transmit powers over all modems and for all tones, for VDSL downstream scenario of Figure 1.14

an improved performance can correspond to increased data rates, but also to other performance metrics such as a larger reach of the DSL service, a more stable behavior, less system power consumption, smaller delay performance, etc.

An overview will now be given of state-of-the-art techniques for spectrum coordination. The system model will be discussed, which relates transmit spectra with data rates and total transmit powers, and which allows the formulation of a number of spectrum coordination problems. Then, an overview is given of state-of-the-art spectrum coordination algorithms, which are optimization procedures designed for tackling the corresponding spectrum coordination problems.

1.6.1 System Model

Within the cable bundle interference environment, each receiver is only interested in the data transmitted by the corresponding transmitter. The data rate of a DSL service is limited by the Gaussian background noise at the corresponding receiver and the crosstalk from the other users. In the case of spectrum coordination, the level of coordination only consists of allocating the transmit powers across all the users and tones. No joint signal coordination is possible among transmitters or receivers. Only single-user encoders and decoders can be employed at each transmitter and receiver, respectively.

Considering (synchronized) DMT-based modems, where transmission is decoupled over multiple independent per-tone channels, the multi-user transmission channel for a particular tone can be viewed as a so-called *Gaussian interference channel* [49].

Establishing the capacity region of a Gaussian interference channel for all possible regimes (strong, moderate and low interference) is a long-standing open problem and remains an active research topic in information theory. For the very strong interference regime, it was first shown in [50] that interference does not reduce the capacity for a two-user Gaussian interference channel, and then subsequently the capacity region was determined in [51, 52]. Recent progress has shown that for the low interference regime, treating interference as noise achieves the sum capacity of a Gaussian interference channel [53, 54, 55].

For a practical DSL cable bundle, the crosstalk channels are typically weaker than the direct channels. It is therefore common to treat the crosstalk interference as additive white Gaussian noise (AWGN). Consequently, the achievable bit rate for line n on tone k can be expressed as follows

$$b_k^n = \log_2\left(1 + \frac{\mathrm{SINR}_k(\mathbf{s}_k)}{\Gamma}\right), \tag{1.18}$$

with the received signal-to-noise-and-interference ratio (SINR) on tone k defined as follows

$$\mathrm{SINR}_k(\mathbf{s}_k) = \frac{|h_k^{n,n}|^2 s_k^n}{\sum_{m \neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n}, \tag{1.19}$$

with $\mathbf{s}_k = [s_k^1, \ldots, s_k^N]^T$, and where the SNR-gap Γ is a constant that depends on the used constellation and modulation, the target probability of error, the channel code, and the noise margin as explained in Section 1.2.3.

The total achievable data rate of a particular DSL service for user n is then given as follows

$$R^n = f_s \sum_{k \in \mathcal{K}} b_k^n. \tag{1.20}$$

This expression is important as it provides the relation between the transmit powers, which are the optimization variables for spectrum coordination, and the data rate, which is an essential performance metric for the DSL system.

Another important performance metric is the per-user total allocated transmit power which is given as follows

$$P^n = \sum_{k \in \mathcal{K}} s_k^n, \qquad n \in \mathcal{N}. \tag{1.21}$$

The transmit power allocations of the different modems are subject to a number of constraints: (i) spectral mask constraints (1.22) so as to ensure spectral compatibility with other DSL services, and (ii), per-user total power constraints (1.23), that arise from limitations on the modem's AFE:

$$s_k^n \leq s_k^{n,\mathrm{mask}} \quad n \in \mathcal{N}, k \in \mathcal{K}, \tag{1.22}$$

$$P^n \leq P^{n,\mathrm{tot}} \quad n \in \mathcal{N}. \tag{1.23}$$

Note that another system model for the achievable data rate exists that takes non-perfect synchronization into account, as proposed in [56]. Such a system model includes intercarrier interference (ICI) which destroys the tone-decoupling assumption. However, in this thesis the focus is on DMT-based DSL systems with perfect synchronization.

1.6.2 Spectrum Coordination Problem Modeling

The management and optimization of the transmit powers across all users and tones provides a means to balance the impact of crosstalk. This can be driven by different objectives and/or constraints, that model the QoS requirements of

the considered cable bundle transmission. For instance, high data rates constitute a popular design objective. The achievable rate region $\mathcal{R}$ characterizes the set of all combinations of data rates that can be achieved, and can now be defined as follows

$$\begin{aligned}\mathcal{R} &= \Big\{(R^n : n \in \mathcal{N}) | R^n = f_s \sum_{k \in K} b_k^n(\mathbf{s}_k), \{\mathbf{s}^n, n \in \mathcal{N}\} \in \mathcal{S}\Big\}, \\ \text{with } \mathcal{S} &= \Big\{(s_k^n : n \in \mathcal{N}, k \in \mathcal{K}) : \sum_{k \in K} s_k^n \leq P^{n,\text{tot}}, 0 \leq s_k^n \leq s_k^{n,\text{mask}}\Big\},\end{aligned} \tag{1.24}$$

with $\mathbf{s}^n = [s_1^n, \ldots, s_K^n]^T$ denoting the transmit spectrum of user n and where $\mathcal{S}$ models the set of constraints on the transmit powers, that are imposed by DSL standards. Note that the rate region is typically not known in advance, and depends on the considered DSL scenario (direct and crosstalk channel frequency responses, noises, etc). The boundary of the rate region characterizes the optimal operating points in terms of data rates.

A spectrum coordination (SC) problem is an optimization problem that models one or more operating points, i.e. a combination of per-user data rates, of the achievable rate region. The relation between the transmit powers, which are the optimization variables in the case of spectrum coordination, and the data rates is given by (1.18), (1.19) and (1.20).

Over the last years, several spectrum coordination problems have been proposed. A first typical SC formulation [19] is as follows

$$\begin{array}{lll} \max\limits_{\{\mathbf{s}^n, n \in \mathcal{N}\}} & R^1 & \\ \text{s.t.} & R^n \geq R^{n,\text{target}}, & n \in \mathcal{N}, n > 1, \\ & P^n \leq P^{n,\text{tot}}, & n \in \mathcal{N}, \\ & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, & n \in \mathcal{N}, k \in \mathcal{K}, \end{array} \tag{1.25}$$

where $R^{n,\text{target}}$ denotes the minimum data rate constraint for user n. (1.25) maximizes the data rate of one modem while satisfying a minimum required data rate for all other modems, and subject to transmit power constraints that are imposed by DSL standards. Note that SC problem (1.25) assumes that the modems $n(> 1)$ can jointly achieve a data rate allocation with $R^n \geq R^{n,\text{target}}, n > 1$. This is not necessarily so in which case the problem is infeasible.

In [19], it is mentioned that the rate region can be assumed to be convex, because of the channel frequency correlation between adjacent tones and the small tone spacing for practical DSL systems. Based on this convexity assumption,

the following SC problem was proposed in [19]:

$$\begin{array}{ll} \max\limits_{\{\mathbf{s}^n, n\in\mathcal{N}\}} & \sum\limits_{n\in\mathcal{N}} w_n R^n \\ \text{s.t.} & P^n \le P^{n,\text{tot}}, \qquad n\in\mathcal{N}, \\ & 0 \le s_k^n \le s_k^{n,\text{mask}}, \quad n\in\mathcal{N}, k\in\mathcal{K}, \end{array} \tag{1.26}$$

where w_n can be regarded as weights that specify the importance of each user n. For a given set of weights $\{w_n, n\in\mathcal{N}\}$, (1.26) models a particular point on the boundary of the rate region. By varying the weights relatively to each other, different boundary points can be obtained, and one can trace out the full achievable rate region.

Data rate constraints can be added to formulation (1.26) as follows:

$$\begin{array}{ll} \max\limits_{\{\mathbf{s}^n, n\in\mathcal{N}\}} & \sum\limits_{n\in\mathcal{N}} w_n R^n \\ \text{s.t.} & R^n \ge R^{n,\text{target}}, \qquad n\in\mathcal{N}, \\ & P^n \le P^{n,\text{tot}}, \qquad n\in\mathcal{N}, \\ & 0 \le s_k^n \le s_k^{n,\text{mask}}, \quad n\in\mathcal{N}, k\in\mathcal{K}, \end{array} \tag{1.27}$$

which can be regarded as a generalization of (1.25), i.e. (1.25) is obtained by fixing the constants in (1.27) to $w_1 = 1, R^{1,\text{target}} = 0, \text{and } \forall n \neq 1 : w_n = 0$.

An alternative SC problem, introduced in [57], is as follows:

$$\begin{array}{ll} \max\limits_{R,\{\mathbf{s}^n, n\in\mathcal{N}\}} & R \\ \text{s.t.} & R^n \ge \beta_n R, \qquad n\in\mathcal{N}, \\ & P^n \le P^{n,\text{tot}}, \qquad n\in\mathcal{N}, \\ & 0 \le s_k^n \le s_k^{n,\text{mask}}, \quad n\in\mathcal{N}, k\in\mathcal{K}, \end{array} \tag{1.28}$$

where β_n is a constant that specifies the minimum required proportion in data rate with respect to a base rate R for user n. This problem allows to model a point on the boundary of the rate region that satisfies a certain proportional data rate allocation over the modems.

In [58], a number of new objectives were proposed that focus on achieving more fairness in the allocations of the data rates over the users. Here, a particular point on the boundary of the rate region is modeled that corresponds to a specific notion of fairness, such as proportional fairness, harmonic-rate fairness, and max-min fairness. The SC problem is as follows

$$\begin{array}{ll} \max\limits_{\{\mathbf{s}^n, n\in\mathcal{N}\}} & H(R^n, n\in\mathcal{N}) \\ \text{s.t.} & P^n \le P^{n,\text{tot}}, \qquad n\in\mathcal{N}, \\ & 0 \le s_k^n \le s_k^{n,\text{mask}}, \quad n\in\mathcal{N}, k\in\mathcal{K}, \end{array} \tag{1.29}$$

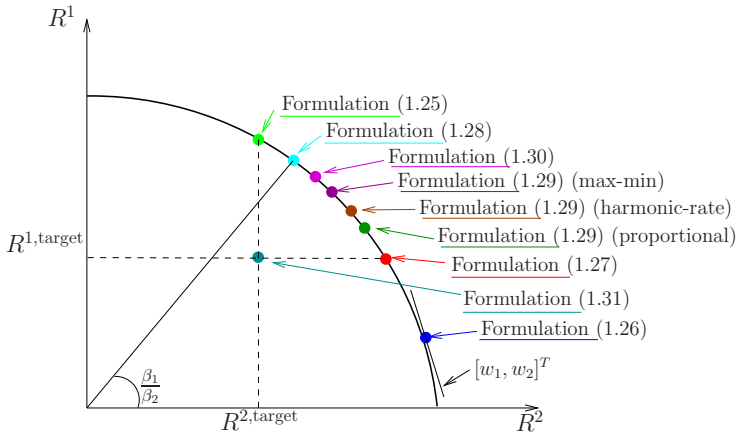


Figure 1.16: Operating points for different SC problem formulations (1.25)-(1.31) for a two-user DSL scenario

with $H(R^n, n \in \mathcal{N})$ having the following particular expressions with an increasing level of fairness:

1. **proportional fairness**: $H(R^n, n \in \mathcal{N}) = \sum_{n \in \mathcal{N}} \ln(R^n)$
2. **harmonic-rate fairness**: $H(R^n, n \in \mathcal{N}) = -\ln\Big(\sum_{n \in \mathcal{N}} (R^n)^{-1}\Big)$
3. **max-min fairness**: $H(R^n, n \in \mathcal{N}) = \min_{n \in \mathcal{N}} R^n$

In [59, 60, 61] another approach was proposed that also focuses on achieving more fairness in the data rate allocations and which is referred to as *balanced capacity*. More specifically, it corresponds to a point on the boundary of the rate region where each user n experiences the same relative loss in data rate performance with respect to its maximum (single-user) capacity $R^{n,\text{SU}}$, i.e. its maximum achievable data rate performance in the absence of crosstalk. The corresponding SC problem is given as follows

$$\begin{aligned} \max_{\{\mathbf{s}^n, n \in \mathcal{N}\}} \quad & \sum_{n \in \mathcal{N}} \frac{1}{N} R^n \\ \text{s.t.} \quad & \frac{R^1}{R^{1,\text{SU}}} = \frac{R^n}{R^{n,\text{SU}}}, \quad n \in \mathcal{N}, n > 1, \\ & P^n \leq P^{n,\text{tot}}, \quad n \in \mathcal{N}, \\ & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \quad n \in \mathcal{N}, k \in \mathcal{K}. \end{aligned} \tag{1.30}$$

For a two-user DSL scenario, the solutions of the SC problems (1.25)-(1.30)

are visualized in the achievable rate region as in Figure 1.16. Note that these SC problems can only characterize a point on the boundary of the achievable rate region, which mostly correspond to transmit power allocations where all users use their full available total transmit power, i.e. $P^n = P^{n,\text{tot}}$. This is not the best strategy in terms of energy efficiency as the energy consumption of a modem is proportional to the per-user total transmit power [62].

In [62, 58, 28, 63, 35, 36], power driven SC problems were proposed that focus on minimizing the transmit powers of the modems subject to minimum data rate constraints, e.g.

$$\begin{array}{lll} \max\limits_{\{\mathbf{s}^n, n\in\mathcal{N}\}} & \sum\limits_{n\in\mathcal{N}} t_n P^n & \\ \text{s.t.} & R^n \geq R^{n,\text{target}}, & n \in \mathcal{N}, \\ & 0 \leq s_k^n & n \in \mathcal{N}, k \in \mathcal{K}, \end{array} \tag{1.31}$$

where the weights t_n indicate the importance of the total transmit power for user n. Unlike the other SC problems, (1.31) characterizes a point inside the rate region as shown in Figure 1.16. This operating point does not correspond to an optimal trade-off in data rates, but rather focuses on improving the energy-efficiency as it minimizes the transmit powers so as to obtain a minimum required data rate for each user.

1.6.3 Spectrum Coordination Algorithms

In this section, an overview will be given of state-of-the-art spectrum coordination algorithms, i.e. algorithms for finding a good, and preferably the best, solution for a particular SC problem. Note that most SC problems are NP-hard nonconvex optimization problems [58], due to the nonconvex relation between the transmit powers and data rates, as given by (1.18), (1.19) and (1.20). Although multiple SC problems exist, as presented in Section 1.6.2, the main focus will be on problem (1.26), which is repeated in (1.32) for convenience. This is the most common problem formulation, where a weighted sum of data rates is maximized subject to power constraints, and which will be referred to as the "constrained weighted rate sum maximization (cWRS)" problem. In Chapters 6 and 7, it will be shown that most SC algorithms designed for the cWRS problem, can easily be modified to tackle the other SC problems of Section 1.6.2.

The primal problem formulation of cWRS is given as follows

$$\begin{array}{lll} \max\limits_{\{\mathbf{s}^n, n\in\mathcal{N}\}} & \sum\limits_{n\in\mathcal{N}} w_n R^n & \\ \text{s.t.} & P^n \leq P^{n,\text{tot}}, & n \in \mathcal{N}, \qquad \text{(cWRS)} \\ & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, & n \in \mathcal{N}, k \in \mathcal{K}, \end{array} \tag{1.32}$$

which corresponds to a nonconvex and NP-hard problem [58]. The number of optimization variables, i.e. transmit powers $\{s_k^n, n \in \mathcal{N}, k \in \mathcal{K}\}$, is equal to KN, where the number of users N ranges between 2-100 and the number of tones K can go up to 4000 (for VDSL2). Depending on the particular DSL scenario, i.e. the particular values for the direct and crosstalk channel frequency responses and the noise levels, the cWRS problem can have multiple locally optimal solutions. Furthermore these locally optimal solutions can significantly differ in optimality, as demonstrated in [31].

Multiple SC algorithms have been proposed to solve the cWRS problem. In this thesis, the main focus is on *dual SC algorithms*. An optimization problem can be regarded from either of two formulations, the primal or the dual problem formulation, which lead to the same solution under certain conditions. Note that the SC problems (1.25)-(1.31) are primal SC problem formulations. Dual SC algorithms are algorithms that focus on the dual SC problem formulations. They constitute the major part of existing SC algorithms, next to a number of heuristic SC algorithms such as for instance the flat power back-off method [64] the reference PSD/length method [64, 65, 66] and the reference noise method [67]. Dual SC algorithms are interesting in that some of them can find the globally optimal solution of the corresponding SC problems, and furthermore because they can be easily analyzed in terms of complexity and performance.

The dual problem formulation of cWRS is given as follows

$$\begin{array}{ll} \min\limits_{\boldsymbol{\lambda}} & g(\boldsymbol{\lambda}) \\ \text{s.t.} & \boldsymbol{\lambda} \geq 0, \end{array} \tag{1.33}$$

where $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_N]^T$ are the *dual variables*, also referred to as *Lagrange multipliers*, and where $g : \mathbb{R}^N \rightarrow \mathbb{R}$ is the *Lagrange dual function* (or just *dual function*), which is defined as an optimization problem with parameter $\boldsymbol{\lambda}$ as follows

$$\begin{array}{l} g(\boldsymbol{\lambda}) = \left\{ \begin{array}{ll} \max\limits_{\{\mathbf{s}_k, k\in\mathcal{K}\}} & \mathcal{L}(\boldsymbol{\lambda}, \mathbf{s}_k, k \in \mathcal{K}) \\ \text{s.t.} & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \qquad n \in \mathcal{N}, k \in \mathcal{K}, \end{array} \right. \\ \text{with} \;\; \mathcal{L}(\boldsymbol{\lambda}, \mathbf{s}_k, k \in \mathcal{K}) = \sum\limits_{n\in\mathcal{N}} w_n R^n - \sum\limits_{n\in\mathcal{N}} \lambda_n (P^n - P^{n,\text{tot}}), \end{array} \tag{1.34}$$

where $\mathcal{L}(\boldsymbol{\lambda}, \mathbf{s}_k, k \in \mathcal{K})$ is called the Lagrangian. The optimal solution of (1.34) depends on the value of the Lagrange multipliers $\boldsymbol{\lambda}$, and will be denoted as $\{\mathbf{s}_k(\boldsymbol{\lambda}), k \in \mathcal{K}\}$.

In [68] it is shown that asymptotic strong duality holds for the dual problem (1.33)(1.34), i.e. the solution of the dual problem (1.33)(1.34) corresponds to the solution of the primal problem (1.32) when the number of tones K goes to infinity. This was also rigorously proven in [58]. For practical DSL systems, the number of tones K is large and so the assumption of strong duality, i.e. zero duality gap, is a valid one.

The main advantage of the dual problem formulation is that it is separable over the tones, which allows to decompose the problem into multiple independent problems for each tone k as follows

$$\begin{aligned} g(\boldsymbol{\lambda}) &= \sum_{k\in K} g_k(\boldsymbol{\lambda}) \\ \text{with } g_k(\boldsymbol{\lambda}) &= \begin{cases} \max\limits_{\mathbf{s}_k} & b_k(\mathbf{s}_k) - \sum\limits_{n\in\mathcal{N}} \lambda_n s_k^n + \sum\limits_{n\in\mathcal{N}} \lambda_n P^{n,\text{tot}}/K \\ \text{s.t.} & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \quad n \in \mathcal{N}, \end{cases} \\ \text{with } \; b_k(\mathbf{s}_k) &= \sum_{n\in\mathcal{N}} w_n f_s b_k^n(\mathbf{s}_k). \end{aligned} \tag{1.35}$$

This is referred to as *dual decomposition*. Each *per-tone problem* $g_k(\boldsymbol{\lambda})$ is of dimension N. The dual problem then consists of (1.33) that is a problem in dimension N, and (1.35) that corresponds to K independent problems of dimension N. Solving $K+1$ problems of dimension N is generally considered to be much more manageable than solving a large-scale problem of dimension KN, such as the primal problem formulation (1.32). Dual SC algorithms are SC algorithms that tackle the dual SC problem formulation (1.33) and (1.35) so as to find the solution of the corresponding primal problem (1.32).

We will refer to problem (1.33) as the *dual master problem* and to problem (1.34), or (1.35), as the *dual slave problem*. In the remainder of this section an overview will be given of state-of-the-art approaches to these two problems together with a discussion of a number of concrete dual SC algorithms.

Approaches for Solving the Dual Master Problem (1.33)

The dual master problem (1.33) is a convex optimization problem [69] in the Lagrange multipliers $\boldsymbol{\lambda}$. Its objective function is convex but non-differentiable in $\boldsymbol{\lambda}$. Several solution approaches have been proposed to tackle this problem:

A. Bisection approach

In [19] a nested bisection search on the Lagrange multipliers is proposed. It is first proven that the Lagrange multiplier λ_n is monotonically decreasing with the corresponding total transmit power P^n. The procedure then consists of finding upper bounds $\boldsymbol{\lambda}^{\text{up}}$ on the optimal Lagrange multipliers $\boldsymbol{\lambda}^*$ that correspond to total transmit powers that are smaller than the per-user total power constraints, i.e. $\sum_{k\in\mathcal{K}} s_k^n(\boldsymbol{\lambda}^{\text{up}}) \leq P^{n,\text{tot}}, n \in \mathcal{N}$. Once these upper bounds are found, and the lower bounds $\boldsymbol{\lambda}^{\text{low}}$ are set to 0, one can perform a bisection search on the Lagrange multipliers in a successive way so as to converge to the optimal Lagrange multipliers $\boldsymbol{\lambda}^*$. The complexity of the bisection approach is exponential in the number of users N.

B. Subgradient approach

In [68] a subgradient update approach is proposed. Since the dual function $g(\boldsymbol{\lambda})$ is convex, a gradient-type search is guaranteed to converge to the global optimum. Although the objective is not necessarily differentiable, it is proven that a subgradient can be defined as follows

$$\mathbf{d} = \sum_{k \in \mathcal{K}} \mathbf{s}_k(\boldsymbol{\lambda}) - \mathbf{P}^{\text{tot}}. \tag{1.36}$$

A subgradient update for the Lagrange multipliers is then proposed as follows

$$\boldsymbol{\lambda}^{l+1} = \left[\boldsymbol{\lambda}^l + r^l \Big(\sum_{k \in \mathcal{K}} \mathbf{s}_k(\boldsymbol{\lambda}) - \mathbf{P}^{\text{tot}} \Big) \right]_0^{\infty}. \tag{1.37}$$

where l refers to the iteration number and r^l is the scalar step size at iteration l. Iteratively applying (1.37), results in convergence to the optimal value $\boldsymbol{\lambda}^*$ if the stepsizes r_l are small enough or satisfy certain conditions [70][68]. For the choice of $r^l = \frac{\beta}{l}$, where β is a constant, the subgradient update approach is always guaranteed to converge. Another stepsize update rule is $r^l = \beta$ for which the convergence of the subgradient update is typically much faster, but it only converges if the constant β is chosen sufficiently small. A tuning of the constant β can be performed so as to improve the convergence speed and to guarantee convergence.

The complexity of the subgradient update approach is polynomial in N [68], and so these algorithms typically converge faster than the bisection approach.

C. Ellipsoid approach

In [68] an ellipsoid method is proposed, which is a so-called cutting-plane method. The idea here is to define an ellipsoid that contains all candidate values for the optimal Lagrange multipliers $\boldsymbol{\lambda}^*$, and to iteratively update the ellipsoid excluding (or cutting out) values that cannot be optimal. More specifically, it starts from an initial ellipsoid that contains all candidate values for $\boldsymbol{\lambda}$. Based on the subgradient that is evaluated in the center of the ellipsoid, half of the ellipsoid is eliminated as these values do not contain the optimum. A new ellipsoid, which is the minimal-volume ellipsoid containing the remaining half, is then formed. This is repeated until the ellipsoid has converged to the desired accuracy.

The computational cost per Lagrange multiplier update is similar to the subgradient approach. Unlike the bisection and the subgradient approach, the ellipsoid method cannot be implemented in a distributed way, in which each user n updates the Lagrange multiplier λ_n autonomously.

Approaches for Solving the Dual Slave Problem (1.35)

The dual slave problem $g(\boldsymbol{\lambda})$ can be decomposed into many independent per-tone problems $g_k(\boldsymbol{\lambda})$ as given in (1.35). Each per-tone problem corresponds however to a highly nonconvex optimization problem in dimension N. Multiple approaches have been proposed to tackle these nonconvex problems:

A. Discrete Exhaustive Search

In [19], the continuous feasible space $\{0 \leq s_k^n \leq s_k^{n,\text{mask}}, k \in \mathcal{K}, n \in \mathcal{N}\}$ is restricted to a discrete feasible space $\{s_k^n \in \{0, \Delta_s, \ldots, s_k^{n,\text{mask}}\}, k \in \mathcal{K}, n \in \mathcal{N}\}$, where Δ_s can be regarded as the desired accuracy on the transmit power. The discretization is motivated by the fact that, in practice, modems can only define their transmit powers up to a certain accuracy. For instance, for current standards Δ_s is set to 0.5 dBm/Hz [71]. For a particular tone k, the number of possible transmit power combinations for all users is then equal to

$$q_k = \prod_{n \in \mathcal{N}} \left(\frac{s_k^{n,\text{mask}}}{\Delta_s} + 1 \right). \tag{1.38}$$

In [19], it is then proposed to solve each per-tone problem $g_k(\boldsymbol{\lambda})$ by an exhaustive search over the q_k discrete transmit power combinations. Note that q_k is exponential in the number of users N, which results in a large computational complexity. Also note that the exhaustive search can be reduced by considering integer bit loadings instead of discrete transmit powers [19].

B. Discrete Coordinate Descent Method

In [22, 21] it is proposed to tackle $g_k(\boldsymbol{\lambda})$ via a discrete coordinate descent approach. More specifically, it first optimizes s_k^1, while keeping $s_k^2, \ldots, s_k^N$ fixed, by a one dimensional discrete exhaustive search over the discrete feasible space $\{s_k^n \in \{0, \Delta_s, \ldots, s_k^{n,\text{mask}}\}, k \in \mathcal{K}, n \in \mathcal{N}\}$. Then it finds the optimal s_k^2 while keeping all other s_k^n fixed. This is repeated multiple times over all users until convergence, which is guaranteed as each iteration strictly increases the objective function. The approach does not necessarily lead to the globally optimal solution of $g_k(\boldsymbol{\lambda})$. For a sufficiently small discretization, it can converge to a locally optimal solution.

C. D.C. Optimization

In [72, 73] the per-tone nonconvex problem $g_k(\boldsymbol{\lambda})$ is reformulated as a difference of convex functions (d.c.) optimization problem as follows

$$\begin{array}{ll} \max\limits_{\mathbf{s}_k} & v(\mathbf{s}_k) - f(\mathbf{s}_k) \\ \text{s.t.} & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \qquad n \in \mathcal{N}, \end{array} \tag{1.39}$$

with

$$v(\mathbf{s}_k) = \sum_{n\in\mathcal{N}} w_n f_s \log_2\left(|h_k^{n,n}|^2 s_k^n + \Gamma\big(\sum_{m\neq n}|h_k^{n,m}|^2 s_k^m + \sigma_k^n\big)\right) - \sum_{n\in\mathcal{N}}\lambda_n s_k^n,$$
$$f(\mathbf{s}_k) = \sum_{n\in\mathcal{N}} w_n f_s \log_2\left(\Gamma\big(\sum_{m\neq n}|h_k^{n,m}|^2 s_k^m + \sigma_k^n\big)\right),$$

where the functions $v(\mathbf{s}_k)$ and $f(\mathbf{s}_k)$ are concave functions in $\mathbf{s}_k$. This is then reformulated as an equivalent concave minimization problem as follows

$$\begin{array}{ll} \min\limits_{\mathbf{s}_k,t} & f(\mathbf{s}_k) - t \\ & t - v(\mathbf{s}_k) \leq 0 \\ \text{s.t.} & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \qquad n \in \mathcal{N} \end{array} \tag{1.40}$$

where the objective is a concave function, and the constraints form a convex set. Then a prismatic branch-and-bound (BnB) algorithm is proposed to solve the corresponding problem to find the globally optimal solution of $g_k(\boldsymbol{\lambda})$. The basic idea of a BnB approach can be summarized as follows. The constraint set is divided into smaller disjunct subsets (branching) for which lower and upper bounds of the minimum value of the objective function are computed (bounding). When the sizes of the subsets decrease, the corresponding bounds can be made tighter. Subsets of the feasible set with lower bounds greater than the best upper bound found so far are deleted from further consideration, because these subsets cannot contain the global optimum.

A branch and bound approach can be regarded as an intelligent exhaustive search that can significantly decrease the search space with respect to the full exhaustive search. The complexity of solving $g_k(\boldsymbol{\lambda})$ however remains exponential in the number of users N.

D. Convex Approximation

Another approach to tackle the nonconvex problem $g_k(\boldsymbol{\lambda})$ is to approximate it with a simpler convex problem and solve that one instead. The performance of the obtained transmit spectra then depends on the chosen convex approximation and the considered DSL scenario.

In [23] a convex approximation is proposed as follows:

$$\forall n \in \mathcal{N}: \quad \begin{array}{ll} \max\limits_{s_k^n} & w_n f_s \log_2\left(1 + \dfrac{1}{\Gamma}\dfrac{|h_k^{n,n}|^2 s_k^n}{\sum_{m\neq n}|h_k^{n,m}|^2 s_k^m + \sigma_k^n}\right) - \lambda_n s_k^n \\ \text{s.t.} & 0 \leq s_k^n \leq s_k^{n,\text{mask}} \end{array} \tag{1.41}$$

With this approximation, each user maximizes its own per-tone performance selfishly, while ignoring the crosstalk impact caused to the other users. The

approximation leads to the optimal solution of $g_k(\boldsymbol{\lambda})$ if no crosstalk is present, and otherwise it leads to a suboptimal solution. The advantage of this convex approximation is that it can be easily solved for each user n in closed form as follows:

$$s_k^n = \left[\frac{w_n f_s}{\ln(2)\lambda_n} - \frac{\Gamma \sum_{m \neq n} |h_k^{n,m}|^2 s_k^m + \Gamma \sigma_k^n}{|h_k^{n,n}|^2} \right]_0^{s_k^{n,\text{mask}}} \tag{1.42}$$

Note that the solution for user n depends on the transmit powers of the other users. Formula (1.42) is iteratively applied over all users and multiple times so as to converge to the solution of (1.41). Although convergence is only proven under limited channel conditions, it is demonstrated that the procedure converges for practical DSL scenarios [23].

In [74] a slightly modified approximation was proposed where λ_n is replaced by $\lambda_n \mu_k^n$, i.e. a tone-dependent scaling of the Lagrange multiplier λ_n. By tuning the values of μ_k^n for the particular DSL scenario, a better approximation and solution can be obtained for $g_k(\boldsymbol{\lambda})$.

The convex approximation approach proposed in [75] is based on geometric programming (GP) [76]. The approximation is based on the following relation

$$\log(1+x) \approx \log(x), \qquad \text{for large } x, \tag{1.43}$$

and is given as follows

$$\begin{aligned} \max_{\mathbf{s}_k} \quad & \sum_{n \in \mathcal{N}} w_n f_s \log_2 \left(\frac{1}{\Gamma} \frac{|h_k^{n,n}|^2 s_k^n}{\sum_{m \neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n} \right) - \sum_{n \in \mathcal{N}} \lambda_n s_k^n \\ \text{s.t.} \quad & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \qquad n \in \mathcal{N} \end{aligned} \tag{1.44}$$

Applying an exponential transformation of variables $s_k^n = e^{\tilde{s}_k^n}$, results in a convex problem in the transformed variables $\tilde{\mathbf{s}}_k$, and more specifically a geometric programming problem. For a high signal-to-interference-and-noise ratio (SINR) the convex approximation, i.e.(1.44) with $s_k^n = e^{\tilde{s}_k^n}$, forms a good approximation and one can expect a near-optimal solution for the original nonconvex problem $g_k(\boldsymbol{\lambda})$.

E. Virtual Reference Line Protection

In [29], an approximation is proposed for $g_k(\boldsymbol{\lambda})$ that is based on the concept of a virtual reference line. Each user defines a virtual reference line, and optimizes its own bit rate performance while taking into account the damage caused on its virtual reference line. This approach avoids a "selfish" behavior leading to a better cable bundle performance. This comes down to the following approx-

imation of $g_k(\boldsymbol{\lambda})$

$$\begin{aligned}\forall n \in \mathcal{N}: \quad \max_{s_k^n} \quad & w_n f_s \log_2\left(1+\frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^n}{\sum_{m\neq n}|h_k^{n,m}|^2 s_k^m+\sigma_k^n}\right) \\ & +w_n^{\text{ref}} f_s \log_2\left(1+\frac{1}{\Gamma}\frac{|h_k^{\text{ref}}|^2 s_k^{\text{ref}}}{|h_k^{\text{ref},n}|^2 s_k^n+\sigma_k^{\text{ref}}}\right)-\lambda_n s_k^n \\ \text{s.t.} \quad & 0\leq s_k^n \leq s_k^{n,\text{mask}}\end{aligned} \tag{1.45}$$

where the parameters $h_k^{\text{ref},n}$, s_k^{ref}, h_k^{ref}, σ_k^{ref}, and w_{ref}, respectively, denote the crosstalk channel frequency response from the nth user into the reference line, the transmit power, the channel frequency response, the noise, and the weight of the virtual reference line. The values of these parameters need to be chosen in advance and are constants.

Note that approximation (1.45) is not convex, but its KKT stationarity condition can be written as a cubic equation that can be easily solved by checking its three roots as well as the boundary conditions $s_k^n = 0$ and $s_k^n = s_k^{n,\text{mask}}$. This generally leads to a near optimal solution of $g_k(\boldsymbol{\lambda})$ if the parameters of the virtual reference line are chosen well. For further details on good parameter choices, we refer to [29].

F. Iterative Convex Approximation

In [77, 38] a number of iterative convex approximation approaches are proposed that can be used to tackle the nonconvex per-tone problem $g_k(\boldsymbol{\lambda})$. This approach consists of approximating the nonconvex problem by a series of improved convex approximations, each of which can be optimally solved in an easy way. Under certain conditions on the particular convex approximation, which are discussed in [77], this approach is guaranteed to converge to a point satisfying the KKT conditions of $g_k(\boldsymbol{\lambda})$.

One popular iterative convex approximation approach is based on the following approximation

$$\alpha\log(x)+\beta \leq \log(1+x) \text{ with } \begin{cases} \alpha = \frac{x_0}{1+x_0} \\ \beta = \log(1+x_0)-\frac{x_0}{1+x_0}\log(x_0), \end{cases} \tag{1.46}$$

that is tight with equality at a chosen value x_0 when the constants α and β are chosen as specified above. This results in the following approximation of per-tone problem $g_k(\boldsymbol{\lambda})$ during iteration l

$$\begin{aligned}&\mathbf{s}_k^*(l): \\ &\begin{cases} \underset{\mathbf{s}_k}{\text{argmax}} & \displaystyle\sum_{n\in\mathcal{N}}\left(w_n f_s \alpha_k^n(l)\log_2\left(\frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^n}{\sum_{m\neq n}|h_k^{n,m}|^2 s_k^m+\sigma_k^n}\right)+\beta_k^n(l)-\lambda_n s_k^n\right) \\ \text{s.t.} & 0\leq s_k^n \leq s_k^{n,\text{mask}}, \qquad n\in\mathcal{N}, \end{cases}\end{aligned} \tag{1.47}$$

where the approximation parameters $\alpha_k^n(l)$ and $\beta_k^n(l)$ are iteratively updated using the following formulas

$$\begin{aligned} x_{0,k}^n &= \frac{1}{\Gamma} \frac{|h_k^{n,n}|^2 s_k^{*,n}(l)}{\sum_{m \neq n} |h_k^{n,m}|^2 s_k^{*,m}(l) + \sigma_k^n}, \\ \alpha_k^n(l+1) &= \frac{x_{0,k}^n}{1 + x_{0,k}^n}, \\ \beta_k^n(l+1) &= \log_2(1 + x_{0,k}^n) - \frac{x_{0,k}^n}{1 + x_{0,k}^n} \log_2(x_{0,k}^n). \end{aligned} \tag{1.48}$$

It can be shown that (1.47) can be transformed into a geometric programming problem, which is a convex optimization problem [76], by an exponential transformation of variables $s_k^n = e^{\tilde{s}_k^n}$. This results in an iterative convex approximation approach that is guaranteed to converge to a point satisfying the KKT conditions of $g_k(\boldsymbol{\lambda})$. Furthermore, in [38], a low-complexity fixed point update formula is proposed to solve the approximation (1.47) for iteration l, which is given as follows,

$$s_k^n = \left[\frac{w_n f_s \alpha_k^n(l)}{\ln(2)\lambda_n + \sum_{m \neq n} \frac{w_m f_s \alpha_k^m(l) \Gamma |h_k^{m,n}|^2}{\sum_{p \neq m} \Gamma |h_k^{m,p}|^2 s_k^p + \Gamma \sigma_k^m}} \right]_0^{s_k^{n,\mathrm{mask}}}, \qquad n \in \mathcal{N} \tag{1.49}$$

and which converges to the optimal solution of (1.47).

Next to this logarithmic approximation, alternative iterative convex approximations are proposed based on a single and double condensation method. For further details we refer to [77].

General Solution Structure for Dual SC algorithms

Most dual SC algorithms have a similar solution structure which is visualized in Figure 1.17. It consists of a dual approach that decouples the problem into a dual master problem and K independent dual slave problems $g_k(\boldsymbol{\lambda})$, and where the solution procedures for both type of problems depend on mutual information, i.e. $\{\mathbf{s}_k, k \in \mathcal{K}\}$ and $\boldsymbol{\lambda}$. These problems are tackled iteratively until the stop condition is satisfied, which corresponds to complementary slackness, i.e. $\lambda_n(\sum_{k \in \mathcal{K}} s_k^n(\boldsymbol{\lambda}) - P^{n,\mathrm{tot}}) = 0, n \in \mathcal{N}$, as strong duality can be assumed [69].

One alternative solution structure for dual SC algorithms is based on an iterative convex approximation approach. This is similar to the iterative convex approximation approach for tackling the per-tone dual slave problem in Section 1.6.3, but differs in that it focuses on the cWRS problem directly, instead of the per-tone problem $g_k(\boldsymbol{\lambda})$. It basically consists of iteratively executing the

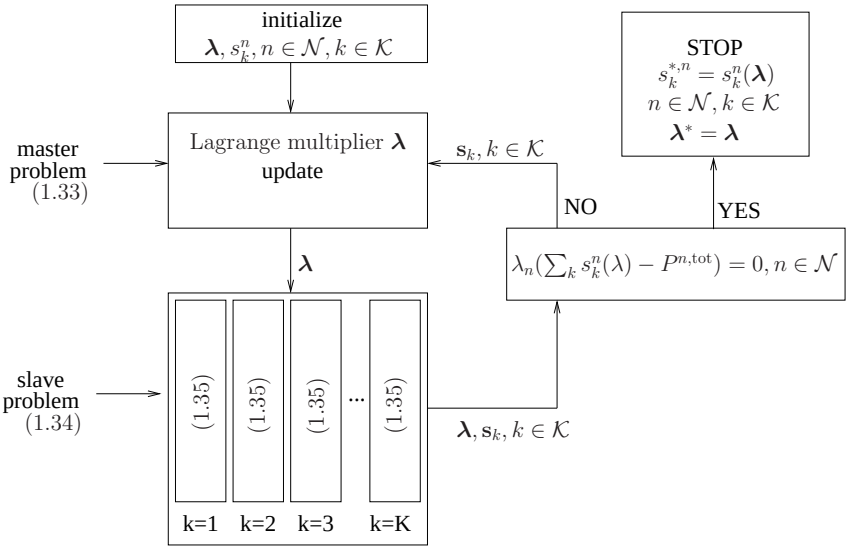


Figure 1.17: General solution structure of dual spectrum coordination algorithms

following two steps: (i) approximating the nonconvex cWRS problem (1.32) by a separable convex problem F_{cvx} and (ii) solving this convex approximation by a dual decomposition approach. Note that under some conditions on the convex approximation, described in [77], iteratively executing these steps results in convergence to a solution that satisfies the KKT conditions of cWRS. This approach is visualized in Figure 1.18, where $\mathrm{F}_{k,\text{cvx}}$ refers to the decomposed per-tone convex problem obtained from the convex approximation F_{cvx}.

Existing dual SC algorithms consist of combinations of the techniques as discussed in Section 1.6.3 for tackling the dual master and slave problems in the solution structures of Figures 1.17 and 1.18.

In [19], the optimal spectrum balancing (OSB) algorithm is proposed. It uses the general solution structure of Figure 1.17, where it uses a bisection update for tackling the dual master problem, and where it tackles the dual slave problems by a discrete exhaustive search. The OSB algorithms succeeds in finding the globally optimal solution within its specified accuracy Δ_s. The computational complexity is exponential in the number of users N. In [68, 8], the solution of the dual master problem is replaced by a subgradient search, and in [68] by an ellipsoid method.

In [22, 21], the iterative spectrum balancing (ISB) algorithm is proposed, which uses the general solution structure of Figure 1.17. The dual master problem is solved using a subgradient update approach and the dual slave problems

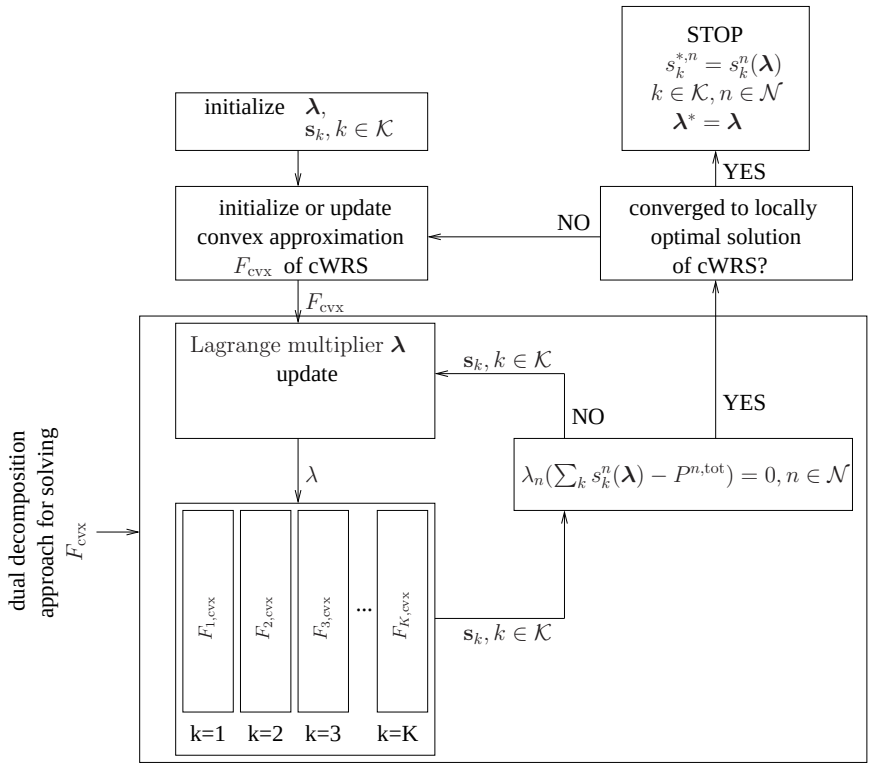


Figure 1.18: Solution structure of iterative convex approximation based SC algorithms

are solved by a discrete coordinate descent method. For a sufficiently small discretization Δ_s, the algorithm can converge to a locally optimal solution, which does not necessarily correspond to the globally optimal solution. The computational complexity is only polynomial in the number of users N [21].

In [73, 72], a prismatic branch-and-bound (PBnB) based SC algorithm is proposed, which uses the general solution structure of Figure 1.17. It uses a subgradient update approach for tackling the dual master problem. The dual slave problems are formulated as D.C. optimization problems and a prismatic branch-and-bound algorithm is used for tackling each of the per-tone dual slave problems. The algorithm succeeds in finding the globally optimal solution and requires an exponential computational complexity in the number of users N.

In [23], the iterative water-filling algorithm (IWF) is proposed. It uses a bisection type of update approach and tackles the nonconvex per-tone problems by a convex approximation. The algorithm performs suboptimally, especially for DSL scenarios with a large amount of crosstalk interference. The computational complexity is only linear in the number of users N.

In [29], the autonomous spectrum balancing (ASB) is proposed, which uses the general solution structure of Figure 1.17. It uses a bisection update approach for the dual master problem and a virtual reference line protection approximation for solving the per-tone dual slave problems. The algorithm performs near-optimally for multiple DSL scenarios, if the reference line is chosen well.

In [38], the successive convex approximation for low-complexity (SCALE) algorithm is proposed. It uses the solution structure of Figure 1.18, where the dual master problem is tackled by a subgradient approach and the per-tone convex dual slave problems by an iterative fixed point update approach as given in (1.49). The convex approximation parameters are determined using the formula (1.48). The SCALE algorithm can however also be viewed from the general solution structure of Figure 1.17, where a subgradient update approach is used for the dual master problem and an iterative convex approximation approach for the nonconvex per-tone dual slave problems. Both approaches for SCALE lead to a locally optimal solution of the cWRS problem. In [38], a scheme is proposed for distributed implementation of the algorithm, i.e. the modems and the SMC perform part of the computations in addition to message-passing between both of them. For further details we refer to [38].

1.7 State-of-the-Art Signal Coordination

Signal coordination is another powerful technique for crosstalk management next to spectrum coordination. It consists of coordinating a set of transmitted and/or received signals in such a way that the crosstalk interference amongst

these signals is removed whilst preserving the signal of interest. The application of signal coordination however requires the colocation of the transmitters and/or the receivers, which is not always the case as was already pointed out in Figure 1.12 of Section 1.5. This excludes for instance the application of signal coordination to mixed CO-RT scenarios.

Signal coordination gives a superior performance with respect to spectrum coordination, as it actually removes the impact of crosstalk, whereas spectrum coordination rather tries to avoid it. However, the required online computational complexity to perform the necessary signal processing operations is also much higher. As such, there is an important trade-off between performance improvement (increased data rates and reach) obtained from applying signal coordination and its corresponding required signal processing computational complexity.

One can distinguish single-sided and two-sided signal coordination. Two-sided signal coordination requires that both sides of the cable bundle are coordinated. The usage of joint linear processing at both sides can then remove the crosstalk interference between the corresponding signals. However, two-sided signal coordination is not common for most practical DSL scenarios since the end-users are usually situated at different locations. Furthermore, it has been shown that single-sided coordination alone can achieve the full theoretical capacity over the corresponding cable bundle transmission.

Single-sided signal coordination techniques are usually distinguished by the particular transmission direction, i.e. upstream or downstream transmission. In the upstream direction, signal coordination can be applied between multiple receivers colocated at the CO or the RT. The received signals are then processed jointly whereas the corresponding transmitters are uncoordinated. This type of channel is also referred to as a multiple access channel (MAC) in information theory. In fact, each individual per-tone transmission can be seen as an individual MAC channel whose sum capacity can be expressed as follows

$$c_k = \log_2 \Big(\det\big(I + H_k S_k H_k^H\big) \Big), \tag{1.50}$$

where $S_k = \mathrm{diag}([\mathrm{s}_\mathrm{k}^1, \ldots, \mathrm{s}_\mathrm{k}^\mathrm{N}])$ which is a diagonal matrix as the transmitters are not coordinated. Receiver signal coordination is also referred to as *crosstalk cancellation*.

Multiple receiver crosstalk cancellation schemes have been proposed. For instance, the minimum mean square error generalized decision feedback equalization (MMSE-GDFE) receiver [78] consists of a feedback-based scheme that successively detects one of the signals and subtracts it from the other received signals so as to remove its crosstalk impact on them. Such a MMSE-GDFE receiver is shown to achieve the sum capacity (1.50). In [79] a QR crosstalk cancellation scheme is proposed that can be seen as a zero forcing GDFE (ZF-GDFE) scheme, and which also employs a feedback-based scheme. In order to

prevent error propagation in these feedback-based schemes, decoding must be done before decisions are fed back, which leads to a high computational complexity and latency. In [80] a linear zero-forcing crosstalk canceller is proposed which significantly reduces the complexity and latency, and does not suffer from error propagation. Furthermore, it is shown that this receiver performs near-optimally in the case of spatially white noise due to the well conditioned structure of the DSL upstream channel matrix.

In the downstream direction, signal coordination can be performed between colocated transmitters. The concept here is to jointly predistort the transmitted signals in such a way that it annihilates with the crosstalk induced during transmission and so that a crosstalk-free signal is received at the uncoordinated receivers. Transmitter signal coordination is also referred to as *crosstalk precoding*.

Multiple crosstalk precoding schemes have been proposed, where some of them consist of feedback-based schemes such as the Tomlinson-Harashima precoder [79], but also linear crosstalk precoders such as the ZF linear precoder [81] and a diagonalizing crosstalk precoder design [82]. Especially, this diagonalizing crosstalk precoder is shown to perform close to optimal due to the well conditioned structure of the DSL downstream channel matrix.

Joint Signal and Spectrum Coordination

One can also combine signal coordination with spectrum coordination, i.e. in addition to applying a crosstalk canceller/precoder one also optimizes the transmit spectra. In [8], water-filling type of procedures are proposed that optimize the transmit spectra for the linear crosstalk canceller [80] and precoder [82].

In [78], a low complexity iterative water-filling procedure for optimizing the transmit spectra is proposed that can achieve the (unweighted) sum capacity of the MAC channel when combined with a MMSE-GDFE receiver. In [83], this is extended to the weighted case where each user's data rate is weighted by a particular weight that emphasizes the relative importance of the corresponding user. Note that the proposed schemes are only optimal if capacity-achieving codes are used. For practical systems, one can use the SNR gap Γ approximation (see also Section 1.2.3), for which the optimality is not guaranteed anymore. For this case, the problem of optimal joint signal and spectrum coordination is still an open problem.

Partial Signal Coordination

Linear crosstalk cancellation or precoding schemes are much less complex with respect to the feedback-based schemes. However their complexity still scales quadratically with the number of coordinated modems N, which results in

an extremely high computational complexity. For example, for a simple linear crosstalk canceller, the coordination of 8 DSL lines transmitting over 4096 tones and at a block rate of 4000 blocks/s, requires more than 1 billion multiplications/s.

In [41, 84], this complexity issue is tackled by a technique referred to as *partial crosstalk cancellation* or *precoding*. The concept here is to tackle only part of the crosstalk impact, instead of trying to remove all crosstalk impact. This is based on the observation that the majority of the crosstalk impact is caused by only a few dominant crosstalkers at a small subset of tones, which is referred to as the *space selectivity* and *frequency selectivity* of crosstalk.

Space selectivity refers to the fact that a particular line in a cable bundle experiences the major part of its crosstalk from a limited set of other lines, such as for instance the neighboring lines as electromagnetic coupling reduces with distance. Furthermore also the near-users in a near-far setting are generally considered as strong crosstalkers. Cancelling only these dominant crosstalkers will remove most of the crosstalk impact.

Frequency selectivity refers to the frequency selective behaviour of crosstalk interference. More specifically, crosstalk coupling typically increases with frequency and can even exhibit nulls in the frequency response.

The space and frequency selectivity of crosstalk results in the fact that most of the impact of a line's crosstalk is caused by a few dominant crosstalkers at a small set of tones. By only tackling these specific crosstalk contributions, one can obtain most of the performance of *full crosstalk cancellation/precoding*, i.e. where all crosstalkers at all tones are removed.

In [41, 84] a set of partial linear crosstalk canceller and precoder designs are proposed that can cancel a specific chosen set of crosstalkers at a specific set of tones, without changing the statistics of the non-cancelled crosstalkers nor the statistics of the noise. In this case, the data rate expression can be given as follows

$$R^{n,\text{partial}} = f_s \sum_{k\in\mathcal{K}} \log_2\left(1 + \frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^n}{\sum_{m\neq n}(1-c_k^{n,m})|h_k^{n,m}|^2 s_k^m + \sigma_k^n}\right) \tag{1.51}$$

with $c_k^{n,m} \in \{0,1\}$, where $c_k^{n,m} = 1$ if user n cancels the crosstalk from user m on tone k, and $c_k^{n,m} = 0$ otherwise. These variables $c_k^{m,n}, m \neq n$, can be seen as complexity resources. In the case that all crosstalk is tackled, i.e. full cancellation, then $\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{N}}\sum_{m\neq n} c_k^{n,m} = C^{\text{full}} = KN(N-1)$. If no crosstalk is cancelled this corresponds to $\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{N}}\sum_{m\neq n} c_k^{n,m} = 0$. Partial crosstalk cancellation linearly scales its computational complexity C^{part} between those two extrema, i.e. $0 \leq \sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{N}}\sum_{m\neq n} c_k^{n,m} = C^{\text{part}} \leq C^{\text{full}}$. For instance, for $C^{\text{total}} = 0.3 \times C^{\text{full}}$, the computational complexity of the

partial crosstalk canceller or precoder is reduced to 30% of that required by the linear crosstalk canceller or precoder.

Given a constraint on the computational complexity, one can optimize the resources $c_k^{n,m}, \forall k, n, m \neq n$, so as to optimize the data rate which can be formulated as follows

$$\begin{array}{ll} \max\limits_{\{c_k^{n,m}, k\in\mathcal{K}, n\in\mathcal{N}, m\neq n\}} & \sum\limits_{n\in\mathcal{N}} R^{n,\text{partial}} \\ \text{s.t.} & \sum\limits_{k\in\mathcal{K}}\sum\limits_{n\in\mathcal{N}}\sum\limits_{m\neq n} c_k^{n,m} \leq C^{\text{budget}} \end{array} \tag{1.52}$$

In [8] several resource optimization algorithms are proposed so as to choose the set of crosstalkers to be tackled, i.e. to choose for which users to cancel which crosstalkers on which tone.

A first algorithm is the *line selection* algorithm. Here, each user cancels the γ largest crosstalkers on each tone, where $\gamma/(N-1)$ is the target reduction in computational complexity.

A second algorithm is the *tone selection* algorithm. Here, each users cancels all crosstalkers on the $\gamma K/(N-1)$ tones for which full cancellation results in the largest bit increase, and cancels no crosstalkers for all other tones. The target complexity reduction is $\gamma/(N-1)$.

Another proposed algorithm is referred to as *simple joint tone-line selection*. For the cable bundle of N lines and K tones, the γNK crosstalkers are cancelled whose cancellation causes the largest bit rate gain in the absence of all other crosstalkers. The target complexity reduction is $\gamma/(N-1)$.

Finally also an optimal algorithm is proposed, which is referred to as *optimum joint tone-line selection*, and consists of a greedy search that succeeds in choosing the optimal set of crosstalkers so as to solve (1.52).

In [8], it has been shown that an optimal allocation of the complexity resources can result in large reductions in computational complexity, while achieving most of the performance gains of full linear signal coordination. For upstream transmission, simulations demonstrate that 90% of the performance, obtained by full linear signal coordination, can be achieved with only 30% of the computational complexity required by applying full signal coordination. For downstream, one achieves a 20%-80% trade-off of computational complexity versus performance (with respect to full linear signal coordination).

1.8 Chapters and Publications Overview

The following publications are included in this thesis:

Chapter 2: Centralized Spectrum Coordination
P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "A Low Complexity Optimal Spectrum Balancing Algorithm for Digital Subscriber Lines", *Signal Processing*, vol. 87, no. 7, Jul. 2007, pp. 1735-1753.

Chapter 3: Distributed Spectrum Coordination
P. Tsiaflakis, M. Diehl, and M. Moonen, "Distributed Spectrum Management Algorithms for Multiuser DSL Networks", *IEEE Transactions on Signal Processing*, vol. 56, no. 10, Oct. 2008, pp. 4825-4843.

Chapter 4: Geometric Programming Based Spectrum Coordination
P. Tsiaflakis, C.W. Tan, Y. Yi, M. Chiang, and M. Moonen, "Optimality Certificate of Dynamic Spectrum Management in Multi-Carrier Interference Channels", in *Proc. of the IEEE International Symposium on Information Theory (ISIT)*, Toronto, Canada, Jul. 2008, pp. 1298-1302.

Chapter 5: Cross-Layer Spectrum Coordination
P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen, "Throughput and Delay of DSL Dynamic Spectrum Management with Dynamic Arrivals", in *Proc. of the IEEE Global Communications Conference (GLOBECOM)*, New Orleans, Louisiana, Nov. 2008, pp. 1-5.

Chapter 6: Green DSL: Energy-Aware Spectrum Coordination
P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen, "Green DSL: Energy-Efficient DSM", in *Proc. of the IEEE International Conference on Communications (ICC)*, Dresden, Germany, Jun. 2009, pp. 1-5.

Chapter 7: General Green DSL: Fair and Energy-Aware Spectrum Coordination
P. Tsiaflakis, Y, Yi, M. Chiang, and M. Moonen, "Fair Greening for DSL Broadband Access", in *Proc. of the GreenMetrics Workshop in conjunction with ACM Sigmetrics/Performance (GreenMetrics)*, Seattle, Washington, Jun. 2009, pp. 1-5.

Chapter 8: Improved Dual Decomposition Based Spectrum Coordination
P. Tsiaflakis, I. Necoara, J.A.K. Suykens, and M. Moonen, "Improved Dual Decomposition Based Optimization for DSL Dynamic Spectrum Management", accepted with mandatory minor revisions for *IEEE Transactions on*

Signal Processing, Sep. 2009.

Chapter 9: Cancellation Resource Optimization for Linear Partial Crosstalk Cancellation
J. Vangorp, P. Tsiaflakis, M. Moonen, J. Verlinden, and G. Ysebaert, "A Dual Decomposition Approach to Partial Crosstalk Cancellation in a Multiuser DMT-xDSL Environment", *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. Article ID 37963, Jun. 2007, pp. 1-11.

Chapter 10: Joint Spectrum and Cancellation Resource Optimization for Linear Partial Crosstalk Cancellation
J. Vangorp, P. Tsiaflakis, M. Moonen, and J. Verlinden, "Joint Spectrum Management and Constrained Partial Crosstalk Cancellation in a Multi-User xDSL Environment", *Signal Processing, Special Section: Information Processing and Data Management in Wireless Sensor Networks*, vol. 87, no. 12, Dec. 2007, pp. 3131-3146.

Chapter 11: Joint Spectrum and Receiver Signal Coordination
P. Tsiaflakis, J. Vangorp, J. Verlinden, and M. Moonen, "Multiple Access Channel Optimal Spectrum Balancing for Upstream DSL Transmission", *IEEE Communications Letters*, vol. 11, no. 4, Apr. 2007, pp. 298-300.

Bibliography

[1] *White paper: Broadband technology overview*, WP6321 ed., Corning, http://www.corning.com, Jun. 2005.

[2] *White paper: Optical Access*, International Engineering Consortium (IEC), http://www.iec.org, 2007.

[3] *White paper: VDSL2, The ideal access technology for delivering video services (Revision 2)*, Aware, Inc., http://www.aware.com, 2006.

[4] F. Vanier, "World broadband statistics: Q1 2009," Point Topic Ltd, http://www.point-topic.com, Tech. Rep., Jun. 2009.

[5] IE Market Research Corp. (http://www.iemarketresearch.com/), "Global DSL Subscribers and Access Revenues Forecast, 2009-2012: Global DSL subscribers to rise to 331 million in 2012 with DSL access revenues reaching $136.4 billion in 2012," Tech. Rep., Feb 2009, internet:http://www.researchandmarkets.com/reports/686876/.

[6] "ADSL2 and ADSL2+ - the new ADSL standards," Aware Inc, Tech. Rep., Mar. 2003.

[7] J. A. C. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Communications Magazine*, vol. 28, no. 5, pp. 5–14, May 1990.

[8] R. Cendrillon, *Multi-user Signal and Spectra Co-ordination for Digital Subscriber Lines.* K.U.Leuven University, Belgium: PhD Thesis, Dec. 2004.

[9] J. M. Cioffi, "EE379C course reader - Advanced digital communication," in *Stanford University. Available at: http://www.stanford.edu/class/ee379c/.*

[10] L. D. Clercq, M. Peeters, S. Schelstraete, and T. Pollet, "Mitigation of radio interference in xDSL transmission," *IEEE Communications Magazine*, vol. 38, no. 3, pp. 168–173, Mar. 2000.

[11] P. Golden, H. Dedieu, and K. S. Jacobsen, *Fundamentals of DSL technology.* Taylor & Francis Group, 2006.

[12] D. G. Mestdagh, M. R. Isaksson, and P. Odling, "Zipper VDSL: a solution for robust duplex communication over telephone lines," *IEEE Communications Magazine*, vol. 38, no. 5, pp. 90–96, May 2000.

[13] *Spectrum Management for Loop Transmission Systems, Issue 2*, ANSI Draft Std. T1E1.4/2003-002R1, 2003.

[14] *Very high speed digital subscriber line (VDSL); functional requirements*, ETSI Std. TS 101 270-1, Rev. V.1.3.1, 2003.

[15] J. Maes, M. Guenach, and M. Peeters, "Statistical channel model for gain quantification of DSL crosstalk mitigation techniques," in *Proc. of the IEEE International Conference on Communications (ICC)*, Jun. 2009.

[16] M. Sorbara, P. Duvaut, F. Shmulyian, S. Singh, and A. Mahadevan, "Construction of a DSL-MIMO channel model for evaluation of FEXT cancellation systems in VDSL2," in *IEEE Sarnoff Symposium*, Apr. 2007, pp. 1–6.

[17] P. Golden, H. Dedieu, and K. S. Jacobsen, Eds., *Implementation and applications of DSL technology.* Auerbach Publications (Taylor & Francis Group), 2008.

[18] S. Jagannathan, *Interference and outage optimization in multi-user multicarrier communication systems.* Stanford University, California, USA: PhD Thesis, Jun. 2008.

[19] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 922–933, May 2006.

[20] R. Cendrillon, M. Moonen, J. Verlinden, T. Bostoen, and W. Yu, "Optimal multi-user spectrum management for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 1, Jun. 2004, pp. 1–5.

[21] R. Cendrillon and M. Moonen, "Iterative spectrum balancing for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 3, May 2005, pp. 1937–1941.

[22] R. Lui and W. Yu, "Low-complexity near-optimal spectrum balancing for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 3, May 2005, pp. 1947–1951.

[23] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[24] K. B. Song, S. T. Chung, G. Ginis, and J. M. Cioffi, "Dynamic spectrum management for next-generation DSL systems," *IEEE Communications Magazine*, vol. 40, no. 10, pp. 101–109, Oct. 2002.

[25] J. M. Cioffi, S. Jagannathan, and W. Lee, "Digital subscriber line (DSL)," *Scholarpedia*, vol. 3, no. 8, p. 3995, 2008.

[26] P. Tsiaflakis, J. Vangorp, M. Moonen, J. Verlinden, and K. V. Acker, "An efficient search algorithm for the lagrange multipliers of optimal spectrum balancing in multi-user xDSL systems," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 4, May 2006.

[27] P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "A low complexity branch and bound approach to optimal spectrum balancing for digital subscriber lines," in *Proc. of the IEEE Global Telecommunications Conference (GLOBECOM)*, Nov. 2006.

[28] ——, "A low complexity optimal spectrum balancing algorithm for digital subscriber lines," *Signal Processing*, vol. 87, no. 7, pp. 1735–1753, Jul. 2007.

[29] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spectrum balancing for digital subscriber lines," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4241–4257, Aug. 2007.

[30] P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "Convex relaxation based low-complexity optimal spectrum balancing for multi-user DSL," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 3, Apr. 2007, pp. III–349 – III–352.

[31] P. Tsiaflakis, M. Diehl, and M. Moonen, "Distributed spectrum management algorithms for multiuser DSL networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4825–4843, Oct. 2008.

[32] P. Tsiaflakis and M. Moonen, "Low-complexity dynamic spectrum management algorithms for digital subscriber lines," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2008, pp. 2769 – 2772.

[33] P. Tsiaflakis, C. Tan, Y. Yi, M. Chiang, and M. Moonen, "Optimality certificate of dynamic spectrum management in multi-carrier interference channels," in *Proc. of the IEEE International Symposium on Information Theory (ISIT)*, Jul. 2008, pp. 1298–1302.

[34] P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen, "Throughput and delay of DSL dynamic spectrum management with dynamic arrivals," in *Proc. of the IEEE Global Telecommunications Conference (GLOBECOM)*, Nov. 2008, pp. 1–5.

[35] ——, "Green DSL: Energy-efficient DSM," in *Proc. of the IEEE International Conference on Communications (ICC)*, Jun. 2009.

[36] ——, "Dynamic spectrum management for green DSL," *submitted*, 2008.

[37] ——, "Fair greening for DSL broadband access," in *Proc. of the GreenMetrics Workshop in conjunction with ACM Sigmetrics/Performance*, Seattle, Washington, Jun. 2009, pp. 1–5.

[38] J. Papandriopoulos and J. S. Evans, "Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 7, Jun. 2006, pp. 3270–3275.

[39] P. Tsiaflakis, I. Necoara, J. A. K. Suykens, and M. Moonen, "An improved dual decomposition approach to DSL dynamic spectrum management," in *Proc. of the 17th European Signal Processing Conference (EUSIPCO)*, Glasgow, Scotland, Aug. 2009.

[40] ——, "Improved dual decomposition based optimization for DSL dynamic spectrum management," *accepted with mandatory minor revisions in IEEE Transactions on Signal Processing*, 2009.

[41] R. Cendrillon, M. Moonen, G. Ginis, K. V. Acker, T. Bostoen, and P. Vandaele, "Partial crosstalk cancellation for upstream VDSL," *EURASIP Journal on Applied Signal Processing*, no. 10, pp. 1520–1535, Aug. 2004.

[42] J. Vangorp, P. Tsiaflakis, M. Moonen, J. Verlinden, and G. Ysebaert, "A dual decomposition approach to partial crosstalk cancellation in a multiuser DMT-xDSL environment," *EURASIP Journal on Advances in Signal Processing*, pp. 1–11, Jun. 2007.

[43] P. Tsiaflakis, J. Vangorp, M. Moonen, J. Verlinden, and G. Ysebaert, "Partial crosstalk cancellation in a multi-user xDSL environment," in *Proc. of the IEEE International Conference on Communications (ICC)*, Jun. 2006.

[44] J. Vangorp, P. Tsiaflakis, M. Moonen, and J. Verlinden, "Joint spectrum management and constrained partial crosstalk cancellation in a multi-user xDSL environment," *Signal Processing*, vol. 87, no. 12, pp. 3131–3146, Dec. 2007.

[45] ——, "Joint spectrum management and constrained partial crosstalk cancellation in a multi-user xDSL environment," in *Proc. of the European Signal Processing Conference (EUSIPCO)*, Sep. 2006.

[46] J. Vangorp, P. Tsiaflakis, M. Moonen, J. Verlinden, and K. V. Acker, "Optimal spectrum balancing in multi-user xDSL systems with ON/OFF power loading," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2006, pp. IV 621–624.

[47] P. Tsiaflakis, J. Vangorp, and M. Moonen, "Multiple access channel optimal spectrum balancing for upstream DSL transmission," *IEEE Communications Letters*, vol. 11, no. 4, pp. 298–300, Apr. 2007.

[48] J. M. Cioffi and M. Mohseni, *Dynamic Spectrum Management - A methodology for providing significantly higher broadband capacity to the users.* Telektronikk, 2004, vol. 4.2004, ch. 3.

[49] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley, Jul. 2006.

[50] A. Carleial, "A case where interference does not reduce capacity (corresp.)," *IEEE Transactions on Information Theory*, vol. 21, no. 5, pp. 569–570, Sep. 1975.

[51] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 49–60, Jan 1981.

[52] H. Sato, "The capacity of the Gaussian interference channel under strong interference (corresp.)," *IEEE Transactions on Information Theory*, vol. 27, no. 6, pp. 786–788, Nov. 1981.

[53] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: Sum capacity in the low interference regime and new outer bounds on the capacity region," *CoRR*, vol. abs/0802.3495, 2008.

[54] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sumrate capacity for Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 689–699, Feb. 2009.

[55] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 620–643, Feb. 2009.

[56] V. M. K. Chan and W. Yu, "Multiuser spectrum optimization for discrete multitone systems with asynchronous crosstalk," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5425–5435, Nov. 2007.

[57] W. Yu, R. Lui, and R. Cendrillon, "Dual optimization methods for multiuser orthogonal frequency-division multiplex systems," in *Proc. of the IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 1, Dec. 2004, pp. 225–229.

[58] Z. Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, Feb. 2008.

[59] T. Sartenaer, L. Vandendorpe, and J. Louveaux, "Balanced capacity of wireline multiuser channels," *IEEE Transactions on Communications*, vol. 53, no. 12, pp. 2029–2042, Dec. 2005.

[60] A. Kalakech, J. Louveaux, and L. Vandendorpe, "A balanced capacity implementation of DSM for DSL systems," in *Proc. of the Symposium on Communications and Vehicular Technology*, Nov. 2006, pp. 44–48.

[61] ——, "Applying the balanced capacity concept to DSL systems," in *Proc. of the IEEE International Conference on Communications*, Jun. 2007, pp. 2717–2721.

[62] M. Wolkerstorfer, D. Statovci, and T. Nordstrom, "Dynamic spectrum management for energy-efficient transmission in DSL," in *Proc. of the 11th IEEE Singapore International Conference on Communication Systems (ICCS)*, Nov. 2008, pp. 1015–1020.

[63] M. Monteiro, N. Lindqvist, and A. Klautau, "Spectrum balancing algorithms for power minimization in dsl networks," in *Proc. of the IEEE International Conference on Communications (ICC)*, Jun. 2009, pp. 1–5.

[64] K. Jacobsen, "Methods of upstream power back-off on very high-speed digital subscriber lines," *IEEE Communications Magazine*, pp. 210–216, Mar. 2001.

[65] V. Oksman, "Optimization of the psd ref for upstream power back-off in VDSL," ANSI Std. Contrib. T1E1.4/2001-102, Tech. Rep., 2001.

[66] S. Schelstraete and V. Oksman, "Proposal for the north-american reference psd," ANSI Std. Contrib. T1E1.4/2001-157, Tech. Rep., 2001.

[67] F. Sjoberg, S. Wilson, and M. Isaksson, "Power back-off in the upstream of VDSL," in *Proc. of the Radiovetenskaplig Koferens (RVK'99)*, vol. 1, Karlskrona, Sweden, Jun. 1999, pp. 436–440.

[68] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1310–1322, Jul 2006.

[69] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge Univ. Press, 2003.

[70] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Athena Scientific, 1999.

[71] "Physical layer management for digital subscriber line (DSL) transceivers," ITU Std. G.997.1, Tech. Rep., 2003.

[72] Y. Xu, S. Panigrahi, and T. Le-Ngoc, "A concave minimization approach to dynamic spectrum management for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 1, Jun. 2006, pp. 84–89.

[73] Y. Xu, T. Le-Ngoc, and S. Panigrahi, "Global concave minimization for optimal spectrum balancing in multi-user DSL networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 2875–2885, Jul. 2008.

[74] W. Lee, Y. Kim, M. H. Brady, and J. M. Cioffi, "Band-preference dynamic spectrum management in a DSL environment," in *Proc. of the IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2006, pp. 1–5.

[75] M. Chiang, *Geometric programming for communication systems*. Foundations and Trends in Communications and Information Theory, Aug. 2005, vol. 2, no. 1.

[76] S. Boyd, S. J. Kim, L. Vandenberghe, and A. Hassibi, "A tutorial on geometric programming," *Optimization and Engineering*, vol. 8, no. 1, pp. 67–127, 2007.

[77] M. Chiang, C. Tan, D. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.

[78] W. Yu and J. M. Cioffi, "Multiuser detection for vector multiple access channels using generalized decision feedback equalization," in *Proc. of the 5th International Conference on Signal Processing*, vol. 3, 2000, pp. 1771–1777.

[79] G. Ginis and J. M. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE Journal on Selected Areas on Communications*, vol. 20, no. 5, pp. 1085–1104, Jun. 2002.

[80] R. Cendrillon, G. Ginis, E. V. den Bogaert, and M. Moonen, "A near-optimal linear crosstalk canceler for upstream VDSL," *IEEE Transactions on Signal Processing*, vol. 54, no. 8, pp. 3136–3146, Aug. 2006.

[81] M. Tsatsanis, *DSL Handbook*. Auerbach, 2004.

[82] R. Cendrillon, G. Ginis, E. V. den Bogaert, and M. Moonen, "A near-optimal linear crosstalk precoder for downstream VDSL," *IEEE Transactions on Communications*, vol. 55, no. 5, pp. 860–863, May 2007.

[83] M. Mohseni, R. Zhang, and J. M. Cioffi, "Optimized transmission for fading multiple-access and broadcast channels with multiple antennas," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1627–1639, Aug. 2006.

[84] R. Cendrillon, G. Ginis, M. Moonen, and K. V. Acker, "Partial crosstalk precompensation in downstream VDSL," *Signal Processing*, vol. 84, no. 11, pp. 2005–2019, 2004.

Chapter 2

Centralized Spectrum Coordination

A Low Complexity Optimal Spectrum Balancing Algorithm for Digital Subscriber Lines

P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden

Published in *Signal Processing*, vol. 87, no. 7, Jul. 2007,
pp. 1735-1753, doi:10.1016/j.sigpro.2007.01.020

Copyright ©2007 Elsevier B.V. All rights reserved.

Contributions of first author

- literature study
- co-development of Lagrange multiplier search algorithm (single-user and multi-user case)
- development of branch and bound procedure for per-tone non-convex problems
- co-analysis of computational complexity Lagrange multiplier search algorithm
- analysis of computational complexity branch and bound procedure
- design of simulation scenarios
- software co-implementation and computer simulations
- co-interpretation of simulation results
- text co-redaction, co-editing and revision

Abstract

In modern DSL systems, multi-user crosstalk is a major source of performance degradation. Optimal Spectrum Balancing (OSB) is a centralized algorithm that mitigates the effect of crosstalk by allocating optimal transmit spectra to all interfering DSL modems. By the use of Lagrange multipliers the algorithm decouples the spectrum management problem into per-tone optimization problems. The remaining issues are then finding the Lagrange multipliers that enforce the constraints and solving the per-tone optimization problems. Finding the optimal Lagrange multipliers can become complex when more than two users are considered. Starting from the single-user case, this paper presents a number of properties, which are then extended to the multi-user case and lead to an efficient search algorithm for the Lagrange multipliers. Simulations show that the number of Lagrange multiplier evaluations is as small as 20-50, independent of the number of users. Secondly, the complexity of the per-tone optimization problems grows exponentially with the number of lines in the binder. For multiple user scenarios this becomes computationally intractable. This paper presents an efficient branch and bound approach for the per-tone optimization problem. Simulations show enormous complexity reductions, especially for a large number of users.

2.1 Introduction

The ever increasing demand for higher data rates forces DSL systems to use higher frequencies, e.g. up to 30 MHz for VDSL2 [1]. At these frequencies, electromagnetic coupling becomes particularly harmful and causes crosstalk between systems operating in the same bundle. This crosstalk, typically 10-15 dB larger than the background noise, is a dominant source of performance degradation in DSL systems currently under development [2].

There are two strategies for dealing with this crosstalk: crosstalk cancellation and spectrum management. Crosstalk cancellation can remove the crosstalk completely with minimal noise enhancement [3][4][5], but requires signal level cooperation between receivers or transmitters. Unfortunately, in an unbundled scenario this cooperation is not available. In this case crosstalk can be mitigated through the use of spectrum management.

Current DSL systems use a Static Spectrum Management (SSM) approach where fixed spectral masks ensure that crosstalk levels remain within an acceptable range [6][7]. Because these spectral masks are designed for worst case loop characteristics, this approach can be extremely suboptimal. Dynamic Spectrum Management (DSM) overcomes this problem by designing the transmit spectrum of each modem according to the topology of the network. In this

way spectra take into account the current requirements of all users, causing as little disturbance as possible.

One of the first DSM algorithms proposed is Iterative Waterfilling (IW) [8]. In this algorithm, each user waterfills its spectrum against the noise and interference. By repeating this in an iterative fashion over the different users, the procedure converges to a "selfish" optimum. IW is a low complexity distributed algorithm, meaning it does not need any form of centralized control. Although IW significantly outperforms SSM, it is found to be suboptimal. This is especially so in heavily unbalanced scenarios, where some lines cause much more crosstalk than others (e.g. near-far scenario).

The Optimal Spectrum Balancing (OSB) algorithm [9][10] provides a computationally tractable way to calculate optimal transmit spectra. By optimizing a weighted rate sum, this algorithm can make every possible trade off between the rates of different users. The damage done to other modems in the network is taken into account explicitly, avoiding the "selfish" optimum and thereby improving on the performance of IW. However, this can only be done when complete information about the channel is available (direct channels as well as crosstalk channels), making OSB only suitable with centralized control in a Spectrum Management Center (SMC).

OSB uses a Lagrange multiplier approach to decouple the spectrum management problem into K per-tone optimization problems, where K is the number of tones. The remaining issues are then finding the Lagrange multipliers that enforce the constraints and solving the non-convex per-tone optimization problems.

Finding the optimal Lagrange multipliers can become complex when more than two users are considered. Starting from the single-user case, this paper presents a number of properties, which are then extended to the multi-user case and lead to an efficient search algorithm for the Lagrange multipliers.

The per-tone optimization problems are non-convex with many local optima. Hence, an exhaustive search was proposed to find the global optimum. Unfortunately the set size of feasible solutions is exponential in the number of users N. This becomes computationally intractable for more than four users. To cope with this complexity, research has been focussing on reducing the complexity without sacrificing performance. In [11][12] the authors present an iterative approach to solve the per-tone optimization problems. It is claimed that these algorithms are computationally tractable for multiple users and lead to near-optimal performance. Unfortunately, it is not possible to compare their performance with the optimal performance for scenarios with more than four users.

In this paper a branch and bound approach is presented to solve the per-tone

optimization problem. Branch and bound [13][14] is a general method for finding optimal solutions of various optimization problems. Branch and bound operations will be proposed which require a limited amount of computation, keeping the total computational complexity low. The complexity of the proposed branch and bound procedure will be compared with the complexity of the exhaustive search.

The paper is organized as follows. Section 2.2 describes the system model for the crosstalk environment. In Section 2.3, different viewpoints on the spectrum management problem are discussed. It is shown that the dual problem decouples the primal spectrum management problem into many per-tone problems. Finally, it is discussed how to solve the dual problem. Section 2.4 starts with a discussion of the single-user spectrum management problem from a dual decomposition viewpoint. Some interesting properties and relations are derived. These relations are then extended to the multi-user case. The major contribution of this section is the derivation of the Lagrange multiplier update strategy based on these relations. This approach gives a certain freedom to define an intelligent step size selection strategy leading to a fast algorithm to find the optimal Lagrange multipliers. Section 2.5 presents the low complexity branch and bound approach for solving the per-tone optimization problem. Finally, in Section 2.6 simulations results are shown for the proposed OSB solution and the overall complexity reduction is evaluated.

2.2 System Model

Most current DSL systems use Discrete Multi-Tone (DMT) modulation. The available frequency band is divided in a number of parallel subchannels or tones. We assume that the modems in the network are synchronized. Each tone is capable of transmitting data independently from other tones, and so the transmit power and the number of bits can be assigned individually for each tone. This gives a large flexibility in optimally shaping the transmit spectrum.

Transmission for a binder of N users can be modelled on each tone k by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k \qquad k = 1 \ldots K.$$

The vector $\mathbf{x}_k = [x_k^1, x_k^2, \ldots, x_k^N]^T$ contains the transmitted signals on tone k for all N users. $[\mathbf{H}_k]_{n,m} = h_k^{n,m}$ is an $N \times N$ matrix containing the channel transfer functions from transmitter m to receiver n. The diagonal elements are the direct channels, the off-diagonal elements are the crosstalk channels. $\mathbf{z}_k$ is the vector of additive noise on tone k, containing thermal noise, alien crosstalk, RFI, etc. The vector $\mathbf{y}_k$ contains the received symbols.

We denote the transmit power as $s_k^n \triangleq \Delta_f E\{|x_k^n|^2\}$, the noise power as $\sigma_k^n \triangleq \Delta_f E\{|z_k^n|^2\}$. The vector containing the transmit power of user n on all tones

is $\mathbf{s}^n \triangleq [s_1^n, s_2^n, \ldots, s_K^n]^T$. The DMT symbol rate is denoted as f_s, the tone spacing as Δ_f.

It is assumed that each modem treats interference from other modems as noise. When the number of interfering modems is large, the interference is well approximated by a Gaussian distribution. Under this assumption the achievable bit loading of user n on tone k, given the transmit spectra $\mathbf{s}_k \triangleq [s_k^1, s_k^2, \ldots, s_k^N]^T$ of all modems in the system, is

$$b_k^n \triangleq \log_2\left(1 + \frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^n}{\sum_{m\neq n}|h_k^{n,m}|^2 s_k^m + \sigma_k^n}\right), \tag{2.1}$$

where Γ denotes the SNR-gap to capacity, which is a function of the desired BER, the coding gain and noise margin. The data rate for user n is

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

Conversely, when the bit loading $\mathbf{b}_k \triangleq [b_k^1, b_k^2, \ldots, b_k^N]^T$ of all the users is given for a specific tone k, the required transmit power for the modems in the system can be calculated by [10]

$$\mathbf{s}_k = \left(\mathbf{D}_k - \mathbf{\Lambda}_k \mathbf{A}_k\right)^{-1} \mathbf{\Lambda}_k \sigma_k \tag{2.2}$$

$$\begin{aligned}
\mathbf{D}_k &\triangleq \operatorname{diag}\{|h_k^{1,1}|^2, |h_k^{2,2}|^2, \ldots, |h_k^{N,N}|^2\} \\
\mathbf{\Lambda}_k &\triangleq \operatorname{diag}\{2^{b_k^1} - 1, 2^{b_k^2} - 1, \ldots, 2^{b_k^N} - 1\} \\
[\mathbf{A}_k]_{n,m} &\triangleq a_k^{n,m} \quad \text{with} \quad a_k^{n,m} \triangleq \begin{cases} 0 & n = m \\ \Gamma|h_k^{n,m}|^2 & n \neq m \end{cases} \\
\sigma_k &\triangleq \Gamma[\sigma_k^1, \sigma_k^2, \ldots, \sigma_k^N]^T
\end{aligned}$$

The total power used by user n is then

$$P^n = \sum_k s_k^n.$$

Note that, based on Formulae (2.1) and (2.2), every $\mathbf{s}_k$ corresponds to a $\mathbf{b}_k$, and vice versa. Finding optimal transmit spectra $\mathbf{s}_k$ (referred to as "power loading") is therefore equivalent to finding optimal bit loads $\mathbf{b}_k$ (referred to as "bit loading"). The Optimal Spectrum Balancing problem (Section 2.3) is mostly presented as a power loading problem. The branch and bound approach in Section 2.5 will be presented as a bit loading problem.

2.3 Optimal Spectrum Balancing

2.3.1 The Spectrum Management Problem

The spectrum management problem amounts to finding optimal transmit spectra for a bundle of interfering DSL lines, following a certain criterion and subject to a number of constraints.

First of all, there is a *total power constraint* $P^{n,tot}$ for each user. This constraint ensures the user's total power does not exceed the maximum allowed total transmit power. On top of this constraint there can be a *spectral mask constraint* $s_k^{n,\text{mask}}$ for each tone to guarantee electromagnetic compatibility with other systems.

A second type of constraint is a *rate constraint* for each user. Typically service providers offer a number of profiles and guarantee a certain Quality of Service. The rate constraint then indicates a minimum data rate required by the user.

The spectrum management problem can be viewed from different angles, each time leading to a different criterion that is then to be optimized by the spectrum management algorithm. Either rate, margin or power of the users can be optimized. An optimal solution has to be found within the domain set out by the various constraints.

- In *rate adaptive mode*, the spectrum management problem is to maximize the sum of the data rates of the users. This will be done by using the available power to load a maximum number of bits on tones. The rate is thus limited by the total power and spectral mask constraints. It is possible that there are many solutions to this problem. Several trade-offs between the rates of individual users may exist, all resulting in the same total sum rate. Rate constraints can be used to select a solution with reasonable rates for all users.

$$\begin{array}{lll} \max\limits_{\mathbf{s}^1,\ldots,\mathbf{s}^N} & \sum\limits_{n=1}^{N} R^n & \\ \text{s.t.} & P^n \leq P^{n,tot} & n = 1 \ldots N \\ & 0 \leq s_k^n \leq s_k^{n,\text{mask}} & n = 1 \ldots N, k = 1 \ldots K \\ & R^n \geq R^{n,target} & n = 1 \ldots N \end{array} \tag{2.3}$$

  In this optimization problem, the first set of constraints indicate total power constraints per user. The second set of constraints are spectral mask constraints and the third set are rate constraints per user.
- In *power adaptive mode*, the spectrum management problem is to minimize the total power needed by all users and still meet the rate constraints. This has to be done without violating the spectral mask con-

straints and total power constraints.

$$\begin{array}{lll} \min\limits_{\mathbf{s}^1,\ldots,\mathbf{s}^N} & \sum\limits_{n=1}^{N} P^n & \\ \text{s.t.} & P^n \leq P^{n,tot} & n = 1 \ldots N \\ & 0 \leq s_k^n \leq s_k^{n,\text{mask}} & n = 1 \ldots N, k = 1 \ldots K \\ & R^n \geq R^{n,target} & n = 1 \ldots N \end{array} \tag{2.4}$$

Total power constraints are represented in the first set of constraints, the second set of constraints are spectral mask constraints and the third set are rate constraints.

- In *margin adaptive mode*, the spectrum management algorithm will use the available power to tightly satisfy the rate constraints. In this mode, the noise margin is maximized, thus minimizing the bit error rate for the requested data rate. In single-user mode, this can be achieved by first using the spectrum management algorithm in power adaptive mode and then assigning all remaining power over the used tones by scaling with a constant factor, while not violating spectral mask constraints. In the multi-user case however, this strategy should be revised because adding power to a user causes more crosstalk to other users.

2.3.2 Dual Decomposition

We will now focus on the *rate adaptive* formulation of the spectrum management problem. The power adaptive form can be treated in a similar way.

The rate adaptive optimization problem (2.3) is a non-convex problem and therefore difficult to solve. To find the global optimum one must exhaustively search through all possible transmit spectra. This leads to an exponential complexity in both the number of users and tones, namely $\mathcal{O}(B^{NK})$ where B is the number of possibilities for the bit or power loading for each tone and each user.

With $K = 256$ in ADSL and $K = 4096$ in VDSL, exhaustively searching all possible transmit spectra is seen to be computationally intractable. The reason behind this exponential complexity in the number of tones is that the total power constraints and rate constraints are coupled across tones. Therefore transmit spectra have to be searched jointly across tones.

In [9][10] it was shown that this complexity can be reduced by focussing on the dual problem to decouple the optimization problem across tones. By using Lagrange multipliers to move constraints coupled over tones into the unconstrained part of the optimization problem, the spectrum management problem can be solved in a per-tone fashion. By choosing appropriate values for the Lagrange multipliers, the constraints can still be enforced.

Following the approach of [15] we first formulate the rate adaptive spectrum management problem (2.3) in a slightly different way. Instead of optimizing the sum rate of the users in a general fashion, while trying to satisfy individual rate constraints, we now enforce a fixed ratio between the rates for all users. Solutions to this problem are more restricted than solutions to (2.3) because now rates in excess of the individual rate constraints cannot be assigned randomly but are divided over the users in proportion to their rate constraints. The resulting spectrum management problem is the maximization of the so-called base rate R while satisfying total power and spectral mask constraints. This set of solutions is further restricted by a rate constraint for the users expressed as a fixed proportion β^n of the base rate. The parameters β^n are given fixed numbers and depend on the importance that we wish to put on the users with respect to each other:

$$\begin{array}{lll} \max\limits_{\mathbf{s}^1,\ldots,\mathbf{s}^N,R} & R & \\ \text{s.t.} & P^n \leq P^{n,tot} & n = 1 \ldots N \\ & 0 \leq s_k^n \leq s_k^{n,\text{mask}} & n = 1 \ldots N, k = 1 \ldots K \\ & R^n \geq \beta^n R & n = 1 \ldots N \end{array} \tag{2.5}$$

Within the framework of [16], optimization problem (2.5) can be referred to as the primal problem. The dual problem corresponding with this primal problem can be formulated as follows

$$\begin{array}{ll} \min\limits_{\boldsymbol{\omega},\boldsymbol{\lambda}} & g(\boldsymbol{\omega},\boldsymbol{\lambda}) \\ \text{s.t.} & \lambda_n \geq 0, \omega_n \geq 0, \quad n = 1 \ldots N \end{array} \tag{2.6}$$

with

$$\begin{array}{rll} g(\boldsymbol{\omega},\boldsymbol{\lambda}) := & \max\limits_{\mathbf{s}^1,\ldots,\mathbf{s}^N,R} & J(\mathbf{s}^1,\ldots,\mathbf{s}^N,R,\boldsymbol{\omega},\boldsymbol{\lambda}) \\ & \text{s.t.} \qquad 0 \leq s_k^n \leq s_k^{n,\text{mask}} & n = 1 \ldots N, k = 1 \ldots K. \end{array} \tag{2.7}$$

where J is called the Lagrangian and can be formulated as follows

$$\begin{aligned} J &= R + \sum_{n=1}^{N} \omega_n \big(R^n - \beta^n R\big) + \sum_{n=1}^{N} \lambda_n \big(P^{n,tot} - \sum_{k=1}^{K} s_k^n\big) \\ &= \big(1 - \sum_{n=1}^{N} \omega_n \beta^n\big) R + \sum_{n=1}^{N} \omega_n R^n + \sum_{n=1}^{N} \lambda_n \big(P^{n,tot} - \sum_{k=1}^{K} s_k^n\big) \end{aligned} \tag{2.8}$$

If $\big(1 - \sum_n \omega_n \beta^n\big) > 0$ the maximization over R results in $R = \infty$. If $\big(1 - \sum_n \omega_n \beta^n\big) < 0$ the maximization over R results in $R = 0$. There only exists a non-trivial solution if

$$\big(1 - \sum_n \omega_n \beta^n\big) = 0. \tag{2.9}$$

This results in the following Lagrangian:

$$J = \sum_{n=1}^{N} \omega_n R^n + \sum_{n=1}^{N} \lambda_n \big(P^{n,tot} - \sum_{k=1}^{K} s_k^n\big). \tag{2.10}$$

This Lagrangian is decoupled across the tones:

$$\begin{aligned} J &= \sum_{k=1}^{K} \Bigg(\underbrace{\sum_{n=1}^{N} \omega_n f_s b_k^n - \sum_{n=1}^{N} \lambda_n s_k^n}_{J_k} \Bigg) + \underbrace{\sum_{n=1}^{N} \lambda_n P^{n,tot}}_{\text{constant}} \\ &= \sum_{k=1}^{K} J_k + \text{constant}. \end{aligned}$$

The constant has no influence on the maximization and can be discarded. Then (for a particular choice of $\lambda_n, \omega_n, n = 1 \dots N$) the problem is reduced to a maximization of a sum across tones, which is equal to the sum of independent maximizations. This is referred to as the dual decomposition into independent per-tone optimization problems. The original complexity of $\mathcal{O}(B^{NK})$, exponential in K, is now reduced to a linear complexity in K: $\mathcal{O}(KB^N)$.

In a three-user VDSL system with 14 possible bit loadings per tone, the complexity is reduced from one maximization over $14^{3\times 4096}$ possibilities to 4096 maximizations over 14^3 possibilities. This is a spectacular reduction in complexity.

Solving the dual problem (see also Section 2.3.3) does not necessarily correspond to solving the original constrained problem. Optimization theory [16] states that the solution to the dual problem provides only an upper bound to the solution of the original problem. The difference between the upper bound and the optimum of the primal problem is called the duality gap. In [15] it is shown that for multi-carrier systems like DMT that satisfy certain conditions, the duality gap is zero. This means the solution to the dual problem is also the solution to the original problem. In practice, it is found that, while the necessary conditions are not strictly satisfied, the dual problem formulation leads to adequate solutions, and so is currently the method of choice.

2.3.3 Solving the Dual Problem

In the previous section the spectrum management problem defined by the primal problem (2.5) was transformed into following dual problem:

$$\begin{array}{ll} \min\limits_{\boldsymbol{\omega},\boldsymbol{\lambda}} & g(\boldsymbol{\omega},\boldsymbol{\lambda}) \\ \text{s.t.} & \lambda_n \geq 0, \omega_n \geq 0, \quad n = 1 \dots N \\ & \big(1 - \sum_n \omega_n \beta^n\big) = 0 \end{array} \tag{2.11}$$

with

$$g(\boldsymbol{\omega},\boldsymbol{\lambda}) := \sum_{k=1}^{K} \underbrace{\max_{\mathbf{s}_k^1,\ldots,\mathbf{s}_k^N} \sum_{n=1}^{N} \omega_n f_s b_k^n - \sum_{n=1}^{N} \lambda_n s_k^n}_{\text{per-tone optimization problem}} \tag{2.12}$$

$$\text{s.t. } 0 \le s_k^n \le s_k^{n,\text{mask}}, \qquad n = 1\ldots N, k = 1\ldots K.$$

Basically the dual problem consists of two optimization problems. First, the optimal Lagrange multipliers ω_n and $\lambda_n (n = 1 \ldots N)$ need to be found such that the total power constraints and the data rate constraints of (2.5) are satisfied. Secondly, for given $\omega_n, \lambda_n (n = 1 \ldots N)$ K independent nonconvex per-tone optimization problems need to be solved to obtain the optimal transmit spectra for all users on all tones.

Given $\omega_n, \lambda_n (n = 1 \ldots N)$ the nonconvex per-tone optimization problem can be solved by performing an exhaustive search over all possible bit or power loading combinations for the users. This results in transmit spectra for all users. For random $\omega_n{}'$s and $\lambda_n{}'$s, the power and rate constraints are generally not satisfied. By choosing appropriate values for the Lagrange multipliers, these constraints can be enforced. Looking at (2.12), it can be seen that the $\lambda_n{}'$s influence the resulting spectra. A larger λ_n for user n results in a larger penalty in the cost function when power is allocated to s_k^n. Therefore the $\lambda_n{}'$s can be viewed as setting a cost for power. The $\omega_n{}'$s have a similar intuitive interpretation. A larger ω_n for user n results in an increased importance attached to its rate. The larger ω_n, the higher the rate allocated to user n compared to other users.

In [10] it is explained that in order to solve the primal problem the Lagrange multipliers $\lambda_n (n = 1 \ldots N)$ need to be found such that the total power constraints are tight ($P^n = P^{n,tot}$) or λ_n is equal to zero (in the case of a non-active power constraint $P^n \le P^{n,tot}$) for each n. The $\omega_n{}'$s will be treated more loosely. The last constraint of (2.11) is related with the last constraint of (2.5) which enforces a fixed proportion β^n to the different users n. This constraint is not so important in our application and we will mainly use the $\omega_n{}'$s to put more importance on some lines. The last constraint of (2.11) is thus removed keeping in mind that the $\omega_n{}'$s can be chosen to satisfy data rate constraints like in (2.3). Thus, the $\omega_n{}'$s can be seen as weights given to the different users.

Starting from the single-user case, Section 2.4 derives a number of properties of the Lagrange multipliers based on simple mathematics. These properties are then extended to the multi-user case and lead to a very efficient and intuitive search algorithm for the Lagrange multipliers. Section 2.5 focuses on the complexity of the discrete per-tone optimization problem. As the dimensionality of the solution set is exponential in the number of users, an exhaustive search

becomes computationally intractable for more than four users. A branch and bound approach will be presented to reduce this complexity significantly.

2.4 An Efficient Lagrange Multiplier Search Algorithm

2.4.1 Single-user Spectrum Management

In this section, the single-user spectrum management problem will be discussed within the framework provided by the dual decomposition method. For the single-user case, the spectrum management problem can be formulated as follows:

$$\begin{array}{ll} \max\limits_{\mathbf{s}} & R \\ \text{s.t.} & \sum_k s_k \leq P^{tot} \\ & 0 \leq s_k \leq s_k^{\text{mask}} \quad k = 1 \ldots K \end{array} \tag{2.13}$$

where now superscript '1' (for user 1) has been omitted. Using a Lagrange multiplier λ to incorporate the total power constraint, the dual problem is formulated as

$$\begin{array}{ll} \max\limits_{\mathbf{s}} & R + \lambda(P^{tot} - \sum\limits_k s_k) \\ \text{s.t.} & 0 \leq s_k \leq s_k^{\text{mask}} \qquad k = 1 \ldots K \end{array} \tag{2.14}$$

Recalling that λ represents a cost for power, larger λ's will result in less power being used, e.g. $\lambda = \infty$ leads to $P = \sum_k s_k = 0$ and thus $R = 0$. For any other λ, the optimal spectrum can be found through a per-tone exhaustive search over all possible bit or power loadings.

This relation can be represented on the power axis shown in Figure 2.1. This axis shows the total power obtained when optimizing the dual problem for a particular λ. If for a particular λ a total power and rate (P^λ, R^λ) is obtained, then optimality of this solution implies that a total power P less than P^λ, i.e. $P < P^\lambda$ must correspond to a rate R smaller than R^λ, i.e. $R < R^\lambda$. This then ensures that if a λ is found that makes the power constraint tight, the primal problem is solved.



Figure 2.1: Power axis

Secondly, it is easily proven that if λ is decreased, then both the achieved bit rate and the consumed power for the optimal solution do not decrease, i.e

$$\begin{aligned} P^{\lambda_B} &\geq P^{\lambda_A} \quad \text{if} \quad \lambda_B < \lambda_A \\ R^{\lambda_B} &\geq R^{\lambda_A} \quad \text{if} \quad \lambda_B < \lambda_A \end{aligned}$$

where $P^{\lambda_A}, P^{\lambda_B}$ and $R^{\lambda_A}, R^{\lambda_B}$ are the optimal total power and rate when $\lambda = \lambda_A$ and $\lambda = \lambda_B$ respectively. This can be proven as follows. Optimality of $(P^{\lambda_A}, R^{\lambda_A})$ for λ_A implies that

$$R^{\lambda_B} - \lambda_A P^{\lambda_B} \leq R^{\lambda_A} - \lambda_A P^{\lambda_A}, \tag{2.15}$$

because the right-hand side indeed uses the spectra that maximize the Lagrangian for λ_A. Any other spectra, e.g. those corresponding to $(P^{\lambda_B}, R^{\lambda_B})$, result in a smaller value for the Lagrangian. A similar statement can be made on the optimality of $(P^{\lambda_B}, R^{\lambda_B})$ for λ_B:

$$R^{\lambda_A} - \lambda_B P^{\lambda_A} \leq R^{\lambda_B} - \lambda_B P^{\lambda_B}. \tag{2.16}$$

Taking the sum of (2.15) and (2.16) results in

$$\underbrace{(\lambda_B - \lambda_A)}_{\Delta\lambda}\underbrace{(P^{\lambda_B} - P^{\lambda_A})}_{\Delta P} \leq 0. \tag{2.17}$$

Assuming $\lambda_B < \lambda_A$, (2.17) implies

$$P^{\lambda_B} \geq P^{\lambda_A}. \tag{2.18}$$

Then by using this in (2.16) we obtain

$$R^{\lambda_B} \geq R^{\lambda_A}. \tag{2.19}$$

Relation (2.17) can be used to construct a simple procedure to find the λ that makes the total power constraint tight. By starting with a random λ, e.g. $\lambda = 0$, we maximize the Lagrangian to obtain a bit and power loading. If the total power then exceeds the total power constraint, λ has to increase ($\Delta\lambda > 0$ so that $\Delta \mathrm{P} \leq 0$). When the total power is below the total power constraint, λ has to decrease ($\Delta\lambda < 0$ so that $\Delta \mathrm{P} \geq 0$). Note that λ has to remain positive. If by decreasing λ we end up with $\lambda = 0$, the total power constraint is effectively made inactive by some other per-tone power constraint, e.g. spectral mask constraints. An update formula for λ is then as follows

$$\Delta\lambda = \mu\Big(\sum_k s_k - P^{tot}\Big) \quad \Rightarrow \quad \lambda^{t+1} = \left[\lambda^t + \mu\Big(\sum_k s_k - P^{tot}\Big)\right]^+, \tag{2.20}$$

where $[x]^+$ means $\max(0, x)$. By varying the step size μ according to Algorithm 2.1 the λ can be found which makes the power constraint tight.

This procedure is illustrated in Figure 2.2. Starting from some initial λ (e.g. $\lambda = 0$), μ is always doubled, creating a trajectory of points on the power axis towards the target power. If for a new point the distance to the target is larger then the distance of the currently best known point, a new trajectory is started by reinitializing (2.20) with the λ and $\sum_k s_k$ of the best known point.

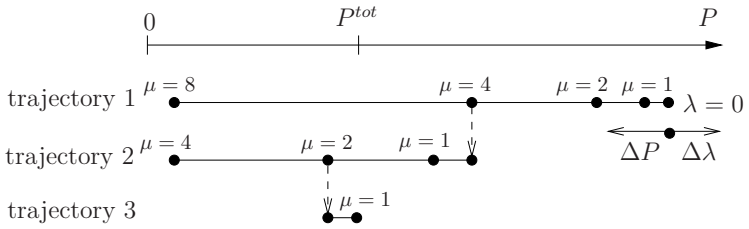


Figure 2.2: Single-user search procedure

Algorithm 2.1 Single-user λ search algorithm

while distance $>$ tolerance **do**
  $\lambda =$ best λ so far
  $\mu = 1$
  while distance $\leq$ previousDistance **do**
    previousDistance = distance
    $\mu = \mu \times 2$
    $\Delta\lambda = -\mu\big(P^{tot} - P^{\lambda}\big)$
    $[P^{\lambda+\Delta\lambda}, \mathbf{s}_k] =$ calculateLoading$([\lambda + \Delta\lambda]^+)$ (per-tone exhaustive search)
    distance $= |P^{tot} - P^{\lambda+\Delta\lambda}|$
  end while
end while

Formula (2.20) also corresponds to the subgradient approach of [15], be it that it was derived here based on a simple problem-specific property (2.17). This in particular allows for alternative stepsize selection strategies, such as Algorithm 2.1. Using the theory provided by [15], a formula similar to (2.17) can also be derived. One can plot the maximum achievable rate as a function of the total power as in Figure 2.3. This plot can be obtained by performing the maximization of the Lagrangian $\max_{\mathbf{s}} R - \lambda P$ for all possible λ's. For each λ this results in an optimal total power usage $P = P^{opt}$ and a corresponding optimal rate $R = R^{opt}$. It is argued in [15] that this function is concave, based

on the assumption that the number of tones is *infinite*. Solving $\max_{\mathbf{s}} R - \lambda P$ for some given λ corresponds to finding the point on the curve where the difference between R^λ and λP^λ is largest. From Figure 2.3 it can be seen that this corresponds to finding the point on the curve where the tangent is equal to λ. Hence for the optimal solution

$$\frac{\partial R(P = P^\lambda)}{\partial P} = \lambda. \tag{2.21}$$

Mathematically, if the maximum achievable rate as a function of the optimal power budget is twice differentiable, the concavity of this curve can be expressed as

$$\frac{\partial^2 R}{\partial^2 P} \leq 0. \tag{2.22}$$

Based on (2.21) and (2.22), for a *small* ΔP the same relation as (2.17) is then derived as follows:

$$\frac{\partial^2 R}{\partial^2 P} \leq 0 \Rightarrow \frac{\partial}{\partial P}\Big(\underbrace{\frac{\partial R}{\partial P}}_{\lambda}\Big) \leq 0 \Rightarrow \frac{\partial \lambda}{\partial P} \leq 0 \overset{\Delta P \Delta P \geq 0}{\Rightarrow} \Big(\underbrace{\frac{\partial \lambda}{\partial P}\Delta P}_{\Delta\lambda}\Big)\Delta P \leq 0.$$

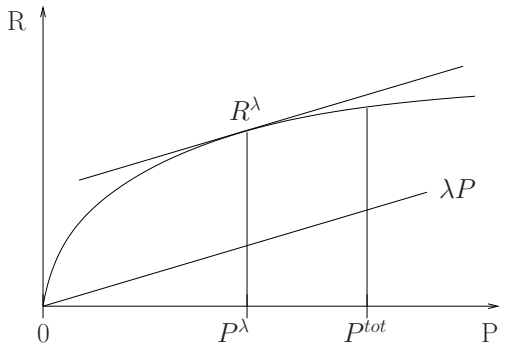


Figure 2.3: Optimal rate vs power

Finally, note that the λ of the dual problem formulation (2.14) without the spectral mask constraint is related to the well-known waterfilling solution[17] in the continuous bit loading case. The waterfilling solution is obtained by differentiating $R + \lambda(P^{tot} - \sum_k s_k)$ with respect to s_k. This leads to

$$s_k + \frac{\Gamma \sigma_k}{|h_k|^2} = \frac{1}{\lambda \ln(2)}.$$

The constant at the right-hand side is the waterfilling level, which is seen to be inversely related to the Lagrange multiplier. A higher waterfilling level

corresponds to more power allocated, which can be obtained by setting a small λ.

2.4.2 Multi-user Spectrum Management

In the previous section, interesting relations (2.17), (2.18) and (2.19) were derived between λ, the total power and the data rate. These relations gave rise to a simple procedure for searching the λ and leading to a total power tightly satisfying the total power constraint. In this section these relations will be investigated for the multi-user case. We will focus on the two-user case first, and then extend the results to the general N-user case.

For the N-user case, the spectrum management problem can be formulated as in Formula (2.3). Formula (2.11) and (2.12) specified the per-tone dual formulation of the generic spectrum management problem. For a two-user case the Lagrange dual function can be formulated as

$$\begin{aligned} \underset{\mathbf{s}^1,\mathbf{s}^2}{\text{argmax}} \quad & \omega_1 R_1 + \omega_2 R_2 + \lambda_1\big(P^{1,tot} - \sum_k s_k^1\big) + \lambda_2\big(P^{2,tot} - \sum_k s_k^2\big) \\ \text{s.t.} \quad & 0 \le s_k^n \le s_k^{\text{mask}} \qquad n = 1,2; k = 1 \dots K \end{aligned} \tag{2.23}$$

Given $\boldsymbol{\omega} = (\omega_1, \omega_2)$ and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$, this Lagrange dual function, decoupled over the tones, can be solved by performing an exhaustive search for each tone over all possible bit or power loading combinations for the users.

$$\begin{aligned} \text{for } k = 1 \dots K, \qquad \{s_k^{1,opt}, s_k^{2,opt}\} = \quad & \underset{s_k^1, s_k^2}{\text{argmax}} \sum_{n=1}^{2} \omega_n f_s b_k^n - \sum_{n=1}^{2} \lambda_n s_k^n \\ & \text{s.t. } 0 \le s_k^n \le s_k^{n,\text{mask}} \qquad n = 1,2 \end{aligned} \tag{2.24}$$

The optimal solution is then a bit and power loading corresponding to total powers and data rates

$$\big(P^{1,\boldsymbol{\omega},\boldsymbol{\lambda}}, R^{1,\boldsymbol{\omega},\boldsymbol{\lambda}}, P^{2,\boldsymbol{\omega},\boldsymbol{\lambda}}, R^{2,\boldsymbol{\omega},\boldsymbol{\lambda}}\big) \tag{2.25}$$

The optimality of this solution implies that for this $\boldsymbol{\lambda}$ and $\boldsymbol{\omega}$ there exists no other bit or power loading giving a larger value to the Lagrangian. This then implies that for a weighted total power budget $\lambda_1 P^1 + \lambda_2 P^2$ smaller than $P^{\boldsymbol{\omega},\boldsymbol{\lambda}} \triangleq \lambda_1 P^{1,\boldsymbol{\omega},\boldsymbol{\lambda}} + \lambda_2 P^{2,\boldsymbol{\omega},\boldsymbol{\lambda}}$, it is impossible to achieve a weighted rate sum (with weights ω_1, ω_2) that is larger than $R^{\boldsymbol{\omega},\boldsymbol{\lambda}} \triangleq \omega_1 R^{1,\boldsymbol{\omega},\boldsymbol{\lambda}} + \omega_2 R^{2,\boldsymbol{\omega},\boldsymbol{\lambda}}$.

This is shown graphically in the power plane of Figure 2.4 (which is the two-user version of Figure 2.1). $\big(P^{1,\boldsymbol{\omega},\boldsymbol{\lambda}}, P^{2,\boldsymbol{\omega},\boldsymbol{\lambda}}\big)$ is the optimal power allocation for a given $\boldsymbol{\lambda}$ and $\boldsymbol{\omega}$. Every loading corresponding to total powers in the marked triangle then has a smaller weighted rate sum.

If a $\boldsymbol{\lambda}$ is found such that $\left(P^{1,\boldsymbol{\omega},\boldsymbol{\lambda}} = P^{1,tot}, P^{2,\boldsymbol{\omega},\boldsymbol{\lambda}} = P^{2,tot}\right)$, this solution satisfies the total power constraints and so has a weighted rate sum larger than every other possible loading in the marked rectangle (i.e. a subset of the marked triangle). Thus the primal problem is solved once the rate constraints are satisfied. If one of the λ_n is zero and the corresponding total power is smaller than the total power constraint, this means that the total power constraint is not active for user n and we should not further try to make this total power tight. This case is handled later in Algorithm 2.2

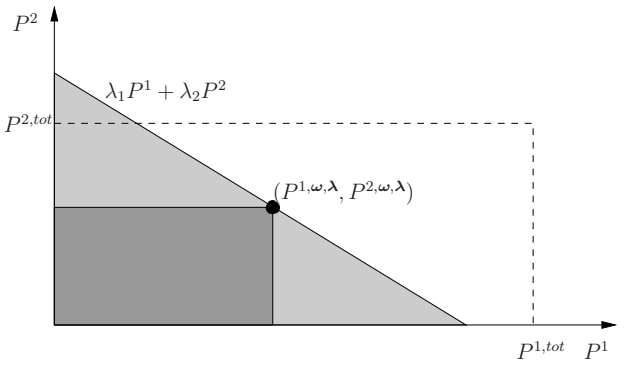


Figure 2.4: Two-user power plane

So our aim is to tune $\boldsymbol{\lambda}$ and $\boldsymbol{\omega}$ such that the total power constraints and rate constraints are satisfied: $P^{1,\boldsymbol{\omega},\boldsymbol{\lambda}} = P^{1,tot}$, $P^{2,\boldsymbol{\omega},\boldsymbol{\lambda}} = P^{2,tot}$, $R^{1,\boldsymbol{\omega},\boldsymbol{\lambda}} \geq R^{1,target}$ and $R^{2,\boldsymbol{\omega},\boldsymbol{\lambda}} \geq R^{2,target}$. The same strategy as in the single-user case can now be followed to first derive a generalization of Formula (2.17), and then an update strategy for $\boldsymbol{\lambda}$ and $\boldsymbol{\omega}$.

Starting from two optimal solutions $(R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}, P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}, R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}, P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A})$ and $(R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B}, P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B}, R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B}, P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B})$ corresponding to $(\boldsymbol{\omega}_A, \boldsymbol{\lambda}_A)$ and $(\boldsymbol{\omega}_B, \boldsymbol{\lambda}_B)$ respectively, optimality for $(\boldsymbol{\omega}_A, \boldsymbol{\lambda}_A)$ implies

$$\begin{aligned} & \omega_{1,A} R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} + \omega_{2,A} R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - \lambda_{1,A} P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - \lambda_{2,A} P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} \\ \leq\ & \omega_{1,A} R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} + \omega_{2,A} R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} - \lambda_{1,A} P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} - \lambda_{2,A} P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} \end{aligned} \tag{2.26}$$

Optimality for $(\boldsymbol{\omega}_B, \boldsymbol{\lambda}_B)$ implies

$$\begin{aligned} & \omega_{1,B} R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} + \omega_{2,B} R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} - \lambda_{1,B} P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} - \lambda_{2,B} P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A} \\ \leq\ & \omega_{1,B} R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} + \omega_{2,B} R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - \lambda_{1,B} P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} - \lambda_{2,B} P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B} \end{aligned} \tag{2.27}$$

Taking the sum of (2.26) and (2.27) results in

$$\begin{aligned}
&-\underbrace{\left(\omega_{1,B}-\omega_{1,A}\right)}_{\Delta\omega_1}\underbrace{\left(R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B}-R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}\right)}_{\Delta R^1}\\
&-\underbrace{\left(\omega_{2,B}-\omega_{2,A}\right)}_{\Delta\omega_2}\underbrace{\left(R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B}-R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}\right)}_{\Delta R^2}\\
&+\underbrace{\left(\lambda_{1,B}-\lambda_{1,A}\right)}_{\Delta\lambda_1}\underbrace{\left(P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B}-P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}\right)}_{\Delta P^1}\\
&+\underbrace{\left(\lambda_{2,B}-\lambda_{2,A}\right)}_{\Delta\lambda_2}\underbrace{\left(P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B}-P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A}\right)}_{\Delta P^2}\leq 0
\end{aligned} \tag{2.28}$$

Relation (2.28) for two users can be extended straightforward to the N-user case:

$$\left[\begin{array}{cc} -(\Delta\boldsymbol{\omega})^T & (\Delta\boldsymbol{\lambda})^T \end{array}\right]\left[\begin{array}{c} \boldsymbol{\Delta R} \\ \boldsymbol{\Delta P} \end{array}\right]\leq 0. \tag{2.29}$$

where $\boldsymbol{\Delta\lambda} = [\Delta\lambda_1,\ldots,\Delta\lambda_N]^T$ and $\boldsymbol{\Delta\omega} = [\Delta\omega_1,\ldots,\Delta\omega_N]^T$ are vectors containing the $\Delta\lambda$'s and $\Delta\omega$'s for the N users, $\boldsymbol{\Delta P} = [\Delta P^1,\ldots,\Delta P^N]^T$ and $\boldsymbol{\Delta R} = [\Delta R^1,\ldots,\Delta R^N]^T$ are vectors with the corresponding total powers and data rates, Δx^n is equal to $x^{n,B} - x^{n,A}$.

Two special cases can be derived from Formula (2.29).

$$\text{fixed } \boldsymbol{\omega} \quad (\Delta\boldsymbol{\omega}=0) \quad \Rightarrow \quad (\Delta\boldsymbol{\lambda})^T\Delta\mathbf{P}\leq 0 \tag{2.30}$$

$$\text{fixed } \boldsymbol{\lambda} \quad (\Delta\boldsymbol{\lambda}=0) \quad \Rightarrow \quad (\Delta\boldsymbol{\omega})^T\Delta\mathbf{R}\geq 0 \tag{2.31}$$

Note that (2.17) is the single-user version of (2.30).

Relation (2.29) can be used to construct a simple procedure to find the $\boldsymbol{\omega}$ and $\boldsymbol{\lambda}$ that enforce the constraints. To simplify the graphical illustration of the procedure, we limit ourselves to a procedure that only updates the $\boldsymbol{\lambda}$ Lagrange multipliers. However, this procedure can be straightforwardly extended to also include the update of the $\boldsymbol{\omega}$ Lagrange multipliers by extending the vectors as in Formula (2.29).

In the two-user power plane, Formula (2.30) can be represented graphically as two vectors with a non-positive inner product, as in Figure 2.5. When searching for the Lagrange multipliers that make the total power constraint tight, we need to make changes to the $\boldsymbol{\lambda}$ such that in the power plane we end up in the point where every user is at maximum power. Because of the non-positive inner product of $\Delta\boldsymbol{\lambda}$ and $\Delta\mathbf{P}$, $\Delta\boldsymbol{\lambda}$ for the desired $\Delta\mathbf{P}$ must be somewhere in the gray half plane opposite to the $\Delta\mathbf{P}$ vector.

This relation between $\Delta\boldsymbol{\lambda}$ and $\Delta\mathbf{P}$ can be used to steer the used power towards the total power constraint. By changing the current $\boldsymbol{\lambda}$ vector by a $\Delta\boldsymbol{\lambda}$ in

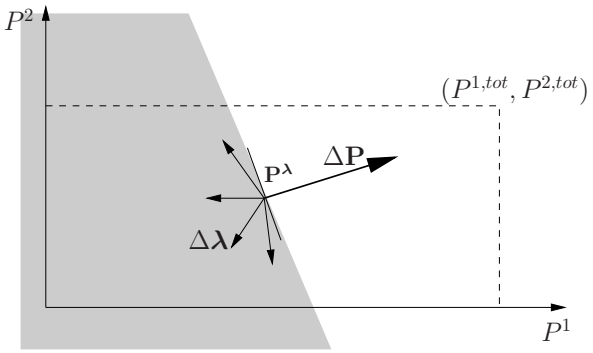


Figure 2.5: Two-user power plane

the opposite direction of the desired $\Delta\mathbf{P}$, Formula (2.30) guarantees that the step taken with $\Delta\mathbf{P}$ will get the used power closer to the power constraint $(P^{1,tot}, P^{2,tot})$, as long as $\Delta\boldsymbol{\lambda}$ is not too large. This is shown in Figure 2.6, where the $\Delta\boldsymbol{\lambda}$ brings $\mathbf{P}^{\boldsymbol{\lambda}}$ to the next point inside the shaded circle.

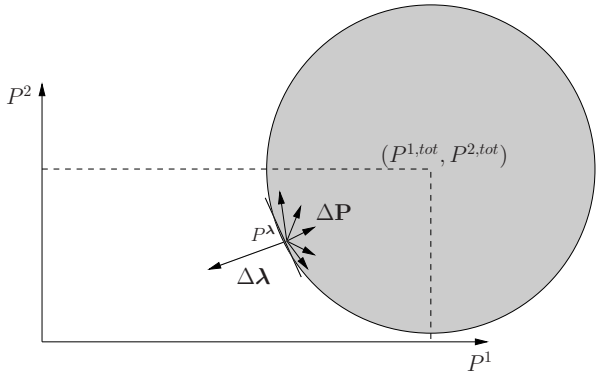


Figure 2.6: Two-user power plane

Mathematically this procedure can be captured in the following update formula:

$$\Delta\boldsymbol{\lambda} = -\mu\big(\mathbf{P}^{tot} - \mathbf{P}^{\boldsymbol{\lambda}}\big) \quad \Rightarrow \quad \boldsymbol{\lambda}^{t+1} = \left[\boldsymbol{\lambda}^{t} - \mu\Big(\mathbf{P}^{tot} - \sum_{k} \mathbf{s}_k\Big)\right]^{+} \tag{2.32}$$

which also applies to the N-user case.

By starting with a small μ, e.g. $\mu = 1$, the used power makes a small step closer to the desired total power. As long as the used power keeps getting closer to the

desired total power, μ can be increased, e.g. doubled. A trajectory of points is then followed, each point with a total power closer to the power constraint. At some point, a $\Delta\boldsymbol{\lambda}$ will be selected taking the used power further from the desired total power than the previous point found along the trajectory. Then this last step has to be discarded and a new trajectory is started using a new direction $\left(\mathbf{P}^{tot}-\mathbf{P}^{\boldsymbol{\lambda}}\right)$. This procedure is formally represented in Algorithm 2.2. The outer loop of this algorithm iterates over the trajectories while the inner loop follows one of the trajectories. A possible evolution of the total power using this strategy is shown in Figure 2.7 for the two-user case.

Remark that if some of the used powers are smaller than the total power constraints ($P^n < P^{n,tot}$) and the corresponding Lagrange multipliers λ_n are zero, the $[\lambda_n + \Delta\lambda_n]^+$ operation keeps these Lagrange multipliers λ_n to zero. The algorithm will so converge to the case that either λ_n becomes zero or $P^n = P^{n,tot}$ for each n. In this case, the Lagrangian objective is identical to the original primal objective, thus solving the original primal problem [10].

Formula (2.32) again corresponds to the subgradient approach of [15], be it that it was derived here based on a simple problem-specific property (2.30). This in particular allows for alternative stepsize selection strategies, such as Algorithm 2.2.

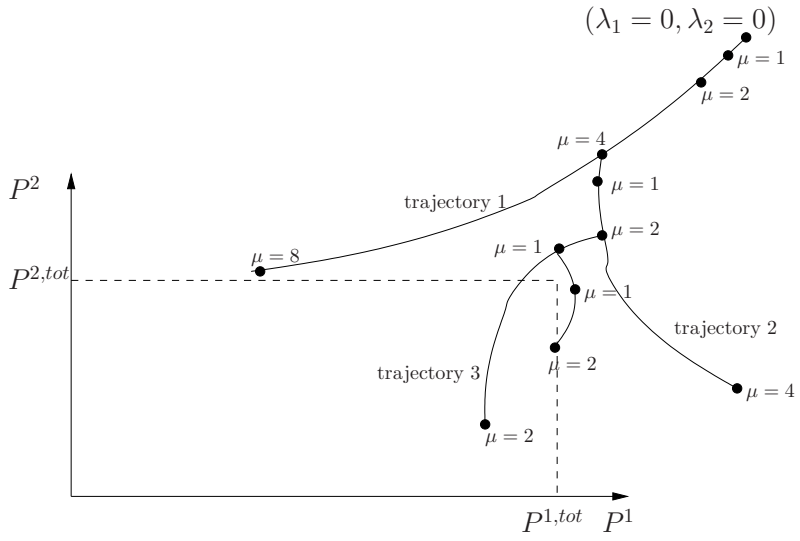


Figure 2.7: Trajectories of total power in the power plane

In the framework of [15], a relation similar to (2.30) results from the concavity of the function that gives the optimal weighted rate $R = \omega_1 R^1 + \omega_2 R^2$ (two-user

Algorithm 2.2 Multi-user $\boldsymbol{\lambda}$ search algorithm

```
while distance > tolerance do
  λ = best λ so far
  μ = 1
  while distance ≤ previousDistance do
    previousDistance = distance
    μ = μ × 2
    Δλ = −μ(P^tot − P^λ)
    [P^{λ+Δλ}, s^n] = calculateLoading([λ + Δλ]^+)        (per-tone exhaustive search)
    distance = ‖P^tot − P^{λ+Δλ}‖
  end while
end while
```

case) in function of the used total power $(\mathrm{P}^1, \mathrm{P}^2)$. Because of this concavity, for *any* x and y the following relation must hold:

$$\begin{bmatrix} x & y \end{bmatrix} \underbrace{\begin{bmatrix} \frac{\partial^2 R}{\partial^2 P^1} & \frac{\partial^2 R}{\partial P^1 \partial P^2} \\ \frac{\partial^2 R}{\partial P^2 \partial P^1} & \frac{\partial^2 R}{\partial^2 P^2} \end{bmatrix}}_{\begin{bmatrix} \frac{\partial \lambda_1}{\partial P^1} & \frac{\partial \lambda_2}{\partial P^1} \\ \frac{\partial \lambda_1}{\partial P^2} & \frac{\partial \lambda_2}{\partial P^2} \end{bmatrix}} \begin{bmatrix} x \\ y \end{bmatrix} \leq 0$$

By choosing $x = \Delta P^1$ and $y = \Delta P^2$, for small ΔP^1, ΔP^2 relation (2.30) is obtained.

2.5 A Low Complexity Branch and Bound Approach to the Per-Tone Optimization Problem

The per-tone optimization problem shown in (2.12) is non-convex with many local optima. In Section 2.3 it was mentioned that an exhaustive search was proposed [10] to find the global optimum. Unfortunately the set size of feasible solutions is exponential in the number of users N. This becomes intractable for more than four users. In this section a branch and bound procedure is presented to solve the per-tone optimization problem. Note that the focus will be on the bit loading instead of the power loading. Every bit loading corresponds to a particular power loading, as can be seen from Formula (2.2). The variables of the optimization problems of the previous sections can be changed from powers

into bits leading to the following equivalent per-tone optimization problem:

$$\begin{aligned} \mathbf{b}_k^{opt} := \quad & \underset{\mathbf{b}_k}{\text{argmax}} \quad \underbrace{\sum_{n=1}^{N} \omega_n f_s b_k^n}_{1} - \underbrace{\sum_{n=1}^{N} \lambda_n s_k^n}_{2} \\ & \text{s.t.} \qquad 0 \le s_k^n \le s_k^{n,\text{mask}}, \qquad n = 1 \dots N. \end{aligned} \tag{2.33}$$

Branch and bound [13][14] is a general method for finding optimal solutions for various optimization problems. A branch and bound procedure has two ingredients. The first ingredient is a smart way of covering the feasible region by several smaller subregions. This leads to a branching operation. The second ingredient is bounding, which is a way of finding upper and lower bounds for the optimal solution within a feasible subregion. The core of the approach is the simple observation that (for a maximization task) if the upper bound of a subregion A is smaller than the lower bound for some other subregion B, then subregion A may be safely discarded from the search. The efficiency of the method depends critically on the effectiveness of the used branching and bounding operations. There is no universal bounding operation that works for all problems. In the remainder of this section a branch and bound procedure will be presented particularly for the per-tone optimization problem (2.33).

The procedure starts from an *initial solution set* containing all feasible solutions. A possible initial solution set is based on a maximum bit loading of every user as follows

$$B = \left\{ \mathbf{b}_k \in \mathbb{R}^N, \forall n : 0 \le b_k^n \le b_k^{n,\text{MAX}} \right\}, \tag{2.34}$$

where $b_k^{n,\text{MAX}}$ is the maximum bit loading for user n on tone k. Based on the fact that the power loading increases with increasing bit loading (see Appendix) a tighter initial solution set can be defined by (2.35)

$$B = \left\{ \mathbf{b}_k \in \mathbb{R}^N, \forall n : 0 \le b_k^n \le \min(\log_2\big(1 + \frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^{n,\text{mask}}}{\sigma_k^n}\big), b_k^{n,\text{MAX}}) \right\}. \tag{2.35}$$

The *branching operation* splits every region in 2^N rectangular subregions. This is done by splitting each dimension (bit loading range for 1 particular user) of the considered region in two. One has several options for splitting a dimension: two equal parts, two parts delimited by integer bit loadings, two parts delimited by fractional bit loadings,.... In the case of integer bit loading the branching can be stopped when all the regions have dimensions smaller than or equal to 1 bit. Figure 2.8 depicts the branching operation starting from an initial solution set for a three-user case. It can be remarked that this "rectangular" branching admits a region (e.g. region I) to be fully characterized by its two extreme points $\mathbf{b}_k^{\min}$ and $\mathbf{b}_k^{\max}$ as shown in Figure 2.8 (with $b_k^{n,\min} \le b_k^n \le b_k^{n,\max}$ for all n, in the considered region).

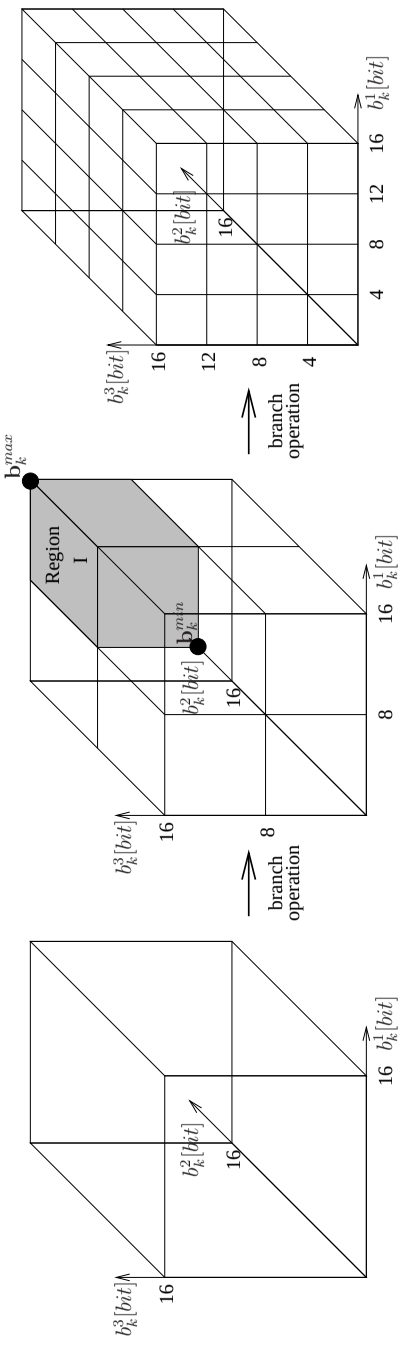


Figure 2.8: Branching operation for a three-user case starting from an initial solution set based on the maximum bit loading with $b_k^{n,\mathrm{MAX}} = 16$. Each axis corresponds to the bit loading of a particular user for tone k. The middle figure emphasizes the fact that a region can be fully characterized by its two points $\mathbf{b}_k^{\min}$ and $\mathbf{b}_k^{\max}$.

After the branching operation, a number of the generated subregions can be discarded from the search because they consist only of power loading combinations that do not satisfy the spectral mask constraints of (2.33). In the Appendix it is proven that the point $\mathbf{b}_k^{\min}$ of Figure 2.8 corresponds to a power loading that is smaller than every other power loading in the considered subregion I. If this point $\mathbf{b}_k^{\min}$ does not satisfy the spectral mask constraints than no point of the considered subregion satisfies the spectral mask constraints and so the subregion can be discarded from the search. Formula (2.2) can be used to check the spectra corresponding to $\mathbf{b}_k^{\min}$.

The *first bound operation* is the calculation of a *lower bound* for the maximum of a subregion. In fact, the cost function value of any bit loading combination of the considered subregion can be used as a lower bound. As the spectra corresponding to $\mathbf{b}_k^{\min}$ have already been calculated to check the spectral mask constraints, the cost function value of this point can be used as a lower bound. The calculation of the lower bound (2.36) then corresponds to a simple sum of products.

$$\text{lower bound}_{\text{region}} = \sum_{n=1}^{N} \omega_n f_s b_k^{n,\min} - \sum_{n=1}^{N} \lambda_n s_k^{n,\min} \tag{2.36}$$

The *second bound operation* is the calculation of an *upper bound* for the maximum of a subregion. Formula (2.33) consists of two terms. Term 1 increases with increasing bit loading. Term 2 decreases with increasing bit loading because the power loading increases with increasing bit loading (see Appendix). An exact upper bound of term 1 is $\sum_{n=1}^{N} \omega_n f_s b_k^{n,\max}$. An exact upper bound of term two is $-\sum_{n=1}^{N} \lambda_n s_k^{n,\min}$. The sum of the upper bounds of term 1 and term 2 is an upper bound for the sum of the two terms.

$$\text{upper bound}_{\text{region}} = \sum_{n=1}^{N} \omega_n f_s b_k^{n,\max} - \sum_{n=1}^{N} \lambda_n s_k^{n,\min}. \tag{2.37}$$

Note that the spectra $s^{n,\min}, \forall n$ have already been calculated. The calculation of the upper bound (2.37) then again corresponds to a simple sum of products. As can be seen, the bound operations represent a small computational complexity. The effectiveness of these bounds in discarding regions is checked by simulations. This is presented in the simulations section.

The complete branch and bound procedure for one tone of optimization problem (2.33) is presented in Algorithm 2.3.

The general spectrum management problem takes total power constraints into account. Suffice it to say that in some cases, the spectral mask constraints are so constraining that the total power constraints are always met, namely

Algorithm 2.3 Branch and bound procedure for the per-tone optimization problem

1: Start from an initial solution set containing all feasible solutions as defined by (2.34) or (2.35)
2: Branching: split the region(s) into subregions as shown in Figure 2.8
3: Discard the subregions which do not satisfy the spectral mask constraints
4: Bounding: determine lower and upper bounds for the maximum of each subregion as defined by (2.36) and (2.37)
5: Determine maximum lower bound MAX_*low* of all subregions.
6: Remove all subregions with an upper bound smaller than MAX_*low*.
7: **if** desired accuracy achieved **then** go to step 9
8: **else** go to step 2
9: Exhaustive search on all remaining subregions

when $\sum_k s_k^{n,\text{mask}} \leq P^{n,tot}, \forall n$. In this case (2.33) reduces to term 1 and the calculation of the lower bound and upper bound is simplified to:

$$\text{lower bound}_{\text{region}} = \sum_{n=1}^{N} \omega_n f_s b_k^{n,\min} \tag{2.38}$$

$$\text{upper bound}_{\text{region}} = \sum_{n=1}^{N} \omega_n f_s b_k^{n,\max}. \tag{2.39}$$

Computational Complexity Analysis

The overall computational complexity of the proposed branch and bound procedure is determined by the number of subregions to be considered and the computational complexity required for each considered subregion.

For general branch and bound procedures the bound operations for a particular subregion require to solve an optimization problem (e.g. linear optimization problem, convex relaxation optimization problem, ...). In the proposed branch and bound procedure the computational complexity required for each subregion is reduced to a minimum. The lower bound computation (2.36) requires one function evaluation of (2.33). The main part of this computation is calculating the spectra corresponding with $b_k^{n,\min}$ (2.2). The multiplications and summations can be neglected with respect to this computation. The upper bound computation (2.37) can reuse this part of the computation of the lower bound (2.33) and requires a few extra multiplications and summations which can be neglected in terms of computational complexity. So the computational complexity for each subregion is approximately one function evaluation of (2.33). Because of this reason the proposed branch and bound scheme is called a low complexity branch and bound scheme.

The number of subregions to be considered is determined by the branching operation and the number of subregions that are discarded. In the worst-case none of the subregions will be discarded, which is very unlikely. In this case the branching operations will generate B^N subregions. This means that in the worst-case when no subregions are discarded the computational complexity consists of calculating B^N function evaluations of (2.33) which is the same as the computational complexity of the exhaustive search. Typically a lot of subregions get discarded based on the bounds and the spectral mask constraints leading to a much more efficient scheme than the exhaustive search.

Speed Up by Use of Correlation

The effectiveness of the presented branch and bound procedure depends on how fast regions are discarded. As already explained a region is discarded from the search when this region has an upper bound smaller than the lower bound of another region. The presence of a region with a large lower bound thus speeds up the convergence. Typically there is a lot of correlation between neighbouring tones , i.e. the optimal bit or power loading of tone k does not change a lot from the optimal bit or power loading of tone $k+1$. When the optimal bit or power loading of tone k is known, the subregion containing this optimal solution for tone k can be used in the branch and bound procedure for tone $k+1$ as a good approximation of the optimum solution for tone $k+1$, having a large lower bound. Because of the presence of this subregion with a large lower bound more subregions will be discarded in a early stage leading to a much smaller set of subregions to consider. This use of tone correlation speeds up the algorithm significantly.

A similar observation is that the optimal bit or power loading of tone k does not change much whenever the Lagrange multipliers λ_n are updated (towards the values that enforce the total power constraints). A similar use of correlation over Lagrange multiplier updates for tone k can then be used to speed up the algorithm.

2.6 Simulation Results

In Section 2.5 we presented a low complexity branch and bound procedure for solving the non-convex per-tone optimization problem. Together with the Lagrange search algorithm presented in Section 2.4, this composes a low complexity solution for Optimal Spectrum Balancing. In this section we will evaluate the complexity reduction of the proposed solution with respect to the OSB algorithm using an exhaustive search [10].

The ADSL downstream scenarios use a line diameter of 0.5 mm (24 AWG). The maximum transmit power is 20.4 dBm [18]. The SNR gap Γ is 12.9 dB,

corresponding to a coding gain of 3 dB, a noise margin of 6 dB and a target symbol error probability of 10^{-7}. The maximum bit loading is 16 bits. The tone spacing Δ_f is 4.3125 kHz and the DMT symbol rate f_s is 4 kHz. The simulations are performed in Matlab on a dual Opteron 250 with 4 GB RAM and a 2.4 GHz processor.

2.6.1 Symmetric Scenario

The first scenario is a symmetric scenario and is shown in Figure 2.9. A slightly modified version of Algorithm 2.2 is used for finding the optimal Lagrange multipliers that enforce the constraints. The intuitive interpretation of Algorithm 2.2 allows for the number of $\boldsymbol{\lambda}$-evaluations to be further reduced. By starting a trajectory with step size $\mu = 1$, a number of $\boldsymbol{\lambda}$-evaluations is wasted to increase the μ to a magnitude for which the total power starts converging towards the total power constraint. Instead, one could start a trajectory with a μ inspired by the best μ of the previous trajectory, avoiding unnecessary $\boldsymbol{\lambda}$-evaluations.

Table 2.1 compares the computational complexity of our branch and bound OSB solution, starting with initial set (2.35), with the computational complexity of the exhaustive search OSB solution. The simulations are done for a two- up to a seven-user scenario. The spectral mask constraints are set to -30dBm/Hz.

The computational complexity of both the branch and bound procedure and the exhaustive search procedure consists of calculating expression (2.33) for given Lagrange multipliers λ_n, ω_n and for a particular bit loading. The main part of this computation corresponds to calculating the spectra for a particular bit loading, as defined by Formula (2.2). The 'Users' column indicates the number of users of the scenario. The 'Nb of Traj' and 'Nb of $\boldsymbol{\lambda}$-ev' columns indicate the number of trajectories and the number of $\boldsymbol{\lambda}$-evaluations respectively that the Lagrange Search algorithm has to perform to find the values that enforce the constraints. The 'Nb ev of (2.2)' indicates the total number of evaluations of (2.2) needed to converge and is a good measure of the complexity of the algorithm. For a two-user case, the total number of evaluations for the exhaustive search is calculated as follows. The maximum number of allocated bits is 16. This means that every user can be assigned 17 possible bit loadings. For two-users there are $289 (= 17^2)$ possible bit loadings. ADSL downstream uses 224 tones. There are 43 λ-evaluations needed to converge. This makes a total of 43 x 224 x $17^2 (= 2783648)$ possibilities that have to be calculated to obtain the optimal bit loading. The ' % ev of (2.2)' column indicates the percentage of evaluations of expression (2.2) with respect to the number of evaluations of expression (2.2) by the exhaustive search. The 'C.r.f.' column gives the complexity reduction factor with respect to the exhaustive search. This is the inverse of the previous column and is a good indication of how much less computation is needed to converge to the optimal solution.

As both procedures use the same Lagrange search, the number of $\boldsymbol{\lambda}$-evaluations and trajectories is equal. It can be seen that this number is smaller than 50, independent of the number of users. This is much faster than the bisection method proposed in [10] where the number of $\boldsymbol{\lambda}$-evaluations is exponential in the number of users. Furthermore it can be seen that the complexity reduction factor increases as the number of users increases. For seven users there is a complexity reduction of 368. Using the exhaustive search, it is impossible to simulate scenarios with more than four users on the standard platform used here. With the proposed branch and bound procedure one can simulate up to *seven-user* scenarios on the same platform within a reasonable time. Typical simulation times are 3 hours for a five-user scenario, 1 day for a six-user scenario and 1 week for a seven-user scenario. The optimal power spectra or bit loads that are obtained can provide insights needed for the design of more practical suboptimal loading schemes.

Table 2.2 shows the results for the spectrum management problem without total power constraints and with spectral mask constraints set to -40dBm/Hz. For the branch and bound procedure, Algorithm 2.3 changes in step 4 using Formula (2.38) and (2.39) instead of Formula (2.36) and (2.37). The computational complexity of the exhaustive search procedure also reduces as the second term of (2.11) is removed. This can be viewed as the evaluation of the (2.36) and (2.37) with $\boldsymbol{\lambda}$ equal to zero. It can be seen that the complexity reduction factors are now even larger than those of Table 2.1. For an *eight-user* scenario there is a complexity reduction of 18518. One very interesting remark is that for all the solutions of Table 2.2 the resulting total powers for every user satisfy the total power constraints. In other words, the spectral mask constraints are so constraining that the total power constraints are not active.

2.6.2 Asymmetric Scenario

The second scenario is an asymmetric scenario and is show in Figure 2.10. The simulations are performed for a two-user case (N = 2) up to an eight-user case (N = 8). The four-user scenario, for example, consists of modem 1 up to modem 4. Table 2.3 shows the results for the spectrum management problem with total power constraints and spectral mask constraints set to -30dBm/Hz. The complexity reduction factors increase with increasing number of users. For a seven-user case there is a complexity reduction with a factor 538. Table 2.4 shows the results for the spectrum management problem without total power constraints and with spectral mask constraints set to -40dBm/Hz. For an eight-user scenario there is a complexity reduction with a factor 12345.

Figures 2.11 and 2.12 show the power loading and bit loading respectively for the seven-user asymmetric scenario with total power constraints and spectral mask constraints set to -30dBm/Hz.

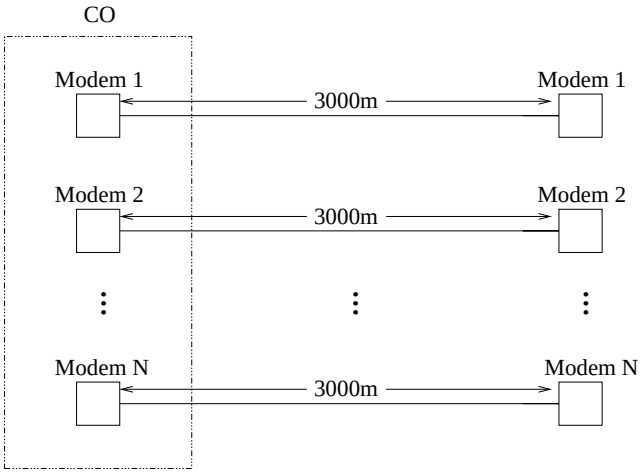


Figure 2.9: Symmetric N-user ADSL scenario

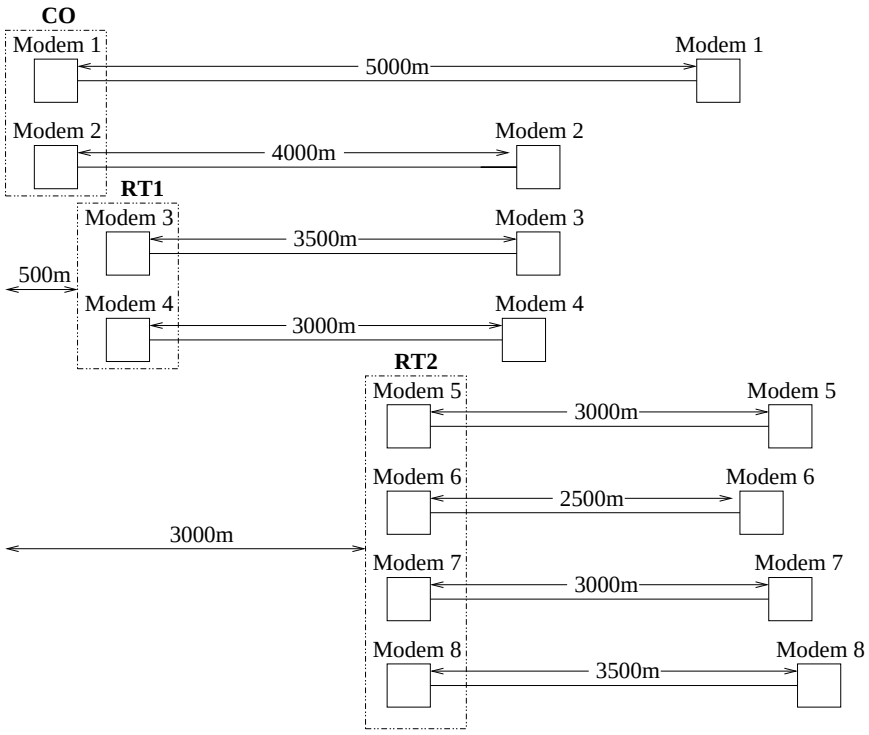


Figure 2.10: Asymmetric N-user ADSL scenario

Table 2.1: Computational complexity comparison of branch and bound procedure and exhaustive search procedure for symmetric ADSL Downstream scenario with power constraints and spectral mask constraints

			Branch and Bound procedure			Exhaustive Search procedure		
Users	Nb of Traj	Nb of λ-ev	Nb ev of (2.2)	% ev of (2.2)	C.r.f.	Nb ev of (2.2)	% ev of (2.2)	C.r.f.
2	11	43	367769	13.2	7.6	2783648	100	1
3	6	30	1840265	5.57	17.9	33015360	100	1
4	9	36	16620093	2.47	40.48	673513344	100	1
5	9	36	120037408	1.053	94.97	11.4e9	100	1
6	14	47	1.49e9	0.586	171	2.54e11	100	1
7	13	44	1.0993e+10	0.272	368	4.044e12	100	1

Table 2.2: Computational complexity comparison of branch and bound procedure and exhaustive search procedure for symmetric ADSL Downstream scenario without power constraints and with spectral mask constraints

			Branch and Bound procedure			Exhaustive Search procedure		
Users	Nb of Traj	Nb of λ-ev	Nb ev of (2.2)	% ev of (2.2)	C.r.f.	Nb ev of (2.2)	% ev of (2.2)	C.r.f.
2	1	1	2638	4.07	25	64736	100	1
3	1	1	8409	0.764	131	1100512	100	1
4	1	1	44929	0.236	423	19e6	100	1
5	1	1	258480	0.081	1231	318e6	100	1
6	1	1	1673325	0.031	3226	5.4e9	100	1
7	1	1	11989628	0.013	7673	92e9	100	1
8	1	1	84312076	0.0054	18518	1.56e12	100	1

Table 2.3: Computational complexity comparison of branch and bound procedure and exhaustive search procedure for asymmetric ADSL Downstream scenario with power constraints and spectral mask constraints

			Branch and Bound procedure			Exhaustive Search procedure		
Users	Nb of Traj	Nb of λ-ev	Nb ev of (2.2)	% ev of (2.2)	C.r.f.	Nb ev of (2.2)	% ev of (2.2)	C.r.f.
2	7	33	149818	7.01	14.3	2136288	100	1
3	3	25	665511	2.42	41.3	27512800	100	1
4	5	29	8161797	1.5	66.7	542552416	100	1
5	11	43	125016093	0.9125	109.6	13.7e9	100	1
6	5	28	681798051	0.45	222	151e9	100	1
7	4	25	4.2708e9	0.186	538	2.297e12	100	1

Table 2.4: Computational complexity comparison of branch and bound procedure and exhaustive search procedure for asymmetric ADSL Downstream scenario without power constraints and with spectral mask constraints

			Branch and Bound procedure			Exhaustive Search procedure		
Users	Nb of Traj	Nb of λ-ev	Nb ev of (2.2)	% ev of (2.2)	C.r.f.	Nb ev of (2.2)	% ev of (2.2)	C.r.f.
2	1	1	1676	2.6	38.5	64736	100	1
3	1	1	5924	0.54	185	1100512	100	1
4	1	1	38619	0.2	500	19e6	100	1
5	1	1	367166	0.11	909	318e6	100	1
6	1	1	3254672	0.06	1667	5.4e9	100	1
7	1	1	22638706	0.0246	4065	92e9	100	1
8	1	1	129249600	0.0081	12345	1.56e12	100	1

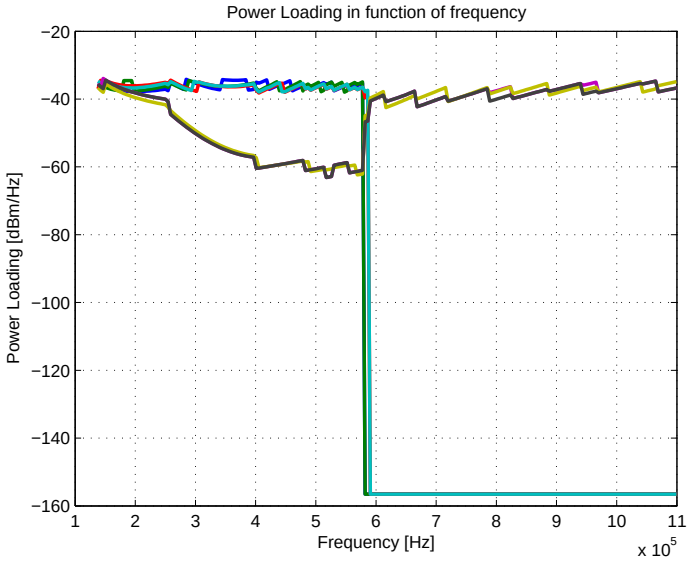


Figure 2.11: Asymmetric seven-user ADSL downstream optimal power loading

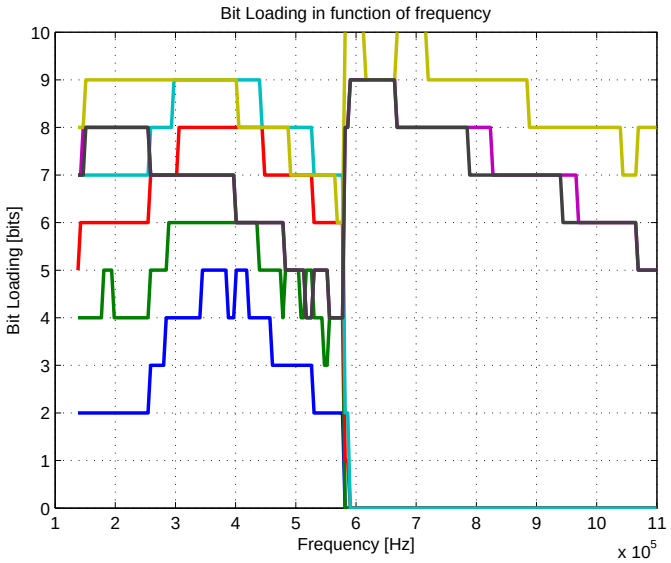


Figure 2.12: Asymmetric seven-user ADSL downstream optimal bit loading

2.7 Conclusion

In modern DSL systems, multi-user crosstalk is a major source of performance degradation. Optimal Spectrum Balancing (OSB) is a centralized algorithm that mitigates the effect of crosstalk by allocating optimal transmit spectra to all interfering DSL modems. By the use of a Lagrange multiplier approach the spectrum management problem is decoupled into non-convex per-tone optimization problems. The remaining issues are finding the Lagrange multipliers that enforce the constraints and solving the per-tone optimization problems. Starting from the single user case, this paper presented a number of properties, which were then extended to the multi-user case and lead to an efficient search algorithm for the Lagrange multipliers. Simulations showed that this algorithm succeeds in finding the optimal Lagrange multipliers in less than 50 iterations, independent of the number of users. Secondly, a low complexity branch and bound procedure is presented to solve the per-tone optimization problems. Simulations showed enormous complexity reductions with respect to the exhaustive search. The combination of the two proposed algorithms creates an OSB solution with a significantly reduced complexity. This reduction makes it possible to simulate realistic *eight-user* scenarios within a reasonable time on a standard platform whereas before it was only possible to simulate up to four-user scenarios on the same platform. The obtained optimal power spectra or bit loads can provide insights needed for the design of practical suboptimal loading schemes.

Appendix: Relation Between Bit Loadings and Power Loadings

In this appendix it will be proven that for a multi-user scenario a nondecreasing bit loading corresponds to a nondecreasing power loading. Mathematically this can be formulated as

Given	$\mathbf{s}_k \geq 0,\ \tilde{\mathbf{s}}_k \geq 0,\ b_k^n = f(\mathbf{s}_k),\ \tilde{b}_k^n = f(\tilde{\mathbf{s}}_k)$
with	f defined by (2.1)
if	$\tilde{\mathbf{b}}_k \geq \mathbf{b}_k$ (a)
then	$\tilde{\mathbf{s}}_k \geq \mathbf{s}_k$

The '$\geq$'-operator applied to vectors means inequalities hold for all elements in the vector.
The proof will be given for a three-user case but it can be extended trivially to N users. It will be proven that condition (a) implies that $\tilde{s}_k^1 \geq s_k^1$. The proofs for $\tilde{s}_k^2 \geq s_k^2$ and $\tilde{s}_k^3 \geq s_k^3$ are similar.

The proof is a proof by contradiction.

Let us assume that $s_k^1 > \tilde{s}_k^1 \geq 0$. This can be reformulated as

$$\tilde{s}_k^1 = \alpha s_k^1 \quad \text{with} \quad 0 < \alpha < 1. \tag{2.40}$$

Condition (a) can be written as

$$\frac{|h_k^{1,1}|^2 s_k^1}{|h_k^{1,2}|^2 s_k^2 + |h_k^{1,3}|^2 s_k^3 + \sigma_k^1} \leq \frac{|h_k^{1,1}|^2 \tilde{s}_k^1}{|h_k^{1,2}|^2 \tilde{s}_k^2 + |h_k^{1,3}|^2 \tilde{s}_k^3 + \sigma_k^1} \tag{2.41}$$

$$\frac{|h_k^{2,2}|^2 s_k^2}{|h_k^{2,1}|^2 s_k^1 + |h_k^{2,3}|^2 s_k^3 + \sigma_k^2} \leq \frac{|h_k^{2,2}|^2 \tilde{s}_k^2}{|h_k^{2,1}|^2 \tilde{s}_k^1 + |h_k^{2,3}|^2 \tilde{s}_k^3 + \sigma_k^2} \tag{2.42}$$

$$\frac{|h_k^{3,3}|^2 s_k^3}{|h_k^{3,1}|^2 s_k^1 + |h_k^{3,2}|^2 s_k^2 + \sigma_k^3} \leq \frac{|h_k^{3,3}|^2 \tilde{s}_k^3}{|h_k^{3,1}|^2 \tilde{s}_k^1 + |h_k^{3,2}|^2 \tilde{s}_k^2 + \sigma_k^3}. \tag{2.43}$$

Using (2.40) in (2.41) gives

$$|h_k^{1,2}|^2 \tilde{s}_k^2 + |h_k^{1,3}|^2 \tilde{s}_k^3 + \sigma_k^1 \leq \alpha(|h_k^{1,2}|^2 s_k^2 + |h_k^{1,3}|^2 s_k^3 + \sigma_k^1). \tag{2.44}$$

We can assume that the noise power σ_k^1 is not equal to zero, which gives:

$$|h_k^{1,2}|^2 \tilde{s}_k^2 + |h_k^{1,3}|^2 \tilde{s}_k^3 < \alpha(|h_k^{1,2}|^2 s_k^2 + |h_k^{1,3}|^2 s_k^3). \tag{2.45}$$

This relation suggests that

$$\tilde{s}_k^2 < \alpha s_k^2 \quad \text{and/or} \quad \tilde{s}_k^3 < \alpha s_k^3.$$

Without loss of generality, let us suppose that $\tilde{s}_k^2 < \alpha s_k^2$. This can be reformulated as

$$\tilde{s}_k^2 = \beta s_k^2 \quad \text{with} \quad 0 < \beta < \alpha < 1. \tag{2.46}$$

Using (2.46) in (2.42) gives

$$|h_k^{2,1}|^2 \tilde{s}_k^1 + |h_k^{2,3}|^2 \tilde{s}_k^3 + \sigma_k^2 \leq \beta(|h_k^{2,1}|^2 s_k^1 + |h_k^{2,3}|^2 s_k^3 + \sigma_k^2). \tag{2.47}$$

We can assume that the noise power σ_k^2 is not equal to zero, which gives:

$$|h_k^{2,1}|^2 \tilde{s}_k^1 + |h_k^{2,3}|^2 \tilde{s}_k^3 < \beta(|h_k^{2,1}|^2 s_k^1 + |h_k^{2,3}|^2 s_k^3). \tag{2.48}$$

Note that $\beta < \alpha$ which results in $\tilde{s}_k^3 < \beta s_k^3$. This can be reformulated as

$$\tilde{s}_k^3 = \gamma s_k^3 \quad \text{with} \quad 0 < \gamma < \beta < \alpha < 1. \tag{2.49}$$

Using (2.49) in (2.43) gives

$$|h_k^{3,1}|^2 \tilde{s}_k^1 + |h_k^{3,2}|^2 \tilde{s}_k^2 + \sigma_k^3 \leq \gamma(|h_k^{3,1}|^2 s_k^1 + |h_k^{3,2}|^2 s_k^2 + \sigma_k^3). \tag{2.50}$$

Based on (2.40), (2.46) and (2.49) we can see that (2.50) is not possible. The initial assumption (2.40) does not hold which proves that $\tilde{s}_k^1 \geq s_k^1 \geq 0$. In a similar way, the proofs can be given for $\tilde{s}_k^2 \geq s_k^2 \geq 0$ and $\tilde{s}_k^3 \geq s_k^3 \geq 0$.

Bibliography

[1] *Transmission and Multiplexing (TM); Access transmission systems on metallic access cables; Very high speed Digital Subscriber Line (VDSL); Functional Requirements*, ETSI Std. TS 101 270-1, Rev. V.1.3.1, 2003.

[2] T. Starr, J. M. Cioffi, P. J. Silverman, *Understanding digital subscriber lines*. Prentice Hall, 1999.

[3] R. Cendrillon, M. Moonen, E. V. den Bogaert, and G. Ginis, "The linear zero-forcing crosstalk canceller is near-optimal in DSL channels," in *Proc. of the IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 4, Dec. 2004, pp. 2334–2338.

[4] R. Cendrillon, M. Moonen, G. Ginis, K. V. Acker, T. Bostoen, and P. Vandaele, "Partial crosstalk cancellation for upstream VDSL," *EURASIP Journal on Applied Signal Processing*, no. 10, pp. 1520–1535, Aug. 2004.

[5] G. Ginis and J. M. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE Journal on Selected Areas on Communications*, vol. 20, no. 5, pp. 1085–1104, Jun. 2002.

[6] *Spectrum management for loop transmission systems*, ANSI Std. T1.417, Issue 2, 2003.

[7] T. Starr, M. Sorbara, J. M. Cioffi, P. J. Silverman, *DSL Advances, ch. 11.* Prentice Hall, 2002.

[8] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[9] R. Cendrillon, M. Moonen, J. Verlinden, T. Bostoen, and W. Yu, "Optimal multi-user spectrum management for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 1, Jun. 2004, pp. 1–5.

[10] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 922–933, May 2006.

[11] R. Cendrillon and M. Moonen, "Iterative spectrum balancing for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 3, May 2005, pp. 1937–1941.

[12] R. Lui and W. Yu, "Low-complexity near-optimal spectrum balancing for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 3, May 2005, pp. 1947–1951.

[13] "Definition: Branch and bound," From Wikipedia, the Free Encyclopedia.

[14] P. M. Pardalos and H. E. Romeijn, *Handbook of global optimization volume 2 (Nonconvex optimization and its applications)*, 1st ed. Springer, Jun. 2002.

[15] W. Yu, R. Lui, and R. Cendrillon, "Dual optimization methods for multiuser orthogonal frequency-division multiplex systems," in *Proc. of the IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 1, Dec. 2004, pp. 225–229.

[16] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge Univ. Press, 2003.

[17] B. S. Krongold, K. Ramchandran, and D. L. Jones, "Computationally efficient optimal power allocation algorithms for multicarrier communication systems," *IEEE Trans. on Communications*, vol. 48, no. 1, pp. 23–27, Jan. 2000.

[18] *Asymmetrical digital subscriber line (ADSL) transceivers*, ITU Std. G.992.1, 1999.

Chapter 3

Distributed Spectrum Coordination

Distributed Spectrum Management Algorithms for Multiuser DSL Networks

P. Tsiaflakis, M. Diehl, and M. Moonen

Published in *IEEE Transactions on Signal Processing*, vol. 56, no. 10, Oct. 2008, pp. 4825-4843.

©2008 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Contributions of first author

- literature study
- development of KKT-framework for state-of-the-art distributed spectrum coordination algorithms
- development of novel distributed spectrum coordination algorithms: ASB2, CA-DSB, DSB and MS-DSB
- interpretation of developed algorithms as per-tone penalty water-filling algorithms
- co-development of complexity analysis
- co-development of convergence analysis
- design of simulation scenarios
- software implementation and computer simulations
- interpretation of simulation results
- text co-redaction, co-editing, co-revision

Abstract

Modern DSL networks suffer from crosstalk among different lines in the same cable bundle. This crosstalk can lead to a major performance degradation. By balancing the transmit power spectra, the impact of crosstalk can be minimized leading to spectacular performance gains. This is referred to as Spectrum Management. In this paper a unifying perspective is presented on distributed spectrum management algorithms based on the Karush-Kuhn-Tucker (KKT) conditions. Furthermore novel distributed algorithms are presented within the same KKT framework. The proposed distributed algorithms consist of local water-filling-like algorithms running in the individual modems, controlled by the spectrum management center. Extensive simulation results show that the proposed algorithms perform very well for several multi-user ADSL and VDSL scenarios.

3.1 Introduction

Digital Subscriber Line (DSL) technology remains by far the most popular broadband access technology with a 65.7% market share. In 2006, the number of DSL subscribers worldwide grew by more than 30% to reach close to 185 million [1]. The increasing demand for higher data rates forces DSL systems to use higher frequencies, e.g. up to 30 MHz for VDSL2. At these high frequencies, electromagnetic coupling becomes particularly harmful and causes interference, also called "crosstalk", among lines operating in the same cable bundle. This crosstalk is the major obstacle for performance improvement in DSL systems currently under development.

Dynamic Spectrum Management (DSM) [2] refers to a set of solutions to the crosstalk problem. Basically these solutions consist of signal level coordination and/or spectrum level coordination. In the signal level coordination solution, the crosstalk is cancelled by signal processing at the transmitter and/or the receiver. Several crosstalk canceller designs have been proposed [3][4][5]. However direct crosstalk cancellation is infeasible in many cases, due to complexity issues (both the amount of computation needed and the requirements of new chip sets) or as a result of unbundling (i.e. multiple operators own different lines of same cable bundle).

In this paper the focus is on spectrum level coordination, which is also referred to as spectrum management, spectrum balancing or multi-carrier power control. Here the transmit power spectrum of each modem is designed to cause minimal disturbance to other lines, while preserving a high data rate. This can significantly improve data rates beyond those of the current static spectrum management where fixed over-conservative standardized transmit power

spectra are used.

The problem of optimally choosing the transmit power spectra in order to maximize the data rates of the network can be formulated as an optimization problem [6][7] which is referred to as the *spectrum management problem*. Unfortunately this optimization problem is nonconvex and can have multiple locally optimal solutions. Three algorithms, Optimal Spectrum Balancing (OSB) [6], Branch and Bound Optimal Spectrum Balancing (BB-OSB) [7] and a prismatic branch and bound algorithm (PBB) [8] respectively, were proposed to compute the globally optimal solution, i.e. the globally optimal transmit spectra for the network of interfering DSL lines. Unfortunately their complexity grows exponentially with the number of modems. A cheaper iterative approach, called Iterative Spectrum Balancing (ISB), was simultaneously proposed in [9] and [10]. Here the complexity is reduced through a series of line-searches. OSB, BB-OSB, PBB and ISB are centralized algorithms which means that they rely on a spectrum management center (SMC) to optimize the transmit power spectra for all modems simultaneously. Note that it is assumed that the SMC has access to estimates of the noise and the crosstalk channels obtained through training procedures or with the aid of crosstalk models and knowledge of the loop topology.

Distributed spectrum management applies a different paradigm where each modem executes a spectrum management algorithm locally to compute its transmit power spectrum. One of the first and most famous distributed algorithms is iterative water-filling (IW) [11]. In this algorithm, each modem maximizes its own bit rate, where the crosstalk from the other modems is simply treated as noise. Although this leads to a 'selfish' solution, the advantage is that the modems do not need any information from the other modems and so the iterative water-filling algorithm is a completely autonomous algorithm.

Another autonomous algorithm is the autonomous spectrum balancing (ASB) algorithm [12], that utilizes the concept of a reference line. Each modem optimizes its own bit rate taking into account the damage caused on the reference line avoiding a "selfish" behavior leading to a better network performance.

Recently another distributed spectrum management paradigm was introduced in [13] and is referred to as SCALE, Successive Convex Approximation for Low-complExity. The algorithm involves iterative convex approximations. Here each modem executes a spectrum management algorithm locally. In addition, limited message-passing is allowed between the modems and the SMC. The resulting method can be seen as a distributed computation across the DSL network.

In this paper, the focus is on distributed spectrum management algorithms. First it is shown how current distributed spectrum management algorithms can be viewed from a similar framework. This framework consists of choosing

an (iterative) approximation of the non-convex spectrum management problem and deriving a distributed solution based on the KKT conditions of the chosen approximation. Furthermore novel distributed spectrum management algorithms are proposed within the same KKT framework. The first algorithm involves an iterative convex approximation approach. This is similar to SCALE, however it chooses a different convex approximation based on a linearization of the non-concave part of the objective function. This choice leads to an interesting water-filling-like structure. The second algorithm is based on the KKT conditions of the spectrum management problem. The third algorithm is a multiple starting point extension of the second algorithm which leads to a better performance for multiple user scenarios with multiple locally optimal solutions. All the proposed algorithms consist of local water-filling-like algorithms running in the modems, steered by a SMC.

The paper is organized as follows. Section 3.2 describes the system model for a network of interfering DSL modems. In Section 3.3 the nonconvex spectrum management problem is briefly formulated. Section 3.4 presents current state-of-the-art distributed spectrum management algorithms from a KKT perspective. In Section 3.5 three novel spectrum management algorithms are proposed and discussed. Finally, in Section 3.6 extensive simulation results compare the proposed algorithms with state-of-the-art distributed and centralized spectrum management algorithms for several multi-user ADSL and VDSL scenarios.

3.2 System Model

Most current DSL systems use Discrete Multi-Tone (DMT) modulation. The basic idea of DMT is to split the available bandwidth into a large number of subcarriers also called tones. Under certain conditions (sufficiently long cyclic prefix and perfect tone synchronization), each tone is capable of transmitting data independently from other tones, and so the transmit power and the number of bits can be assigned individually for each tone. This gives a large flexibility in optimally shaping the transmit spectrum. The transmission for a binder of N modems can be modeled on each tone k by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k \qquad , k = 1 \dots K.$$

where K is the number of tones. The vector $\mathbf{x}_k = [x_k^1, x_k^2, \dots, x_k^N]^T$ contains the transmitted symbols on tone k for all N modems. $[\mathbf{H}_k]_{n,m} = h_k^{n,m}$ is an $N \times N$ matrix containing the channel transfer functions from transmitter m to receiver n on tone k. The diagonal elements are the direct channels, the off-diagonal elements are the crosstalk channels. $\mathbf{z}_k$ is the vector of additive noise on tone k, containing thermal noise, alien crosstalk, RFI,.... The vector $\mathbf{y}_k$ contains the received symbols.

The transmit power is denoted as $s_k^n \triangleq \Delta_f E\{|x_k^n|^2\}$, the noise power as $\sigma_k^n \triangleq \Delta_f E\{|z_k^n|^2\}$. The vector containing the transmit power of modem n on all tones is $\mathbf{s}^n \triangleq [s_1^n, s_2^n, \ldots, s_K^n]^T$. The DMT symbol rate is denoted as f_s, the tone spacing as Δ_f.

In this model, no signal coordination is assumed between modems and the signals from other modems are treated as noise: such a multi-user channel is referred to as an "interference channel". When the number of interfering modems is large, the interference is well approximated by a Gaussian distribution. Under this assumption the achievable bit loading for modem n on tone k, given the transmit spectra $\mathbf{s}_k \triangleq [s_k^1, s_k^2, \ldots, s_k^N]^T$ of all modems in the system, is

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

where Γ denotes the signal-to-noise ratio (SNR) gap to capacity, which is a function of the desired bit error ratio (BER), the coding gain and noise margin [14].

The data rate for modem n and the total power used by modem n are respectively

$$R^n = f_s \sum_k b_k^n \qquad \text{and} \qquad P^n = \sum_k s_k^n.$$

3.3 Spectrum Management Problem

The problem of optimally balancing the transmit power spectra in order to maximize the data rates of the DSL network is referred to as the rate-adaptive spectrum management problem. The objective is to find the optimal transmit spectra for a bundle of interfering DSL lines, maximizing a weighted sum of data rates subject to per-modem total power constraints and spectral mask constraints. This can be formulated as the following non-convex optimization problem:

$$\begin{aligned} \max_{\mathbf{s}^1,\ldots,\mathbf{s}^N} \quad & \sum_{n=1}^{N} w_n R^n \\ \text{s.t.} \quad & \sum_k s_k^n \leq P^{n,\text{tot}} \quad , \quad \forall n, \\ \text{s.t.} \quad & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \quad \forall n, \forall k, \end{aligned} \tag{3.2}$$

where $P^{n,\text{tot}}$ denotes the total power budget for modem n and $s_k^{n,\text{mask}}$ denotes the spectral mask for modem n on tone k. The weights w_n are used to put more emphasis on some modems. In [7] it is explained how these weights can be adjusted when extra data rate constraints have to be satisfied. All

the algorithms explained in this paper can be easily extended with data rate constraints.

3.4 A KKT Perspective on State-of-the-Art Distributed Spectrum Management Algorithms

In this section it will be shown how state-of-the-art distributed spectrum management algorithms can be viewed from a similar framework. This framework consists of two steps. The first step is approximating the non-convex spectrum management problem (3.2). This can be done by a (non-)convex approximation or by iterative (non-)convex approximations. The second step is deriving a distributed solution based on the KKT conditions of these approximations. This unifying perspective simplifies the comparison of the different algorithms to better understand their strengths and limitations. Furthermore it shows that a key design parameter is choosing a good (iterative) approximation. Possible design criteria that can be considered in the choice of approximations are complexity, (sub-)optimality, convergence properties, communication overhead, etc. It will be shown how state-of-the-art distributed algorithms fit in this framework by explaining their approximation and how they are derived based on the KKT conditions of their approximation. However there are more possible (iterative) approximations to tackle the spectrum management problem each of them having different properties. In Section 3.5 it will be shown how other approximations lead to algorithms with interesting properties.

3.4.1 KKT perspective on Iterative Water-filling

The iterative water-filling algorithm [11] is an extension of the single-user water-filling algorithm [15] to multiple users. Each modem iteratively applies the single-user water-filling algorithm while the crosstalk of the other modems is fixed. In this way each modem maximizes its own bit rate without taking into account the damage caused to the other modems. This approach generally does not lead to the solution of the spectrum management problem (3.2). In fact, it corresponds to solving the following optimization problem for each modem n:

$$\begin{aligned} \max_{s_1^n,\ldots,s_K^n} \quad & w_n f_s \sum_k \log_2 \left(1 + \tfrac{1}{\Gamma} \frac{|h_k^{n,n}|^2 s_k^n}{\sum_{m \neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n} \right) \\ \text{s.t.} \quad & \sum_k s_k^n \leq P^{n,\text{tot}} \\ \text{s.t.} \quad & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \forall k. \end{aligned} \tag{3.3}$$

The objective function is a concave function as the crosstalk of the other modems $\sum_{m\neq n}|h_k^{n,m}|^2 s_k^m$ is fixed. Problem (3.3) can be viewed as a convex approximation of (3.2). If the crosstalk is negligible, approximation (3.3) will lead to a solution similar to the solution of (3.2). In the case of large crosstalk, solving approximation (3.3) leads to a different and suboptimal solution. The advantage of the approximation is its convexity. The globally optimal (and distributed) solution of this convex problem (for the n-th modem, with fixed crosstalk from the other modems) can be found by solving the system obtained by its KKT conditions. The KKT conditions for modem n are as follows:

$$\frac{w_n f_s |h_k^{n,n}|^2}{\log(2)(|h_k^{n,n}|^2 s_k^n + \sum_{m\neq n}\Gamma|h_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n)} - \lambda_n + \mu_k^n - \nu_k^n = 0, \forall k, \quad (3.4)$$

$$\begin{aligned}
\sum_{k=1}^{K} s_k^n - P^{n,\mathrm{tot}} &\leq 0, && (3.5)\\
-s_k^n &\leq 0, \forall k, && (3.6)\\
s_k^n &\leq s_k^{n,\mathrm{mask}}, \forall k && (3.7)\\
\lambda_n &\geq 0, && (3.8)\\
\mu_k^n &\geq 0, \forall k, && (3.9)\\
\nu_k^n &\geq 0, \forall k, && (3.10)\\
\lambda_n\left(\sum_{k=1}^{K} s_k^n - P^{n,\mathrm{tot}}\right) &= 0, && (3.11)\\
-\mu_k^n s_k^n &= 0, \forall k, && (3.12)\\
\nu_k^n(s_k^n - s_k^{n,\mathrm{mask}}) &= 0, \forall k. && (3.13)
\end{aligned}$$

Using (3.4) the optimal transmit powers s_k^n for fixed dual variables $\lambda_n, \nu_k^n, \mu_k^n$ are given by the following closed form formula

$$s_k^n = \frac{w_n f_s}{\log(2)(\lambda_n + \nu_k^n - \mu_k^n)} - \frac{\sum_{m\neq n}\Gamma|h_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n}{|h_k^{n,n}|^2}, \quad (3.14)$$

where λ_n corresponds to constraint $\sum_k s_k^n \leq P^{n,\mathrm{tot}}$, ν_k^n corresponds to constraint $s_k^n \leq s_k^{n,\mathrm{mask}}$ and μ_k^n corresponds to constraint $0 \leq s_k^n$. The KKT conditions are then equivalent to the following conditions:

1. If $\dfrac{w_n f_s}{\log(2)\lambda_n} - \dfrac{\sum_{m\neq n}\Gamma|h_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n}{|h_k^{n,n}|^2} \leq 0,$
   then $s_k^n = 0, \nu_k^n = 0,$
   $\mu_k^n = \lambda_n - \dfrac{w_n f_s |h_k^{n,n}|^2}{\log(2)(\sum_{m\neq n}\Gamma|h_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n)}.$

$$
\begin{aligned}
&2.\ \text{If } 0 \leq \frac{w_n f_s}{\log(2)\lambda_n} - \frac{\sum_{m \neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma \sigma_k^n}{|h_k^{n,n}|^2} \leq s_k^{n,\text{mask}}, \\
&\quad \text{then } s_k^n = \frac{w_n f_s}{\log(2)\lambda_n} - \frac{\sum_{m \neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma \sigma_k^n}{|h_k^{n,n}|^2}, \\
&\quad \nu_k^n = 0, \mu_k^n = 0. \\
&3.\ \text{If } \frac{w_n f_s}{\log(2)\lambda_n} - \frac{\sum_{m \neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma \sigma_k^n}{|h_k^{n,n}|^2} \geq s_k^{n,\text{mask}}, \\
&\quad \text{then } s_k^n = s_k^{n,\text{mask}}, \mu_k^n = 0, \\
&\quad \nu_k^n = \frac{w_n f_s}{\log(2)\left(s_k^{n,\text{mask}} + \frac{\sum_{m \neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma \sigma_k^n}{|h_k^{n,n}|^2}\right)} - \lambda_n. \\
&4.\ \lambda_n \left(\textstyle\sum_{k=1}^K s_k^n - P^{n,\text{tot}}\right) = 0
\end{aligned}
\tag{3.15}
$$

This can be summarized as follows

$$
s_k^n = \left[\frac{w_n f_s}{\log(2)\lambda_n} - \frac{\Gamma(\sum_{m \neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n)}{|h_k^{n,n}|^2} \right]_0^{s_k^{n,\text{mask}}}, \tag{3.16}
$$

$$
\lambda_n \left(\sum_{k=1}^K s_k^n - P^{n,\text{tot}} \right) = 0, \tag{3.17}
$$

where $[x]^a$ means $\min(x, a)$ and $[x]_b$ means $\max(x, b)$. Note that the spectral mask constraints are easily taken into account. Furthermore λ_n has to be chosen so that the complementary slackness condition (3.17) is satisfied. Note that the total power constraints will always be active. This follows from the particular choice of convex approximation (3.3). A modem always benefits from allocating more power and therefore will use its total available power. The λ_n can be found using a distributed approach [7][13][16][17] for updating λ_n:

$$
\lambda_n = \left[\lambda_n + \epsilon \left(\sum_k s_k^n - P^{n,\text{tot}} \right) \right]_0. \tag{3.18}
$$

where ϵ is a step size parameter. Alternatively, one can use a simple bisection search on the λ_n for every modem n.

This leads to a distributed algorithm where every modem n iteratively applies formula (3.16) and bisects on λ_n to converge to the optimal solution of approximation (3.3). The IW algorithm is shown in Algorithm 3.1, where δ is a

small value indicating the required accuracy in total power, γ is a small value for stopping the bisection search on λ_n in the case of an inactive total power constraint and $\Lambda^{\max}$ is a maximum value for the λ_n.

Algorithm 3.1 IW

Modem n algorithm:
Initialize $s_k^n = 0, \forall k$
loop {execute at regular intervals}
  $\lambda_n^{\min} = 0, \lambda_n^{\max} = \Lambda^{\max}$, $\lambda_n = (\lambda_n^{\max} + \lambda_n^{\min})/2$
  while $|\sum_k s_k^n - P^{n,\text{tot}}| > \delta$ and $\lambda_n > \gamma$ **do**
    $\lambda_n = (\lambda_n^{\max} + \lambda_n^{\min})/2$
    Update s_k^n using (3.16), $\forall k$
    if $\sum_k s_k^n > P^{n,\text{tot}}$ **then**
      $\lambda_n^{\min} = \lambda_n$
    else
      $\lambda_n^{\max} = \lambda_n$
    end if
  end while
end loop

To summarize, it is shown that the iterative water-filling algorithm can be seen as a convex approximation of the spectrum management problem (3.2). The choice of this approximation leads to near-optimal performance for small crosstalk scenarios. The KKT conditions are then used to derive the distributed (and fully autonomous) iterative water-filling procedure.

3.4.2 KKT Perspective on SCALE

The SCALE algorithm, proposed in [13], is based on iterative convex approximations of the spectrum management problem (3.2). The following lower bound on $\log(1+z)$ is used

$$\alpha_k^n \log z + \beta_k^n \leq \log(1+z) \tag{3.19}$$

$$\text{where} \quad \begin{array}{l} \alpha_k^n = \frac{z_0}{1+z_0} \\ \beta_k^n = \log(1+z_0) - \frac{z_0}{1+z_0}\log z_0 \end{array} .$$

This bound is tight with equality at a chosen value $z_0 > 0$ when the constants $\{\alpha_k^n, \beta_k^n\}$ are chosen as specified above. This leads to the following approxima-

tion of problem (3.2):

$$\begin{aligned} \max_{\mathbf{s}^1,\ldots,\mathbf{s}^N} \quad & \sum_n w_n f_s \sum_k \alpha_k^n \log_2 \left(\frac{1}{\Gamma} \frac{|h_k^{n,n}|^2 s_k^n}{\sum_{m\neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n} \right) + \beta_k^n \\ \text{s.t.} \quad & \sum_k s_k^n \leq P^{n,\text{tot}} \quad , \quad \forall n, \\ \text{s.t.} \quad & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \quad \forall n, \forall k. \end{aligned} \tag{3.20}$$

Using the transformation $\tilde{s}_k^n = \log(s_k^n)$, the approximation (3.20) is equivalent to the following convex optimization problem

$$\begin{aligned} \max_{\tilde{\mathbf{s}}^1,\ldots,\tilde{\mathbf{s}}^N} \quad & \sum_n w_n f_s \sum_k \alpha_k^n \log_2 \left(\frac{1}{\Gamma} \frac{|h_k^{n,n}|^2 e^{\tilde{s}_k^n}}{\sum_{m\neq n} |h_k^{n,m}|^2 e^{\tilde{s}_k^m} + \sigma_k^n} \right) + \beta_k^n \\ \text{s.t.} \quad & \sum_k e^{\tilde{s}_k^n} \leq P^{n,\text{tot}} \quad , \quad \forall n, \\ \text{s.t.} \quad & e^{\tilde{s}_k^n} \leq s_k^{n,\text{mask}}, \quad \forall n, \forall k. \end{aligned} \tag{3.21}$$

where the constraint $0 \leq s_k^n$ is always satisfied if $\tilde{s}_k^n$ is transformed back to s_k^n.

Now a similar approach (see (3.4)-(3.15)) can be followed as in the previous section to derive a distributed algorithm based on the KKT conditions. In particular, based on the KKT stationarity condition and the complementarity conditions of approximation (3.21), the following fixed point equation can be derived:

$$s_k^n = \left\lfloor \frac{w_n f_s \alpha_k^n}{\log(2)\lambda_n + \underbrace{\sum_{m\neq n} \frac{w_m f_s \alpha_k^m \Gamma |h_k^{m,n}|^2}{\sum_{p\neq m} \Gamma |h_k^{m,p}|^2 s_k^p + \Gamma \sigma_k^m}}_{P_k^{\text{SCALE},n}}} \right\rfloor_0^{s_k^{n,\text{mask}}}. \tag{3.22}$$

Note that the transformation $s_k^n = e^{\tilde{s}_k^n}$ is used to transform back to s_k^n as explained in [13] (where the constraint $0 \leq s_k^n$ is added as $\tilde{s}_k^n$ is transformed back to s_k^n).

Every modem iteratively applies formula (3.22) and bisects on λ_n to converge to the solution of the approximated problem (3.21). By regularly updating the approximation parameters α_k^n with formula (3.19) the network converges to a locally optimal solution of problem (3.2). Note that the algorithm cannot be executed in a fully distributed manner as the term $P_k^{\text{SCALE},n}$ (3.22) depends on information which is not available to each modem. However this information can be acquired with some limited message-passing as explained in [13]. This leads to Algorithm 3.2 adopted from [13].

Algorithm 3.2 SCALE

Modem n algorithm:
Initialize $s_k^n = 0, \alpha_k^n = 1, \forall k$
loop {execute at regular intervals}
  $\lambda_n^{\min} = 0, \lambda_n^{\max} = \Lambda^{\max}$, $\lambda_n = (\lambda_n^{\max} + \lambda_n^{\min})/2$
  Receive messages $P_k^{\text{SCALE},n}$ from SMC, $\forall k$
  while $|\sum_k s_k^n - P^{n,\text{tot}}| > \delta$ and $\lambda_n > \gamma$ **do**
    $\lambda_n = (\lambda_n^{\max} + \lambda_n^{\min})/2$
    Update s_k^n using (3.22), $\forall k$
    if $\sum_k s_k^n > P^{n,\text{tot}}$ **then**
      $\lambda_n^{\min} = \lambda_n$
    else
      $\lambda_n^{\max} = \lambda_n$
    end if
  end while
  Measure $\text{int}_k^n = \sum_{p \neq n} |h_k^{n,p}|^2 s_k^p + \sigma_k^p$, $\forall k$
  Update α_k^n with (3.19), $\forall k$
  Transmit $\alpha_k^n / \text{int}_k^n$ to SMC, $\forall k$
end loop
SMC algorithm:
loop
  Receive messages $\alpha_k^n / \text{int}_k^n$ from modems n, $\forall k$
  Compute messages $P_k^{\text{SCALE},n}$ and send to each modem n, $\forall k$.
end loop

In [18] a slightly different approximation is proposed based on a high SNR approach as follows

$$\log z \approx \log(1+z). \tag{3.23}$$

This leads to an approximation which is similar to the approximation of SCALE where the parameters are fixed to $\alpha_k^n = 1$ and $\beta_k^n = 0$. Unlike the SCALE algorithm, this high SNR approach does not necessarily converge to a local optimum of problem (3.2). If the high SNR assumption holds, this approach leads to a near-optimal performance. Note that for this approximation, no approximation parameters need to be updated, however this reduction in complexity is negligible.

3.4.3 KKT Perspective on ASB

The ASB algorithm, proposed in [12], utilizes the concept of a reference line. This reference line should be chosen so as to represent a statistical average of all crosstalk victim lines within a typical network. In [12] heuristics are provided for choosing the reference line. Once the reference line is fixed each modem optimizes its own bit rate taking into account the damage caused on the reference line.

This can be viewed as the following approximation (3.24) of problem (3.2) where $h_k^{n,\mathrm{ref}}, h_k^{\mathrm{ref}}, s_k^{\mathrm{ref}}, \sigma_k^{\mathrm{ref}}, w_n^{\mathrm{ref}}$ are respectively the crosstalk channel from the n-th modem into the reference line, the channel attenuation, the power, the noise and the weight of the reference line. These values need to be chosen in advance and are constants. Every modem n has to solve optimization problem (3.24).

$$\begin{aligned}
\max_{s_1^n,\ldots,s_K^n} \quad & w_n f_s \sum_k \log_2\left(1+\frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^n}{\sum_{m\neq n}|h_k^{n,m}|^2 s_k^m+\sigma_k^n}\right) \\
& \underbrace{+w_n^{\mathrm{ref}} f_s \sum_k \log_2\left(1+\frac{1}{\Gamma}\frac{|h_k^{\mathrm{ref}}|^2 s_k^{\mathrm{ref}}}{|h_k^{n,\mathrm{ref}}|^2 s_k^n+\sigma_k^{\mathrm{ref}}}\right)}_{T_n} \\
\text{s.t.} \quad & \sum_k s_k^n \leq P^{n,\mathrm{tot}}, \\
\text{s.t.} \quad & 0 \leq s_k^n \leq s_k^{n,\mathrm{mask}}, \forall k.
\end{aligned} \tag{3.24}$$

Although approximation (3.24) is not convex, the main observation of ASB is that it can be solved in an easy way. The KKT stationarity condition of

approximation (3.24) leads to the following equation:

$$\begin{aligned}\forall n,k: & \frac{w_n f_s |h_k^{n,n}|^2}{\log(2)(|h_k^{n,n}|^2 \underbrace{s_k^n}_{Q} + \sum_{m\neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n)} \\ & - \frac{w_n^{\text{ref}} f_s |h_k^{\text{ref}}|^2 s_k^{\text{ref}} \Gamma |h_k^{n,\text{ref}}|^2}{\log(2)(\Gamma |h_k^{n,\text{ref}}|^2 s_k^n + \Gamma\sigma_k^{\text{ref}})(|h_k^{\text{ref}}|^2 s_k^{\text{ref}} + \Gamma |h_k^{n,\text{ref}}|^2 s_k^n + \Gamma\sigma_k^{\text{ref}})} \\ & -\lambda_n - \nu_k^n + \mu_k^n = 0.\end{aligned} \tag{3.25}$$

In [12] the authors propose to write this equation as a cubic equation and to solve it by checking the three roots as well as the boundary solutions $s_k^n = 0$ and $s_k^n = s_k^{n,\text{mask}}$. In [12] an extension is proposed to include multiple reference lines. However for M reference lines, the ASB algorithm needs to solve a polynomial equation of degree $2M+1$ for each modem n on each tone k.

To reduce this complexity, we propose an alternative solution similarly to the approach of the previous sections. This will be referred to as ASB-2. Taking into account the KKT complementarity conditions of (3.24), equation (3.25) can be reformulated as a fixed point equation such that s_k^n, indicated by the letter Q in formula (3.25), is isolated as follows:

$$s_k^n = \left[\frac{w_n f_s / \log(2)}{\lambda_n + P_k^{\text{ASB}-2,n}} - \frac{\sum_{m\neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n}{|h_k^{n,n}|^2} \right]_0^{s_k^{n,\text{mask}}} \tag{3.26}$$

$$P_k^{\text{ASB}-2,n} = \frac{w_n^{\text{ref}} f_s |h_k^{\text{ref}}|^2 s_k^{\text{ref}} |h_k^{n,\text{ref}}|^2}{\log(2)(|h_k^{n,\text{ref}}|^2 s_k^n + \sigma_k^{\text{ref}})(|h_k^{\text{ref}}|^2 s_k^{\text{ref}} + \Gamma |h_k^{n,\text{ref}}|^2 s_k^n + \Gamma\sigma_k^{\text{ref}})}$$

$P_k^{\text{ASB}-2,n}$ is here expressed for 1 reference line but it is easy to extend to M reference lines. Note that this fixed point equation does not necessarily converge to the same solution as the solution proposed in [12]. Every modem iteratively applies formula (3.26) and bisects on λ_n. If the reference lines are fixed for every modem, the algorithm can be executed in an autonomous manner which is a great asset for practical implementation. This leads to Algorithm 3.3. If line 8 of Algorithm 3.3 is changed by checking the three roots of cubic equation (3.25) and the boundary solutions $s_k^n = 0$ and $s_k^n = s_k^{n,\text{mask}}$, the ASB algorithm from [12] is obtained.

Algorithm 3.3 ASB-2

Modem n algorithm:
Initialize all $s_k^n = s_k^{n,\text{mask}}/2$, $\forall k$
Initialize all reference line parameters
loop {execute at regular intervals}
  $\lambda_n^{\min} = 0, \lambda_n^{\max} = \Lambda^{\max}$, $\lambda_n = (\lambda_n^{\max} + \lambda_n^{\min})/2$
  while $|\sum_k s_k^n - P^{n,\text{tot}}| > \delta$ and $\lambda_n > \gamma$ **do**
    $\lambda_n = (\lambda_n^{\max} + \lambda_n^{\min})/2$
    Update s_k^n using (3.26), $\forall k$
    if $\sum_k s_k^n > P^{n,\text{tot}}$ **then**
      $\lambda_n^{\min} = \lambda_n$
    else
      $\lambda_n^{\max} = \lambda_n$
    end if
  end while
end loop

3.5 Novel Distributed Spectrum Management Algorithms

In this section three novel distributed spectrum management algorithms are proposed within the same KKT framework introduced in the previous section. The focus is on a distributed approach where the modems execute a local spectrum management algorithm and where in addition, some limited message-passing is allowed between the modems and the SMC. Such an approach is interesting from a practical point of view. It combines a fast local algorithm adapting fast to channel and noise changes with the operation of the SMC to infrequently steer the local algorithms so that the network converges to a better performance. A similar approach is also used in [19],[20] and [13].

As explained in the beginning of Section 3.4, the choice of the type of (iterative) approximations has an influence on many factors like complexity, (sub-)optimality, convergence properties, communication overhead, convex rate region, practical implementation, ... One constraint for our spectrum management algorithms is that we want the local algorithms to be water-filling type procedures based on formulas (3.27) or (3.28)

$$s_k^n = \frac{w_n f_s}{\log(2)\lambda_n \delta_k^n} - \frac{\Gamma(\sum_{m\neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n)}{|h_k^{n,n}|^2} \tag{3.27}$$

$$s_k^n = \frac{w_n f_s}{\log(2)\lambda_n + \epsilon_k^n} - \frac{\Gamma(\sum_{m\neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n)}{|h_k^{n,n}|^2} \tag{3.28}$$

where δ_k^n and ϵ_k^n are constants and where (3.28) can be obtained from (3.27)

through a simple transformation as follows

$$\delta_k^n = 1 + \frac{\epsilon_k^n}{\log(2)\lambda_n}. \tag{3.29}$$

Water-filling formulas are well-established and understood in the communications field. Many discrete bit loading algorithms are based on water-filling type of procedures. In Section 3.5.4 it will be shown how the MS-DSB algorithm proposed in Section 3.5.3 can be used together with existing (water-filling inspired) fast bit loading algorithms [19].

Finally the convergence and complexity of the proposed algorithms is discussed.

3.5.1 Convex Approximation Distributed Spectrum Balancing (CA-DSB)

The first spectrum management algorithm is based on an iterative convex approximation approach similar to SCALE. However it chooses a different convex approximation so that a water-filling type of formula is obtained. This approach is also described in [16]. Here it will be presented from a KKT perspective.

Note that the spectrum management problem (3.2) can be reformulated as follows

$$\begin{aligned}
\max_{\mathbf{s}^1,\ldots,\mathbf{s}^N} \quad & \sum_n w_n f_s \sum_k \log_2\Big(\sum_{m=1}^{N} |\tilde{h}_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n\Big) \\
& \underbrace{- \sum_n w_n f_s \sum_k \log_2\Big(\Gamma\Big(\sum_{m\neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n\Big)\Big)}_{\text{A=nonconcave part}} \\
\text{s.t.} \quad & \sum_k s_k^n \leq P^{n,\text{tot}} \quad , \quad \forall n, \\
\text{s.t.} \quad & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \quad \forall n, \forall k, \\
\text{with} \quad & |\tilde{h}_k^{n,m}|^2 \begin{cases} = \Gamma |h_k^{n,m}|^2 & , n \neq m \\ = |h_k^{n,n}|^2 & , n = m. \end{cases}
\end{aligned} \tag{3.30}$$

The basic idea is to approximate the nonconcave (convex) part A of the objective function of (3.30) in a point $\mathbf{s}_{k,\text{ap}}$ by a lower bound hyperplane as follows

$$-\Big(\sum_{m\neq n} a_k^{m,n} s_k^m + c_k^n\Big) \leq \underbrace{-\log_2\Big(\Gamma\Big(\sum_{m\neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n\Big)\Big)}_{A_k^n} \tag{3.31}$$

with equality in point $\mathbf{s}_{k,\text{ap}}$.

The approximation parameters $c_k^n, a_k^{m,n}(\forall m)$ are readily obtained based on solving a linear system of N equations in N unknowns or based on the gradient

and function value of A_k^n (3.31) at $\mathbf{s}_{k,\mathrm{ap}}$. This lower bound approximation leads to the following convex optimization problem:

$$\begin{aligned}\max_{\mathbf{s}^1,\ldots,\mathbf{s}^N} \quad & \sum_n w_n f_s \sum_k \log_2\Big(\sum_{m=1}^{N} |\tilde{h}_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n\Big) \\ & \qquad -\sum_n w_n f_s \sum_k \Big(\sum_{m\neq n} a_k^{m,n} s_k^m + c_k^n\Big) \\ \text{s.t.} \quad & \sum_k s_k^n \leq P^{n,\mathrm{tot}} \quad , \quad \forall n, \\ \text{s.t.} \quad & 0 \leq s_k^n \leq s_k^{n,\mathrm{mask}}, \quad \forall n, \forall k.\end{aligned} \tag{3.32}$$

By iteratively solving problem (3.32), obtaining $\mathbf{s}_k^* \ \forall k$, and updating the approximation parameters $a_k^{m,n}, c_k^n$ for $\mathbf{s}_{k,\mathrm{ap}} = \mathbf{s}_k^*$ using (3.31), the sequence of convex subproblems produces a monotonically increasing objective value which converges to a locally optimal solution of the spectrum management problem (3.2).

The KKT conditions of (3.32), in particular the KKT stationarity condition and complementarity conditions, can be used to derive the following fixed point equation to find the spectra that solve the convex approximation (3.32):

$$s_k^n = \left[\frac{w_n f_s/\log(2)}{\underbrace{\lambda_n + \sum_{m\neq n} \omega_m f_s a_k^{n,m} - \sum_{m\neq n} \frac{w_m f_s \Gamma |h_k^{m,n}|^2/\log(2)}{\sum_p |\tilde{h}_k^{m,p}|^2 s_k^p + \Gamma\sigma_k^m}}_{P_k^{\mathrm{CA-DSB},n}}} - \frac{\sum_{m\neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n}{|h_k^{n,n}|^2} \right]_0^{s_k^{n,\mathrm{mask}}} . \tag{3.33}$$

Note that formula (3.33) is a water-filling type of equation, more specifically it corresponds to formula (3.28) where $\epsilon_k^n = P_k^{\mathrm{CA-DSB},n}$. Although the term $P_k^{\mathrm{CA-DSB},n}$ depends on $s_k^n, \forall n$ it is viewed as a constant which is updated only infrequently by the SMC. Every modem iteratively applies formula (3.33) where $P_k^{\mathrm{CA-DSB},n}$ is a constant and bisects on λ_n, as shown in Algorithm 3.4. Note that only one fixed point iteration is performed for every λ_n. In addition, the approximation parameters $a_k^{m,n}, c_k^n$ and the term $P_k^{\mathrm{CA-DSB},n}$ are updated infrequently (3.31) by the SMC to converge to a locally optimal solution of the spectrum management problem. The algorithm will be referred to as CA-DSB, i.e. convex approximation distributed spectrum balancing.

Algorithm 3.4 CA-DSB

Modem n algorithm:
Initialize $s_k^n = s_k^{n,\text{mask}}/2$, $\forall k$
loop {execute at regular intervals}
  $\lambda_n^{\min} = 0, \lambda_n^{\max} = \Lambda^{\max}$, $\lambda_n = (\lambda_n^{\max} + \lambda_n^{\min})/2$
  Receive messages $P_k^{\text{CA}-\text{DSB},n}$ from SMC, $\forall k$
  while $|\sum_k s_k^n - P^{n,\text{tot}}| > \delta$ and $\lambda_n > \gamma$ **do**
    $\lambda_n = (\lambda_n^{\max} + \lambda_n^{\min})/2$
    Update s_k^n using (3.33), $\forall k$
    if $\sum_k s_k^n > P^{n,\text{tot}}$ **then**
      $\lambda_n^{\min} = \lambda_n$
    else
      $\lambda_n^{\max} = \lambda_n$
    end if
  end while
  Measure $\text{int}_k^n = \sum_{p \neq n} |h_k^{n,p}|^2 s_k^p + \sigma_k^p$, $\forall k$
  Transmit int_k^n, s_k^n to SMC, $\forall k$
end loop
SMC algorithm:
loop
  Receive messages int_k^n, s_k^n from modems n, $\forall k$
  Update $a_k^{n,m}$ by solving linear system (3.31) using int_k^n, s_k^n, $\forall k$
  Compute messages $P_k^{\text{CA}-\text{DSB},n}$ and send to each modem n, $\forall k$.
end loop

3.5.2 Distributed Spectrum Balancing (DSB)

The second distributed spectrum management algorithm starts from the non-convex spectrum management problem (3.2) directly without involving any (iterative) approximation. Based on the KKT conditions of (3.2) a distributed solution is derived. It is known that the KKT conditions are necessary for a solution in nonlinear programming to be locally optimal. Here this will be used the other way around. There will be looked for a solution satisfying the KKT conditions which may correspond to a local or a global optimum.

The starting point is the KKT stationarity condition of the reformulated spectrum management problem (3.30):

$$\begin{aligned}\forall n,k: \frac{w_n f_s |h_k^{n,n}|^2}{\log(2)(|h_k^{n,n}|^2 \underbrace{s_k^n}_{Q} + \sum_{m\neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n)} \\ + \sum_{m\neq n} \frac{w_m f_s \Gamma |h_k^{m,n}|^2/\log(2)}{\sum_p |\tilde{h}_k^{m,p}|^2 s_k^p + \Gamma\sigma_k^m} \\ - \sum_{m\neq n} \frac{w_m f_s \Gamma |h_k^{m,n}|^2/\log(2)}{\sum_{p\neq m} \Gamma |h_k^{m,p}|^2 s_k^p + \Gamma\sigma_k^m} - \lambda_n - \nu_k^n + \mu_k^n = 0.\end{aligned} \tag{3.34}$$

By using the KKT complementarity conditions, equation (3.34) can be reformulated as a fixed point equation such that s_k^n, indicated by the letter Q in formula (3.34), is isolated as follows:

$$s_k^n = \left[\frac{w_n f_s/\log(2)}{\lambda_n + P_k^{\mathrm{DSB}_1,n} - P_k^{\mathrm{DSB}_2,n}} - \frac{\sum_{m\neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n}{|h_k^{n,n}|^2} \right]_0^{s_k^{n,\mathrm{mask}}} \tag{3.35}$$

with $P_k^{\mathrm{DSB}_1,n} = \sum_{m\neq n} \frac{w_m f_s \Gamma |h_k^{m,n}|^2/\log(2)}{\sum_{p\neq m} \Gamma |h_k^{m,p}|^2 s_k^p + \Gamma\sigma_k^m}$

and $P_k^{\mathrm{DSB}_2,n} = \sum_{m\neq n} \frac{w_m f_s \Gamma |h_k^{m,n}|^2/\log(2)}{\sum_p |\tilde{h}_k^{m,p}|^2 s_k^p + \Gamma\sigma_k^m}$.

Note that formula (3.35) is a water-filling type of equation, more specifically it corresponds to formula (3.28) where $\epsilon_k^n = P_k^{\mathrm{DSB}_1,n} - P_k^{\mathrm{DSB}_2,n}$. Although the term $P_k^{\mathrm{DSB}_1,n} - P_k^{\mathrm{DSB}_2,n}$ depends on $s_k^n, \forall n$ it is viewed as a constant which is updated only infrequently by the SMC. Every modem iteratively applies formula (3.35) and bisects on λ_n.

Note that formula (3.35) can be reformulated as follows

$$s_k^n = \left[\frac{w_n f_s / \log(2)}{\lambda_n + \sum_{m \neq n} \frac{w_m f_s \Gamma |h_k^{m,n}|^2}{\log(2)} (\frac{1}{\text{int}_k^m} - \frac{1}{\text{rec}_k^m})} - \frac{\text{int}_k^n}{|h_k^{n,n}|^2} \right]_0^{s_k^{n,\text{mask}}}$$

where (3.36)

$$\text{int}_k^n = \sum_{m \neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma \sigma_k^n$$

$$\text{rec}_k^n = \sum_m |\tilde{h}_k^{n,m}|^2 s_k^m + \Gamma \sigma_k^n$$

The quantity int_k^n is the received interference and noise of modem n on tone k and the quantity rec_k^n is the received signal of modem n on tone k. Both quantities are already measured by current state-of-the-art modems. The algorithm is shown in Algorithm 3.5 and it will be referred to as DSB, i.e. distributed spectrum balancing. A very similar algorithm, called Modified Iterative Water-filling (MIW), is recently proposed in [21]. MIW and DSB have the same transmit power update formulas and so use the same concept of water-filling in addition with a tone dependent penalty, called taxation term in [21]. Furthermore, MIW extends the focus to the asynchronous DSL transmission case [22], where it is shown that MIW, and so also DSB, can simply be extended to incorporate intercarrier interference (ICI). DSB incorporates spectral mask constraints and so MIW can similarly be extended to incorporate these extra constraints.

Note that there are a number of correspondences between SCALE, CA-DSB and DSB. Term $P_k^{\text{DSB}_1,n}$ of (3.35) is very similar to term $P_k^{\text{SCALE},n}$ of the SCALE formula (3.22). Furthermore term $P_k^{\text{DSB}_2,n}$ of (3.35) is the same as the second term of $P_k^{\text{CA-DSB},n}$ of CA-DSB (3.33). In fact if we use $a_k^{n,m} = \frac{\Gamma |h_k^{m,n}|^2 / \log(2)}{\sum_{p \neq m} \Gamma |h_k^{m,p}|^2 s_k^p + \Gamma \sigma_k^m}$ in (3.33) then term $P_k^{\text{DSB}_1,n} - P_k^{\text{DSB}_2,n}$ of (3.35) equals term $P_k^{\text{CA-DSB},n}$ of (3.33).

Finally note that if in ASB the T_n of (3.24) is chosen as $\sum_{m \neq n} w_m R^m$, the DSB algorithm is obtained. This fact can be used to develop good heuristics for choosing the reference line. One should tune the parameters of the reference line for ASB so that it optimally follows $\sum_{m \neq n} w_m R^m$ for varying s_k^n , $\forall k$.

3.5.3 Multiple Starting Point Distributed Spectrum Balancing (MS-DSB)

The third spectrum management algorithm is an extension of DSB with multiple starting points. The distributed algorithms SCALE, CA-DSB and DSB are algorithms that converge to a locally optimal solution of problem (3.2). The

Algorithm 3.5 DSB

Modem n algorithm:
Initialize all $s_k^n = s_k^{n,\text{mask}}/2$, $\forall k$
loop {execute at regular intervals}
  $\lambda_n^{\min} = 0, \lambda_n^{\max} = \Lambda^{\max}$, $\lambda_n = (\lambda_n^{\max} + \lambda_n^{\min})/2$
  Receive messages $(P_k^{\text{DSB}_1,n} - P_k^{\text{DSB}_2,n})$ from SMC, $\forall k$
  while $|\sum_k s_k^n - P^{n,\text{tot}}| > \delta$ and $\lambda_n > \gamma$ **do**
    $\lambda_n = (\lambda_n^{\max} + \lambda_n^{\min})/2$
    Update s_k^n using (3.35), $\forall k$
    if $\sum_k s_k^n > P^{n,\text{tot}}$ **then**
      $\lambda_n^{\min} = \lambda_n$
    else
      $\lambda_n^{\max} = \lambda_n$
    end if
  end while
  Measure $\text{int}_k^n = \sum_{p \neq n} |h_k^{n,p}|^2 s_k^p + \sigma_k^p$, $\forall k$
  Compute messages $\left(\frac{1}{\text{int}_k^n} - \frac{1}{\text{rec}_k^n}\right)$,$\forall k$, and send to SMC
end loop
SMC algorithm:
loop
  Receive messages $\left(\frac{1}{\text{int}_k^n} - \frac{1}{\text{rec}_k^n}\right)$ from modems n, $\forall k$
  Compute messages $(P_k^{\text{DSB}_1,n} - P_k^{\text{DSB}_2,n})$ and send to each modem n, $\forall k$.
end loop

spectrum management problem (3.2) is a nonconvex optimization problem that can have multiple locally optimal solutions. These solutions can differ significantly in obtained weighted data rate sum. This means that locally optimal algorithms like SCALE, CA-DSB and DSB can sometimes perform quite suboptimally compared to the global optimum, depending on the DSL scenario.

An important factor that determines the obtained locally optimal solution is the initial transmit power spectrum. For SCALE as presented in [13], it was suggested to start from initial transmit power spectra equal to zero $s_k^n = 0,\ \forall n, k$. For CA-DSB and DSB, it is suggested to start from initial transmit power spectra $s_k^n = s_k^{n,\text{mask}}/2,\ \forall n, k$. The heuristic for this initial choice is that if there is a uniform distribution for the possible power loadings, the maximum difference between the initial power loading and the obtained power loading is minimized on average. Note that if there are no spectral mask constraints, one can put $s_k^{n,\text{mask}} = P^{n,\text{tot}}$.

In simulation Section 3.6.1 it will be observed that CA-DSB and DSB, with the same initial transmit power spectra, generally converge to the same solution. In some cases the SCALE algorithm gives a different solution, meaning it converges to a different locally optimal solution. If the initial power spectra are chosen equally for the three algorithms, they generally converge to the same solutions. This is confirmed by many simulations. This means that SCALE, CA-DSB and DSB converge to locally optimal solutions which depend on the initial power spectra.

Based on the above observations, we propose to combine DSB with a multiple starting point approach. For each tone, multiple initial transmit powers are chosen and for each choice the transmit powers are updated iteratively so as to converge to a local optimum. The best of the converged results is then chosen for each tone independently. By using multiple initial transmit powers it is more likely that the global optimum for that tone will be found. In order to keep the complexity low, we propose to use $N + 2$ initial transmit power sets. These initial transmit power sets are given in Table 3.1 where $\mathbf{0}$ is the zero vector of dimension N, $\mathbf{s}_k^{\text{mask}}$ is a vector of dimension N with the corresponding spectral masks on tone k for the N modems and $\mathbf{e_n}.s_k^{n,\text{mask}}$ is the unit vector in the n-th dimension multiplied by the corresponding spectral mask. The first transmit power set corresponds to the initial transmit power set of the SCALE algorithm. The second transmit power set corresponds to the initial transmit power set of the CA-DSB and DSB algorithms. This choice ensures that the obtained local optimum on tone k is at least as good as those obtained by the SCALE, CA-DSB and DSB algorithms. The remaining N transmit power sets are chosen to correspond to initial transmit power sets where only one modem is active in each tone k. The reasoning behind this choice is that we observe that for large crosstalk the locally optimal solutions can get isolated along the axes. A fully random approach would then fail to find these isolated locally

optimal points unless many (i.e. many more than N+2) initial points are chosen randomly. A recent work [23] actually proves that in the case of large crosstalk, the optimal solution converges to an FDMA solution where only one user is active in each tone k.

Table 3.1: $N+2$ initial transmit powers for each tone k used by MS-DSB

1	$\mathbf{0}$
2	$\mathbf{s}_k^{\text{mask}}/2$
$3 \ldots (N+2)$	$\mathbf{e_n}.s_k^{n,\text{mask}}$, for $n = 1 \ldots N$

Algorithm 3.6 MS-DSB

```
1: SMC algorithm:
2: loop {Execute at regular intervals}
3:    Receive messages (1/int_k^n − 1/rec_k^n) from modems n, ∀k
4:    Compute messages (P_k^{DSB1,n} − P_k^{DSB2,n}) ∀n, k.
5:    while complementary slackness conditions (3.11) not satisfied do
6:       update Lagrange multipliers using (3.18)
7:       for all tones k do
8:          for each initial transmit power set do
9:             for iterations do
10:               for each modem do
11:                  Apply formula (3.35)
12:               end for
13:            end for
14:         end for
15:         Take best of all converged solutions
16:      end for
17:   end while
18:   Compute per-tone dependent penalties P_k^{DSB1,n} − P_k^{DSB2,n}, ∀n, k
19:   Send P_k^{DSB1,n} − P_k^{DSB2,n} to all modems n
20: end loop
21: Modem n algorithm:
22: loop
23:   Apply modem part of DSB algorithm
24: end loop
```

This leads to Algorithm 3.6. Each modem executes the modem part of the DSB algorithm locally so that it can react fast to channel and noise changes. The SMC executes the multiple starting point extension. The SMC regularly receives messages $\left(\frac{1}{\text{int}_k^n} - \frac{1}{\text{rec}_k^n}\right)$ from the N modems and computes the quantities $P_k^{\text{DSB}_1,n} - P_k^{\text{DSB}_2,n}$. The loop at line 5 will search for the Lagrange multipliers so that the active power constraints are satisfied, using a subgradient approach

at line 6 as in formula (3.18) (see also [7][17]). Note that all the Lagrange multipliers are updated in parallel. Then for each tone and for each initial transmit power set, the modems iterate using the DSB formula (3.35). The best of the converged results is used. When the loop is converged the quantities $P_k^{\mathrm{DSB}_1,n} - P_k^{\mathrm{DSB}_2,n}$ are communicated to the modems so that they can steer their transmit spectra to improve the network performance. This algorithm will be referred to as MS-DSB, i.e. multiple starting point distributed spectrum balancing.

3.5.4 Per-Tone Penalty Water-filling Algorithms

The water-filling formula (3.16) of Section 3.4.1 is a well-established formula in the communications field. In this formula, the Lagrange multiplier λ_n can be seen as a penalty for modem n: the larger λ_n, the smaller the power allocated to the n-th modem. The interesting part is that this penalty λ_n is equal for all the tones k of modem n. Formulas (3.27) and (3.28) can be seen as water-filling formulas with the addition of a *per-tone penalty*. For each modem n, each tone k is penalized differently. In formula (3.27) this is done by a penalizing scaling factor whereas in formula (3.28) this is done by a penalizing term. The CA-DSB, DSB and ASB-2 formulas, (3.33), (3.35) and (3.26) respectively, can be viewed as water-filling formulas with the addition of a *per-tone penalty term*. These penalties are summarized in Table 3.2.

In [19] the scaled iterative water-filling (SIW) algorithm is proposed which is based on per-tone scaling factors. Furthermore a fast bit loading implementation is proposed. The MS-DSB algorithm can be used in combination with this fast bit loading implementation as follows. Based on the per-tone penalty terms the equivalent per-tone scaling factors can be obtained by following closed form formula

$$\delta_k^n = 1 + \frac{(P_k^{\mathrm{DSB}_1,n} - P_k^{\mathrm{DSB}_2,n})}{\lambda_n} \quad \forall k, n. \tag{3.37}$$

Note that this per-tone scaling factor depends on the Lagrange multiplier λ_n. This information is available at the SMC and so the corresponding scaling factors can be computed. These scaling factors are regularly updated and communicated to the modems where the fast bit loading algorithms [19] are implemented. This fits well in the band-preference framework proposed in [19] and [20]. In Algorithm 3.7 this bit loading version of the MS-DSB algorithm is shown.

To reduce the number of communication messages between the SMC and the modems a single scaling factor can be computed for a group of adjacent tones which are likely to have similar properties in a DSL channel. This was also proposed in [19]. The total number of tones K can be grouped into M bands. For each band only one scaling factor is computed which represents all the K/M tones in this band. This approximation can be improved if a curve of higher

Table 3.2: Water-filling penalties for distributed algorithms

	per-user penalty	per-tone penalty term
IW	λ_n	0
CA-DSB	λ_n	$P_k^{\mathrm{CA-DSB},n}$
DSB	λ_n	$P_k^{\mathrm{DSB}_1,n} - P_k^{\mathrm{DSB}_2,n}$
ASB-2	λ_n	$P_k^{\mathrm{ASB-2},n}$

degree is fitted for the tones within one band, leading to a slightly increased message-passing.

Algorithm 3.7 MS-DSB (bit loading version)

SMC algorithm:
loop {Execute at regular intervals}
  Receive messages $\left(\frac{1}{\mathrm{int}_k^n} - \frac{1}{\mathrm{rec}_k^n}\right)$ from modems n, $\forall k$
  Compute messages $(P_k^{\mathrm{DSB}_1,n} - P_k^{\mathrm{DSB}_2,n})$ $\forall n, k$.
  while complementary slackness conditions (3.11) not satisfied **do**
    update Lagrange multipliers using (3.18)
    for all tones k **do**
      for each initial transmit power set **do**
        for iterations **do**
          for each modem **do**
            Apply formula (3.35)
          end for
        end for
      end for
      Take best of all converged solutions
    end for
  end while
  Compute DSB scaling factors δ_k^n (3.37)
  Send δ_k^n to all modems n, $\forall k$
end loop
Modem n algorithm:
loop
  Apply fast scaled bit loading algorithm [19]
  At regular intervals:
  Receive scaling factors $\delta_k^n, \forall k$
  Transmit $\left(\frac{1}{\mathrm{int}_k^n} - \frac{1}{\mathrm{rec}_k^n}\right)$ $\forall k$ to SMC
end loop

3.5.5 Complexity Analysis

The complexity of the proposed algorithms can be separated into the message-passing communication overhead and the complexity of the power updates by the modems.

The complexity of the power updates refers to the complexity of the modems to converge to their transmit powers satisfying the total power constraints for fixed per-tone penalties, multiplied with the total number of updates of the per-tone penalties. For fixed per-tone penalties the CA-DSB, DSB and ASB-2 algorithms consist of two levels of cycles. An outer cycle iterates over the modems and an inner cycle bisects on λ_n for each modem until total power constraints are satisfied. The number of outer cycles ν is typically only 3 for the transmit powers to converge to within 1% of the previous cycle. The number of inner iterations (bisections on Lagrange multipliers λ_n) is denoted as I_b. In order to achieve an accurate precision, I_b can be fixed at 34 similarly to the approach in [12]. Furthermore K transmit powers have to be computed for a given λ_n. In order to converge to the powers satisfying the total power constraints for fixed per-tone penalties, this leads to a complexity of $\nu I_b NK$ times updating powers using (3.33),(3.35) or (3.26) for CA-DSB, DSB or ASB-2 respectively.

For the CA-DSB and DSB algorithms this complexity has to be multiplied with the total number of updates of the per-tone penalties which will be referred to as L. We assume that the per-tone penalties of all the modems are updated at the same time. L depends on the scenario. For small crosstalk scenarios L is typically only 2. Furthermore this number does not increase with the number of modems if small crosstalk modems are added. For large crosstalk scenarios L is typically 20-40 and there is a small increase in L if strong crosstalk modems are added to the cable bundle. Note that the updating of the per-tone penalties corresponds to the act of iteratively linearizing the nonconcave part of (3.30). From the optimization perspective, such strategy is known to have sublinear convergence. However for the considered DSL scenarios the convergence is quite satisfying. This leads to a total complexity of $\nu I_b LNK$ times updating powers using (3.33) or (3.35) for CA-DSB or DSB respectively.

For the MS-DSB algorithm the main computation happens in the SMC. A difference is that the Lagrange multipliers λ_n $\forall n$ are updated in parallel which typically requires less than 50 updates [7]. This leads to a complexity of $\mu ILKN(N+2)$ times (3.35) where $(N+2)$ refers to the number of initial transmit power sets, I refers to the number of Lagrange multiplier updates which is typically less than 50 and μ is the number of iterations over the users within one tone as shown in Algorithm 3.6 which is typically only 3.

Secondly there is a message-passing overhead for CA-DSB, DSB and MS-DSB.

Table 3.3: Algorithm complexity and communication overhead for distributed spectrum management algorithms where ν:no. of outer cycles over modems , I_b:no. of bisections on Lagrange multiplier λ_n, K:no. of tones, N:no. of modems, L:no. of updates of per-tone penalties, μ:no. of iterations over modems within one tone, I:no. of centralized Lagrange multiplier updates. The unity of the communication overhead is the number of messages. The unity of the algorithm complexity is the number of evaluations of the corresponding power update formula (x) in order to converge.

Algorithm	Complexity [no. of power updates using (x)]	Communic. Overhead [no. of messages]
IW	$\nu I_b KN$ times (3.16)	0
SCALE	$\nu I_b LKN$ times (3.22)	$2LKN$
ASB-2	$\nu I_b KN$ times (3.26)	0
CA-DSB	$\nu I_b LKN$ times (3.33)	$3LKN$
DSB	$\nu I_b LKN$ times (3.35)	$2LKN$
MS-DSB	$\mu ILKN(N+2)$ times (3.35)	$2LKN$

In a system where all the modems update their per-tone penalties at the same time the total communication complexity for DSB and MS-DSB will be $2LNK$ messages. This can be explained as follows. Between each modem and the SMC K messages have to be communicated in both directions for updating the K per-tone penalties, leading to $2K$ messages. For N users this leads to $2NK$ messages. Furthermore there are typically L updates of the per-tone penalties as explained in the previous paragraphs, leading to $2LNK$ messages. For CA-DSB there are two messages to the SMC instead of one (line 16 Algorithm 3.4) leading to a total of $3LNK$ messages.

This is summarized in Table 3.3. Note that for IW and SCALE the implementations of Algorithm 3.1 and 3.2 are used respectively.

3.5.6 Convergence Analysis

First the convergence of the CA-DSB algorithm is discussed. This algorithm basically consists of two parts. The first part is a message-passing protocol that updates the approximation parameters leading to iterative convex approximations. The second part is an iterative update of each users' transmit powers for fixed per-tone penalties which corresponds to lines 6-14 of Algorithm 3.4. In the following theorems the convergence is discussed of these two parts.

Theorem 3.1 *By iteratively solving convex subproblem (3.32) and updating the approximation parameters $a_k^{m,n}$ using (3.31), the sequence of convex subproblems has a monotonically increasing objective value which converges to a locally optimal solution of the spectrum management problem (3.2).*

Proof : [Proof of Theorem 3.1] Reformulation (3.30) of the spectrum management problem can be shortly formulated as follows

$$\max_{\mathbf{s}\in\Omega} h_1(\mathbf{s}) - h_2(\mathbf{s}) \quad \equiv \quad \min_{\mathbf{s}\in\Omega} g_1(\mathbf{s}) - g_2(\mathbf{s}) \quad \equiv \quad \min_{\mathbf{s}\in\Omega} f(\mathbf{s}), \tag{3.38}$$

where $h_1(\mathbf{s})$ and $h_2(\mathbf{s})$ are concave and respectively correspond to the first and second part of the objective function of (3.30). Furthermore $g_1(\mathbf{s})$ and $g_2(\mathbf{s})$ are convex where $g_i(\mathbf{s}) = -h_i(\mathbf{s})$ and $f(\mathbf{s})$ is nonconvex where $f(\mathbf{s}) = g_1(\mathbf{s}) - g_2(\mathbf{s})$. The constraint set is a convex set $\Omega = \{\mathbf{s} \in \mathbf{R}^{NK} | \sum_{k=1}^{K} s_k^n \leq P^{n,\text{tot}} \ \forall n \ ; \ 0 \leq s_k^n \leq s_k^{n,\text{mask}} \ \forall n, k\}$.

The proposed lower bound convex approximation by CA-DSB (3.32) can be formulated as

$$\min_{\mathbf{s}\in\Omega} g_1(\mathbf{s}) - (g_2(\mathbf{s}_{\text{ap}}) + \nabla g_2(\mathbf{s}_{\text{ap}})^T(\mathbf{s} - \mathbf{s}_{\text{ap}})),$$

where $\mathbf{s}_{\text{ap}}$ is the approximation point.

We make the technical assumption that $g_1(\mathbf{s})$ has an upper bound on its curvature, i.e. $\|\nabla^2 g_1(\mathbf{s})\|_2 \leq C$ for all $\mathbf{s} \in \Omega$. This implies that $g_1(\mathbf{s}+\mathbf{p}) \leq g_1(\mathbf{s}) + \nabla g_1(\mathbf{s})^T\mathbf{p} + \frac{C}{2}\|\mathbf{p}\|_2^2$. Due to convexity of $g_2(\mathbf{s})$, we also have

$$f(\mathbf{s}+\mathbf{p}) \leq f(\mathbf{s}) + \nabla f(\mathbf{s})^T\mathbf{p} + \frac{C}{2}\|\mathbf{p}\|_2^2. \tag{3.39}$$

The key goal is to prove that the sequence $\mathbf{s}(t+1)$ for $t = 1, 2, \ldots$ produces a monotonically decreasing objective value which converges to a local optimum of (3.38), where $\mathbf{s}(t+1)$ is defined as follows

$$\mathbf{s}(t+1) = \arg\min_{\mathbf{s}\in\Omega} g_1(\mathbf{s}) - \Big(g_2\big(\mathbf{s}(t)\big) + \nabla g_2\big(\mathbf{s}(t)\big)^T\big(\mathbf{s} - \mathbf{s}(t)\big)\Big),$$

and where $\mathbf{s}(t)$ is the value of the transmit powers at iteration t.

To this end we prove that (a) the sequence is monotonically decreasing $f(\mathbf{s}(t+1)) \leq f(\mathbf{s}(t))$ and that (b) it converges to a local optimum of (3.38).

(a):

$$\begin{aligned}
&f(\mathbf{s}(t)) \\
&= g_1\big(\mathbf{s}(t)\big) - g_2\big(\mathbf{s}(t)\big) \\
&= g_1\big(\mathbf{s}(t)\big) - \Big(g_2\big(\mathbf{s}(t)\big) + \frac{\partial g_2\big(\mathbf{s}(t)\big)}{\partial \mathbf{s}}\big(\mathbf{s}(t) - \mathbf{s}(t)\big)\Big) \\
&\geq \min_{\mathbf{s}\in\Omega} g_1(\mathbf{s}) - \Big(g_2\big(\mathbf{s}(t)\big) + \frac{\partial g_2\big(\mathbf{s}(t)\big)}{\partial \mathbf{s}}\big(\mathbf{s} - \mathbf{s}(t)\big)\Big) \\
&= g_1\big(\mathbf{s}(t+1)\big) - \Big(g_2\big(\mathbf{s}(t)\big) + \frac{\partial g_2\big(\mathbf{s}(t)\big)}{\partial \mathbf{s}}\big(\mathbf{s}(t+1) - \mathbf{s}(t)\big)\Big)
\end{aligned}$$

$$\geq g_1\big(\mathbf{s}(t+1)\big) - g_2\big(\mathbf{s}(t+1)\big) = f(\mathbf{s}(t+1)) \tag{3.40}$$

where (3.40) uses the lower bound property (3.31).

If the sequence is bounded below, we have $f(\mathbf{s}(t)) \to f^*$. Due to continuity of f and compactness of Ω, the sequence $(\mathbf{s}(t))_{t\in\mathbb{N}}$ has at least one accumulation point. Let us call this $\mathbf{s}^*$.

(b): We now show that $\mathbf{s}^*$ is a local optimum for (3.38). Due to convexity of Ω, the necessary conditions for a point $\mathbf{s}^*$ to be a local optimum for (3.38) are

$$\forall \mathbf{y} \in \Omega: \quad \nabla f(\mathbf{s}^*)^T(\mathbf{y}-\mathbf{s}^*) \geq 0. \tag{3.41}$$

We now show that (3.41) holds. For this aim let us assume that (3.41) is NOT satisfied in the limit, i.e.

$$\exists \mathbf{y} \in \Omega: \quad \nabla f(\mathbf{s}^*)^T(\mathbf{y}-\mathbf{s}^*) < 0 \tag{3.42}$$

For this fixed $\mathbf{y} \in \Omega$, due to continuity of $\nabla f(\mathbf{s})^T(\mathbf{y}-\mathbf{s})$ w.r.t. $\mathbf{s}$, there is a neighborhood $\mathcal{N}$ of $\mathbf{s}^*$ and an $\epsilon > 0$ and $M > \frac{\epsilon}{C}$ so that for all $s \in \mathcal{N}$ holds that $\nabla f(\mathbf{s})^T(\mathbf{y}-\mathbf{s}) < -\epsilon$ and $\|\mathbf{y}-\mathbf{s}\|_2^2 < M$. Let us now regard the convex subproblem at any iterate $\mathbf{s}(t)$ in this neighborhood. It holds that

$$\begin{aligned}
&f(\mathbf{s}(t+1)) \\
&\leq \min_{\mathbf{s}\in\Omega} g_1(\mathbf{s}) - (g_2(\mathbf{s}(t)) + \nabla g_2(\mathbf{s}(t))^T(\mathbf{s}-\mathbf{s}(t))) \\
&\leq \min_{\mathbf{s}\in\Omega} f(\mathbf{s}(t)) + \nabla f(\mathbf{s}(t))^T(\mathbf{s}-\mathbf{s}(t))) + \frac{C}{2}\|\mathbf{s}-\mathbf{s}(t)\|_2^2 \\
&\leq f(\mathbf{s}(t)) + \nabla f(\mathbf{s}(t))^T(\mathbf{z}-\mathbf{s}(t)) + \frac{C}{2}\|\mathbf{z}-\mathbf{s}(t)\|_2^2
\end{aligned}$$

Where $\mathbf{z}$ can be an arbitrary point in Ω.

By taking a point $\mathbf{z} = \mathbf{s}(t) + \mu(\mathbf{y}-\mathbf{s}(t))$ for the $\mathbf{y}$ above, we obtain

$$\begin{aligned}
&f(\mathbf{s}(t+1)) \\
&\leq f(\mathbf{s}(t)) + \mu\nabla f(\mathbf{s}(t))^T(\mathbf{y}-\mathbf{s}(t)) + \frac{C}{2}\mu^2\|\mathbf{y}-\mathbf{s}(t)\|_2^2 \\
&\leq f(\mathbf{s}(t)) - \mu\epsilon + \frac{CM}{2}\mu^2
\end{aligned}$$

and we can minimize w.r.t. to μ, on the line, to yield

$$\begin{aligned}
f(\mathbf{s}(t+1)) &\leq \min_{\mu} f(\mathbf{s}(t)) - \mu\epsilon + \frac{CM}{2}\mu^2 \\
&= f(\mathbf{s}(t)) - \frac{\epsilon}{2CM}
\end{aligned}$$

where the minimum is attained at $\mu = \frac{\epsilon}{CM} \leq 1$. This constant decrease at any iterate $\mathbf{s}(t)$ in the neighborhood $\mathcal{N}$ (which contains infinitely iterates $\mathbf{s}(t)$ because $\mathbf{s}^*$ is an accumulation point) is a contradiction to $f(\mathbf{s}(t)) \rightarrow f^*$. This means (3.41) holds and thus the iterative approach converges to a stationary point of (3.38). □

Theorem 3.2 *For fixed per-tone penalties* $P_k^{\mathrm{CA-DSB},n}$, *if* $\max_{m \neq n,k} \Gamma |h_k^{n,m}|^2/|h_k^{n,n}|^2 < 1/(N-1)$, *then the sequential or parallel update of each users' transmit powers using (3.33) converges to the unique fixed point in an N-user system.*

Proof : [Proof of Theorem 3.2] The main observation here is that the update formula (3.33) of CA-DSB for fixed per-tone penalties has exactly the same structure as the update formula of the ASB-S2 algorithm of [12] for some arbitrary reference line. In [12] convergence conditions for the ASB-S2 are derived based on contraction mapping in the transmit power updates. For an N-user system it is proven that sequential or parallel updates of each users' transmit powers converge to the unique fixed point when $\max_{m \neq n,k} \Gamma |h_k^{n,m}|^2/|h_k^{n,n}|^2 < 1/(N-1)$. These conditions do not depend on the values of the per-tone penalties. So ASB-S2 and CA-DSB have a similar convergence behavior and theorem 3.2 is proven. □

In Section 3.5.2 it was mentioned that for $a_k^{n,m} = \frac{\Gamma |h_k^{m,n}|^2 / \log(2)}{\sum_{p \neq m} \Gamma |h_k^{m,p}|^2 s_k^p + \Gamma \sigma_k^m}$ the per-tone penalties of DSB and CA-DSB are exactly the same. In fact this is the closed form solution of the system of N equations that the SMC part of CA-DSB has to solve. So the DSB algorithm is similar to the CA-DSB algorithm except that part of the calculation of the approximation parameters is done in the modems. CA-DSB and DSB lead to the same per-tone penalties and so have similar convergence properties. MS-DSB is a multiple starting point extension of DSB and so has the same convergence properties.

As explained in [12], the convergence conditions are not satisfied for every possible DSL scenario. However it should be mentioned that convergence problems have never been encountered during extensive simulations of multi-user ADSL and VDSL scenarios (Section VI).

3.6 Simulation Results

In this section simulation results are shown for the distributed spectrum management algorithms. The following parameter settings are assumed for the DSL scenarios. The twisted pair lines have a diameter of 0.5 mm (24 AWG).

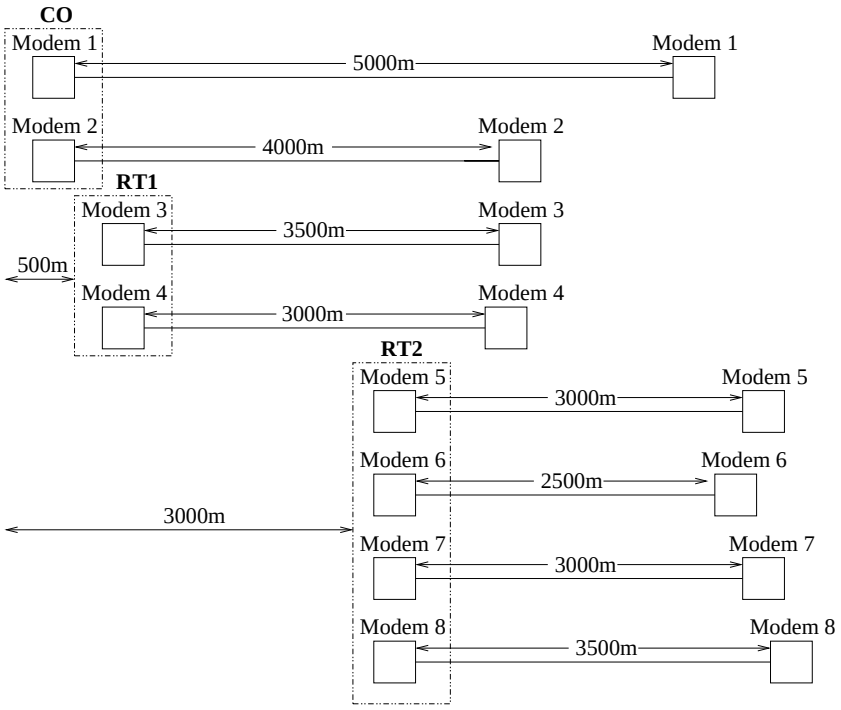


Figure 3.1: Asymmetric N-user ADSL simulation scenario with different loop lengths and different CO-RT distances

The maximum transmit power is 20.4 dBm [24] for the ADSL scenarios and 11.5 dBm for the VDSL scenarios. The SNR gap Γ is 12.9 dB, corresponding to a coding gain of 3 dB, a noise margin of 6 dB and a target symbol error probability of 10^{-7}. The tone spacing Δ_f is 4.3125 kHz. The DMT symbol rate f_s is 4 kHz. For all scenarios, the weights w_n are chosen equal for all users $(n = 1 \ldots N)$, namely $w_n = \frac{1}{N}$. The simulations are performed in Matlab.

3.6.1 ADSL performances

The asymmetric and symmetric ADSL Downstream (DS) scenarios considered here are shown in Figure 3.1 and Figure 3.2 respectively. The simulations are performed for a two-user case ($N = 2$) up to a seven-user case ($N = 7$). The four-user scenario, for example, consists of active modems 1,2,3,4 where modems 5,6,7 are inactive.

In Table 3.4 simulation results are shown for the asymmetric downstream scenarios for 2 up to 7 modems without spectral mask constraints. The 'OSB'

Table 3.4: Comparison DSM algorithms for asymmetric N-user ADSL Downstream scenario Figure 3.1 without spectral mask constraints: columns indicate DSM algorithms, rows indicate no. of modems, three cells per row and column indicating the simulation time, percental performance in weighted data rate sum and no. of Lagrange multiplier updates respectively. For ISB, two no. in each cell are given for two different iteration orders as explained in Section 3.6.1.

Modems	OSB	BB-OSB	ISB	IW	SCALE	ASB-2	CA-DSB	DSB	MS-DSB
2	3min	30s	14.00s / 14.03s	0.03s	0.56s	0.04s	0.08s	0.05s	1.18s
	100.0%	99.0%	100.0% / 100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	20	25	23 / 23	-	-	-	-	-	23
3	3h	3min	25.97s / 26.10s	0.08s	1.09s	0.08s	0.17s	0.1s	1.44s
	100.0%	99.8%	99.9% / 99.9%	99.8%	100.0%	99.9%	100.0%	100.0%	100.0%
	20	17	23 / 23	-	-	-	-	-	21
4	5d	20min	48.23s / 81.05s	0.13s	1.69s	0.10s	0.33s	0.19s	2.19s
	100.0%	99.6%	99.8% / 99.9%	99.5%	100.0%	99.6%	100.0%	100.0%	100.0%
	21	21	25 / 43	-	-	-	-	-	24
5	-	3h	76.56s / 493.06s	0.24s	2.75s	0.66s	3.13s	1.66s	5.20s
	-	100%	91.2% / 99.4%	71.8%	98.3%	86.9%	92.8%	92.8%	100.0%
	-	35	25 / 63	-	-	-	-	-	35
6	-	1d	102.49s / 167.00s	0.28s	3.48s	0.80s	5.21s	2.61s	7.30s
	-	99.4%	94.1% / 96.6%	86.9%	89.5%	98.3%	99.4%	99.4%	100.0%
	-	20	24 / 39	-	-	-	-	-	23
7	-	7d	137.79s / 321.06s	0.36s	7.15s	0.94s	5.89s	3.24s	6.21s
	-	99.6%	90.6% / 98.4%	92.7%	83.6%	99.8%	100.0%	100.0%	100.0%
	-	17	23 / 53	-	-	-	-	-	15

Table 3.5: Comparison DSM algorithms for asymmetric N-user ADSL Downstream scenario Figure 3.1 with spectral mask constraints: columns indicate DSM algorithms, rows indicate no. of modems, three cells per row and column indicating the simulation time, percental performance in weighted data rate sum and no. of Lagrange multiplier updates respectively. For ISB, two no. in each cell are given for two different iteration orders as explained in Section 3.6.1.

Modems	OSB	BB-OSB	ISB	IW	SCALE	ASB-2	CA-DSB	DSB	MS-DSB
2	9s	0.76s	0.63s / 0.63s	0.04s	0.05s	0.05s	0.04s	0.02s	0.10s
	100.0%	79.9%	100.0% / 100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	1	1	1	-	-	-	-	-	1
3	9min	2.19s	1.17s / 1.19s	0.05s	0.06s	0.08s	0.10s	0.08s	0.11s
	100.0%	89.3%	100.0% / 100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	1	1	1	-	-	-	-	-	1
4	6h	10.81s	1.93s / 1.97s	0.09s	4.69s	0.14s	0.33s	0.29s	0.13s
	100.0%	92.2%	100.0% / 100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	1	1	1	-	-	-	-	-	1
5	-	70.69s	2.96s / 3.02s	0.22s	5.32s	0.60s	2.83s	1.69s	0.18s
	-	93.0%	100.0% / 100.0%	68.5%	100.0%	81.3%	89.8%	89.8%	100.0%
	-	1	1	-	-	-	-	-	1
6	-	10.96min	4.33s / 4.47s	0.38s	5.30s	0.75s	6.39s	4.56s	0.36s
	-	94.6%	92.6% / 98.0%	88.5%	93.8%	97.7%	99.1%	99.2%	100.0%
	-	1	1	-	-	-	-	-	1
7	-	1.4h	6.08s / 6.30s	0.47s	8.17s	0.87	9.34s	5.99s	0.46s
	-	95.6%	90.3% / 97.9%	95.5%	86.4%	99.9%	100.0%	100.0%	100.0%
	-	1	1	-	-	-	-	-	1

Table 3.6: Comparison DSM algorithms for symmetric N-user ADSL Downstream scenario Figure 3.2 without spectral mask constraints: columns indicate DSM algorithms, rows indicate no. of modems, three cells per row and column indicating the simulation time, percental performance in weighted data rate sum and no. of Lagrange multiplier updates respectively. For ISB, two no. in each cell are given for two different iteration orders as explained in Section 3.6.1

Modems	OSB	BB-OSB	ISB	IW	SCALE	ASB-2	CA-DSB	DSB	MS-DSB
2	3min	30s	17.70s / 18.09s	0.05s	1.11s	0.05s	0.07s	0.03s	1.91s
	100.0%	99.3%	100.0% / 100.0%	99.6%	100.0%	100.0%	100.0%	100.0%	100.0%
	35	35	29 / 29	-	-	-	-	-	33
3	3h	3min	19.77s / 19.95s	0.10s	1.60s	0.07s	0.12s	0.05s	1.17s
	100.0%	99.3%	99.9% / 100.0%	99.9%	100.0%	100.0%	100.0%	100.0%	100.0%
	22	22	17 / 17	-	-	-	-	-	17
4	5d	20min	33.17s / 33.63s	0.13s	1.76s	0.10s	0.25s	0.15s	1.77s
	100.0%	99.2%	100.0% / 100.0%	99.7%	100.0%	100.0%	100.0%	100.0%	100.0%
	28	28	17 / 17	-	-	-	-	-	19
5	-	3h	25.03s / 42.80s	0.13s	2.59s	0.13s	0.47s	0.24s	1.85s
	-	99.3%	99.9% / 100.0%	99.6%	100.0%	99.8%	100.0%	100.0%	100.0%
	-	28	8 / 14	-	-	-	-	-	13
6	-	1d	71.28s / 72.10s	0.25s	3.09s	0.25s	0.63s	0.32s	5.04s
	-	99.1%	99.7% / 99.7%	99.4%	100.0%	99.8%	100.0%	100.0%	100.0%
	-	39	16 / 16	-	-	-	-	-	16
7	-	7d	131.31s / 132.47s	0.30s	3.95s	0.30s	0.84s	0.47s	7.47s
	-	98.8%	100.0% / 100.0%	99.4%	100.0%	99.8%	100.0%	100.0%	100.0%
	-	36	21 / 21	-	-	-	-	-	18

Table 3.7: Comparison DSM algorithms for symmetric N-user ADSL Downstream scenario Figure 3.2 with spectral mask constraints: columns indicate DSM algorithms, rows indicate no. of modems, three cells per row and column indicating the simulation time, percental performance in weighted data rate sum and no. of Lagrange multiplier updates respectively. For ISB, two no. in each cell are given for two different iteration orders as explained in Section 3.6.1

Modems	OSB	BB-OSB	ISB	IW	SCALE	ASB-2	CA-DSB	DSB	MS-DSB
2	9s	8.01s	0.63s / 0.65s	0.02s	0.09s	0.07s	0.11s	0.06s	0.05s
	100.0%	89.1%	100.0% / 100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	1	1	1	-	-	-	-	-	1
3	9min	2.49s	1.18s / 1.25s	0.04s	0.10s	0.09s	0.09s	0.08s	0.07s
	100.0%	88.9%	100.0% / 100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	1	1	1	-	-	-	-	-	1
4	6h	12.17s	1.94s / 1.97s	0.05s	0.08s	0.14s	0.12s	0.11s	0.09s
	100.0%	88.9%	100.0% / 100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	1	1	1	-	-	-	-	-	1
5	-	77.05s	3.01s / 3.07s	0.07s	0.14s	0.19s	0.19s	0.12s	0.15s
	-	88.6%	100.0% / 100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	-	1	1	-	-	-	-	-	1
6	-	9.14min	4.40s / 4.50s	0.07s	0.12s	0.23s	0.24s	0.16s	0.32s
	-	88.6%	100.0% / 100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	-	1	1	-	-	-	-	-	1
7	-	1.13h	6.18s / 6.28s	0.18s	0.16s	0.28s	0.48s	0.21s	0.45s
	-	88.6%	100.0% / 100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	-	1	1	-	-	-	-	-	1

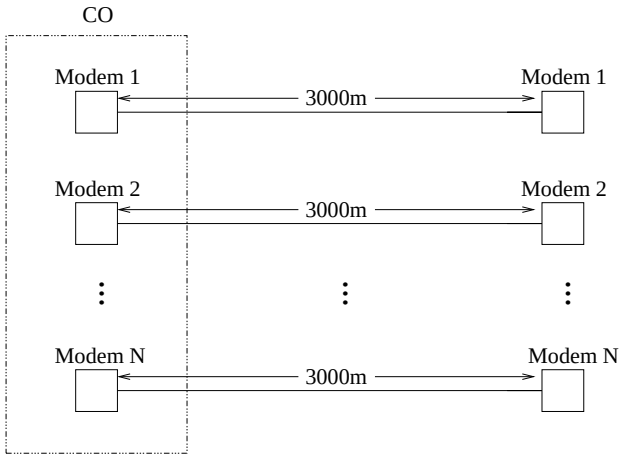


Figure 3.2: Symmetric N-user ADSL simulation scenario with equal loop lengths of 3000m

column refers to the Optimal Spectrum Balancing algorithm proposed in [6] where the power loading is discretized in 100 equal dBm steps. This fine discretization leads to a good approximation of the continuous solution. A finer discretization would increase the complexity too much here. The 'BB-OSB' column refers to the branch and bound OSB algorithm proposed in [7] where discrete bit loading is considered ranging from 0 up to 16 bits in steps of 1 bit. Note that the discrete nature of BB-OSB leads to a weighted sum rate loss with respect to the OSB algorithm. The 'ISB' column refers to the ISB algorithm proposed in [10][9] with power loading. For the Lagrange multiplier search of OSB and ISB, the algorithms described in [7] [25] are used.

The 'IW', 'SCALE', 'ASB-2', 'CA-DSB', 'DSB' and 'MS-DSB' columns refer to the methods described in Sections 3.4.1, 3.4.2, 3.4.3, 3.5.1, 3.5.2 and 3.5.3 respectively. Note that the parameters of the reference line for the ASB-2 algorithm are chosen based on the guidelines provided in [12].

For each number of modems and for each spectrum management algorithm there are three cells representing the simulation time, the performance in weighted data rate sum and the number of Lagrange multiplier updates to converge to the total power constraints respectively. Note that the ISB column has two numbers for each cell. The reason is that the performance of the ISB algorithm depends on the iteration order. The number of possible iteration orders is $N!$ for an N-modem scenario. However in order to reduce the total number of simulated iteration orders, just N iteration orders are simulated for every scenario. For instance, for the four-user case, ISB is performed for the iteration orders 1-2-3-4, 2-3-4-1, 3-4-1-2 and 4-1-2-3. The left number in the

cell then represents the smallest value in simulation time, percental weighted data rate sum and number of Lagrange multiplier updates of the four simulated iteration orders. The right number represents the largest value in simulation time, percental weighted data rate sum and number of Lagrange multiplier updates of the four simulated iteration orders. Note that the difference between these two values increases even more if more iteration orders are simulated but we have restricted ourselves to simulating N iteration orders just to show the dependence of ISB on the iteration order.

The OSB algorithm takes too long to simulate for more than 4 users and so only simulation results are shown up to 4 users. Therefor the performance in weighted data rate sum is measured with respect to the MS-DSB algorithm. This is the reason why the performance of the MS-DSB algorithm is always 100%. Note that this does not mean that the MS-DSB algorithm performs globally optimal for every DSL scenario.

From Table 3.4 it can be seen that OSB and BB-OSB have an exponential complexity in the number of modems. This large computational burden excludes practical implementation. However the results of OSB and BB-OSB can be used to evaluate the performance of the other algorithms. The ISB algorithm is much faster but shows suboptimality for scenarios with more than four users. The performance also strongly depends on the iteration order. Note that the number of Lagrange multiplier updates to converge to the total power constraints are less than 65 for all the scenarios.

The distributed algorithms of Table 3.4 are significantly faster than the centralized algorithms. Note that these algorithms (except for MS-DSB) update their Lagrange multipliers individually, so there is no total number of joint Lagrange multipliers updates.

IW performs suboptimally for the scenarios with more than four users. The reference line approach of ASB-2 performs clearly better than IW. CA-DSB and DSB perform similarly because their initial transmit power spectra are equal, as explained in Section 3.5.3. SCALE performs better for the five-user scenario but worse for the six- and seven-user scenarios. In the seven-user case it performs worse than IW. This is due to the fact that it converges to a bad locally optimal solution. Note that it is verified by simulations that SCALE leads to the same solutions as CA-DSB and DSB when it uses the same initial transmit power spectra.

It can be seen that the MS-DSB algorithm performs at least as good as the SCALE, CA-DSB and DSB algorithms. Note that MS-DSB is not a globally optimal algorithm but its probability to converge to the globally optimal solution is larger than for SCALE, CA-DSB and DSB. Furthermore its increase in simulation time with respect to the other distributed algorithms is small. Note that MS-DSB updates its Lagrange multipliers in parallel whereas for SCALE,

Table 3.8: Data rate performance comparison obtained by IW and DSB algorithms for upstream VDSL scenario Figure 3.3

	IW	DSB
weighted sum rate [Mb/s]	14.5	14.8
Modem 1 [Mb/s]	0.2	1.1
Modem 2 [Mb/s]	19.2	19.2
Modem 3 [Mb/s]	19.1	19.4
Modem 4 [Mb/s]	19.3	19.3

CA-DSB and DSB the Lagrange multipliers are updated individually by the modems.

Table 3.5 shows the simulation results for the downstream asymmetric scenario with spectral mask constraints. Note that there is no Lagrange multiplier search needed in this case. The first choice of the Lagrange multipliers equal to zero leads to a solution already satisfying the total power constraints. The total power constraints are thus not active. Furthermore the same conclusions can be drawn as for Table 3.4. Table 3.6 and Table 3.7 show the simulation results for the downstream symmetric scenario without and with spectral mask constraints respectively. It can be seen that all the algorithms perform similarly. So symmetric DSL scenarios are not really a challenge and one should opt for one of the low-complexity algorithms like IW, SCALE, CA-DSB or DSB.

Summarizing, it can be stated that SCALE, CA-DSB and DSB generally lead to good locally optimal solutions. However their performance depends on the initial transmit power spectra. For some DSL scenarios the non-convex spectrum management problem can have several locally optimal solutions. In this case SCALE, CA-DSB and DSB can perform quite suboptimally. MS-DSB tackles this issue by carefully choosing multiple initial transmit spectra. Although this approach does not guarantee a globally optimal solution, it leads to an improved performance with a small increase in complexity.

3.6.2 VDSL performances

In Figure 3.3 a four-user VDSL upstream scenario is shown. This scenario consists of one long line and three short lines causing a lot of crosstalk into the long line. If the spectrum is not managed properly the performance of the long line will be very poor. In Table 3.8 the results are shown for IW and DSB. It can be seen that the weighted data rate sum is slightly larger for DSB. However in terms of bit rates of the individual modems, DSB succeeds in increasing the bit rate of the long line up to 550% with respect to IW (that achieves a 'selfish optimum') while keeping the same bit rates on the short lines.

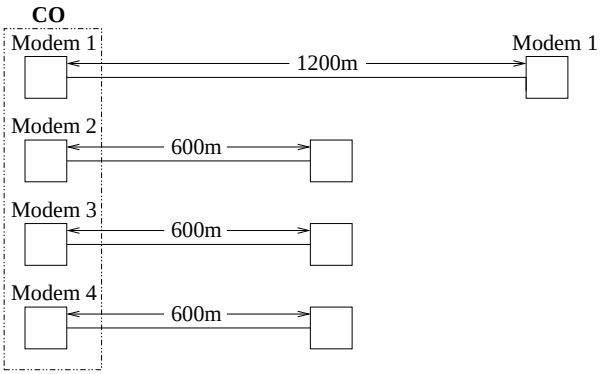


Figure 3.3: 4-user upstream VDSL simulation scenario 1 with 3 near users (600m) and 1 far user (1200m)

Table 3.9: Data rate performance comparison obtained by IW and DSB algorithms for upstream VDSL scenario Figure 3.4

	IW	DSB
weighted sum rate [Mb/s]	14.8	18.4
Modem 1 [Mb/s]	0.026	1.10
Modem 2 [Mb/s]	0.026	1.10
Modem 3 [Mb/s]	29.3	28.6
Modem 4 [Mb/s]	29.6	42.4

In Figure 3.4 an even more asymmetric VDSL upstream scenario is shown. The comparison between the performance for IW and DSB is shown in Table 3.9 . The bit rates of the long lines are increased by a factor 42. The bit rate of the second short line is increased by a factor 1.4. The bit rate of the first short line is decreased by only 2.7%. For these heavily unbalanced VDSL scenarios appropriate spectrum management can thus indeed lead to significant performance improvements.

Note that all users for the IW algorithm use their full power budget for the scenarios of Figure 3.3 and 3.4. For given equal weights, we have verified that for these scenarios this solution leads to the maximum weighted sum rate obtained by IW and to that end the comparison is done in a fair way.

Finally in Figure 3.5 a six-user upstream VDSL scenario is shown. The transmit power spectra obtained by IW and DSB are shown in Figures 3.6 and 3.7 respectively where the numbers indicate the modems. It can be seen that the IW transmit spectra are very selfish with full power transmission in frequency bands that are also used by other modems. The transmit power spectra of

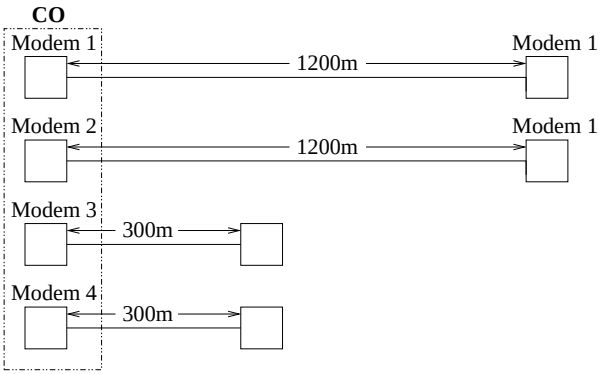


Figure 3.4: 4-user upstream VDSL simulation scenario 2 with 2 near users (300m) and 2 far users (1200m)

DSB are very different. The line of 1200m uses tones 1-166 while the other modems decrease their transmit spectra in this frequency band. This decrease in transmit spectrum is called a back-off. The 600m line transmits at full power on tones 167-337 while it backs-off on tones 1-166. Modem 3 uses tones 338-732 and backs-off on the other tones. Modem 6 uses tones 733-1147 and backs-off on the other tones. Modems 4 and 5 do not dominate any tones but back-off on tones 1-732. These modems do not use their full power budget. When the crosstalk interference increases further, the optimal power loading will converge to a FDMA solution where only one modem is active in each tone interval. This fact was also recently proved in [23].

3.7 Conclusion

In this paper, it is shown that state-of-the-art distributed spectrum management algorithms can be viewed from a similar framework. This framework consists of choosing an (iterative) (non-)convex approximation of the spectrum management problem and deriving a distributed solution based on the KKT conditions of the chosen approximation. Based on this unifying perspective novel distributed spectrum management algorithms are proposed. A first algorithm ASB-2 is a variant of the ASB algorithm. It is fully autonomous and can be simply extended to include multiple reference lines without a significant complexity increase. A second algorithm CA-DSB involves an iterative convex approximation approach. A third algorithm DSB is based on the KKT conditions of the spectrum management problem. The fourth algorithm MS-DSB is a multiple starting point extension of the third algorithm which leads to a better performance for scenarios with multiple locally optimal solutions.

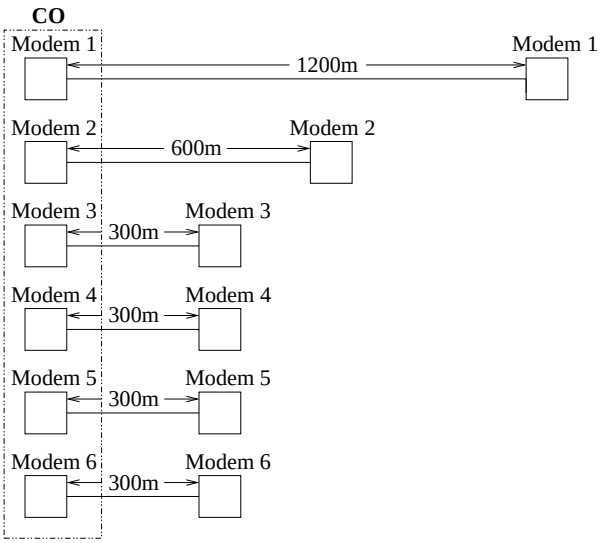


Figure 3.5: 6-user upstream VDSL simulation scenario with 4 users at 300m, 1 user at 600 m and 1 user at 1200 m from the Central Office (CO)

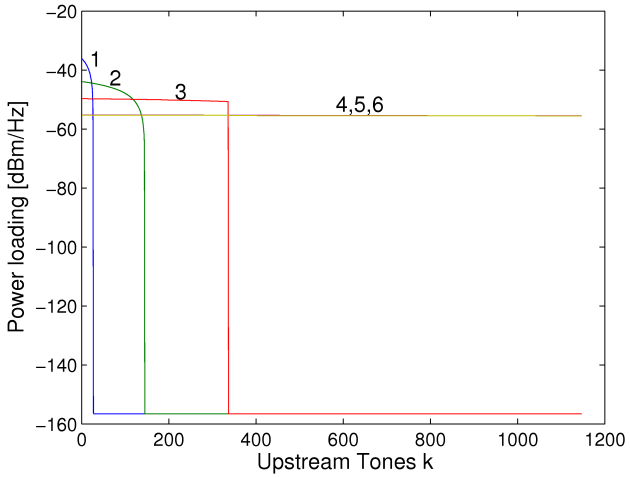


Figure 3.6: Transmit power loading obtained by the IW algorithm for the 6 users over the 1147 upstream tones of the VDSL upstream scenario of Figure 3.5. Transmit power of each user is indicated by a different number. Transmit powers of users 4, 5 and 6 coincide.

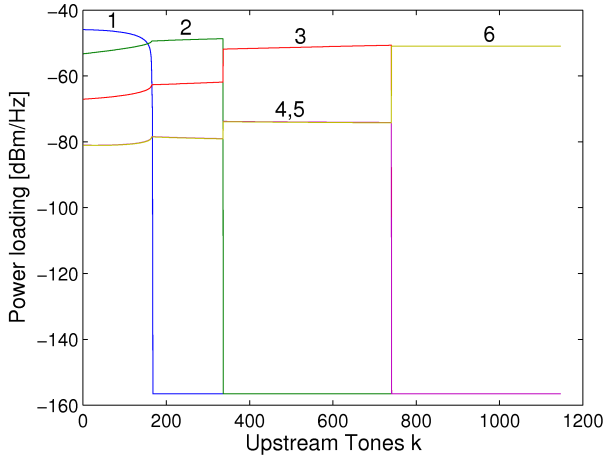


Figure 3.7: Transmit power loading obtained by the DSB algorithm for the 6 users over the 1147 upstream tones of the VDSL upstream scenario of Figure 3.5. Transmit power of each user is indicated by a different number. Transmit powers of users 4 and 5 coincide.

CA-DSB, DSB and MS-DSB are locally optimal algorithms. They consist of water-filling-like algorithms running in the individual modems and controlled by a spectrum management center to improve the overall performance of the network. A convergence and complexity analysis is provided for the proposed algorithms. Extensive simulation results for several DSL scenarios show that these algorithms perform very well in mitigating the effect of crosstalk.

Bibliography

[1] DSL Forum, "News release - DSL dominates global broadband subscriber growth," Tech. Rep., Mar. 2007.

[2] K. B. Song, S. T. Chung, G. Ginis, and J. M. Cioffi, "Dynamic spectrum management for next-generation DSL systems," *IEEE Communications Magazine*, vol. 40, no. 10, pp. 101–109, Oct. 2002.

[3] G. Taubock and W. Henkel, "Mimo systems in the subscriber-line network," in *Proc. of the 5th International OFDM Workshop*, Sep. 2000, pp. 18.1–18.3.

[4] W. Yu and J. M. Cioffi, "Multiuser detection in vector multiple access channels using generalized decision feedback equalization," in *Proc. of*

the 5th International Conference on Signal Processing, World Computer Congress, vol. 20, Aug. 2000, pp. 1771–1777.

[5] G. Ginis and J. M. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE Journal on Selected Areas on Communications*, vol. 20, no. 5, pp. 1085–1104, Jun. 2002.

[6] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 922–933, May 2006.

[7] P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "A low complexity optimal spectrum balancing algorithm for digital subscriber lines," *Signal Processing*, vol. 87, no. 7, pp. 1735–1753, Jul. 2007.

[8] Y. Xu, S. Panigrahi, and T. Le-Ngoc, "A concave minimization approach to dynamic spectrum management for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 1, Jun. 2006, pp. 84–89.

[9] R. Lui and W. Yu, "Low-complexity near-optimal spectrum balancing for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 3, May 2005, pp. 1947–1951.

[10] R. Cendrillon and M. Moonen, "Iterative spectrum balancing for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 3, May 2005, pp. 1937–1941.

[11] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[12] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spectrum balancing for digital subscriber lines," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4241–4257, Aug. 2007.

[13] J. Papandriopoulos and J. S. Evans, "Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 7, Jun. 2006, pp. 3270–3275.

[14] T. Starr, J. M. Cioffi, P. J. Silverman, *Understanding digital subscriber lines.* Prentice Hall, 1999.

[15] J. M. Cioffi, "EE379C course reader - Advanced digital communication," in *Stanford University. Available at: http://www.stanford.edu/class/ee379c/.*

[16] P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "Convex relaxation based low-complexity optimal spectrum balancing for multi-user DSL," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 3, Apr. 2007, pp. III–349 – III–352.

[17] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1310–1322, Jul 2006.

[18] M. Chiang, *Geometric programming for communication systems*. Foundations and Trends in Communications and Information Theory, Aug. 2005, vol. 2, no. 1.

[19] W. Lee, Y. Kim, M. H. Brady, and J. M. Cioffi, "Band-preference dynamic spectrum management in a DSL environment," in *Proc. of the IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2006, pp. 1–5.

[20] J. M. Cioffi, W. Rhee, M. Mohseni, and M. Brady, "Band preference in dynamic spectrum management," in *Proc. of the European Signal Processing Conference (EUSIPCO)*, Sep. 2004.

[21] W. Yu, "Multiuser water-filling in the presence of crosstalk," in *Proc. of the Information Theory and Applications (ITA)*, Feb. 2007.

[22] V. M. K. Chan and W. Yu, "Multiuser spectrum optimization for discrete multitone systems with asynchronous crosstalk," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5425–5435, Nov. 2007.

[23] S. Hayashi and Z. Q. Luo, "Dynamic spectrum management: When is FDMA sum-rate optimal?" in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 3, Apr. 2007, pp. III–609 – III–612.

[24] *Asymmetrical digital subscriber line (ADSL) transceivers*, ITU Std. G.992.1, 1999.

[25] P. Tsiaflakis, J. Vangorp, M. Moonen, J. Verlinden, and K. V. Acker, "An efficient search algorithm for the lagrange multipliers of optimal spectrum balancing in multi-user xDSL systems," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 4, May 2006.

Chapter 4

Geometric Programming Based Spectrum Coordination

Optimality Certificate of Dynamic Spectrum Management in Multi-Carrier Interference Channels

P. Tsiaflakis, C.W. Tan, Y. Yi, M. Chiang, and M. Moonen

Published in *Proc. of the IEEE International Symposium on Information Theory (ISIT)*, Toronto, Canada, Jul. 2008, pp. 1298-1302.

©2008 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Contributions of first author

- literature study
- co-development of GP reformulation and optimality certificate
- co-analysis of optimality certificate
- co-development of DSM-GP algorithm
- design of simulation scenarios
- software implementation and computer simulations
- co-interpretation of simulation results
- co-formulation of conclusion
- text co-redaction, co-editing, co-revision

Abstract

The multi-carrier interference channel where interference is treated as additive white Gaussian noise, is a very active topic of research, particularly important in the area of Dynamic Spectrum Management (DSM) for Digital Subscriber Lines (DSL). Here, multiple users optimize their transmit power spectra so as to maximize the total weighted sum of data rates. The corresponding optimization problem is however nonconvex and thus computationally intractable, i.e. a certificate of global optimality requires exponential time complexity algorithms. This paper shows that under certain channel conditions, this nonconvex problem can be solved in polynomial time with a certificate of global optimality. The channel conditions are discussed consisting of different interference models including synchronous and asynchronous DSL transmission. Simulations demonstrate its applicability to realistic DSL scenarios.

4.1 Introduction

Digital Subscriber Line (DSL) technology remains a popular choice of broadband access technology in the world [1]. The corresponding increased demand for greater bandwidth and faster connection has led to several technological approaches to mitigate the channel impairments. This has become one of the most important applications of the multi-carrier interference channel model in today's Internet.

One of the major limitations of DSL performance today is crosstalk among different lines operating in the same cable bundle. Dynamic Spectrum Management (DSM) refers to a set of solutions to the crosstalk problem. Basically these solutions consist of signal level coordination and/or spectrum level coordination. For example, receiver and transmitter signal level coordination correspond to a multiple access channel and broadcast channel respectively. In this paper the focus is on spectrum level coordination, which is also referred to as spectrum management, spectrum balancing, or multi-carrier power control.

Spectrum level coordination corresponds to a multi-carrier interference channel where a number of different users try to communicate their separate information over a coupled (wired) channel. The capacity region of the interference channel is still an open problem, but for practical systems where the crosstalk channels are typically weaker than the direct channel, treating interference as additive white Gaussian noise has been the most practical communication strategy in operational networks. In this case the transmit power spectrum of each user is designed so as to minimize interference to other lines, while maintaining a satisfactory data rate. Such a technique reduces the impact of crosstalk and significantly improves the overall data rate beyond those of the

current approach of static spectrum management where fixed over-conservative transmit power spectra are used.

The problem of optimally choosing the transmit power spectra in order to maximize the total weighted sum of data rates of the users turns out to be a *nonconvex optimization problem*. Several solutions have been proposed ranging from centralized to distributed algorithms, e.g., [2][3][4][5][6]. However, there is an undesirable gap in the analysis of these DSM algorithms: we either have algorithms with optimality certificate but exponential running-time, or polynomial-time solutions with only numerical verification of small suboptimality gap but no theoretical characterization of optimality condition.

Indeed, on the one hand, polynomial-time algorithms like IW[5], ISB[3] and ASB[4] have only sufficient conditions for convergence but not for optimality. On the other hand, Optimal Spectrum Balancing (OSB) [2], Branch and Bound Optimal Spectrum Balancing (BB-OSB)[7] and a prismatic branch and bound algorithm (PBB)[8], provide a certificate of finding the globally optimal solution, up to a certain accuracy depending on the used discretization of the search space. Unfortunately, the complexity of these algorithms grows exponentially with the number of users. Furthermore most of the aforementioned algorithms assume that the number of frequency tones is infinitely large so that the Lagrangian duality gap can be viewed as zero [9].

In this paper, we give conditions on the channel, noise, and power constraints, under which this difficult nonconvex problem of DSM can be solved in polynomial time with an optimality certificate. To the best of our knowledge, our results provide the first sufficient condition for optimality for a finite number of carriers, and it is applicable to a general model including synchronous as well as asynchronous DSL transmission [10]. There are two key innovations in our analysis: the use of Geometric Programming (GP) and M-matrix theory. Furthermore, the optimality analysis leads to a new DSM algorithm, with simulations demonstrating the usefulness of this optimality certificate to realistic DSL scenarios.

The following notations are used. Boldface uppercase letters denote matrices, boldface lowercase letters denote column vectors and italics denote scalars. Furthermore $\mathbf{x} \geq \mathbf{y}$ denotes componentwise inequality between vectors $\mathbf{x}$ and $\mathbf{y}$ and $\mathbf{X} \geq \mathbf{Y}$ denotes componentwise inequality between matrices $\mathbf{X}$ and $\mathbf{Y}$. We also let $(\mathbf{x})_l$ denote the lth element of $\mathbf{x}$ and $[\mathbf{X}]_{ij}$ denote the element of matrix X on row i and column j. We denote the identity matrix by $\mathbf{I}$ and the spectral radius of $\mathbf{T}$ by $\rho(\mathbf{T})$.

4.2 System Model

Most current DSL systems use Discrete Multi-Tone (DMT) modulation. The basic idea of DMT is to split the available bandwidth into a large number of frequency bands, also called tones. For each tone, a transmit power can be allocated individually which corresponds to a number of transmitted bits. The total data rate for each user is then obtained by adding its transmitted bits over all tones.

In our model, no signal coordination is assumed between the transmitting and/or receiving modem for each user. Each user regards the signals from the other users as noise. The only degree of freedom is choosing the transmit powers of the different users over the different tones.

The transmit power is denoted as s_k^n, where n is the user index ranging from 1 to N and k is the tone index ranging from 1 to K. The noise power is denoted as σ_k^n. The vector containing the transmit powers of user n on all tones is $\mathbf{s}^n \triangleq [s_1^n, s_2^n, \ldots, s_K^n]^T$. $h_k^{n,m}$ denotes the squared channel gain magnitude from user m into user n on tone k, where it refers to the squared channel gain magnitude of line n when $m = n$. Furthermore int_k^n is the interference caused to user n on frequency tone k. The achievable bit loading for user n on tone k can then be expressed as

$$b_k^n \triangleq \log_2\left(1 + \frac{h_k^{n,n} s_k^n}{\Gamma(\text{int}_k^n + \sigma_k^n)}\right) \text{ bits/ s/ Hz}, \tag{4.1}$$

where Γ denotes the signal-to-noise ratio (SNR) gap to capacity, which is a function of the desired bit error ratio (BER), the coding gain and noise margin [11].

The DMT symbol rate is denoted as f_s. The total data rate for user n is given by

$$R^n = f_s \sum_k b_k^n. \tag{4.2}$$

Such a model can be used to analyze the following special cases.

The first case we consider is synchronized DSL transmission. Here it is assumed that all users are aligned in frequency so that each tone is capable of transmitting data independently from the other tones. The interference caused to user n on tone k can then be expressed as $\text{int}_k^n = \sum_{m \neq n} h_k^{n,m} s_k^m$.

The second case is the asynchronous DSL transmission case [10]. Here there is not only crosstalk within one tone but also between different tones of different users. The interference caused to user n on tone k can then be expressed as $\text{int}_k^n = \sum_{m \neq n} (\sum_{p=1}^{K} h_{k,p}^{n,m} s_p^m)$ where $h_{k,p}^{n,m}$ refers to the squared channel gain magnitude from user m tone p into user n tone k.

The third case includes all the previous cases and equals $\text{int}_k^n = \left(\sum_{m=1}^{N} (\sum_{p=1}^{K} h_{k,p}^{n,m} s_p^m) \right) - h_{k,k}^{n,n}$, where $h_{k,p}^{n,m}$ is the squared channel gain magnitude from user m tone p into user n tone k. All these cases can be summarized with the following bit loading expression for user n on tone k

$$b_k^n = \log_2 \left(1 + \frac{h_{k,k}^{n,n} s_k^n}{\sum_{m=1}^{N} \sum_{p=1}^{K} h_{k,p}^{n,m} s_p^m - h_{k,k}^{n,n} s_k^n + \Gamma \sigma_k^n} \right). \tag{4.3}$$

For each of the above cases the parameters $h_{k,p}^{n,m}$ have their specific values. In the rest of this paper we will focus on this general case (4.3) so that the results can be used for all three cases. Note that for notational simplicity, we absorb Γ into the definition of $h_{k,p}^{n,m}$.

4.3 Spectrum Management Problems

This paper aims to maximize the data rates of DSL copper lines (in a bundle). In particular, the objective is to find the optimal transmit power spectra that maximize a weighted sum of data rates subject to individual power constraints. Assuming a per-user total power constraint P^n for each user n, this is formulated as the following problem:

$$\begin{aligned} \max_{\mathbf{s}^1,\ldots,\mathbf{s}^N} \quad & \sum_{n=1}^{N} w_n R^n \\ \text{s.t.} \quad & \sum_{k=1}^{K} s_k^n \le P^n \quad , \quad \forall n, \\ & 0 \le s_k^n, \quad \forall n, \forall k. \end{aligned} \tag{4.4}$$

Instead of per-user total power constraints, we will consider one total power constraint P on all users, where $P = \sum_{n=1}^{N} P^n$. Using (4.2) this can be formulated as follows:

$$\begin{aligned} \max_{\mathbf{s}^1,\ldots,\mathbf{s}^N} \quad & \sum_{n=1}^{N} w_n f_s \sum_{k=1}^{K} b_k^n \\ \text{s.t.} \quad & \sum_{k=1}^{K} \sum_{n=1}^{N} s_k^n \le P \quad , \\ & 0 \le s_k^n, \quad \forall n, \forall k, \end{aligned} \tag{4.5}$$

where b_k^n is given by the general interference model (4.3). Note that (4.4) and (4.5) are both nonconvex, and thus intractable. Also note that in (4.5)

the inequality total power constraint can be replaced by an equality. This is justified by lemma 1.

Lemma 1 *At optimality, we have* $\sum_{k=1}^{K}\sum_{n=1}^{N} s_k^n = P$ *in (4.5).*

Proof : Suppose the optimal solution to (4.5) results in a total power P^{t} where $P^{\mathrm{t}} < P$. As this is the optimal solution it means that no better solution can exist. This is not true because by scaling all the powers s_k^n by a factor P/P^{t}, the objective function in (4.5) increases. Thus, at optimality, $P^{\mathrm{t}} = P$ in (4.5). □

Note that Problem (4.5) can be seen as a relaxation of problem (4.4). For symmetric crosstalk scenarios with not too large crosstalk (cf. Section 4.4) the optimal solution of (4.5) allocates equal powers to all users and so solving (4.5) and (4.4) leads to exactly the same solution, i.e. zero relaxation gap. For asymmetric crosstalk scenarios, the users will not necessarily be allocated the same amount of total transmit powers and there may be a relaxation gap between the solution of (4.4) and (4.5). The solution of (4.5) can however provide a useful upper bound for problem (4.4).

Finally, we note that the nonconvex spectrum management problem with one total power constraint on all users (4.5) has also been considered in other work [12].

4.4 Optimality Certificate for Polynomial Time Spectrum Management

4.4.1 Polynomial Time Spectrum Management

Since (4.5) is nonconvex, it can have many locally optimal solutions depending on the values of the channel and noise coefficients. In order to solve (4.5) with a certificate of global optimality, i.e., the solution is globally optimal, one often has to resort to exponential time complexity algorithms such as [12][2]. In Theorem 1, we provide sufficient conditions on the channel and noise coefficients and the total power constraint for which (4.5) can be solved in polynomial time with a certificate of global optimality. Before describing Theorem 1, we first define the matrix $\mathbf{B}$ as follows:

$$[\mathbf{B}]_{d,l} = \left(\frac{h_{k,p}^{n,m}}{h_{k,k}^{n,n}} + \frac{\Gamma\sigma_k^n}{P h_{k,k}^{n,n}}\right), \tag{4.6}$$

where

$$d = n + (k-1)N,$$
$$l = m + (p-1)N.$$

Theorem 4.1 *For the general interference system model (4.3), (4.5) can be solved optimally in polynomial time when*

$$\mathbf{B} \text{ is nonsingular and } \mathbf{C} \triangleq \mathbf{I} - \mathbf{B}^{-1} \geq \mathbf{0}. \tag{4.7}$$

Proof : We start from b_k^n of (4.5) by incorporating the total power equality constraint $\sum_{k=1}^{K}\sum_{n=1}^{N} s_k^n = P$ (see Lemma 1) as follows:

$$b_k^n = \log_2\left(1 + \frac{h_{k,k}^{n,n} s_k^n}{\sum_{m=1}^{N}\sum_{p=1}^{K} h_{k,p}^{n,m} s_p^m - h_{k,k}^{n,n} s_k^n + \Gamma\sigma_k^n \frac{\sum_{m=1}^{N}\sum_{p=1}^{K} s_p^m}{P}}\right) \tag{4.8}$$

This is reorganized as follows:

$$b_k^n = \log_2\left(\frac{\sum_{m=1}^{N}\sum_{p=1}^{K}(\frac{h_{k,p}^{n,m}}{h_{k,k}^{n,n}} + \frac{\Gamma\sigma_k^n}{h_{k,k}^{n,n}P}) s_p^m}{\sum_{m=1}^{N}\sum_{p=1}^{K}(\frac{h_{k,p}^{n,m}}{h_{k,k}^{n,n}} + \frac{\Gamma\sigma_k^n}{h_{k,k}^{n,n}P}) s_p^m - s_k^n}\right). \tag{4.9}$$

Using the symbol $\tilde{B}_{k,p}^{n,m} = \left(\frac{h_{k,p}^{n,m}}{h_{k,k}^{n,n}} + \frac{\Gamma\sigma_k^n}{Ph_{k,k}^{n,n}}\right)$, this can be simplified by

$$b_k^n = \log_2\left(\frac{\sum_{m=1}^{N}\sum_{p=1}^{K}\tilde{B}_{k,p}^{n,m} s_p^m}{\sum_{m=1}^{N}\sum_{p=1}^{K}\tilde{B}_{k,p}^{n,m} s_p^m - s_k^n}\right). \tag{4.10}$$

As it is not so easy to represent four-dimensional matrices, we go to a two-dimensional matrix by a bijection map $\{l,k\} \leftrightarrow \{n,m,k,p\}$ as follows:

$$\begin{aligned} d &= n + (k-1)N \\ l &= m + (p-1)N \\ [\mathbf{B}]_{d,l} &= \tilde{B}_{k,p}^{n,m} \end{aligned} \tag{4.11}$$

$$\sum_{m=1}^{N}\sum_{p=1}^{K}\tilde{B}_{k,p}^{n,m}s_p^m = \sum_{l=1}^{L=KN}[\mathbf{B}]_{d,l}s_l, \tag{4.12}$$

Now, (4.5) can be reformulated as follows:

$$\begin{array}{ll}\max\limits_{s_1\ldots s_L} & \sum\limits_{l=1}^{L=KN} w_l \log_2\left(\dfrac{\sum\limits_{q=1}^{L}[\mathbf{B}]_{l,q}s_q}{\sum\limits_{q=1}^{L}[\mathbf{B}]_{l,q}s_q - s_l}\right)\\ \text{s.t.} & \sum\limits_{l=1}^{L} s_l = P\\ \text{s.t.} & s_l \geq 0 \quad \forall l\end{array} \tag{4.13}$$

where $w_l = w_{\mathrm{rem}(l-1,N)+1}$ and $\mathrm{rem}(a,b)$ refers to the remainder after dividing a by b.

Using condition (4.7), we can make the following change of variables:

$$\mathbf{y} = \mathbf{B}\mathbf{s}, \qquad \mathbf{s} = \mathbf{B}^{-1}\mathbf{y} = \mathbf{y} - \mathbf{C}\mathbf{y}. \tag{4.14}$$

If (4.14) is substituted in (4.13) the cost function reduces to

$$\sum_{l=1}^{L} w_l \log(y_l/(\mathbf{C}\mathbf{y})_l) = \log\left(\prod_{l=1}^{L}(y_l^{w_l}/(\mathbf{C}\mathbf{y})_l^{w_l})\right), \tag{4.15}$$

and the last constraint to

$$y_l \geq (\mathbf{C}\mathbf{y})_l, \qquad l = 1,\ldots,n. \tag{4.16}$$

We thus obtain the following optimization problem:

$$\begin{array}{ll}\min\limits_{\mathbf{y}} & \prod\limits_{l=1}^{L}((\mathbf{C}\mathbf{y})_l y_l^{-1})^{w_l}\\ \text{s.t.} & (\mathbf{C}\mathbf{y})_l y_l^{-1} \leq 1, \qquad l = 1,\ldots,n\\ & \sum_l(\mathbf{B}^{-1}\mathbf{y})_l = P.\end{array} \tag{4.17}$$

Note that (4.15) and (4.16) are homogeneous in $\mathbf{y}$ (and so in $\mathbf{s}$). Thus the constraint $\sum_l(\mathbf{B}^{-1}\mathbf{y})_l = P$ or, equivalently, $\sum_l^L s_l = P$ acts as a normalization instead of a constraint. As a consequence it can be replaced by another normalization on $\mathbf{y}$ without changing the optimal value of (4.17), e.g.,

$$\prod_{l=1}^{L} y_l = 1. \tag{4.18}$$

This leads us to the following equivalent optimization problem:

$$\begin{array}{ll} \min\limits_{\mathbf{y}} & \prod\limits_{l=1}^{L}((\mathbf{C}\mathbf{y})_l y_l^{-1})^{w_l} \\ \text{s.t.} & (\mathbf{C}\mathbf{y})_l y_l^{-1} \leq 1, \qquad l = 1, \ldots, L \\ & \prod\limits_{l=1}^{L} y_l = 1. \end{array} \tag{4.19}$$

This can be reformulated as the following standard GP:

$$\begin{array}{ll} \min\limits_{\mathbf{y},\mathbf{t}} & \prod\limits_{l=1}^{L}(t_l)^{w_l} \\ \text{s.t.} & (\mathbf{C}\mathbf{y})_l y_l^{-1} t_l^{-1} \leq 1, \qquad l = 1, \ldots, L \\ & (\mathbf{C}\mathbf{y})_l y_l^{-1} \leq 1, \qquad l = 1, \ldots, L \\ & \prod\limits_{l=1}^{L} y_l = 1. \end{array} \tag{4.20}$$

For a particular obtained $\mathbf{y}$ for (4.20), we can recover the optimizer $\mathbf{s}$ of (4.5) by first performing

$$\mathbf{s} = (\mathbf{I} - \mathbf{C})\mathbf{y} \tag{4.21}$$

and then normalize (scale) this $\mathbf{s}$ such that it satisfies $\mathbf{1}^T\mathbf{s} = P$.

So when condition (4.7) is satisfied, the solution of the non-convex spectrum optimization problem (4.5) can be found by solving a GP given by (4.20). A GP can be turned into a convex optimization problem so that a local optimum is also a global optimum. Furthermore the global optimum can be computed very efficiently. Numerical efficiency holds both in theory and in practice: interior point methods applied to GP have provably polynomial time complexity [13]. This concludes the proof of Theorem 1. □

4.4.2 Optimality Analysis

In Section 4.4.1, we provided condition (4.7) under which nonconvex problem (4.5) can be solved in polynomial time with a certificate of global optimality. In this section, we further investigate this sufficient condition. We will first introduce some definitions from matrix theory.

Definition 4.2 Any matrix $\mathbf{Q}$ of the following form:

$$\mathbf{Q} = t\mathbf{I} - \mathbf{T}, \;\; t > \rho(\mathbf{T}), \; \mathbf{T} \geq \mathbf{0}, \tag{4.22}$$

is called an M-matrix [14]. □

Definition 4.3 A nonsingular nonnegative matrix $\mathbf{Q}$ is said to be an inverse M-matrix if $\mathbf{Q}^{-1}$ is an M-matrix.

□

Using the above definitions, condition (4.7) can be reformulated as follows: $\mathbf{B}$ is an inverse M-matrix with $t = 1$ where t is defined in Definition 1. Generally this means that the off-diagonal elements of $\mathbf{B}^{-1}$ have to be negative and the diagonal elements of $\mathbf{B}^{-1}$ have to be smaller than 1.

Characterizing all matrices whose inverses are M-matrices is an active research field in the mathematical society, but there are only limited results, even for symmetric matrices [15]. In [15] one special type of matrix is introduced, called *generalized ultrametric matrix*, which is an inverse M-matrix. However this matrix does not necessarily satisfy the extra condition that $t = 1$. Thus for an N-user case with K frequency tones, it is difficult to provide an exact meaning of these conditions on the channel and noise coefficients.

However, for a 2-user case with a single frequency tone, we can develop a sufficient condition on the channel and noise coefficients, and the total power, in order for the condition (4.7) to hold, i.e., $\mathbf{B}$ is an inverse M-matrix with $t = 1$. It also gives us an intuitive understanding of condition (4.7) for the 2-user case.

Theorem 4.4 *For a 2-user DSL scenario with a single frequency tone, if*

$$h_{1,1}^{1,2}\, h_{1,1}^{2,1} + h_{1,1}^{2,1}\, \frac{\Gamma\sigma_1^1}{P} + h_{1,1}^{1,2}\, \frac{\Gamma\sigma_1^2}{P} \leq \min\left(h_{1,1}^{1,1} \frac{\Gamma\sigma_1^2}{P}, h_{1,1}^{2,2} \frac{\Gamma\sigma_1^1}{P} \right), \tag{4.23}$$

then condition (4.7) is satisfied.

Proof: For a 2-user DSL scenario with a single frequency tone, $\mathbf{B}$ is a 2×2 matrix. Now, (4.7) can be expressed in terms of the elements of $\mathbf{B}$, i.e. $[\mathbf{B}]_{i,j} = b_{ij}$, as

$$\mathbf{I} - \mathbf{B}^{-1} \geq 0 \quad \Leftrightarrow \quad \begin{aligned} \frac{b_{12}}{b_{11}b_{22} - b_{12}b_{21}} &\geq 0, \\ \frac{b_{21}}{b_{11}b_{22} - b_{12}b_{21}} &\geq 0, \\ 1 - \frac{b_{11}}{b_{11}b_{22} - b_{12}b_{21}} &\geq 0, \\ 1 - \frac{b_{22}}{b_{11}b_{22} - b_{12}b_{21}} &\geq 0. \end{aligned} \tag{4.24}$$

Since the elements of $\mathbf{B}$ are all positive, (4.24) implies

$$\max(b_{11}, b_{22}) \leq b_{11}b_{22} - b_{12}b_{21}, \tag{4.25}$$

which can be rewritten in terms of the channel coefficients, noise and total power P:

$$h_{1,1}^{1,2}\, h_{1,1}^{2,1} + h_{1,1}^{2,1}\, \frac{\Gamma\sigma_1^1}{P} + h_{1,1}^{1,2}\, \frac{\Gamma\sigma_1^2}{P} \leq \min\left(h_{1,1}^{1,1}\frac{\Gamma\sigma_1^2}{P}, h_{1,1}^{2,2}\frac{\Gamma\sigma_1^1}{P}\right). \tag{4.26}$$

□

Intuitively, condition (4.7) is satisfied if both the crosstalk (i.e. $h_{1,1}^{2,1}, h_{1,1}^{1,2}$) and the total power constraint P are not exceedingly large. The same intuition can be expected to hold for the case where there are multiple users and frequency tones, where developing similar sufficient conditions is left as a future work.

4.4.3 Algorithm DSM-GP

Next, we turn from the optimality analysis based on GP and M-matrix theory to the design of our DSM algorithm, Algorithm DSM-GP. Given the channel, noise and power constraint parameters of the DSL scenario, (4.7) in Theorem 4.1 can be checked easily. If (4.7) in Theorem 4.1 is true, we can solve (4.5) with global optimality by first solving (4.20), then transforming back to the transmit powers (4.21) and scale so that the total power of all users is equal to P.

Suppose that (4.7) in Theorem 4.1 is violated, but there are only a small percentage of negative C_{ij} elements. In this case, we can set the negative elements to be zero, resulting in the matrix $\tilde{C}_{ij} = \max\{C_{ij}, 0\}$. Intuitively, solving (4.20) using $\tilde{\mathbf{C}}$ should still lead to a near-optimal performance when the percentage of negative elements is reasonably small. We show numerically in the next section that this approach is near optimal. However, the continuity of optimized objective function with respect to the perturbation of C_{ij} remains an open issue. The obtained solution is then projected back to the feasible domain of (4.5) by proportionally adjusting the positive transmit power to satisfy Lemma 1. However, the objective in (4.5) thus obtained using $\tilde{\mathbf{C}}$ may no longer be globally optimal, and does not upper bound the global optimal objective of (4.4). We propose the following algorithm to solve (4.5):

Algorithm 4.1 DSM-GP Algorithm

Step 1 Calculate $\mathbf{C} = \mathbf{I} - \mathbf{B}^{-1}$.
Step 2 Force negative elements of $\mathbf{C}$ to zero leading to $\tilde{\mathbf{C}}$.
Step 3 Replace $\mathbf{C}$ in (4.20) by $\tilde{\mathbf{C}}$ and solve it using a GP solver.
Step 4 Transform GP solution $\mathbf{y}$ to transmit powers $\mathbf{s}$ by $\mathbf{s} = (\mathbf{I} - \tilde{\mathbf{C}})\mathbf{y}$ where negative entries of $\mathbf{s}$ are set to zero.
Step 5 Scale the positive entries of $\mathbf{s}$ proportionally such that $\sum_l s_l = P$ (cf. Lemma 1).

4.5 Simulation Results

In this section simulation results will be shown for Algorithm 4.1 proposed in Section 4.4.3 for synchronous as well as asynchronous symmetric N-user ADSL downstream scenarios.

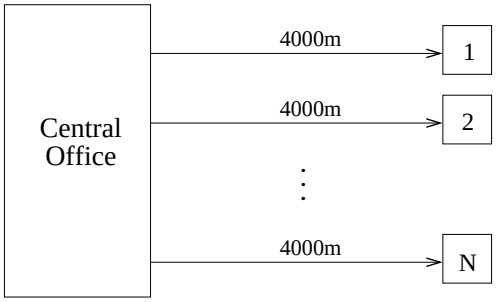


Figure 4.1: Symmetric N-user ADSL downstream scenario

The ADSL downstream scenario is shown in Figure 4.1. The simulations are performed for a two-user case ($N = 2$) up to an eight-user case ($N = 8$). The four-user scenario, for example, consists of active users 1,2,3,4 where users 5,6,7,8 are inactive. The twisted pair lines have a diameter of 0.5 mm (24 AWG). The maximum transmit power is 20.4 dBm [16]. The SNR gap Γ is 12.9 dB, corresponding to a coding gain of 3 dB, a noise margin of 6 dB and a target symbol error probability of 10^{-7}. The tone spacing Δ_f is 4.3125 kHz. The DMT symbol rate f_s is 4 kHz. The simulations are performed in Matlab on a TravelMate 4002WLMi with 768 MB of RAM and an Intel Pentium M processor 1.60 GHz.

In Table 4.1 the simulation results are shown for the synchronous DSL transmission case, i.e. no intercarrier interference (ICI). The first column denotes the number of users. The second column denotes the performance of Algorithm 4.1 with respect to the global optimal algorithm OSB with a very fine granularity. The third and fourth column denote the number and percentage of negative C_{ij} elements for condition (4.7) respectively.

The simulation time of OSB requires exponential complexity, i.e. 3 minutes, 3 hours and 5 days for 2,3 and 4 users scenarios respectively. For more than four users it is infeasible to execute within acceptable time. Therefore the performance for more than 4 users can not be compared to OSB. The simulation time of Algorithm 4.1 just requires a few seconds even for the 8-user scenario.

Note that condition (4.7) is satisfied up to the 6-user scenario. This means that Algorithm 4.1 succeeds in finding the optimal transmit spectra with a certificate of global optimality for up to the 6-user scenario in a few seconds.

Although the conditions are not satisfied for the 7- and 8-user scenarios, it has been verified that Algorithm 4.1 leads to near-optimal transmit spectra. That is because the percentage of negative elements is extremely small and forcing these to zero does not significantly change the problem. However there is no guarantee of global optimality. If the percentage of negative C_{ij} elements is small, it is reasonably to expect the suboptimal transmit spectra to be near optimal.

In Table 4.2 the simulation results are shown for the asynchronous DSL transmission case [10], i.e. with intercarrier interference (ICI). Condition (4.7) is satisfied up to the 6-user scenario. For the 7- and 8-user scenarios there is a small percentage of negative C_{ij} elements. This percentage is larger than for the synchronous case because there is a larger amount of crosstalk caused by ICI. As there exists no globally optimal method for asynchronous DSL transmission, the performance of Algorithm 4.1 cannot be compared to such a benchmark. However for the 2- up to the 6-user case we still have a certificate of global optimality for Algorithm 4.1. This shows another powerful application of an optimality certificate in DSM algorithm analysis. In fact, since the channel conditions are symmetric, the optimal solution computed using Algorithm 4.1 for (4.5) is equal to (4.4).

Table 4.1: Analysis performance Algorithm 4.1: Synchronous DSL transmission

Users	Performance wrt OSB	Number of negative **C** entries	% of negative **C** entries
2	100.0%	0	0
3	100.0%	0	0
4	100.0%	0	0
5	/	0	0
6	/	0	0
7	/	140	0.0057
8	/	352	0.011

4.6 Conclusion

Maximizing the weighted sum rate of all users in a multi-carrier interference channel is an important problem that finds application in DSL. DSL spectrum management has attracted many researchers who produced a wide variety of algorithms. However, since the optimization problem is nonconvex, all the proposed DSM algorithms in previous work that produce an optimality certificate have exponential-time complexity. Also, most previous DSM algorithms give only sufficient conditions for convergence to a suboptimal solution. This paper leverages on GP and M-matrix theory to generate a polynomial-time optimal-

Table 4.2: Analysis performance Algorithm 4.1: Asynchronous DSL transmission

Users	Number of negative **C** entries	% of negative **C** entries
2	0	0
3	0	0
4	0	0
5	0	0
6	0	0
7	203	0.0083
8	392	0.012

ity certificate for a set of conditions on channel, noise and power constraints. Our results shed new light on the structure of power allocation across multiple carriers in DSL interference channels. Furthermore, we propose Algorithm DSM-GP that can compute the global optimal transmit power allocation. This paper provides another step towards understanding the tradeoff between DSM algorithm's complexity and performance guarantee.

Acknowledgement

We would like to acknowledge collaboration on the GP reformulation with Lieven Vandenberghe at University of California Los Angeles.

Bibliography

[1] DSL Forum (www.dslforum.org), "DSL dominates global broadband subscriber growth," Tech. Rep., Mar. 2007, internet: http://www.broadband-forum.org/news/download/pressreleeases/YE06_Release.pdf.

[2] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 922–933, May 2006.

[3] R. Cendrillon and M. Moonen, "Iterative spectrum balancing for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 3, May 2005, pp. 1937–1941.

[4] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spec-

trum balancing for digital subscriber lines," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4241–4257, Aug. 2007.

[5] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[6] P. Tsiaflakis, M. Diehl, and M. Moonen, "Distributed spectrum management algorithms for multiuser DSL networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4825–4843, Oct. 2008.

[7] P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "A low complexity optimal spectrum balancing algorithm for digital subscriber lines," *Signal Processing*, vol. 87, no. 7, pp. 1735–1753, Jul. 2007.

[8] Y. Xu, S. Panigrahi, and T. Le-Ngoc, "A concave minimization approach to dynamic spectrum management for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 1, Jun. 2006, pp. 84–89.

[9] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1310–1322, Jul 2006.

[10] V. M. K. Chan and W. Yu, "Multiuser spectrum optimization for discrete multitone systems with asynchronous crosstalk," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5425–5435, Nov. 2007.

[11] T. Starr, M. Sorbara, J. M. Cioffi, P. J. Silverman, *DSL Advances, ch. 11.* Prentice Hall, 2002.

[12] Y. Xu, S. Panigrahi, and T. Le-Ngoc, "Optimal spectrum management with group power constraint for digital subscriber loops," in *Proc. of the Canadian Conference on Electrical and Computer Engineering*, May 2005, pp. 1598–1601.

[13] Y. Nesterov and A. Nemirovsky, *Interior Point Polynomial Methods in Convex Programming.* SIAM Press, 1994.

[14] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences.* SIAM Classics in Applied Mathematics, 1994.

[15] J. J. McDonald, M. Neumann, H. Schneider, and M. J. Tsatsomeros, "Inverse m-matrix inequalities and generalized ultrametric matrices," *Linear Algebra and its Applications*, vol. 220, pp. 321–341, Apr. 1995.

[16] *Asymmetrical digital subscriber line (ADSL) transceivers*, ITU Std. G.992.1, 1999.

Chapter 5

Cross-Layer Spectrum Coordination

Throughput and Delay of DSL Dynamic Spectrum Management with Dynamic Arrivals

P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen

Published in *Proc. of the IEEE Global Communications Conference (GLOBECOM)*, New Orleans, Louisiana, Nov. 2008, pp. 1-5.

©2008 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Contributions of first author

- literature study
- co-development of extended cross-layer system model
- co-development of throughput-optimal joint scheduling and spectrum coordination algorithms
- co-development of δ-randomized spectrum coordination algorithms: R1-/R2-/R3-DSM
- design of simulation scenarios
- software implementation and computer simulations
- co-analysis and co-interpretation of simulation results
- co-formulation of conclusion
- text co-redaction, co-editing, co-revision

Abstract

In modern DSL networks, crosstalk among different lines (i.e., users) is the major source of performance degradation. Dynamic spectrum management (DSM) refers to a set of techniques to mitigate the effect of crosstalk leading to spectacular performance gains. However the main research efforts in DSM aim at only physical layer performance whereas the true end user experience depends on what they see at the application rather than the physical layer. Upper layer performance metrics like throughput and delay may be much more important to improve the user satisfaction. To that end, we provide a framework to study upper layer performance by looking at scheduling and DSM together. We show how optimal scheduling can be combined with optimal DSM and provide throughput-optimal scheduling algorithms which require only polynomial complexity. We furthermore present extensions that significantly improve delay performance by using the specific structure of the underlying problem.

5.1 Introduction

Digital subscriber line (DSL) technology provides broadband access over twisted pair copper wires of the existing telephone network. Nowadays it is still the most popular broadband access technology worldwide. The major obstacle for performance improvement remains the excessive electromagnetic interference, also called crosstalk, generated among different lines in the same cable bundle.

One promising set of techniques for tackling the crosstalk problem is dynamic spectrum management (DSM). The main idea of DSM is to prevent and/or remove crosstalk by spectrum and/or signal coordination, respectively. In this paper, we focus on spectrum level coordination, also referred to as spectrum management, spectrum balancing, or multi-carrier power control. From an information-theoretic point of view, this is known as a multi-carrier interference channel where the interference from the other users is treated as noise. Many DSM algorithms[1] are proposed in literature ranging from fully autonomous [1, 2] and distributed [3, 4, 5] to centralized algorithms [6, 7, 8].

The major research efforts in DSM algorithms aim at only physical layer performance, which is to maximize the data rates subject to power constraints. However the end user experience depends on what they see at the application rather than physical layer. The application workload naturally arrives with only a finite workload instead of the infinite workload assumed by physical layer DSM design. Furthermore an increasing amount of applications (e.g. video and voice) are delay-sensitive. Motivated by the above two that have

[1] In this text DSM refers to spectrum balancing.

not been treated in existing and vast DSM literature, we look at how upper layer scheduling and physical layer DSM can work together so as to incorporate upper layer performance metrics such as throughput and delay. As we show later, both optimal scheduling and DSM are computationally intractable, i.e., NP-hard. This paper tries to address the questions of how relaxations on both problems work well together towards good upper layer performance.

The main contributions of this paper are as follows:

(i) We provide a generalized framework to study the upper layer performance over DSL networks, motivated by the recent research in other areas, e.g., wireless scheduling. This framework is a useful tool to understand throughput and delay performance provided by DSM algorithms at the physical layer, as well as gives us practical implications on the design of future DSM algorithms. The proposed framework facilitates characterization of throughput and delay properties, and also discloses the generic trade-off between complexity, and throughput or delay.

(ii) Using this framework, we connect throughput-optimal scheduling to the globally optimal DSM algorithms that require exponential computational complexity. However, somewhat surprisingly, we show that it is possible to achieve optimal throughput by using (randomized) sub-optimal DSM algorithms with only polynomial time complexity. However, the price to pay is increasing delay.

(iii) Then, we provide algorithms which significantly improve delay performance with only small extra complexity, by exploiting the specific structure of our problem, namely correlation in frequency and time domain, and also single user optimality for large crosstalk DSL scenarios.

5.2 System Model and Performance Metrics

5.2.1 System Model

Network and Traffic Models. We consider a discrete time-slotted system, indexed by t, consisting of N interfering DMT-DSL modems or users. We denote by K the number of frequency bands or tones available for each user. We abuse the notations N and K to refer to the index set of users and tones. Each user has an infinite-size buffer which is fed by exogenous arrivals. We denote by $A^n(t)$ the number of arrivals (in bits) to the user $n \in N$, at time t. The arrival process is assumed to be i.i.d. across time-slots, where $\mathbb{E}[A^n(t)] = \lambda^n$. We assume that the duration of one time-slot is small enough so that λ^n is upper bounded by a constant, denoted by Ω, i.e., $A^n(t) \leq \Omega$, $\forall n \in N$, $\forall t \geq 0$. Denote by $\boldsymbol{\lambda} = (\lambda^n : \ n \in N)$ the (mean) arrival vector across users.

Resource Model. The network resources are represented by a finite set $\mathcal{R}$ of

the feasible rate vectors, referred to as *rate region* describing the simultaneous achievable rates (in bits/slot) of the users. The data rates of the users in turn depend on their transmit powers and the resulting interference across tones and users. This results in the following characterization of the achievable rate region $\mathcal{R}$:

$$\mathcal{R} = \Big\{(R^n : n \in N) | R^n = f_s \sum_k b_k^n, \sum_{k \in K} s_k^n \leq P^n, 0 \leq s_k^n \leq s_k^{n,\text{mask}}\Big\},$$

where s_k^n denotes the transmit power of user n on tone k, P^n denotes the total power budget available to user n, $s_k^{n,\text{mask}}$ denotes the spectral mask constraint for user n on tone k, and f_s denotes the DMT symbol rate. The b_k^n, which depends on the transmit powers, denotes the bit rate of user n on tone k, given by:

$$b_k^n \triangleq \log_2\left(1 + \frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^n}{\sum_{m \neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n}\right) \text{ bits/ s/ Hz}, \tag{5.1}$$

where $[\mathbf{H}_k]_{n,m} = h_k^{n,m}$ is an $N \times N$ matrix containing the channel transfer functions from transmitter m to receiver n on tone k. The diagonal elements are the direct channels, the off-diagonal elements are the crosstalk channels. σ_k^n denotes the noise power that contains thermal noise, alien crosstalk and radio frequency interference (RFI). Note that the channel and noise are characterized as constant over time but it can simply be extended to be varying in function of time slots. Γ denotes the signal-to-noise ratio (SNR) gap to capacity, which is a function of the desired bit error ratio (BER), the coding gain and noise margin [9].

Resource Allocation Scheme. Then, a resource allocation scheme chooses a sequence of rate schedules $(\boldsymbol{R}(t) = (R^n(t) : n \in N))_{t=0}^{\infty}$, $\boldsymbol{R}(t) \in \mathcal{R}$. Since the main resources of our system are the transmit power spectra $(s_k^n(t) : n \in N, k \in K)$ and the rate schedule $\boldsymbol{R}(t)$ is determined by the allocated power spectra, we use "resource allocation" and "(power) scheduling" algorithm interchangeably throughout the paper.

Queueing Dynamics. Denote by $Q^n(t)$ the queue length of the buffer of user n at time t. The queueing dynamics are then defined by the following recursion:

$$Q^n(t+1) \quad = \quad \big[Q^n(t) - R^n(t)\big]^+ + A^n(t+1), \tag{5.2}$$

where $[x]^+ = \max(x, 0)$.

5.2.2 Performance Metrics

Performance metrics considered in this paper are throughput and delay that may be the two performance goals in networking and communication systems.

To define throughput, we first define the notion of stability that essentially represents the condition that queue-lengths remain finite.

Definition 5.1 [Stability] The system is said to be stable, if

$$\limsup_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t} \mathbb{E}\left[\sum_{n\in N} Q^n(\tau) \right] < \infty.$$

□

A primary performance objective of any scheduling is to guarantee stability whenever possible, i.e., whenever the arrival rate vector $\boldsymbol{\lambda}$ belongs to the throughput region defined as:

Definition 5.2 [Throughput-region] The throughput-region $\Lambda \subset \mathbb{R}_+^N$ is the set of all arrival vectors for which there exists a scheduling algorithm stabilizing the system. □

Throughput-region can be regarded as the maximum throughput that one can achieve. In other words, no scheduling can stabilize the systems for the arrival rate vector outside of the throughput-region. The throughput-region can be characterized as: $\Lambda = \text{convex-hull}(\mathcal{R})$ [10]. We say that a scheduling which can stabilize the system for any arrival rate vector in Λ is *throughput-optimal.*

In addition to throughput, we also consider the total (stationary) queue length over users, given by: $\sum_{n\in N} \mathbb{E}[Q^n(\tau)]$. This naturally relates to delay from Little's law.

Our objective is to develop a resource allocation scheme for this coupled N-user DSL system, that achieves throughput-optimality *without* knowledge of the mean arrival rates.

5.3 Throughput-Optimal Scheduling and DSM Algorithms

Conventional DSM algorithms aim at optimizing physical layer performance, i.e. allocate transmit powers so to maximize data rates subject to power constraints. Here we will extend the objective to upper layer performance metrics such as throughput and delay.

We first describe a scheduling that is throughput-optimal, and then show its connection to the conventional DSM problem formulation. To that end, we

first define the *weight* of the rate schedule $\boldsymbol{R}(t)$ as: $W(t) \triangleq W(\boldsymbol{R}(t)) \triangleq \sum_{n\in N} Q^n(t)R^n(t)$. Consider the following scheduling, referred to as Max-Weight (MW): at time-slot t, it schedules $\boldsymbol{R}^\star(t)$ that maximizes the weight, i.e.,

$$\boldsymbol{R}^\star(t) = \arg\max_{\boldsymbol{R}\in\mathcal{R}} \sum_{n\in N} Q^n(t)R^n.$$

It has been proved that MW scheduling is throughput-optimal under slightly different system models (e.g., [11]), and extension to our system model is straightforward. By incorporating the DSM physical layer resources introduced in Section 5.2.1, it is easy to observe that MW scheduling comes down to solving the following highly nonconvex optimization problem at every time-slot t:

$$\begin{aligned} \max_{s_1^1,\ldots,s_K^N} \quad & \sum_{n=1}^{N} Q^n(t)R^n(t) \\ \text{s.t.} \quad & \sum_{k=1}^{K} s_k^n(t) \le P^n, \qquad \forall n, \\ & 0 \le s_k^n(t) \le s_k^{n,\text{mask}}, \ \forall n,k. \end{aligned} \tag{5.3}$$

Note that the objective of the existing (rate-adaptive) DSM algorithms in literature is to solve the optimization problem in (5.3), where the queue length of the user n is used as its weight[2], i.e., $w^n = Q^n(t)$ (see e.g., [6]). In other words, existing DSM algorithms can be reused as an underlying building block of Max-Weight scheduling, where the DSM problem has to be solved at each time slot, replacing w^n by the queue lengths $Q^n(t)$ at time slot t.

At this point we distinguish three different types of DSM algorithms that are (i) globally optimal, (ii) locally optimal, and (iii) heuristic. Globally optimal DSM algorithms, such as OSB [6], BB-OSB[7] and PBB[8], succeed in finding the globally optimal solution of (5.3) where the feasible set of transmit power spectra is discretized up to accuracy Δ_s. Locally optimal DSM algorithms, such as SCALE[3], DSB[4] and MIW[5], only guarantee a locally optimal solution to (5.3). Heuristic DSM algorithms, such as IW[1] and ASB[2], do not necessarily ensure globally or locally optimal solutions to (5.3), but have practically strong merits in near-optimality with reasonably low complexity.

Thus, scheduling based on globally optimal DSM algorithms can be regarded as throughput-optimal algorithms, i.e., achieving maximum throughput performance. However, these algorithms require computationally intractable complexity (i.e., NP-hard in terms of N), leading to difficulty in practical implementation. Our question is whether we can achieve throughput-optimality even

[2]This weight differs from the weight of a "schedule", but for simplicity we use the same term for both cases.

with sub-optimal DSM algorithms having lower, practical complexity. In this paper, due to space limitations, we restrict our focus on just locally optimal algorithms, and further we do not differentiate centralized and distributed algorithms. To formally study, we first introduce randomized scheduling algorithms in the next section.

5.4 Polynomial Complexity Scheduling Achieving Throughput Optimality

Locally optimal DSM algorithms [3, 4, 5] require only polynomial complexity. However they sometimes fail to find the globally optimal solution of the nonconvex problem (5.3). Indeed, their final solution strongly depends on the chosen *initial starting point.* Some initial starting points lead to the globally optimal solution whereas others lead to only a locally optimal solution. We note that there always exist at least one initial starting point or a region of starting points leading to global optimality. We will exploit this starting point dependence in combination with an appropriate scheduling scheme to design polynomial complexity throughput-optimal scheduling algorithms.

5.4.1 δ-Randomized DSM Algorithms

To achieve our goal, we first introduce a notion of δ-randomized DSM algorithms, defined as follows:

Definition 5.3 [δ-randomized DSM] A DSM algorithm, which produces a random schedule $\boldsymbol{R}'(t)$, is δ-randomized for some $0 < \delta \leq 1$, if at each time-slot t,

$$\mathbb{P}[\boldsymbol{R}'(t) \geq \boldsymbol{R}^{\star}(t) \mid \boldsymbol{Q}(t)] \geq \delta, \tag{5.4}$$

where $\boldsymbol{R}^{\star}(t)$ is the optimal objective value of (5.3). □

In other words, a δ-randomized DSM algorithm randomly generates a rate schedule $\boldsymbol{R}'(t)$ that is guaranteed to be as good as the optimal rate schedule with non-zero probability. Clearly, globally optimal DSM algorithms are 1-randomized DSM. Note that, as mentioned in Section 5.3, optimality is defined in the discretized setting.

Now, consider the scheduling algorithm, presented in Algorithm 5.1, using a δ-randomized DSM, referred to as *Pick-and-Compare*. Pick-and-Compare scheduling can be interpreted as a randomized scheduling that produces a "reasonably good" schedule in terms of its non-zero probability of finding a

Algorithm 5.1 Pick-and-Compare Scheduling: at time-slot t

Step 1 "Pick" a random rate schedule $\boldsymbol{R}'(t)$ using δ-randomized DSM.
Step 2 Compute the weight of $\boldsymbol{R}'(t)$, i.e., $W(\boldsymbol{R}'(t))$.
Step 3 "Compare" $W(\boldsymbol{R}'(t))$ and $W(\boldsymbol{R}(t-1))$, and take the rate schedule with larger weight for the rate schedule at slot t, i.e., $\boldsymbol{R}(t) = \arg\max_{\boldsymbol{S}\in\{\boldsymbol{R}'(t),\boldsymbol{R}(t-1)\}} W(\boldsymbol{S})$.
Step 4 Apply $\boldsymbol{R}(t)$ to transmit data.

globally optimal schedule (**Step 1**), in conjunction with progressively selecting better schedules by comparing the previous schedule and the current randomized schedule (**Step 3**). Again, when $\delta = 1$, the Pick-and-Compare scheduling recovers the Max-Weight scheduling algorithm. Theorem 5.4 states the throughput property of the Pick-and-Compare scheduling.

Theorem 5.4 *For any $0 < \delta \leq 1$, a Pick-and-Compare scheduling with δ-randomized DSM is throughput-optimal.*

The same result has been proved under different system models (see e.g., [12, 13] in wireless scheduling). Thus, we omit the complete proof due to space limitations, but provide only sketch of the proof: The stochastic stability of the considered queueing system can be proved by defining a certain Lyapunov function and then proving that the system has a negative Lyapunov drift whenever the aggregate queue-length is greater than some constant B. The throughput-optimality of MW scheduling can be proved by defining a quadratic Lyapunov function, $L(t) \triangleq \sum_{n\in N} Q^n(t)^2$. In fact, we can easily show that Pick-and-Compare scheduling satisfies the following: $W(t) \geq W^\star(t) - C(t)$, where $E[C(t)] \leq K$, for some constant K. The (random) suboptimality $C(t)$ is due to the fact that Pick-and-Compare can find the maximum-weight scheduling only every $1/\delta$ slot on average, and this additive suboptimality does not affect the stability region.

These δ-randomized algorithms indeed represent a parameterized family of scheduling algorithms that achieve throughput-optimality. The parameter δ can range from 1 to a very small number, where typically, a smaller δ leads to algorithms with lower complexity at the cost of increasing delays, as discussed in Sections 5.4.2 and 5.4.3.

Now, we provide a simple way of realizing δ-randomized DSM using the existing locally optimal DSM algorithms. This is based on random selection of an initial starting point. We first denote by $\mathcal{T}$ the set of all possible transmit power combinations discretized up to accuracy Δ_s:

$$\mathcal{T} = \{(y_{ij} : i \in N, j \in K) | y_{ij} \in \{0, \Delta_s, \ldots, s_j^{i,\max}\}\}, \tag{5.5}$$

where $s_j^{i,\max} \triangleq \min(P^i, s_j^{i,\text{mask}})$. Note that the globally optimal solution $\boldsymbol{s}^\star$, obtained by the globally optimal DSM algorithms, belongs to this set. By randomly choosing an element of $\mathcal{T}$ and using it as the initial transmit power vector for the locally optimal DSM algorithm, there should be a non-zero probability to converge to globally optimal performance. More specifically, if the optimal solution is chosen from $\mathcal{T}$ (i.e., we allow only discretized powers in the system), it converges to the global-optimal solution as the locally optimal DSM algorithms are monotonically increasing over their successive iterations. This leads to a δ-randomized DSM algorithm, which will be referred to as **R1-DSM**. Then, **R1-DSM** can be used in combination with the Pick-and-Compare scheduling of Algorithm 5.1 to obtain a throughput optimal DSM scheduling scheme with only polynomial complexity, where the lower-bound on δ is $\frac{1}{|\mathcal{T}|}$.

R1-DSM (Random Initial Point):

1. Pick a random initial transmit power $\boldsymbol{x} \in \mathcal{T}$ uniformly,
2. Apply a locally optimal DSM algorithm starting with $\boldsymbol{x}$.

5.4.2 Delay Performance of δ-Randomized DSM

As shown in Section 5.4.1, it is possible to achieve throughput optimality with randomized locally optimal DSM algorithms. What price should be paid for this complexity reduction from exponential to polynomial without losing throughput region? Next, in Section 5.4.3 we will show that the price is *delay*. Further, we use this characterization of price to develop other scheduling algorithms with polynomial complexity that improve delay performance with small extra complexity.

Calculating the exact delay performance in our system is known to be very difficult, still an open problem mainly due to complex coupling of queueing dynamics across users, tones, and stochastic arrivals. Therefore we rely on a delay bound. Although this delay bound may not be tight in some scenarios, this bound is quite helpful to understanding how delay performance scales with δ, which in turn relates to the complexity of scheduling algorithms.

The delay performance should depend on the input arrival rate $\boldsymbol{\lambda}$, e.g., as $\boldsymbol{\lambda}$ approaches the boundary of the throughput-region, then delay will correspondingly increase. To quantify this intuition, we use the notion of distance between the arrival vector and the throughput region as follows:

Definition 5.5 [Distance]

$$d(\boldsymbol{\lambda}) = \sup\{\delta : \boldsymbol{\lambda} \in (1-\delta)\Lambda\}. \tag{5.6}$$

□

This distance essentially represents how heavily the system is loaded, where smaller $\boldsymbol{\lambda}$ leads to larger $d(\boldsymbol{\lambda})$. Using this definition, the following bound on the delay can be obtained for randomized scheduling with δ-randomized DSM algorithm and $\boldsymbol{\lambda} \in \Lambda$.

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{n\in N} \mathbb{E}\Big[Q^n(t)\Big] \quad \leq \quad \frac{N^2\Omega(2/\delta+1)}{d(\boldsymbol{\lambda})}.$$

Recall that Ω is the maximum number of arrivals at one slot. We can observe that as δ decreases, the delay bound increases. We skip the proof for space limitation. The idea is to define a quadratic Lyapunov function, and to telescope its drift across time-slots (see [10] for details).

Now, let us investigate the delay performance of Pick-and-Compare with **R1-DSM**. As discussed earlier, $\delta \geq \frac{1}{|\mathcal{T}|}$, where the delay bound becomes very large, i.e., delay bound $\approx O(|\mathcal{T}|)$. This indicates that the delay of **R1-DSM** is pretty large in the worst-case sense. However, note that in many cases, δ may be much larger since depending on the scenario, we may have many possible initial starting points converging to the global optimal solution. Next, we will propose a modified DSM that improves the delay performance significantly with small extra complexity.

5.4.3 Significant Delay Decrease with Small Extra Complexity

Locally optimal DSM algorithms reduce complexity, which, instead, guarantees finding the global optimal solution only probabilistically (in conjunction with a random initial point) with negative effect on delay. Can we improve delay performance significantly (i.e., increase the probability of finding the globally optimal solution), by exploiting the specific structure of our problem? The answer is positive, and we will use the following specific features to achieve the goal:

1. **Time-slot correlation**: When queue lengths are large (i.e., the input arrival vector is close to the boundary of throughput-region), the system does not observe much difference in weights (i.e., queue lengths) over subsequent time-slots.
2. **Tone correlation**: Subsequent tones have similar channel characteristics and so also have similar solutions.

These problem-specific correlations across time-slot and tones can actually be exploited to choose the initial starting points more intelligently. One can add two additional initial starting point to the random point in each tone: (1) the best local optimum at that tone from the previous time slot using the idea of *time-slot correlation*, and (2) the best local optimum from the previous tone

using the idea of *tone correlation*. The inclusion of these extra initial points will increase δ, resulting in better delay performance. This leads to the following δ-randomized DSM algorithm that can be combined with the Pick-and-Compare in Algorithm 5.1 to obtain a throughput-optimal scheduling with significantly improved delay performance, compared to **R1-DSM**.

R2-DSM (Time-slot and tone correlation):

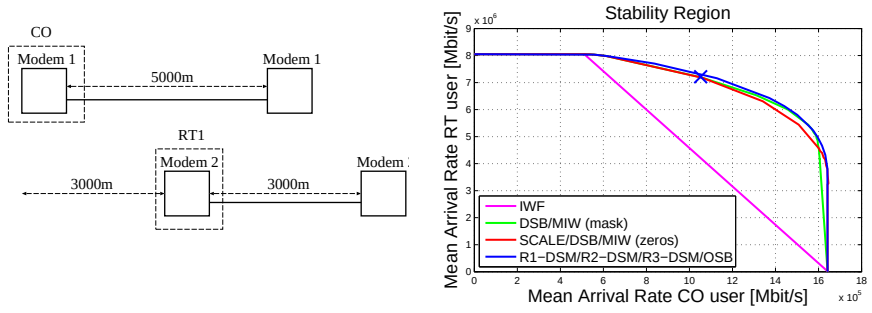
1. For each tone, choose three initial starting points: *i)* random point, *ii)* the best local optimum at that tone from the previous time slot, *iii)* the best local optimum from the previous tone,
2. Apply a locally optimal algorithm for all initial starting points and retain best converged local optimum in each tone.

A third δ-randomized DSM is inspired by the recent result in [14], where it is stated that for large crosstalk scenarios the solution of (5.3) is an FDMA solution, i.e. a solution where only one user is active at each tone. We indeed have observed that for large crosstalk scenarios the locally optimal solutions get isolated along the axes [4]. Therefore in addition to the random initial point, we propose to extend the number of initial points in each tone $k \in K$ so that it includes all the solutions where only one user transmits at spectral mask $s_k^{n,\text{mask}}$. By adding these initial points, the likelihood that one of these initial starting points leads to the globally optimal solution of that tone, increases significantly. This results in the following δ-randomized DSM algorithm:

R3-DSM (N single-user initial points):

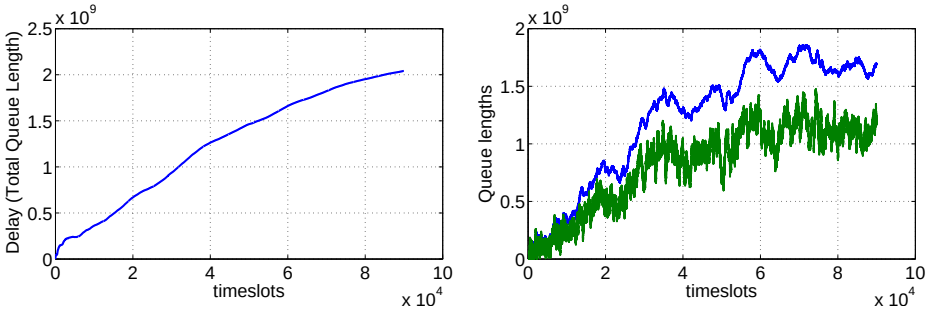
1. For each tone, choose $N+1$ initial starting points: *i)* random point, *ii)* N single user initial starting points at spectral mask,
2. Apply locally optimal algorithm for all initial starting points and retain best converged local optimum in each tone.

Note that all of **R1-DSM**, **R2-DSM**, and **R3-DSM** provably achieve the maximum throughput. We observe that there is a clear trade-off between complexity and delay. In terms of complexity, **R2-DSM** requires 2 times more complexity than **R1-DSM**, and **R3-DSM** requires N times more complexity than **R1-DSM**. In terms of δ, inverse proportional to the delay, we have that $\delta_{\text{R1-DSM}} \leq \delta_{\text{R2-DSM}} \simeq \delta_{\text{R3-DSM}}$. In fact for very nonconvex DSL scenarios we observe that $\delta_{\text{R1-DSM}} \ll \delta_{\text{R2-DSM}} \simeq \delta_{\text{R3-DSM}} \simeq 1$ which means that the delay performance is much better for **R2-DSM** and **R3-DSM** for just a slight increase in complexity.



(a) (b)

Figure 5.1: Simulation Scenario and Stability Region. The curve indicated by DSB/MIW (mask) refers to the stability region of the DSB and MIW algorithms with fixed initial starting point equal to spectral mask, i.e. $s_k^n = s_k^{n,\text{mask}}, \forall n, k$. The curve indicated with SCALE/DSB/MIW (zeros) starts from the all zero starting point.

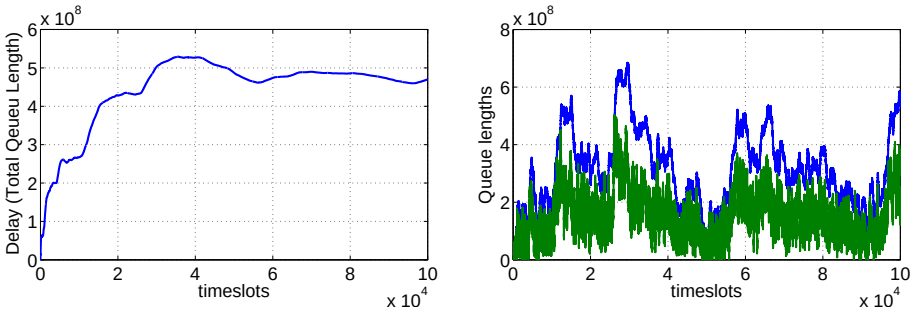


(a) Aggregate total queue length (b) Instantaneous queue lengths [bits]

Figure 5.2: Stability and delay behaviour of Pick-and-Compare with DSB (zeros) algorithm: (blue: CO user, green: RT user)

5.5 Simulation Results

This section presents simulation results of throughput and delay properties of the proposed algorithms for a realistic ADSL downstream scenario shown in Figure 5.1(a). The twisted pair lines have a diameter of 0.5 mm (24 AWG). The maximum transmit power is 20.4 dBm. The SNR gap Γ is 12.9 dB. The tone spacing Δ_f is 4.3125 kHz. The DMT symbol rate f_s is 4 kHz. Note that for most of the DSL scenarios, locally optimal algorithms perform globally optimal. For this reason we intentionally chose to focus on the CO-RT scenario of Figure 5.1(a), which is known as a very difficult scenario where locally optimal



(a) Aggregate total queue length (b) Instantaneous queue lengths [bits]

Figure 5.3: Stability and delay behaviour of Pick-and-Compare with **R2-DSM**: (blue: CO user, green: RT user)

algorithms generally fail to find the globally optimal solution.

In Figure 5.1(b) we show the stability regions for various algorithms including randomized algorithms proposed in this paper. Note that the throughput of the randomized DSM algorithms is optimal in contrary to the algorithms with fixed initial points, whose stability regions are sub-optimal and depend on the initial starting points. In Figures 5.2 and 5.3 we show the delay performance and instantaneous queue lengths of Pick-and-Compare scheduling combined with DSB (or similarly for MIW) with fixed zeros initial points and **R2-DSM**, respectively, for the mean arrival rate indicated with the cross in Figure 5.1(b). One can see that **R2-DSM** succeeds in stabilizing the system with only small delay, whereas the queues for DSB keep increasing after 100000 time slots. Finally we would like to remark that the delay performance of **R2-DSM** and **R3-DSM** is much better with respect to **R1-DSM**, as we analyzed in the delay bound.

5.6 Conclusion

It is crucial to understand and design algorithms for DSL systems from the perspective of QoS that will be experienced by the users. In this paper, we show that the results from physical-layer DSM, under assumption of infinite backlog, may be significantly changed. More specifically, by jointly considering upper layer scheduling and the physical-layer DSM algorithm, somewhat surprisingly, even sub-optimal DSM algorithms can achieve throughput maximization, when intelligent scheduling algorithms are employed, with an interesting trade-off between complexity and delay.

Bibliography

[1] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[2] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spectrum balancing for digital subscriber lines," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4241–4257, Aug. 2007.

[3] J. Papandriopoulos and J. S. Evans, "Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 7, Jun. 2006, pp. 3270–3275.

[4] P. Tsiaflakis, M. Diehl, and M. Moonen, "Distributed spectrum management algorithms for multiuser DSL networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4825–4843, Oct. 2008.

[5] W. Yu, "Multiuser water-filling in the presence of crosstalk," in *Proc. of the Information Theory and Applications (ITA)*, Feb. 2007.

[6] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 922–933, May 2006.

[7] P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "A low complexity optimal spectrum balancing algorithm for digital subscriber lines," *Signal Processing*, vol. 87, no. 7, pp. 1735–1753, Jul. 2007.

[8] Y. Xu, T. Le-Ngoc, and S. Panigrahi, "Global concave minimization for optimal spectrum balancing in multi-user DSL networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 2875–2885, Jul. 2008.

[9] T. Starr, J. M. Cioffi, P. J. Silverman, *Understanding digital subscriber lines*. Prentice Hall, 1999.

[10] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Foundations and Trends in Networking*, vol. 1, no. 1, pp. 1–144, 2006.

[11] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling for maximum throughput in multihop radio networks," *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936–1949, Dec. 1992.

[12] L. Tassiulas, "Scheduling and performance limits of networks with constantly changing topology." *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 1067–1073, 1997.

[13] S. Sanghavi, L. Bui, and R. Srikant, "Distributed link scheduling with constant overhead," in *Proceedings of ACM Sigmetrics*, 2007.

[14] S. Hayashi and Z. Q. Luo, "Dynamic spectrum management: When is FDMA sum-rate optimal?" in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 3, Apr. 2007, pp. III–609 – III–612.

Chapter 6

Green DSL: Energy-Aware Spectrum Coordination

Green DSL: Energy-Efficient DSM

P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen

Published in *Proc. of the IEEE International Conference on Communications (ICC)*, Dresden, Germany, Jun. 2009, pp. 1-5.

©2009 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Contributions of first author

- literature study
- co-development of Green DSL framework
- co-development of Green DSL general solution procedure
- co-analysis of Green DSL problem formulations
- co-analysis of Green DSL performance impact
- design of simulation scenarios
- software implementation and computer simulations
- co-analysis and co-interpretation of simulation results
- co-formulation of Green DSL core messages
- co-formulation of conclusion
- text co-redaction, co-editing, co-revision

Abstract

Dynamic spectrum management (DSM) has been recognized as a key technology for tackling multi-user crosstalk interference for DSL broadband access. Up to now, DSM design has mainly been focusing on maximization of data rates. However, recently, reducing the total power has become a main target, as IT power consumption has been identified as a significant contributor to global warming. In this paper we extend traditional DSM design towards a much wider energy-efficient scope and show how to tackle the corresponding optimization problems. The impact of this 'green DSL' approach is evaluated for practice with some surprisingly good numerical results. Furthermore bounds are provided on the trade-off between data rate performance and power saving.

6.1 Introduction

Digital subscriber line (DSL) technology refers to a family of technologies that provides digital broadband access over the local telephone network. It is still the dominating broadband access technology with 66% of all broadband access subscribers worldwide using DSL to access the Internet [1]. In order to cope with the increasing demands of the users (end-users as well as service providers) and to stay competitive with other broadband access technologies, DSL technology is continuously extended to encounter the corresponding technological issues. One of the major challenges, is to overcome the electromagnetic interference, also called *crosstalk*, generated among different lines operating in the same cable bundle. Different lines (i.e. users) indeed interfere with each other, leading to a very challenging interference environment where proper management is required to prevent a huge performance degradation.

Dynamic spectrum management (DSM) is recognized as a key technology for tackling this crosstalk problem [2]. The main idea of DSM is to prevent and/or remove crosstalk by using spectrum and/or signal coordination of all users, respectively. We will focus on spectrum coordination, also referred to as spectrum balancing or multi-carrier power control.

The major research efforts in DSM[1] algorithm design have been focusing on maximizing the data rates (i.e. rate-adaptive DSM [3]) without any power minimizing design objective (margin-adaptive [3]), except for the power constraints defined by DSL standards. However recently power consumption has started to gain a lot of importance (ITU-T Study Group 15, European Code of Conduct for broadband equipment). Information and Communication Technologies (ICTs) have been identified as significant contributors to *global warming* [4]. Broadband equipment contributes to the electricity consumption and

[1] In this text, DSM refers to spectrum coordination.

depending on the penetration level, the specifications of the equipment and the requirements of the service provider, a total European consumption of up to 50 TWh per year can be estimated for the year 2015 [5]. Therefore the European Code of Conduct for Broadband Equipment takes initiative in setting up general principles and actions, and targets to limit the (maximum) electricity consumption to 25 TWh per year which is equivalent to 5.5 Million ton of oil equivalent (TOE) and to a total saving of about € 7.5 Billions per year. DSL, as the most deployed broadband technology, plays an important role in this setup [6]. The DSL Forum encourages international standards bodies to develop techniques for power reduction within the scope of their activities and to maximize the savings while preserving and enhancing quality of service [6]. One technology that fits well in this framework is DSM.

This gave us the motivation to revisit DSM and extend its design with power related objectives. This approach would benefit from the traditional pure data rate maximizing approach in two ways [7]: (i) Adding objectives and/or constraints for limiting transmit power reduces overall power consumption by DSL systems, as power consumed by DSL modems is often dominated by circuits used for transmitting power, and (ii) smaller transmit powers encourage a 'polite' behaviour as less crosstalk is radiated into other DSL modems. These benefits of making the copper 'greener' in an evolving and increasingly energy-efficient world, were recently also pointed out by the authors in [7].

The main contributions of this paper are as follows:

(i) We extend the traditional rate-adaptive DSM design to a much wider 'green' setting, incorporating power limiting objectives and/or constraints. This unifying 'green DSL' framework leads to a much larger freedom and potential of DSM in managing QoS for DSL broadband access.
(ii) We provide a systematic procedure for tackling this extended set of (non-convex) DSM problem formulations by introducing extended Lagrange multipliers and slightly modifying existing DSM algorithms.
(iii) We demonstrate the substantial potential of 'green DSL' in achieving large power savings of up to 50% while still achieving 85% of full-power data rate performance for realistic DSL scenarios. Furthermore lower bounds are provided on the data rate performance versus power saving.

6.2 System Model

We consider a system model consisting of N interfering (synchronous) DMT-DSL modems or users. We denote by K the number of frequency bands or tones available for each user. We abuse the notations N and K to refer to the index set of users and tones. Modems employ single-user encoding and decoding

(treating interference as noise). This results in the following characterization of the achievable rate region $\mathcal{R}$:

$$\mathcal{R} = \Big\{(R^n : n \in N) | R^n = f_s \sum_{k \in K} b_k^n(\mathbf{s}_k), \mathbf{s} \in \mathcal{S}\Big\},$$

where

$$\mathcal{S} = \Big\{\mathbf{s} \in \mathbf{R}^{N \times K} : s_k^n = [\mathbf{s}]_{n,k}, \sum_{k \in K} s_k^n = P^n, P^n \leq P^{n,\text{tot}}, 0 \leq s_k^n \leq s_k^{n,\text{mask}}\Big\}, \tag{6.1}$$

$$b_k^n(\mathbf{s}_k) \triangleq \log_2\left(1 + \frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^n}{\sum_{m \neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n}\right) \text{bits/Hz}, \tag{6.2}$$

and where R^n denotes the data rate of user n, f_s denotes the DMT symbol rate, s_k^n denotes the transmit power of user n on tone k, b_k^n, which depends on the transmit powers $\mathbf{s}_k = [s_k^1, \ldots, s_k^N]^T$, denotes the bit rate of user n on tone k, $P^{n,\text{tot}}$ denotes the total power budget available to user n and $s_k^{n,\text{mask}}$ denotes the spectral mask constraint for user n on tone k. Furthermore $[\mathbf{H}_k]_{n,m} = h_k^{n,m}$ is an $N \times N$ matrix containing the channel transfer functions from transmitter m to receiver n on tone k. The diagonal elements are the direct channels, the off-diagonal elements are the crosstalk channels. σ_k^n denotes the noise power on tone k in receiver n that contains thermal noise, alien crosstalk and radio frequency interference (RFI). Γ denotes the signal-to-noise ratio (SNR) gap to capacity, which is a function of the desired bit error ratio (BER), the coding gain and noise margin [8]. Note that set $\mathcal{S}$ constitutes the set of constraints that are imposed by DSL standards, including total power constraints as well as spectral mask constraints.

6.3 Green DSL

6.3.1 Unifying Green DSL Formulation

The most common approach [9, 10, 11, 12, 13, 14] to DSM design is a fully data rate driven approach where a weighted sum of data rates is to be optimized as follows:

$$\begin{array}{ll} \max\limits_{\{s_k^n, k \in K, n \in N\}} & \sum\limits_{n \in N} w_n R^n \\ \text{s.t.} & P^n \leq P^{n,\text{tot}}, n \in N \\ & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, k \in K, n \in N, \end{array} \tag{6.3}$$

where $w_n, \forall n$, are weights to specify the importance of each user n. Extra lower bounds on the data rates $R^n \geq R^{n,\text{target}}, \forall n$, are sometimes also added, where $R^{n,\text{target}}$ corresponds to the minimum data rate requirement for user n.

A first relevant extension towards an energy-minimizing approach would be to impose a constraint on the sum of all allocated transmit powers so as to reduce the total consumed power by factor α, leading to the following DSM design:

$$\begin{array}{ll} \max\limits_{\{s_k^n, k\in K, n\in N\}} & \sum\limits_{n\in N} w_n R^n \\ \text{s.t.} & P^n \leq P^{n,\text{tot}}, n \in N \\ & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, k \in K, n \in N \\ & R^n \geq R^{n,\text{target}}, \forall n \in N \\ & \sum\limits_{n\in N} P^n \leq \alpha \sum\limits_{n\in N} P^{n,\text{tot}}, \end{array} \tag{6.4}$$

where α is a chosen constant smaller than 1, and denotes the required power reduction with respect to full power usage.

Another relevant formulation could be to drive the full objective towards energy minimization subject to minimum data rate constraints as follows:

$$\begin{array}{ll} \min\limits_{\{s_k^n, k\in K, n\in N\}} & \sum\limits_{n\in N} P^n \\ \text{s.t.} & P^n \leq P^{n,\text{tot}}, n \in N \\ & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, k \in K, n \in N \\ & R^n \geq R^{n,\text{target}}, \forall n \in N \end{array} \tag{6.5}$$

A general green DSL formulation can be obtained by introducing both data rates and powers into the objective as follows:

$$\begin{array}{ll} \max\limits_{\{s_k^n, k\in K, n\in N\}} & \sum\limits_{n\in N} w_n R^n - \sum\limits_{n\in N} t_n P^n \\ \text{s.t.} & P^n \leq P^{n,\text{tot}}, n \in N \\ & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, k \in K, n \in N \\ & R^n \geq R^{n,\text{target}}, \forall n \in N \\ & \sum\limits_{n\in N} P^n \leq \alpha \sum\limits_{n\in N} P^{n,\text{tot}}, \end{array} \tag{6.6}$$

Starting from general problem formulation (6.6), we can derive a suite of DSM problem formulations, including (6.4) and (6.5), by removing constraints and/or parts of the objective function. This leads to the tree structure of Figure 6.1 where the nodes correspond to DSM problem formulations and the arrows indicate a removal of a constraint or part of the objective function. More specifically, **a** and **b** indicate the presence in the objective of a weighted sum of data rates $\sum_{n\in N} w_n R^n$ and weighted sum of powers $-\sum_{n\in N} t_n P^n$, respectively. **c** and **d** indicate the presence of target data rate constraints $R^n \geq R^{n,\text{target}}$ and global total power constraints $\sum_{n\in N} P^n \leq \alpha \sum_{n\in N} P^{n,\text{tot}}$, respectively. As an example, problem formulations (6.3), (6.4), (6.5) and (6.6) correspond to nodes **a**, **acd**, **bc** and **abcd** respectively.

All the problem formulations in the tree are non-convex optimization problems because of the non-convex relation between the data rates R^n and the transmit powers s_k^n. Note that only the bold nodes correspond to DSM problem formulations that have been studied up to now. The different formulations in the tree can be interesting for different practical scenarios and this leads to an increased DSM configuration potential where it is the task of the service provider to choose the good formulation to satisfy its corresponding QoS requirements.

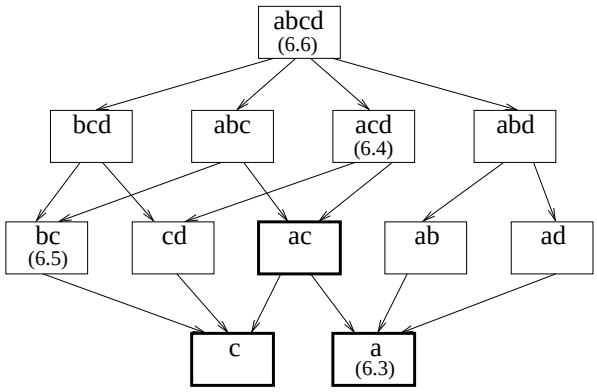


Figure 6.1: Tree structure of DSM problem formulations

6.3.2 General Systematic Solution Procedure

In this section we will propose a procedure to tackle the general non-convex optimization problem (6.6), i.e. node **abcd** in Figure 6.1, and show that the other formulations can be tackled similarly by redefining the Lagrange multipliers.

The procedure uses a *dual decomposition* approach similar to [15], where it was proved that the duality gap for this type of problem formulations can be assumed to be zero (as the number of tones is very large for typical DSL scenarios). More specifically, the procedure consists of solving two main problems, corresponding to a master problem (6.7) and a slave problem (6.8), namely:

$$\begin{array}{ll} \min\limits_{\theta} & L(\theta) \\ \text{s.t.} & \theta \geq 0 \end{array} \tag{6.7}$$

where $\theta = [\nu_1, \ldots, \nu_N, \lambda_1, \ldots, \lambda_N, \lambda_a]^T$, '$\geq$' corresponds to component-wise inequality and the dual function $L(\theta)$ corresponds to the slave optimization

problem, which can be decomposed in K independent subproblems as follows:

$$L(\theta) = \sum_k L_k(\theta)$$
$$\text{with } L_k(\theta) = \begin{cases} \max\limits_{\mathbf{s}_k} & \sum\limits_n \tilde{w}_n f_s b_k^n - \sum\limits_n \tilde{\lambda}_n s_k^n \\ \text{s.t.} & 0 \le s_k^n \le s_k^{n,\text{mask}}, \quad \forall n \in \mathcal{N}, \end{cases} \tag{6.8}$$

where $\tilde{w}_n = \omega_n + \nu_n$, $\tilde{\lambda}_n = t_n + \lambda_n + \lambda_\alpha$ and where Lagrange multipliers $\nu_n, \lambda_n, \lambda_a$ correspond to constraints $R^n \ge R^{n,\text{target}}, P^n \le P^{n,\text{tot}}$ and $\sum_n P^n \le \alpha \sum_n P^{n,\text{tot}}$ respectively. Note that (6.8) is similar to the slave problems of the pure data rate driven DSM formulations, e.g. (6.3), where the Lagrange multipliers $\tilde{w}_n, \tilde{\lambda}_n, \forall n$, are defined differently.

The solution of (6.7), i.e. the optimal Lagrange multipliers so that all constraints are satisfied, can be obtained using the following subgradient updates [15], until the corresponding KKT complementarity conditions are satisfied:

$$\nu_n = [\nu_n + \delta(R^{n,\text{target}} - R^n)]^+, \ \forall n, \tag{6.9}$$
$$\lambda_n = [\lambda_n + \epsilon(P^n - P^{n,\text{tot}})]^+, \ \forall n, \tag{6.10}$$
$$\lambda_a = [\lambda_a + \phi(\sum_n P^n - \alpha \sum P^{n,\text{tot}})]^+, \tag{6.11}$$

where δ, ϵ, ϕ are stepsizes that can chosen using different approaches [15][16] and where $[x]^+ = \max(x, 0)$.

For given Lagrange multipliers θ, the K slave subproblems (6.8) can be solved using existing DSM algorithms (IW, ASB(2), SCALE, MIW, DSB, MS-DSB, OSB) by replacing the usual Lagrange multipliers by the extended Lagrange multipliers $\tilde{w}_n, \tilde{\lambda}_n, \forall n$. Each of these algorithms has its own properties trading-off complexity and performance (sub optimal, locally optimal, globally optimal) as explained in [14].

This procedure for problem formulation **abcd** in Figure 6.1 can also be used for all the other problem formulations in the tree by redefining the extended Lagrange multipliers and/or removing some Lagrange multiplier updates. This is summarized in Table 6.1 where $\mathcal{B}$ consists of the set of Lagrange multiplier update formulas. Note that for the traditional DSM formulation (6.3), i.e. node **a**, the extended Lagrange multipliers $\tilde{w}_n, \tilde{\lambda}_n$, correspond to the weights w_n and the usual Lagrange multipliers λ_n, respectively.

Table 6.1: Definition of extended Lagrange multipliers $\tilde{w}_n, \tilde{\lambda}_n, \forall n$, for different problem statements

Problem	$\tilde{w}_n$	$\tilde{\lambda}_n$	$\mathcal{B}$
abcd	$w_n + \nu_n$	$\lambda_n + t_n + \lambda_a$	(6.9)(6.10)(6.11)
bcd	ν_n	$\lambda_n + t_n + \lambda_a$	(6.9)(6.10)(6.11)
abc	$w_n + \nu_n$	$\lambda_n + t_n$	(6.9)(6.10)
acd	$w_n + \nu_n$	$\lambda_n + \lambda_a$	(6.9)(6.10)(6.11)
abd	w_n	$\lambda_n + t_n + \lambda_a$	(6.10)(6.11)
bc	ν_n	$\lambda_n + t_n$	(6.9)(6.10)
cd	ν_n	$\lambda_n + \lambda_a$	(6.9)(6.10)(6.11)
ac	$w_n + \nu_n$	λ_n	(6.9)(6.10)
ab	w_n	$\lambda_n + t_n$	(6.10)
ad	w_n	$\lambda_n + \lambda_a$	(6.10)(6.11)
c	ν_n	λ_n	(6.9)(6.10)
a	w_n	λ_n	(6.10)

6.4 Performance Impact of Green DSL

What is the impact of the proposed energy-minimizing approaches for realistic DSL scenarios in terms of data rate performance and power saving? Is there a bound on the data rate performance loss for a given total power saving? Answering these questions will give us a clear picture of the potential of energy-minimizing DSM in the green DSL setting. For this, we will start from simulations of realistic DSL scenarios and theoretically analyze the obtained results.

The following parameter settings are assumed for the DSL scenarios. The twisted pair lines have a diameter of 0.5 mm (24 AWG). The maximum transmit power is 20.4 dBm. The SNR gap Γ is 12.9 dB, corresponding to a coding gain of 3 dB, a noise margin of 6 dB and a target symbol error probability of 10^{-7}. The tone spacing Δ_f is 4.3125 kHz. The DMT symbol rate f_s is 4 kHz. Note that we use the MS-DSB algorithm [14] with modified Lagrange multipliers as in Table 6.1 to solve the DSM problems in this section. In [14] it is shown that this algorithm is very effective in achieving globally optimal performance for most practical DSL scenarios.

In order to quantify the performance impact, we introduce the following power efficiency performance measure:

Definition 6.1 [Power usage c] The power usage c for the unifying green DSL formulation (6.6) is defined as the ratio of the sum of all allocated transmit powers and the sum of all available power budgets defined by DSL standards,

as follows

$$c(\Omega) \triangleq \frac{\sum_{n\in N} P^{n,*}(\Omega)}{\sum_{n\in N} P^{n,\text{tot}}} \leq \alpha \leq 1, \tag{6.12}$$

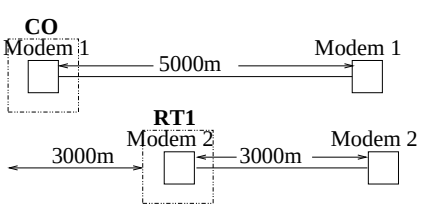
where $P^{n,*}(\Omega)$ refers to the optimal allocated power of user n for (6.6) with given parameters $\Omega = (w_1, \ldots, w_N, t_1, \ldots, t_N, R^{1,\text{target}}, \ldots, R^{N,\text{target}}, \alpha)$. □

One relevant DSL scenario is shown in Figure 6.2(a). This is a so-called near-far scenario which is known to be challenging, where DSM can make a substantial difference. Its corresponding rate region is shown in Figure 6.2(c) where the blue curve is the rate region at full power, obtained by solving the traditional DSM formulation (6.3), and the green curve is the rate region at half power, obtained by solving the proposed DSM formulation (6.4) with $\alpha = 0.5$. In Figure 6.2(d) the blue curve shows the percental data rate performance as a function of the power usage c for this near-far scenario. This curve is obtained by solving the proposed DSM formulation (6.5) for the different target rates $R^{n,\text{target}}, \forall n$, indicated in Figure 6.2(c) with red circles, corresponding to 40%, 50%, 60%, 70%, 80%, 90%, 95% and 100% of the achievable data rate performance. More formally this curve corresponds to the percental data rate performance β as a function of $c(0, 0, 1, 1, \beta R^{1,o}, \beta R^{2,o}, 1)$, where $(R^{1,o}, R^{2,o})$ is indicated in Figure 6.2(c) and β ranges from 0 to 1. One can observe that saving 50% of total power leads to only a small decrease in achievable rate region (green curve in Figure 6.2(c)). More specifically, it corresponds to 85% of data rate performance as can be seen in Figure 6.2(d) (blue curve).

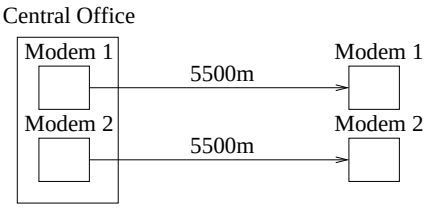
To further understand this phenomenon, the evolution of the bit loading is shown in Figure 6.2(e) for a linearly increasing power budget ranging from 20% to 100% (i.e. full power) in steps of 10%. One can observe a law of diminishing returns, i.e. a linear increase of power leads to a less than linear increase in data rate. In other words, the higher the bit rate, the less effective that power becomes.

In Figure 6.2(b) a symmetric DSL scenario is shown. Its corresponding trade-off between data rate performance and power saving is shown as the red curve in Figure 6.2(d). One can observe a data rate performance of 72% for 50% of power saving. Furthermore the evolution of the bit loading is shown in Figure 6.3(a) for a linearly increasing power budget ranging from 10% to 100% in steps of 10%. One can observe here that the effect of diminishing returns is less obvious than for Figure 6.2(e). The data rate performance for given power usage thus depends on the type of scenario.

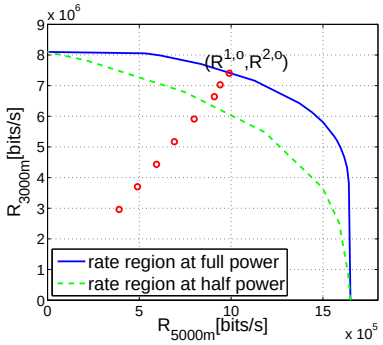
The following Theorem 6.2 provides a lower bound on the optimal data rate performance for different levels of power saving, irrespective of the scenario.



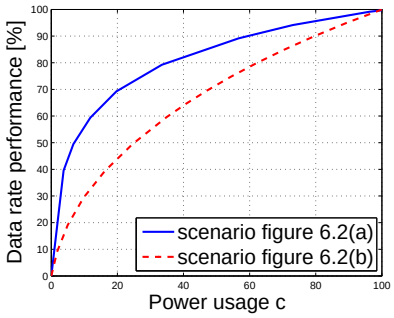
(a) Near-far ADSL scenario



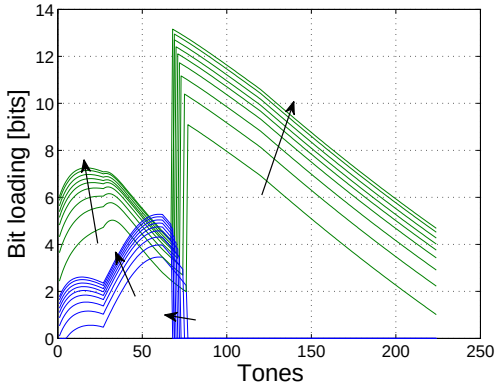
(b) Symmetric ADSL scenario



(c) Achievable rate region for scenario figure 6.2(a) at different total powers

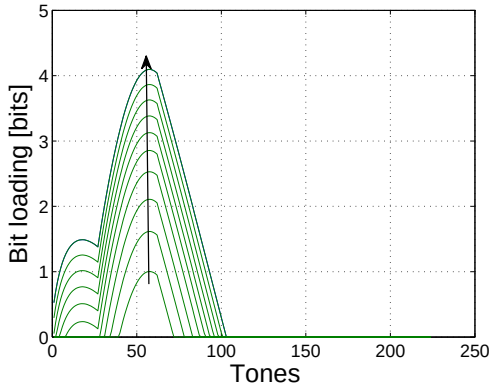


(d) Percental data rate performance versus power usage for different scenarios

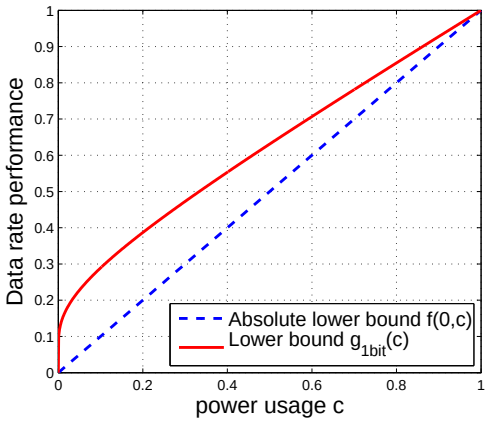


(e) Evolution of bit loading for scenario figure 6.2(a) for increasing powers

Figure 6.2: Simulation Scenarios and Results



(a) Evolution of bit loading for scenario figure 6.2(b) for increasing powers



(b) Lower bounds on percental data rate performance vs power usage

Figure 6.3: Simulation Scenarios and Results

Theorem 6.2 *In the worst case the optimal data rate performance decreases linearly as a function of a decreasing power usage c. Furthermore, for a minimum bit loading of 1 bit after power reduction, the lower bound on the optimal data rate performance $g_{1\text{bit}}$ as a function of the power usage c is given by following relation*

$$g_{1\text{bit}}(c) = \log_2\left(1 + \frac{1}{c}\right)^{-1}. \tag{6.13}$$

Proof : A first observation is that the worst case corresponds to a zero-crosstalk case. This is easy to understand if we know that the more crosstalk is present, the more power we need to increase the data rate and so the less effective that power becomes, leading to a smaller slope of the data rate performance as a function of the power usage. For this zero-crosstalk worst case scenario the percental data rate performance in function of the power usage c on a tone for a user can be expressed as

$$f(\text{SNR}, c) = \frac{\log_2(1 + SNR \times c)}{\log_2(1 + SNR)}, \tag{6.14}$$

where SNR denotes the signal-to-noise ratio. Note that although this focuses on only one tone and user, the global effect can be seen as an average of these effects on all tones and users. The function f is increasing in SNR. Using l'Hôpital's rule, one can verify the following:

$$\begin{array}{l} \lim_{\text{SNR}\to 0} f(\text{SNR}, c) = c \\ \lim_{\text{SNR}\to \infty} f(\text{SNR}, c) = 1 \end{array} \tag{6.15}$$

This illustrates two points. As the SNR becomes very large, the data rate performance loss will vanish. Secondly, the percental data rate performance is minimized when SNR is zero and this leads to a *linear* absolute lower bound on the optimal data rate performance in function of the power usage c.

However zero SNR has no practical meaning and therefore it is better to put a certain lower bound on the SNR. One practical intuitive way of lower bounding is by enforcing a minimum bit loading after power reduction of one bit so that at least one bit can be transmitted on the line. This corresponds to the constraint $\text{SNR} \times c = 1$, which leads to the following relation:

$$g_{1\text{bit}}(c) = f(\frac{1}{c}, c) = \frac{\log_2(1+1)}{\log_2(1+\frac{1}{c})} = \log_2\left(1 + \frac{1}{c}\right)^{-1}. \tag{6.16}$$

□

Note that (6.13) corresponds to a lower bound on data rate performance of 63% for 50% power saving. The lower bounds of Theorem 6.2 are summarized

in Figure 6.3(b)

This phenomenon of diminishing returns is also quite intuitive as the relation between bits and powers is a logarithm (6.2). Adding one extra bit requires at least a doubling of the signal-to-noise ratio (SNR), and leads to an increase of the bit rate by a factor $(m+1)/m$, where m is the bit rate before doubling the power. This means that the higher the bit rate, the less effective power becomes.

In Figure 6.2(e), the bitloadings range from 0 to 13 bits, with an average of 6 bits. This large average bit rate leads to poor power efficiency. This means that the last added bits require a large amount of power and so by reducing power to 50% of the full power budget, only a small decrease in data rate is incurred, i.e. 85%. In Figure 6.3(a), the bitloadings range from 0 to 4 bits, with an average of 2 bits. This small average leads to better power efficiency and so also a smaller data rate performance for 50% power usage, i.e. 72%.

Power efficiency depends on the SNR, which depends on the line attenuation, and this in turn depends on the length of the lines. So in scenarios with long line lengths, larger data rate decreases are observed when the total power is reduced. In the limit when SNR goes to zero (very large attenuation or very large noise), the logarithm behaves as a linear function and decreasing power by a factor 2 leads to a decrease in rate of factor 2. However in practice, typical scenarios have an average SNR which is much larger than zero, leading to only small data rate performance losses under large power savings. This result is quite promising for the future of green DSL.

6.5 Conclusion

Energy saving is becoming an important goal in broadband access design. Here we provide the first unifying framework for green DSL formulations and provide a systematic procedure for tackling the corresponding optimization problems. Furthermore we evaluate the impact of this energy-efficient approach for realistic DSL scenarios. We show that the worst case corresponds to a linear decrease of the optimal data rate performance for decreasing power usage. In practice, the data rate performance losses are typically much smaller, which is a promising result for this green DSL approach.

Bibliography

[1] DSL Forum (www.dslforum.org), "DSL dominates global broadband subscriber growth," Tech. Rep., Mar. 2007, internet: http://www.broadband-forum.org/news/download/pressreleeases/YE06_Release.pdf.

[2] K. B. Song, S. T. Chung, G. Ginis, and J. M. Cioffi, "Dynamic spectrum management for next-generation DSL systems," *IEEE Communications Magazine*, vol. 40, no. 10, pp. 101–109, Oct. 2002.

[3] J. M. Cioffi, EE 379C Advanced Digital Communication, Course reader: Chapter 4: "Multi-channel Modulation" [On-line] Available: http://www.stanford.edu/group/cioffi/ee379c/, Feb. 2009.

[4] ITU-T Study Group 15, "ICTs and Climate Change: ITU background report," Feb. 2008, Internet: http://www.itu.int/dms_pub/itu-t/oth/09/05/T09050000010001PDFE.pdf.

[5] E. C. D. J. Paolo Bertoldi, "European Code of Conduct for Broadband Equipment," Feb. 2008, Internet: http://www.itu.int/dms_pub/itu-t/oth/09/05/T09050000010004PDFE.pdf.

[6] Les Humphrey BT, UK,, "Opportunities and techniques for power saving in DSL," Feb. 2008, Internet: http://www.itu.int/dms_pub/itu-t/oth/09/05/T09050000010003PDFE.pdf.

[7] J. M. Cioffi, S. Jagannathan, W. Lee, H. Zou, A. Chowdhery, W. Rhee, G. Ginis, and P. Silverman, "Greener copper with dynamic spectrum management," in *AccessNets*, Las Vegas, NV, USA, Oct. 2008.

[8] T. Starr, J. M. Cioffi, P. J. Silverman, *Understanding digital subscriber lines*. Prentice Hall, 1999.

[9] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 922–933, May 2006.

[10] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[11] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spectrum balancing for digital subscriber lines," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4241–4257, Aug. 2007.

[12] J. Papandriopoulos and J. S. Evans, "Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 7, Jun. 2006, pp. 3270–3275.

[13] W. Yu, "Multiuser water-filling in the presence of crosstalk," in *Proc. of the Information Theory and Applications (ITA)*, Feb. 2007.

[14] P. Tsiaflakis, M. Diehl, and M. Moonen, "Distributed spectrum management algorithms for multiuser DSL networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4825–4843, Oct. 2008.

[15] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1310–1322, Jul 2006.

[16] P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "A low complexity optimal spectrum balancing algorithm for digital subscriber lines," *Signal Processing*, vol. 87, no. 7, pp. 1735–1753, Jul. 2007.

Chapter 7

General Green DSL: Fair and Energy-Aware Spectrum Coordination

Fair Greening for DSL Broadband Access

P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen

Published in *Proc. of the GreenMetrics Workshop in conjunction with ACM Sigmetrics/Performance (GreenMetrics)*, Seattle, Washington, Jun. 2009, pp. 1-5.

ACM COPYRIGHT NOTICE. Copyright ©2009 by the Association for Computing Machinery, Inc. Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Publications Dept., ACM, Inc., fax +1 (212) 869-0481, or permissions@acm.org.

Contributions of first author

- literature study
- co-development of fair greening concept
- co-development of four fair greening problem formulations
- co-development of unified Green DSL problem formulation
- co-development of unified solution algorithm: General-Green-DSL
- design of simulation scenarios
- software implementation and computer simulations
- co-development of greening fairness index
- co-analysis and co-interpretation of simulation results
- co-formulation of conclusion
- text co-redaction, co-editing and co-revision

Abstract

Given that broadband access networks are an integral part of the ICT infrastructure and that DSL is the most widely deployed broadband access technology, greening DSL has become important. Our recent work demonstrated a promising tradeoff between data rate performance and energy conservation. However, more greening still implies possibly lower data rate, and allocating this "price of greening" across interfering users needs to be fair. This paper proposes four formulations of fair greening in interference-limited networks, unifies them into one general representation, and develops a unified algorithm to solve them effectively. Simulations quantify the intuitions on fairness in greening DSL, as these four alternative approaches offer a range of choices between maintaining a high sum data rate and enforcing various definitions of fairness. Fairness of allocating the price of greening is also interesting in its own right.

7.1 Introduction

Greening the broadband access network is an essential part of the emerging trend towards Green Information and Communication Technology (Green ICT). The total number of broadband access lines worldwide is expected to grow from 393 mln in 2008 to 635 mln by the end of 2013. One of the most widely deployed broadband access technologies is digital subscriber line (DSL) that has a market share of over 65% and more than 200 million subscribers worldwide. DSL refers to a family of technologies that are capable of delivering broadband data rates over the copper wires of local telephone networks. One of the major obstacles to further performance improvement remains to be the electromagnetic interference, referred to as crosstalk, generated among different lines operating in the same cable bundle. This crosstalk is typically 10-20 dB larger than the background noise.

Dynamic spectrum management (DSM) has been developed as a key solution for tackling the crosstalk problem. The main idea of DSM is to prevent and/or remove crosstalk by spectrum and/or signal coordination of all users. In this paper we focus on spectrum coordination, also referred to as multi-carrier power control or spectrum balancing. The concept here is to jointly optimize the transmit power spectra of multiple interfering users so as to prevent the destructive impact caused by crosstalk interference. The last five years have seen a variety of powerful DSM algorithms to significantly increase the data rates of DSL networks. These algorithms range from fully autonomous [1, 2], to distributed [3, 4] and fully centralized DSM algorithms [5], where the main objective was all on data rate maximization without considering energy consumption.

As part of the push towards Green ICT (by regulatory organizations, e.g., ITU-T Study Group 15 and European Code of Conduct for broadband equipment [6]), reducing energy consumption of broadband equipment has now become an important consideration. In particular, the DSL Forum encourages international standards bodies to develop techniques for power reduction while preserving performance.

Energy-aware DSM is one of the most effective and readily deployable ways to enable *Green DSL.* It offers a general framework to manage transmit spectra so as to find a good trade-off between data rate performance and transmit energy consumption. In our recent work [7], it is shown that large power savings can be obtained with only minor degradations in data rate performance: a typical DSL deployment can have an $85-50$ tradeoff with 85% of the data rate maintained by spending only 50% of energy. In related work such as [8], the special case of total energy minimizing DSM is also studied.

However, what is left under-studied in green broadband access is the issue of *fairness*. Greening, either by imposing energy consumption constraints or by adding energy-aware terms in the objective function, will reduce performance metrics. This is the "price of greening". In an interference-limited physical layer such as DSL, how should this price of greening be allocated across the interfering users so that it satisfies some notion of fairness? This is the subject of the current paper.

Imposing fairness can change the shape of the tradeoff curve between performance (sum data rate) and greening (energy conservation). Furthermore, there is an interesting tradeoff between how fair a greening approach is and how much energy it can save. Among the four ways of fair greening proposed, one (method 3) can maintain the highest data rate and be greened systematically with fairness installed, while another (method 4) can provide completely proportional fair greening. These discussions on fairness of allocating the price of greening among competing users may be of interest to other areas of Green ICT.

After reviewing the DSL physical layer interference model in Section 7.2, and our recently developed Green DSL framework in Section 7.3, we develop four notions of fairness in allocating the price of greening, resulting in four members of a new family of optimization problems in Section 7.4. We then show how a unified problem statement and solution algorithm can tackle this family of problems by leveraging and extending state-of-the-art DSM algorithms. This algorithm is tested using realistic DSL channel simulators, with key results summarized in Section 7.5.

7.2 DSL System Model

We provide a brief introduction to DSL and refer the readers to [9] for details. Consider a DSL network with a set of $\mathcal{N} = \{1, \ldots, N\}$ transmitting modems (i.e., users, lines) and $\mathcal{K} = \{1, \ldots, K\}$ tones (i.e., frequency carriers). We make the standard assumption of perfect synchronization and discrete multi-tone modulation. Modems employ single-user encoding and decoding (treating interference as noise), resulting in the following expression for the achievable bit rate of modem n on tone k:

$$b_k^n(\mathbf{s}_k) \triangleq \log_2\left(1 + \frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^n}{\sum\limits_{m\neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n}\right)\text{bits/Hz}, \tag{7.1}$$

where s_k^n denotes the transmit power of modem n on tone k, $\mathbf{s}_k = [s_k^1, \ldots, s_k^N]^T$ denotes the transmit power of all modems n on tone k, $[\mathbf{H}_k]_{n,m} = h_k^{n,m}$ is an $N \times N$ matrix containing the channel transfer functions from transmitter m to receiver n on tone k. The diagonal elements are the direct channels, the off-diagonal elements are the crosstalk channels. σ_k^n denotes the noise power on tone k in receiver n that contains thermal noise, alien crosstalk and radio frequency interference. Γ denotes the gap to capacity, which is a function of the desired bit error ratio (BER), the coding gain and noise margin [9].

The data rate of modem n, denoted by R^n, is the total sum of the data rate over tones considering the symbol rate, i.e., $R^n = f_s \sum_{k\in\mathcal{K}} b_k^n(\mathbf{s}_k)$, where f_s denotes the symbol rate. Let $\mathbf{R} = [R^1, \ldots, R^N]^T$. Denote by $P^n = \sum_{k\in\mathcal{K}} s_k^n$ the total transmit power of modem n. DSL standards impose the constraints on the transmit powers in terms of the total powers (7.2) as well as spectral masks (7.3), given by:

$$\mathbf{P} \preceq \mathbf{P}^{\text{tot}} \tag{7.2}$$

$$\mathbf{s}_k \in \mathcal{S}_k = \{\mathbf{s}_k \in^N : 0 \leq s_k^n \leq s_k^{n,\text{mask}}, n \in \mathcal{N}\}, \tag{7.3}$$

where '$\preceq$' denotes the element-wise inequality. The $\mathbf{P} = [P^1, \ldots, P^N]^T$ and $\mathbf{P}^{\text{tot}} = [P^{1,\text{tot}}, \ldots, P^{N,\text{tot}}]^T$ with $P^{n,\text{tot}}$ being the total power budget in modem n. The $s_k^{n,\text{mask}}$ corresponds to the spectral mask constraint for user n on tone k. Then, the set of all achievable combinations of data rates can be modelled by the achievable rate region $\mathcal{R}$:

$$\mathcal{R} = \Big\{(R^n : n \in \mathcal{N}) | R^n = f_s \sum_{k\in\mathcal{K}} b_k^n(\mathbf{s}_k), \text{s.t. (7.2) and (7.3)}\Big\}.$$

7.3 Greening DSL

In our preliminary study [7], we proposed a unifying green DSL framework, formulated by the following problem:

$$\begin{aligned} \max_{\{s_k^n, k\in\mathcal{K}, n\in\mathcal{N}\}} \quad & \sum_{n\in\mathcal{N}} w_n R^n - \sum_{n\in\mathcal{N}} t_n P^n \\ \text{s.t.} \quad & P^n \le P^{n,\text{tot}} && n\in\mathcal{N}, \\ & 0 \le s_k^n \le s_k^{n,\text{mask}} && k\in\mathcal{K}, n\in\mathcal{N}, \\ & R^n \ge R^{n,\text{target}}, && n\in\mathcal{N}, \\ & \sum_{n\in\mathcal{N}} P^n \le \alpha \sum_{n\in\mathcal{N}} P^{n,\text{tot}}. \end{aligned} \tag{7.4}$$

The above problem is a general physical layer resource allocation formulation that can model different operating points jointly in terms of data rates ($R^n : n\in\mathcal{N}$) and total transmit powers ($P^n : n\in\mathcal{N}$).

Choosing the proper values of $\{w_n, t_n, R^{n,\text{target}}, \alpha\}$ enables network operators to steer the efficient configuration between data rates and energy consumption. In particular, there are two alternative ways to green DSM to different degrees: (1) change the ratio between rate weight w and power weight t, and (2) change the energy consumption constraint α. Based on the generalized green DSL framework in (7.4), we showed that significant power savings (e.g., 50%) are feasible only with minor degradation in data rate (e.g., 15%) [7]. This desirable tradeoff is in part due to the details in the interference structure of DSL and the dependence of rate on the Signal-to-Interference-Ratio.

7.4 Fairness in Greening DSL

However, what remains under-studied is the issue of fairness. The interfering users of a DSL network emit different amount of interference to other users, causing different degrees of difficulty in achieving a desirable "green vs. fast" tradeoff. How to allocate the price of greening, in terms of rate reduction, among the competing users now becomes important if green DSL is going to be standardized and deployed. In this and the next section, we study four different notions of fairness, which lead to four optimization problems that present different energy-performance tradeoff curves and yet can be solved in one unified way.

7.4.1 Fair Greening Formulations

There is no universally agreed notion of fairness, especially when applied to an emerging topic like Green ICT. Based on a range of reasonable views on what

Table 7.1: Reduction of the unified formulation (7.9) into the four special cases

	Energy-fair	Rate-fair	Sum-rate-fair	Energy/Rate proportional
U	$\boldsymbol{w}^T\hat{\mathbf{R}}$	$-\mathbf{1}^T\hat{\mathbf{P}}$	$\boldsymbol{w}^T\hat{\mathbf{R}} + \delta\sum_n G_n(\hat{P}^n)$	$\boldsymbol{w}^T\hat{\mathbf{R}}$
$\boldsymbol{I}$	$\hat{\mathbf{P}} - \beta\mathbf{P}^{\text{tot}}$	$\begin{bmatrix}\hat{\mathbf{P}} - \mathbf{P}^{\text{tot}} \\ \beta\mathbf{R}^o - \hat{\mathbf{R}}\end{bmatrix}$	$\begin{bmatrix}\hat{\mathbf{P}} - \mathbf{P}^{\text{tot}} \\ \mathbf{1}^T\hat{\mathbf{P}} - \beta\mathbf{1}^T\mathbf{P}^{\text{tot}}\end{bmatrix}$	$\hat{\mathbf{P}} - \beta\mathbf{P}^{\text{tot}}$
$\boldsymbol{E}$	/	/	/	$\hat{\mathbf{R}} - \beta\mathbf{R}^o(\hat{\mathbf{P}}/\mathbf{P}^o)$
$\mathbf{f}$	$\boldsymbol{w}$	$[\lambda_{N+1}, \ldots, \lambda_{2N}]^T$	$\boldsymbol{w}$	$\boldsymbol{w} + \boldsymbol{\mu}$
$\mathbf{h}$	$\boldsymbol{\lambda}$	$[\lambda_1, \ldots, \lambda_N]^T + 1$	$\boldsymbol{t} + \epsilon(\mathbf{P} - \hat{\mathbf{P}})$	$\boldsymbol{\lambda} + \boldsymbol{\mu}\beta\mathbf{R}^o/\mathbf{P}^o$
$\hat{\mathbf{R}}^n$	/	/	/	/
$\hat{\mathbf{P}}^n$	/	/	$G^{-1}_{n,\hat{P}^n}((\lambda_n + \lambda_{N+1} - t_n)/\delta)$	/

fairness is, we develop four formulations of Green DSL in this section, each parameterized by a "degree of greening" parameter $\beta \in [0,1]$. Throughout this subsection, for a given greening formulation, we denote by $\{\boldsymbol{R}^o, \boldsymbol{P}^o\}$ the rate-power point somewhere on the border of the rate region $\mathcal{R}$ when $\beta = 1$, i.e., the point when no greening is applied.

(1) Energy-fair greening
In this approach, we target "perfect power fairness" by having weighted rate maximization as the objective, and proportionally reducing the available transmit power per modem:

$$\textbf{Greening 1:} \quad \max_{\{\mathbf{s}_k \in \mathcal{S}_k, k \in \mathcal{K}\}} \quad \sum_{n \in \mathcal{N}} w_n R^n \qquad \text{subject to} \quad P^n \leq \beta P^{n,\text{tot}},\ n \in \mathcal{N}. \tag{7.5}$$

(2) Rate-fair greening
In contrast to the previous approach, this approach targets "perfect rate fairness" by having power minimization as the objective and by proportionally reducing the minimum target data rate of users:

$$\textbf{Greening 2:} \quad \max_{\{\mathbf{s}_k \in \mathcal{S}_k, k \in \mathcal{K}\}} \quad -\sum_{n \in \mathcal{N}} P^n \qquad \text{subject to} \quad \begin{aligned} &P^n \leq P^{n,\text{tot}}, \quad n \in \mathcal{N}, \\ &R^n \geq \beta R^{n,o}, \quad n \in \mathcal{N}. \end{aligned} \tag{7.6}$$

(3) Sum-rate greening with fair regularization
Yet another alternative formulation is to start from the weighted sum rate formulation with a system total power constraint parameterized by β. However, this formulation does not impose any fairness consideration and indeed can return solutions that favor some users substantially. Therefore we augment the objective function with a fairness term G:

$$\textbf{Greening 3:} \quad \max_{\{\mathbf{s}_k \in \mathcal{S}_k, k \in \mathcal{K}\}} \quad \sum_{n \in \mathcal{N}} w_n R^n + \delta \sum_{n \in \mathcal{N}} G_n(P^n) \qquad \text{subject to} \quad \begin{aligned} &P^n \leq P^{n,\text{tot}}, \quad n \in \mathcal{N} \\ &\sum_{n \in \mathcal{N}} P^n \leq \beta \sum_{n \in \mathcal{N}} P^{n,\text{tot}}, \end{aligned} \tag{7.7}$$

where δ is a weighting factor tuned to emphasize the importance of greening fairness relative to the maximal weighted data rate performance. The function $G_n(P^n)$ may take different forms depending on the desired energy fairness. One possibility is α-fairness: $G_n(\cdot) = (\cdot)^{1-\alpha}/(1-\alpha)$, for $\alpha > 0$, and $\log(\cdot)$, for $\alpha = 1$, which includes max-min ($\alpha \to \infty$) and proportional fairness ($\alpha = 1$) as special cases. Another possibility is to use second-moment as a measure of fairness: $G_n(.) = -(.)^2$. Therefore, we have the following three special cases:

- **Greening 3A:** (7.7) with $\delta = 0$;
- **Greening 3B:** (7.7) with $\delta > 0$ and $G_n(.) = -(.)^2$;
- **Greening 3C:** (7.7) with $\delta > 0$ and $G_n(.) = \log(.)$.

(4) Energy/Rate proportional greening
In the fourth approach, the ratios of their data rate decrease and their total transmit power decrease are kept the same across the users: those that transmit at higher rate takes a proportionally larger share of the price of greening:

$$\textbf{Greening 4:} \quad \begin{array}{ll} \max\limits_{\{\mathbf{s}_k \in \mathcal{S}_k, k \in \mathcal{K}\}} & \sum\limits_{n \in \mathcal{N}} w_n R^n \\ \text{subject to} & P^n \leq P^{n,\text{tot}}, \quad n \in \mathcal{N}, \\ & \dfrac{R^n}{R^{n,o}} / \dfrac{P^n}{P^{n,o}} = \beta, n \in \mathcal{N}. \end{array} \tag{7.8}$$

This formulation requires no weight to be tuned as in (7.7). Note that $\beta = 1$ corresponds to no greening and increasing β corresponds to greening. One can easily extend (7.8) with an extra bisection search to find the value of β that corresponds to a particular system power usage. This is because system power usage is monotone decreasing with increasing β.

7.4.2 Unified Formulation and Solution Algorithm

Mathematically, the four fair greening methods in (7.5), (7.6), (7.7) and (7.8) can be shown as special cases of the generalized fair greening formulation:

$$\begin{array}{ll} \max\limits_{\hat{\mathbf{R}}, \hat{\mathbf{P}}, \{\mathbf{s}_k \in \mathcal{S}_k, k \in \mathcal{K}\}} & U(\hat{\mathbf{R}}, \hat{\mathbf{P}}) \\ \text{subject to} & \boldsymbol{I}(\hat{\mathbf{R}}, \hat{\mathbf{P}}) \leq 0, \quad \boldsymbol{E}(\hat{\mathbf{R}}, \hat{\mathbf{P}}) = 0, \\ & \hat{\mathbf{R}} = \mathbf{R}, \quad \mathbf{P} = \hat{\mathbf{P}}, \end{array} \tag{7.9}$$

where the specific forms for $U(\hat{\mathbf{R}}, \hat{\mathbf{P}})$, $I(\hat{\mathbf{R}}, \hat{\mathbf{P}})$ and $E(\hat{\mathbf{R}}, \hat{\mathbf{P}})$ corresponding to the four formulations, are given in Table 7.1.

Solving these problems is challenging since they are NP-hard non-convex problems. To tackle computational difficulty, we use Lagrange relaxation. The power of this approach lies in the fact that (asymptotic) strong duality holds despite the nonconvexity of the optimization problem [10] and that the overall problem can be easily decomposed into the many much simpler independent subproblems. We present the following three subproblems to be solved, based on the Lagrange relaxation of the original problem (7.9), that are reduced in dimension and are thus much simpler to solve. More details of this decomposition methodology can be found, e.g., in [11, 1].

P1. Find the optimal Lagrange multipliers $\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \boldsymbol{t}$, that enforce the constraints. This can be solved by standard subgradient based updates of the Lagrange multipliers where the subgradients are $\boldsymbol{I}(\hat{\mathbf{R}}, \hat{\mathbf{P}})$ and $\boldsymbol{E}(\hat{\mathbf{R}}, \hat{\mathbf{P}})$.

P2. Solve the convex unconstrained problem in $\hat{\mathbf{R}}$ and $\hat{\mathbf{P}}$:

$$\max_{\hat{\mathbf{R}},\hat{\mathbf{P}}} \; U(\hat{\mathbf{R}},\hat{\mathbf{P}}) \; - \; \boldsymbol{\lambda}^T \boldsymbol{I}(\hat{\mathbf{R}},\hat{\mathbf{P}}) \; - \; \boldsymbol{\mu}^T \boldsymbol{E}(\hat{\mathbf{R}},\hat{\mathbf{P}}) \; - \; \boldsymbol{\nu}^T \hat{\mathbf{R}} \; + \; \boldsymbol{t}^T \hat{\mathbf{P}} \quad (7.10)$$

Using first order necessary conditions, $\hat{\mathbf{R}}$ and $\hat{\mathbf{P}}$ are given by solving the following system:

$$U_{\hat{R}^n}(\hat{\mathbf{R}},\hat{\mathbf{P}}) - \boldsymbol{\lambda}^T \boldsymbol{I}_{\hat{R}^n}(\hat{\mathbf{R}},\hat{\mathbf{P}}) - \boldsymbol{\mu}^T \boldsymbol{E}_{\hat{R}^n}(\hat{\mathbf{R}},\hat{\mathbf{P}}) - \nu_n = 0, \quad n \in \mathcal{N} \quad (7.11)$$

$$U_{\hat{P}^n}(\hat{\mathbf{R}},\hat{\mathbf{P}}) - \boldsymbol{\lambda}^T \boldsymbol{I}_{\hat{P}^n}(\hat{\mathbf{R}},\hat{\mathbf{P}}) - \boldsymbol{\mu}^T \boldsymbol{E}_{\hat{P}^n}(\hat{\mathbf{R}},\hat{\mathbf{P}}) + t_n = 0, \quad n \in \mathcal{N} \quad (7.12)$$

where $F_x(.)$ refers to the derivative of a function $F(.)$ w.r.t. the variable x, e.g., $U_{\hat{P}^n}(\hat{\mathbf{R}},\hat{\mathbf{P}}) = \partial U(\hat{\mathbf{R}},\hat{\mathbf{P}})/\partial \hat{P}^n$.

P3. Solve the per-tone problems for each tone k:

$$\max_{\{\mathbf{s}_k \in \mathcal{S}_k\}} \sum_{n \in \mathcal{N}} \nu_n b_k^n(\mathbf{s}_k) - \sum_{n \in \mathcal{N}} t_n s_k^n \quad (7.13)$$

This non-convex problem in dimension N can be solved using existing per-tone solutions, e.g., [5, 12, 4, 1, 3].

From the three decomposed subproblems above, we can readily develop the procedure for solving the unified green DSL formulation (7.9), as is described in the algorithm **General-Green-DSL**. The functions $\boldsymbol{f}$ and $\boldsymbol{h}$ can be standard subgradient updates, i.e., $\boldsymbol{\nu} = \boldsymbol{\nu} + \epsilon(\hat{\mathbf{R}} - \mathbf{R})$ and $\boldsymbol{t} = \boldsymbol{t} + \epsilon(\mathbf{P} - \hat{\mathbf{P}})$, where ϵ refers to the step size for the subgradient updates. We prove that the algorithm converges to a point which is either a locally or globally optimum, depending on the particular DSM algorithm used in **P3**. We skip its convergence proof due to space limitation.

General-Green-DSL

1: Initialize $\hat{\mathbf{R}}, \hat{\mathbf{P}}, \{\mathbf{s}_k \in \mathcal{S}_k, k \in \mathcal{K}\}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{t}, \boldsymbol{\nu}$
2: **repeat**
3: $\quad \boldsymbol{\lambda} = [\boldsymbol{\lambda} + \epsilon(\boldsymbol{I}(\hat{\mathbf{R}},\hat{\mathbf{P}}))]^+$
4: $\quad \boldsymbol{\mu} = \boldsymbol{\mu} + \epsilon(\boldsymbol{E}(\hat{\mathbf{R}},\hat{\mathbf{P}}))$
5: $\quad \boldsymbol{\nu} = \boldsymbol{f}(\hat{\mathbf{R}},\hat{\mathbf{P}}, \{\mathbf{s}_k \in \mathcal{S}_k, k \in \mathcal{K}\}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{t}, \boldsymbol{\nu})$
6: $\quad \boldsymbol{t} = \boldsymbol{h}(\hat{\mathbf{R}},\hat{\mathbf{P}}, \{\mathbf{s}_k \in \mathcal{S}_k, k \in \mathcal{K}\}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{t}, \boldsymbol{\nu})$
7: $\quad$ Obtain $\hat{\mathbf{R}}$ and $\hat{\mathbf{P}}$ by solving system (7.11) and (7.12)
8: $\quad \forall k \in \mathcal{K}$: Solve per-tone problem (7.13)
9: **until** convergence

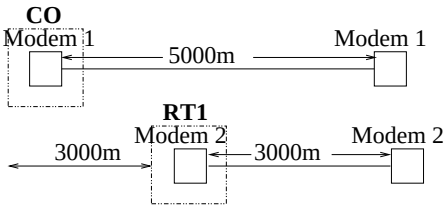
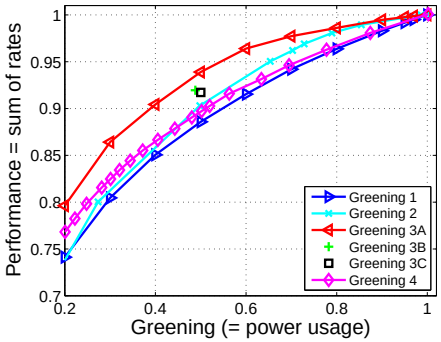


Figure 7.1: 2-user near-far ADSL scenario

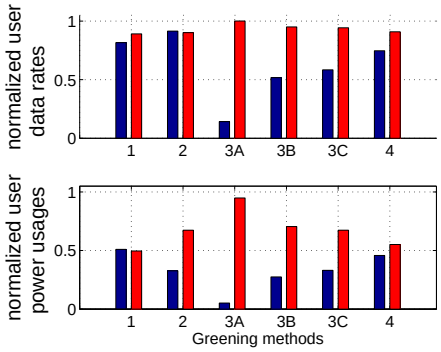
7.5 Simulation Results

We consider a practically highly relevant ADSL scenario as shown in Figure 7.1, where for illustrative purpose we present the 2-user dynamics. This is a so-called near-far scenario, which is known to be a challenging scenario, where DSM can make a substantial difference. More extensive scenario simulations will be presented in the journal version. The following realistic parameters settings are assumed in the channel simulator. The twisted pair lines have a diameter of 0.5 mm (24 AWG). The maximum modems' total transmit power $P^{n,\text{tot}}$ is 20.4 dBm. The SNR gap Γ is 12.9 dB. The tone spacing Δ_f is 4.3125 kHz. The DMT symbol rate f_s is 4 kHz. For solving the per-tone problem (7.13), we used the MS-DSB algorithm that is shown to converge to the globally optimal solution for this scenario [4].

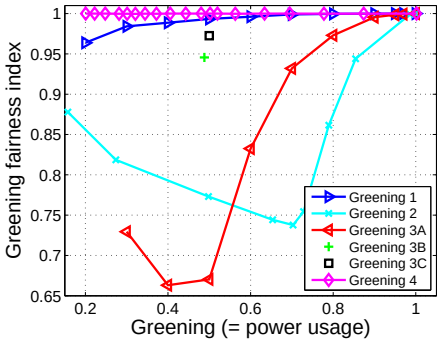
Figure 7.2(a) shows the trade-off between the *normalized* sum data rate performance (w.r.t the maximum unweighted sum rate $\sum_n R^{n,o}$ for full power usage), and the *normalized* greening measured by the actual consumed power divided by the total maximum available power (i.e., $\sum_n P^n / \sum_n P^{n,\text{tot}}$), for the six different greening methods in Section 7.4.1. We generally observe logarithmic curves, implying that significant power savings can be achieved with only small degradations in data rate performance. Different greening methods pay a different price of greening (i.e., loss in sum rate data rate performance due to greening). Greening 3A is the best in terms of sum rate performance, i.e., 50% of system power is saved while still achieving 93% of data rate performance. However, Greening 3A is intuitively the most unfair, since it contains no mechanisms to steer towards fair transmit power and/or data rate allocations. Other greening methods have worse sum rate performance, where in particular that of Greening 1 (energy-fair) gives the smallest sum rate. Note that for Greenings 3B and 3C, only one simulation point is shown rather than a parametric curve, due to their dependency on the chosen weight δ in (7.7). Different values for this weight lead to different trade-offs, and it is not straightforward to tune this weight δ in order to arrive at a specific tradeoff point. This extra tuning for a good trade-off can be regarded as a pitfall of Greenings 3B and 3C from the perspective of green system design.



(a) *Normalized* sum data rate performance vs. greening for different greening methods



(b) *Normalized* data rates (top) and *normalized* transmit powers (bottom) for two modems (blue: modem 1, red: modem 2) for different greening methods when 50% greening is applied



(c) Greening fairness index $\mathcal{F}$ (7.14) vs. greening for different greening methods

Figure 7.2: Simulation Results

Figure 7.2(b) shows the distribution of the *normalized* data rates (w.r.t. the full power data rate $R^{n,o}$ for each user n) and *normalized* transmit powers (w.r.t. the full power $P^{n,\text{tot}}$ budget of each user n) across the two users for different greening methods when 50% greening is applied. We observe that Greening 3A allocates the data rates and the transmit powers very unfairly over the users, i.e., user 2 dominates over user 1 in terms of transmit power and data rate. By definition, Greening 1 equalizes the normalized energy as it proportionally allocates its transmit powers over the users. Greening 2 equalizes the normalized data rate, but has uneven allocation of transmit powers. Greenings 3B and 3C succeed in obtaining relatively better fairness in terms of transmit powers as well as data rates than Greening 3A. This is due to the addition of the fairness term into their objective functions. However, the price of this is a reduced sum data rate performance as seen in Figure 7.2(a), compared to Greening 3A. Finally, Greening 4 reduces its data rates proportionally to its transmit powers for all modems as it was designed for.

To further quantitatively evaluate the fairness of greening methods, a definition of fairness measure is needed. A good candidate is a measure that jointly considers energy and rate fairness in its definition. To that end, we propose the following definition of *greening fairness index* $\mathcal{F}$:

$$\mathcal{F} = \frac{1}{N-1}\left(\frac{(\sum_{n\in\mathcal{N}} x_n)^2}{\sum_{n\in\mathcal{N}} x_n^2} - 1\right), \tag{7.14}$$

with $x_n = (R^n/R^{n,o})/(P^n/P^{n,o})$ and $\{\mathbf{R}^o, \mathbf{P}^o\}$ denotes the point on the boundary of the rate region without greening. $\mathcal{F} = 1$ when all users have the same ratio between data rate decrease and power decrease, and approaches 0 as these ratios start to deviate from each other. Figure 7.2(c) shows the trade-off between the greening fairness index $\mathcal{F}$ and power reduction due to greening for the proposed six methods. The key messages are as follows:

- Greening 4 is 100% fair w.r.t. $\mathcal{F}$, since it was constructed by considering both energy and data rate in the first place.
- Energy-fair Greening 1 is also quite fair, whereas rate-fair Greening 2 is unfair.
- Sum rate greening 3A is very unfair. However, by adding the fairness terms to Greening 3A, i.e. Greening 3B and 3C, the fairness behavior, w.r.t. $\mathcal{F}$, becomes much better, i.e. from 67% to more than 94%.

In summary, Green DSL can generally lead to large energy savings with only minor data rate performance degradation. However one should carefully choose its greening strategy that significantly affects fairness in the data rates and the amount of energy savings. Some greening strategies can lead to very unfair allocations in terms of data rates and/or transmit powers, which can even put some users on weak DSL lines out-of-service. The application of fair greening technology can prevent that, but the price to pay is possibly a further loss in

sum rate performance.

7.6 Conclusion

In order for green DSL to be deployed, Internet Service Providers must ensure certain notions of fairness across the interfering users, when allocating the tradeoff between high data rate and low energy consumption. There is no single universally agreed notion of fairness in greening, and this paper explores four alternative approaches. Their unified representation can be solved effectively by the General-Green-DSL algorithm. They present a range of choices in the three way tradeoff among sum data rate, total energy consumption, and the fairness index. In particular, Method 4 guarantees a proportional relationship between performance and greening across the users, while maintaining a reasonable data rate.

Bibliography

[1] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spectrum balancing for digital subscriber lines," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4241–4257, Aug. 2007.

[2] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[3] J. Papandriopoulos and J. S. Evans, "Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 7, Jun. 2006, pp. 3270–3275.

[4] P. Tsiaflakis, M. Diehl, and M. Moonen, "Distributed spectrum management algorithms for multiuser DSL networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4825–4843, Oct. 2008.

[5] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 922–933, May 2006.

[6] E. C. D. J. Paolo Bertoldi, "European Code of Conduct for Broadband Equipment," Feb. 2008, Internet: http://www.itu.int/dms_pub/itu-t/oth/09/05/T09050000010004PDFE.pdf.

[7] P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen, "Green DSL: Energy-efficient DSM," in *Proc. of the IEEE International Conference on Communications (ICC)*, Jun. 2009.

[8] M. Wolkerstorfer, D. Statovci, and T. Nordstrom, "Dynamic spectrum management for energy-efficient transmission in DSL," in *Proc. of the 11th IEEE Singapore International Conference on Communication Systems (ICCS)*, Nov. 2008, pp. 1015–1020.

[9] T. Starr, J. M. Cioffi, P. J. Silverman, *Understanding digital subscriber lines*. Prentice Hall, 1999.

[10] Z. Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, Feb. 2008.

[11] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1310–1322, Jul 2006.

[12] R. Lui and W. Yu, "Low-complexity near-optimal spectrum balancing for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 3, May 2005, pp. 1947–1951.

Chapter 8

Improved Dual Decomposition Based Spectrum Coordination

Improved Dual Decomposition Based Optimization for DSL Dynamic Spectrum Management

P. Tsiaflakis, I. Necoara, J.A.K. Suykens, and M. Moonen

Accepted with mandatory minor revisions for *IEEE Transactions on Signal Processing*, September, 2009

This work has been submitted to the IEEE for publication and it has been accepted with mandatory minor revisions for *IEEE Transactions on Signal Processing*. Copyright may be transferred without notice. The final version may have some minor modifications with respect to the current version. A portion of this paper has appeared in the Proceedings of the 17th European Signal Processing Conference (EUSIPCO), August 2009 [1].

Contributions of first author

- literature study
- co-development of improved dual decomposition approach for iterative convex approximation based spectrum coordination algorithms
- co-development of improved dual decomposition approach for direct spectrum coordination algorithms
- co-development of interleaving procedure for recovering the primal solution from the dual solution
- design of simulation scenarios
- software implementation and computer simulations
- co-analysis and co-interpretation of simulation results
- text co-redaction, co-editing, co-revision

Abstract

Dynamic spectrum management (DSM) has been recognized as a key technology to significantly improve the performance of digital subscriber line (DSL) broadband access networks. The basic concept of DSM is to coordinate transmission over multiple DSL lines so as to mitigate the impact of crosstalk interference amongst them. Many algorithms have been proposed to tackle the nonconvex optimization problems appearing in DSM, many of them relying on a standard subgradient based dual decomposition approach. In practice however, this approach is often found to lead to extremely slow convergence or even no convergence at all, one of the reasons being the very difficult tuning of the stepsize parameters. In this paper we propose a novel improved dual decomposition approach inspired by recent advances in mathematical programming. It uses a smoothing technique for the Lagrangian combined with an optimal gradient based scheme for updating the Lagrange multipliers. The stepsize parameters are furthermore selected optimally removing the need for a tuning strategy. With this approach we show how the convergence of current state-of-the-art DSM algorithms based on iterative convex approximations (SCALE, CA-DSB) can be improved by one order of magnitude. Furthermore we apply the improved dual decomposition approach to other DSM algorithms (OSB, ISB, ASB, (MS)-DSB, MIW) and propose further improvements to obtain fast and robust DSM algorithms. Finally, we demonstrate the effectiveness of the improved dual decomposition approach for a number of realistic multi-user DSL scenarios.

8.1 Introduction

Digital subscriber line (DSL) technology refers to a family of technologies that provide digital broadband access over the local telephone network. It is currently the dominating broadband access technology with more than 66% of all broadband access subscribers worldwide using DSL to access the Internet. It is forecasted that the number of DSL subscribers will even rise to 331 million in 2012 with DSL access revenues reaching $136.4 billion in 2012 [2]. The major obstacle for further performance improvement in modern DSL networks is the so-called crosstalk, i.e. the electromagnetic interference amongst different lines in the same cable bundle. Different lines (i.e. users) indeed interfere with each other, leading to a very challenging interference environment where proper management of the resources is required to prevent a huge performance degradation.

Dynamic spectrum management (DSM) has been recognized as a key technology to significantly improve the performance of DSL broadband access networks [3]. The basic concept of DSM is to coordinate transmission over multiple DSL

lines so as to mitigate the impact of crosstalk interference amongst them. There are two types of coordination referred to as spectrum level and signal level coordination. Here, we will focus on spectrum level coordination, also referred to as spectrum management, spectrum balancing or multi-carrier power control. Spectrum management aims to allocate transmit spectra, i.e. transmit powers over all available frequencies (tones), to the different users so as to achieve some design objective. This generally corresponds to an optimization problem, where typically a weighted sum of user data rates is maximized subject to power constraints [4, 5, 6], which will be referred to as "constrained weighted rate sum maximization (cWRS)". Recently this has been extended to other design objectives as well, such as power driven designs (green DSL[7],[8], [9]) and other utility driven designs [10, 11]. As shown in [7], the key component to these designs is an efficient solution for the cWRS problem. Therefore we will mainly focus on this problem and aim to find a robust and efficient solution for it.

The cWRS problem is known to be an NP-hard, separable nonconvex optimization problem, that can have many locally optimal solutions [10][12]. Even for moderately sized problems (with 5-20 users and 200-4000 tones), finding the globally optimal solution is computationally prohibitive. In [5] and [13] the authors proposed to use a dual decomposition approach with a standard subgradient based updating of the Lagrange multipliers. Many DSM algorithms [4, 6, 14, 13, 12, 15, 16, 17, 18] have been proposed recently, that use the standard subgradient based dual decomposition approach. In practice, however, this approach is often found to lead to extremely slow convergence or even no convergence at all, especially so for large DSL scenarios with large crosstalk. One of the reasons is the very difficult tuning of the stepsize parameters so as to guarantee fast convergence. We would like to remark here that the subgradient based dual decomposition algorithms are not the only algorithms in use for spectrum management and that a number of alternative approaches exist, e.g. the ellipsoid dual update approach proposed in [5], Spectrum Balancing Levin-Campello (SBLC) proposed in [19, 20], etc.

In this paper we propose a novel improved dual decomposition approach inspired by recent advances in mathematical programming, more specifically the proximal center based decomposition method recently proposed in [21]. This method uses a smoothing technique [22] for the Lagrangian and combines it with an accelerated scheme for smooth optimization [23]. Moreover, the stepsize parameter is determined automatically so as to obtain fast convergence, removing the need for a stepsize tuning strategy. However, the proximal center based method is proposed for separable convex problems, it is only derived for two users under equality constraints, and it is presented from a rather high-level point of view, where the concrete steps of the algorithms are general optimization problems to be solved (see Algorithm 3.2 in [21]). DSM optimization problems are, however, highly nonconvex problems with multiple user scenarios

and multiple tones under inequality constraints. In this paper we extend the proximal center based decomposition method to a concrete improved dual decomposition approach for particular application in the context of DSM. More specifically, it consists of an extension of the proximal center based method to the nonconvex cWRS problem, where it is derived for multiple users and multiple tones under inequality constraints and with concrete efficient implementations for specific types of DSM problems. With the proposed approach, we show how the convergence of current state-of-the-art DSM algorithms based on iterative convex approximations (e.g. successive convex approximation for low-complexity (SCALE) [15], convex approximation distributed spectrum balancing (CA-DSB) [12]) can be speed up by one order of magnitude, without increasing the computational complexity for each iteration. Furthermore we apply the improved dual decomposition approach to other DSM algorithms (optimal spectrum balancing (OSB) [4], modified prismatic branch-and-bound algorithm (PBnB) [14], iterative spectrum balancing (ISB) [13], autonomous spectrum balancing (ASB) [17], (multiple starting point) distributed spectrum balancing ((MS-)DSB) [12], modified iterative water-filling (MIW) [16], branch-and-bound optimal spectrum balancing (BB-OSB) [6]), again leading to much faster converging DSM algorithms. Then we demonstrate an important pitfall of applying the dual decomposition approach to nonconvex DSM problems and propose an effective solution that further improves the robustness of current DSM algorithms. Finally we demonstrate the effectiveness of the improved dual decomposition approach for a number of realistic multi-user DSL scenarios.

This paper is organized as follows. In Section 8.2, the system model is introduced for the DSL multi-user environment. In Section 8.3, the basic cWRS problem is described and existing DSM algorithms for this problem are reviewed, that rely on a subgradient based dual decomposition approach. In Section 8.4.1 an improved dual decomposition approach is proposed for DSM algorithms based on iterative convex approximations. The improved dual decomposition approach is furthermore applied to other DSM algorithms in Section 8.4.2. In Section 8.5, the problem of obtaining a primal solution from the dual solution is described and an effective solution for it is proposed. Finally in Section 8.6, simulation results are shown.

8.2 System Model

We consider a system consisting of $\mathcal{N} = \{1, \ldots, N\}$ interfering DSL users (i.e., lines, modems) with standard synchronous discrete multi-tone (DMT) modulation with $\mathcal{K} = \{1, \ldots, K\}$ tones (i.e., frequencies or carriers). The transmission can be modeled independently on each tone k by

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

The vector $\mathbf{x}_k = [x_k^1, \ldots, x_k^N]^T$ contains the transmitted signals on tone k, where x_k^n refers to the signal transmitted by user n on tone k. Vectors $\mathbf{z}_k$ and $\mathbf{y}_k$ have similar structures; $\mathbf{z}_k$ refers to the additive noise on tone k, containing thermal noise, alien crosstalk, radio frequency interference (RFI), etc, and $\mathbf{y}_k$ refers to the received signals on tone k. $\mathbf{H}_k$ is an $N \times N$ channel matrix with $[\mathbf{H}_k]_{n,m} = h_k^{n,m}$ referring to the channel gain from transmitter m to receiver n on tone k. The diagonal elements are the direct channels and the off-diagonal elements are the crosstalk channels.

The transmit power of user n on tone k, also referred to as transmit power spectral density, is denoted as $s_k^n \triangleq \Delta_f E\{|x_k^n|^2\}$, where Δ_f refers to the tone spacing. The vector $\mathbf{s}_k \triangleq \{s_k^n, n \in \mathcal{N}\}$ denotes the transmit powers of all users on tone k. The vector $\mathbf{s}^n \triangleq \{s_k^n, k \in \mathcal{K}\}$ denotes the transmit powers of user n on all tones. The received noise power by user n on tone k, also referred to as noise spectral density, is denoted as $\sigma_k^n \triangleq \Delta_f E\{|z_k^n|^2\}$.

Note that we assume no signal coordination at the transmitters and at the receivers, and that the interference is treated as additive white Gaussian noise. Under this standard assumption the bit loading for user n on tone k, given the transmit spectra $\mathbf{s}_k$ of all users on tone k, is

$$b_k^n \triangleq b_k^n(\mathbf{s}_k) \triangleq \log_2\left(1 + \frac{1}{\Gamma}\frac{|h_k^{n,n}|^2 s_k^n}{\sum_{m \neq n} |h_k^{n,m}|^2 s_k^m + \sigma_k^n}\right) \text{ bits/Hz}, \tag{8.1}$$

where Γ denotes the SNR-gap to capacity, which is a function of the desired BER, the coding gain and noise margin [24]. The DMT symbol rate is denoted as f_s. The achievable total data rate for user n and the total power used by user n are equal to, respectively:

$$R^n \triangleq f_s \sum_{k \in \mathcal{K}} b_k^n \quad \text{and} \quad P^n \triangleq \sum_{k \in \mathcal{K}} s_k^n. \tag{8.2}$$

8.3 Dynamic Spectrum Management

8.3.1 Dynamic Spectrum Management Problem

The basic goal of DSM through spectrum level coordination is to allocate the transmit powers dynamically in response to physical channel conditions (channel gains and noise) so as to pursue certain design objectives and/or satisfy certain constraints. The constraints are mostly per-user total power constraints and so-called spectral mask constraints, i.e.

$$\begin{array}{lll} P^n \leq P^{n,\text{tot}} & n \in \mathcal{N} & \text{(total power constraints)} \\ 0 \leq s_k^n \leq s_k^{n,\text{mask}} & n \in \mathcal{N}, k \in \mathcal{K} & \text{(spectral mask constraints)} \end{array} \tag{8.3}$$

where $P^{n,\text{tot}}$ refers to the total available power budget for user n and $s_k^{n,\text{mask}}$ refers to the spectral mask constraint for user n on tone k. The user total power constraints can also be written in vector notation as $\mathbf{P} \leq \mathbf{P}^{\text{tot}}$, where $\mathbf{P} = [P^1, \ldots, P^N]^T$ and $\mathbf{P}^{\text{tot}} = [P^{1,\text{tot}}, \ldots, P^{N,\text{tot}}]^T$, and where '$\leq$' denotes a component-wise inequality.

The set of all possible data rate allocations that satisfy the constraints (8.3) can be characterized by the achievable rate region $\mathcal{R}$:

$$\{R, \ldots, =\} \Big\{ (R^n : n \in \mathcal{N}) | R^n = f_s \sum_{k \in K} b_k^n(\mathbf{s}_k), \text{ s.t. } (8.3) \Big\}.$$

A typical design objective is to achieve some Pareto optimal allocation of the data rates R^n [18, 4, 13, 25, 14, 6, 17, 26, 15, 16, 12]. This results in the following typical DSM optimization problem, which will be referred as the constrained weighted rate sum maximization (cWRS) formulation, where w_n is the weight given to user n:

$$\begin{array}{lll} \max\limits_{\{\mathbf{s}^n, n \in \mathcal{N}\}} & \sum\limits_{n \in \mathcal{N}} w_n R^n & \\ \text{s.t.} & P^n \leq P^{n,\text{tot}}, & n \in \mathcal{N}, \qquad \text{(cWRS)} \\ & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, & n \in \mathcal{N}, k \in \mathcal{K}. \end{array} \tag{8.4}$$

However, many other DSM formulations are possible. We refer to [7] containing a collection of other relevant DSM formulations. As shown in [7], the key component to tackling these is an efficient solution for cWRS problem (8.4). Therefore we will focus on this problem and aim to find a robust and efficient solution for it.

8.3.2 Dynamic Spectrum Management Algorithms

cWRS problem (8.4) is an NP-hard separable nonconvex optimization problem [10]. The number of optimization variables is equal to KN, where the number of users N ranges between 2-100 and the number of tones K can go up to 4000. Depending on the specific values of the channel and noise parameters, there can be many locally optimal solutions, that can differ significantly in value, as shown in [12]. In [5] the authors show that strong duality holds for the continuous (frequency range) formulation, and in [10] the authors prove asymptotic strong duality for the discrete (frequency range) formulation, i.e. the duality gap goes to zero as $K \to \infty$. These results suggest that a Lagrange dual decomposition approach is a viable way to reach approximate optimality for the discrete formulation (8.4), if the frequency range is finely discretized, as it is indeed the case in practical DSL scenarios where K is large [5]. Many dual decomposition based DSM algorithms [6, 14, 13, 12, 15, 16, 17, 18] have been

proposed for solving (8.4), that use a standard subgradient based updating of the Lagrange multipliers.

The dual problem formulation of (8.4) consists of two subproblems, namely a master problem

$$\begin{array}{ll} \min\limits_{\boldsymbol{\lambda}} & g(\boldsymbol{\lambda}) \\ \text{s.t.} & \boldsymbol{\lambda} \geq 0 \end{array} \tag{8.5}$$

where $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_N]^T$, and a slave problem defined by the dual function $g(\boldsymbol{\lambda})$:

$$\begin{array}{l} g(\boldsymbol{\lambda}) = \left\{ \begin{array}{ll} \max\limits_{\{\mathbf{s}_k, k\in\mathcal{K}\}} & \mathcal{L}(\boldsymbol{\lambda}, \mathbf{s}_k, k \in \mathcal{K}) \\ \text{s.t.} & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \qquad n \in \mathcal{N}, k \in \mathcal{K}, \end{array} \right. \\ \text{with } \mathcal{L}(\boldsymbol{\lambda}, \mathbf{s}_k, k \in \mathcal{K}) = \sum\limits_{n\in\mathcal{N}} w_n R^n - \sum\limits_{n\in\mathcal{N}} \lambda_n (P^n - P^{n,\text{tot}}) \end{array} \tag{8.6}$$

where $\mathcal{L}(\boldsymbol{\lambda}, \mathbf{s}_k, k \in \mathcal{K})$ is the Lagrangian. This can be reformulated as:

$$\begin{array}{l} g(\boldsymbol{\lambda}) = \left\{ \begin{array}{ll} \max\limits_{\{\mathbf{s}_k, k\in\mathcal{K}\}} & \sum\limits_{k\in\mathcal{K}} \Big\{ b_k(\mathbf{s_k}) - \sum\limits_{n\in\mathcal{N}} \lambda_n s_k^n + \sum\limits_{n\in\mathcal{N}} \lambda_n P^{n,\text{tot}}/K \Big\} \\ \text{s.t.} & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \qquad n \in \mathcal{N}, k \in \mathcal{K} \end{array} \right. \\ \text{with } b_k(\mathbf{s}_k) = \sum\limits_{n\in\mathcal{N}} w_n f_s b_k^n(\mathbf{s}_k) \end{array} \tag{8.7}$$

The slave optimization problem (8.7) can then be decomposed into K independent nonconvex subproblems (dual decomposition):

$$\begin{array}{l} g(\boldsymbol{\lambda}) = \sum\limits_{k\in K} g_k(\boldsymbol{\lambda}) \\ \text{with } g_k(\boldsymbol{\lambda}) = \left\{ \begin{array}{ll} \max\limits_{\mathbf{s}_k} & b_k(\mathbf{s}_k) - \sum\limits_{n\in\mathcal{N}} \lambda_n s_k^n + \sum\limits_{n\in\mathcal{N}} \lambda_n P^{n,\text{tot}}/K \\ \text{s.t.} & 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \quad n \in \mathcal{N} \end{array} \right. \end{array} \tag{8.8}$$

The master problem (8.5), also called the dual problem, is a convex optimization problem. Its objective function, i.e. the dual function $g(\boldsymbol{\lambda})$, is however non-differentiable. The reason for this non-differentiability is that the underlying slave optimization problem (8.6) can have multiple globally optimal solutions for some values of the Lagrange multipliers $\boldsymbol{\lambda}$. In [5][13] a subgradient updating approach is proposed for this dual master problem, where the subgradient is defined as,

$$dg(\boldsymbol{\lambda}) \triangleq \sum_{k\in\mathcal{K}} \mathbf{s}_k(\boldsymbol{\lambda}) - \mathbf{P}^{\text{tot}} \tag{8.9}$$

with $\mathbf{s}_k(\boldsymbol{\lambda})$ referring to the optimal solution of (8.8) for given Lagrange multipliers $\boldsymbol{\lambda}$, also called dual variables, and the corresponding subgradient update

is:

$$\boldsymbol{\lambda} = \Big[\boldsymbol{\lambda} + \delta(\sum_{k\in\mathcal{K}} \mathbf{s}_k(\boldsymbol{\lambda}) - \mathbf{P}^{\text{tot}})\Big]^+, \tag{8.10}$$

where $[\mathbf{x}]^+$ denotes the projection of $\mathbf{x} \in \mathcal{R}^N$ onto $\mathcal{R}^N_+$, and where the stepsize δ can be tuned using different procedures [5, 6], e.g. $\delta = q$ or $\delta = q/i$ where q is the initial stepsize and i is the iteration counter. By iteratively applying (8.10) and (8.8), convergence to an optimal solution of (8.5) can be achieved, i.e. $\boldsymbol{\lambda} \rightarrow \boldsymbol{\lambda}^*$, for which the complementary conditions, $\lambda_n(\sum_{k\in\mathcal{K}} s^n_k(\boldsymbol{\lambda}) - P^{n,\text{tot}}) = 0,\ n \in \mathcal{N}$, are satisfied when strong duality "holds" ($K \rightarrow \infty$). This general standard subgradient based dual decomposition approach is visualized in Figure 8.1.

Note that the per-tone subproblems (8.8) are nonconvex optimization problems. Many existing DSM algorithms differ only in the way these subproblems are solved, where strategies are proposed such as exhaustive discrete search (OSB)[4], branch and bound search (PBnB[14], BB-OSB[6]), coordinate descent discrete search (ISB)[13][25], solving the KKT system (DSB[12], MIW[16], MS-DSB[12]), and heuristic approximation (ASB[17], ASB2[12]).

An alternative approach is based on iterative convex approximations such as in SCALE[15] and CA-DSB[12]. This approach basically consists of iteratively executing the following two steps: (i) approximating the nonconvex cWRS problem (8.4) by a separable convex optimization problem F_{cvx}, and (ii) solving this convex approximation by using a subgradient based dual decomposition approach. Note that under some conditions on the approximation, described in [27], iteratively executing these steps results in asymptotic convergence to a locally optimal solution of cWRS (8.4). The convex approximations used by CA-DSB and SCALE both satisfy these conditions. This approach is visualized in Figure 8.2, where $F_{k,\text{cvx}}$ refers to the per-tone convex problem obtained from the convex approximation F_{cvx}. We emphasize that these DSM algorithms also use a subgradient based dual decomposition approach to solve a convex optimization problem in each iteration. This step requires the major part of the computational cost.

8.4 Improved Dual Decomposition

In practice, the standard subgradient based dual decomposition approach is often found to lead to extremely slow convergence or even no convergence at all, especially so for large DSL scenarios (6-20 users) with large crosstalk (VDSL(2)). This is because of different reasons: (i) subgradient methods are generally known not to be efficient, i.e. showing worst case convergence of order $\mathcal{O}(\frac{1}{\epsilon^2})$ with ϵ referring to the required accuracy of the approximation of the optimum [23], (ii) the stepsize used by subgradient methods is quite difficult

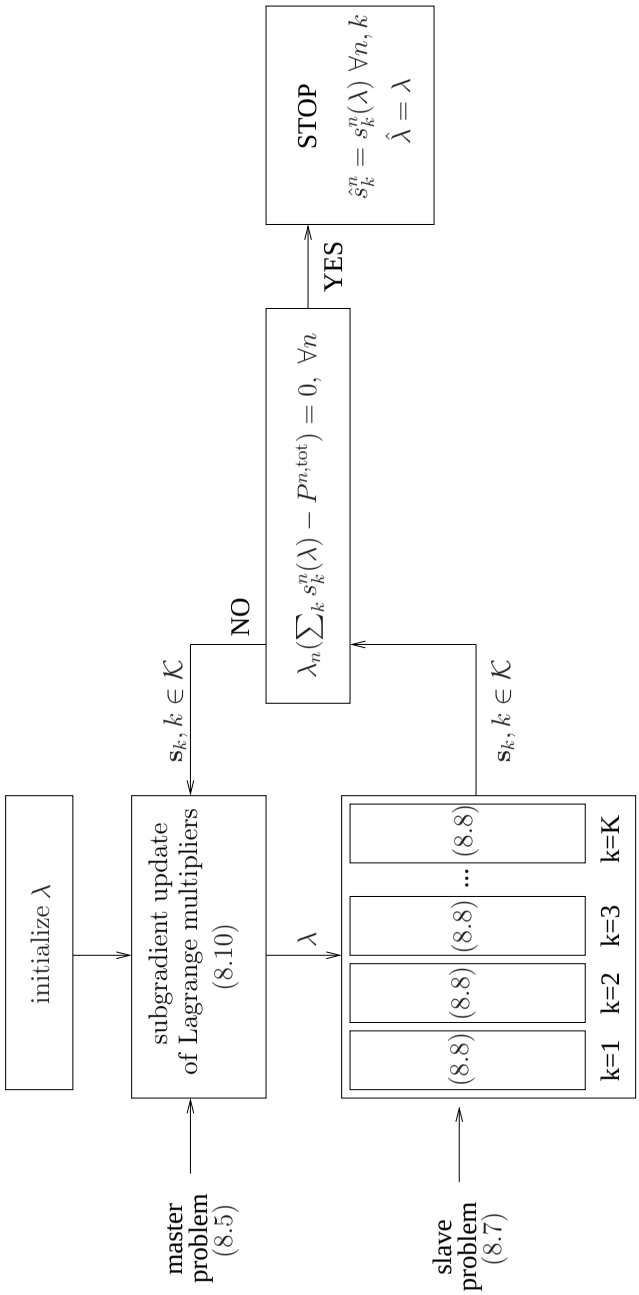


Figure 8.1: General structure of subgradient based dual decomposition approach for DSM

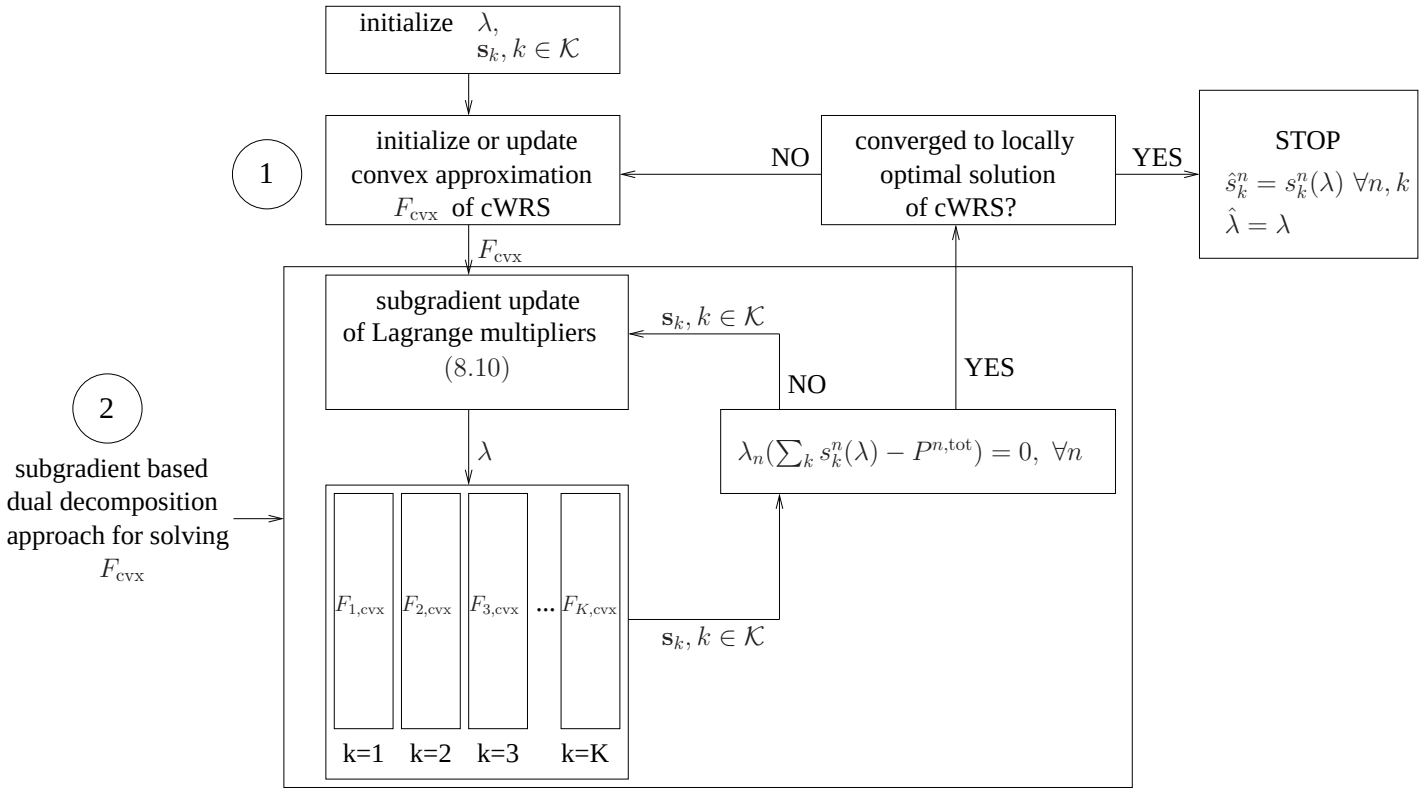


Figure 8.2: Structure of iterative convex approximation approach for DSM

to tune in order to guarantee fast convergence, (iii) the nonconvex nature of the problem implies that special care should be taken in obtaining the optimal primal variables from the optimal dual variables.

Several alternative dual decomposition approaches have been proposed such as the alternating direction method [28], proximal method of multipliers [29], partial inverse method [30], etc. However these approaches only apply to separable convex problems, i.e. with a separable convex objective function and convex coupling constraints. Furthermore they destroy the separability of the problem, they cannot deal with inequality constraints in general, and they are sensitive to the chosen parameter values. Here, we focus on a recently proposed dual decomposition approach in [21], referred to as the proximal center based decomposition method. This method shows interesting properties, namely it preserves the separability of the problem, it uses an optimal gradient based scheme, and it uses an optimal stepsize, thus removing the need for a tuning strategy. However, this method is proposed for convex separable problems, it is mainly derived for two users under equality constraints, and it is presented from a rather high-level point of view, where the steps consist of general optimization problems to be solved (see Algorithm 3.2 in [21]). In this section we extend this method to a concrete improved dual decomposition approach for solving the non-convex problem cWRS (8.4). This approach will be used first in Section 8.4.1 to improve the convergence speed of DSM algorithms using iterative convex approximations (SCALE, CA-DSB) with one order of magnitude. In Section 8.4.2 this will be extended to other DSM algorithms such as OSB, ISB, PBnB, BB-OSB, ASB, (MS-)DSB, MIW, etc. We will refer to these DSM algorithms that are not based on iterative convex approximations as "direct DSM algorithms".

8.4.1 An Improved Dual decomposition Approach for Iterative Convex Approximation Based DSM Algorithms

Two state-of-the-art DSM algorithms that are based on iterative convex approximations are SCALE and CA-DSB. These basically consist of two steps as explained in Section 8.3.2, which are iteratively executed. In this section we will propose an improved dual decomposition approach for solving the convex optimization problem in the second step. We will elaborate this for CA-DSB and prove that its convergence speed is improved by one order of magnitude, i.e. from $\mathcal{O}(\frac{1}{\epsilon^2})$ to $\mathcal{O}(\frac{1}{\epsilon})$, with similar computational complexity with respect to the subgradient based dual decomposition approach. The improved dual decomposition approach can similarly be applied to the SCALE algorithm to obtain a similar speed up, but requires more complicated notation because of the inherent exponential transformation of variables. The content of this section has also appeared in [1].

For CA-DSB, the convex approximation in each iteration is obtained by reformulating the objective of cWRS, as a sum of a concave part and a convex part, and then approximating the convex part by a first order Taylor expansion. The resulting convex approximation, its dual formulation, dual function, and Lagrangian are given in (8.11), (8.12), (8.13), and (8.14), respectively.

$$f^*_{\text{cvx}} = \Big\{ \max_{\{\mathbf{s}_k \in \mathcal{S}_k, k\in\mathcal{K}\}} \sum_{k\in\mathcal{K}} b_{k,\text{cvx}}(\mathbf{s}_k) \;\; \text{s.t.} \sum_{k\in\mathcal{K}} s_k^n \le P^{n,\text{tot}},\; n \in \mathcal{N} \Big\} \tag{8.11}$$

$$\min_{\boldsymbol{\lambda}\ge 0} g_{\text{cvx}}(\boldsymbol{\lambda}) \tag{8.12}$$

$$g_{\text{cvx}}(\boldsymbol{\lambda}) = \max_{\{\mathbf{s}_k \in \mathcal{S}_k, k\in\mathcal{K}\}} \mathcal{L}_{\text{cvx}}(\mathbf{s}_k, k\in\mathcal{K}, \boldsymbol{\lambda}) \tag{8.13}$$

$$\mathcal{L}_{\text{cvx}}(\mathbf{s}_k, k\in\mathcal{K}, \boldsymbol{\lambda}) = \sum_{k\in\mathcal{K}} b_{k,\text{cvx}}(\mathbf{s}_k) - \sum_{k\in\mathcal{K}} \sum_{n\in\mathcal{N}} \lambda_n s_k^n + \sum_{n\in\mathcal{N}} \lambda_n P^{n,\text{tot}} \tag{8.14}$$

where $\mathcal{S}_k = \{\mathbf{s}_k \in \mathcal{R}^n : 0 \le s_k^n \le s_k^{n,\max}, n \in \mathcal{N}\}$ is a compact convex set with $s_k^{n,\max} := \min(s_k^{n,\text{mask}}, P^{n,\text{tot}})$ and $P^{n,\text{tot}} < \infty$, and where $b_{k,\text{cvx}}(\mathbf{s}_k)$ is concave and given as:

$$b_{k,\text{cvx}}(\mathbf{s}_k) = \sum_{n\in\mathcal{N}} w_n f_s \log_2\Big(\sum_{m\in\mathcal{N}} |\tilde{h}_k^{n,m}|^2 s_k^m + \Gamma\sigma_k^n \Big) - \sum_{n\in\mathcal{N}} w_n f_s \Big(\sum_{m\ne n} a_k^{m,n} s_k^m + c_k^n \Big), \tag{8.15}$$

with $a_k^{n,m}, c_k^n, \forall n, m, k$ constant approximation parameters, obtained by a closed-form formula in the approximation step [12], and with

$$|\tilde{h}_k^{n,m}|^2 \begin{cases} = \Gamma |h_k^{n,m}|^2, & n \ne m \\ = |h_k^{n,m}|^2, & n = m. \end{cases} \tag{8.16}$$

The convex problem (8.11) has a separable structure and so the standard way to solve it is by focusing on the dual problem (8.12) and using a subgradient update approach for the dual variables. This subgradient based dual decomposition approach is however known [23] to have a convergence speed of order $\mathcal{O}(\frac{1}{\epsilon^2})$, where ϵ is the required accuracy for the approximation of the optimum. In the sequel, it will be shown how the "proximal center based decomposition" method from [21] can be adapted for solving the convex approximation, leading to a scheme with convergence speed of order $\mathcal{O}(\frac{1}{\epsilon})$, i.e. one order of magnitude faster but with the same computational complexity. The basic steps in this result are as follows. First an approximated (smoothed) dual function $\bar{g}_{\text{cvx}}(\boldsymbol{\lambda})$ is defined that can be chosen to be arbitrarily close to the original dual function $g_{\text{cvx}}(\boldsymbol{\lambda})$. Then it is proven that this smoothed dual function $\bar{g}_{\text{cvx}}$ is differentiable and has a Lipschitz continuous gradient. Finally an optimal gradient scheme is applied to the smoothed dual problem.

We introduce the following functions $d_k(\mathbf{s}_k), k \in \mathcal{K}$, which are called prox-functions in [21] and are defined as follows:

Definition 8.1 A prox-function $d_k(\mathbf{s}_k)$ has the following properties:

- $d_k(\mathbf{s}_k)$ is a non-negative continuous and strongly convex function[1] with convexity parameter $\sigma_{\mathcal{S}_k}$
- $d_k(\mathbf{s}_k)$ is defined for the compact convex set $\mathcal{S}_k$

□

An example of a valid prox-function is $d_k(\mathbf{s}_k) = \frac{1}{2}\|\mathbf{s}_k\|^2$, which is also used in our concrete implementations (see Section 8.6). As many other valid prox-functions exist, and in order not to loose generality, we continue with $d_k(\mathbf{s}_k)$. Since $\mathcal{S}_k, k \in \mathcal{K}$, are compact and $d_k(\mathbf{s}_k)$ are continuous, we can choose finite and positive constants such that

$$D_{\mathcal{S}_k} \geq \max_{\mathbf{s}_k \in \mathcal{S}_k} d_k(\mathbf{s}_k), k \in \mathcal{K}. \tag{8.17}$$

This upper bound for $d_k(\mathbf{s}_k)$ can be easily computed. For instance, for the choice of prox-function $d_k(\mathbf{s}_k) = \frac{1}{2}\|\mathbf{s}_k\|^2$, $D_{\mathcal{S}_k}$ can be computed in closed-form as $D_{\mathcal{S}_k} = \frac{1}{2}\sum_{n\in\mathcal{N}}(s_k^{n,\max})^2$.

The prox-functions can be used to smoothen the dual function $g_{\text{cvx}}(\boldsymbol{\lambda})$ to obtain a smoothed dual function $\bar{g}_{\text{cvx}}(\boldsymbol{\lambda})$ as follows:

$$\bar{g}_{\text{cvx}}(\boldsymbol{\lambda}) = \max_{\{\mathbf{s}_k \in \mathcal{S}_k, k\in\mathcal{K}\}} \sum_{k\in\mathcal{K}} \Big\{ b_{k,\text{cvx}}(\mathbf{s}_k) - \sum_{n\in\mathcal{N}} \lambda_n (s_k^n - P^{n,\text{tot}}/K) - c d_k(\mathbf{s}_k) \Big\}, \tag{8.18}$$

where c is a positive smoothness parameter that will be defined in closed-form in Theorem 8.3 later this section. The value of this parameter c is defined sufficiently small, so as to make the smoothed dual function arbitrarily close to the original dual function.

One useful property of the particular choice of prox-functions is that they do not destroy the separability of the objective function in (8.18), i.e.

$$\bar{g}_{\text{cvx}}(\boldsymbol{\lambda}) = \sum_{k\in\mathcal{K}} \Big\{ \max_{\mathbf{s}_k \in \mathcal{S}_k} b_{k,\text{cvx}}(\mathbf{s}_k) - \sum_{n\in\mathcal{N}} \lambda_n (s_k^n - P^{n,\text{tot}}/K) - c d_k(\mathbf{s}_k) \Big\}. \tag{8.19}$$

Denote by $\bar{\mathbf{s}}_{k,\text{cvx}}(\boldsymbol{\lambda}), k \in \mathcal{K}$, the optimal solution of the maximization problem in (8.19). The following theorem describes the properties of the smoothed dual function $\bar{g}_{\text{cvx}}(\boldsymbol{\lambda})$:

[1] A continuously differentiable function $f(x)$ is called strongly convex on $\mathbb{R}^N$ if there exists a constant μ, called the convexity parameter of f such that for any $x, y \in \mathbb{R}^N$ we have $f(y) \geq f(x) + \nabla f(x)^T(y-x) + \frac{1}{2}\mu\|y-x\|^2$ [23].

Theorem 8.2 ([21]) *The function $\bar{g}_{\text{cvx}}(\boldsymbol{\lambda})$ is convex and continuously differentiable at any $\boldsymbol{\lambda} \in \mathcal{R}^n$. Moreover, its gradient $\nabla \bar{g}_{\text{cvx}}(\boldsymbol{\lambda}) = \sum_{k\in\mathcal{K}} \bar{\mathbf{s}}_{k,\text{cvx}}(\boldsymbol{\lambda}) - \mathbf{P}^{\text{tot}}$ is Lipschitz continuous with Lipschitz constant $L_c = \sum_{k\in\mathcal{K}} \frac{1}{c\sigma_{\mathcal{S}_k}}$. The following inequalities also hold:*

$$\bar{g}_{\text{cvx}}(\boldsymbol{\lambda}) \leq g_{\text{cvx}}(\boldsymbol{\lambda}) \leq \bar{g}_{\text{cvx}}(\boldsymbol{\lambda}) + c\sum_{k\in\mathcal{K}} D_{\mathcal{S}_k} \qquad \boldsymbol{\lambda} \in \mathcal{R}^n \tag{8.20}$$

The addition of the prox-functions thus leads to a convex differentiable dual function with Lipschitz continuous gradient. Now instead of solving the original dual problem (8.12), we focus on the problem:

$$\min_{\boldsymbol{\lambda}\geq 0} \bar{g}_{\text{cvx}}(\boldsymbol{\lambda}) \tag{8.21}$$

Note that, by defining c sufficiently small in (8.19), the solution of (8.21) can be brought arbitrarily close to the solution of (8.12). Taking the particular structure of (8.21) into account, i.e. a differentiable objective function with Lipschitz continuous gradient, we propose the optimal gradient based scheme given in Algorithm 8.1, derived from [21], for solving (8.11). This algorithm will be referred to as the *improved dual decomposition algorithm* for solving the convex approximation of CA-DSB (8.11). The specific value for the parameters c, $D_{\mathcal{S}_k}$, L_c and $\sigma_{\mathcal{S}_k}$ are fully defined by the choice of the prox-function $d_k(\mathbf{s}_k)$ and also the required accuracy ϵ, that depends on the application. For instance, for the choice of prox-function $d_k(\mathbf{s}_k) = \frac{1}{2}\|\mathbf{s}_k\|^2$, the following simple closed-form expressions can be derived for the parameters

$$D_{\mathcal{S}_k} = \tfrac{1}{2}\sum_{n\in\mathcal{N}} (s_k^{n,\max})^2, \quad c = \frac{\epsilon}{\frac{1}{2}\sum_{k\in\mathcal{K}}\sum_{n\in\mathcal{N}} (s_k^{n,\max})^2}, \quad \sigma_{\mathcal{S}_k} = 1, \quad L_c = \tfrac{K}{c}. \tag{8.22}$$

Note that lines 6, 7 and 8 of Algorithm 8.1 constitute the major part of the computational complexity, and these computations are also done by subgradient based dual decomposition algorithms. Lines 9, 10 and 11 are the extra lines of the improved approach with respect to the subgradient approach, but the computational complexity of these lines is negligible with respect to that of lines 6, 7 and 8. So in terms of complexity per Lagrange multiplier update, we can state that these are similar for both approaches.

The remaining issue is to prove that $\hat{\mathbf{s}}_k, k \in \mathcal{K}$, i.e. the result of Algorithm 8.1, converges to an ϵ-optimal solution in $i_{\max}$ iterations where $i_{\max}$ is of the order $\mathcal{O}(\frac{1}{\epsilon})$. For this we define the following lemmas that will be used in the sequel.

Lemma 2 *For any $\mathbf{y} \in \mathcal{R}^n$ and $\mathbf{z} \geq 0$, the following inequality holds[2]:*

$$\mathbf{y}^T\mathbf{z} \ \leq \|[\mathbf{y}]^+\|\|\mathbf{z}\| \tag{8.23}$$

[2]For the sake of an easy exposition we consider in the paper only the Euclidian norm $\|\cdot\|$, although other norms can also be used (see [21] for a detailed exposition).

Algorithm 8.1 Improved dual decomposition algorithm for solving (8.11) for CA-DSB

$i := 0$, tmp $:= 0$
initialize $i_{\max}$, $\boldsymbol{\lambda}^i$
initialize required application accuracy ϵ (= upper bound on the dual gap)
$c := \frac{\epsilon}{\sum_{k\in\mathcal{K}} D_{\mathcal{S}_k}}$, $L_c := \sum_{k\in\mathcal{K}} \frac{1}{c\ \sigma_{\mathcal{S}_k}}$
for $i = 0 \dots i_{\max}$ **do**
  $\forall k : \mathbf{s}_k^{i+1} = \underset{\{\mathbf{s}_k\in\mathcal{S}_k\}}{\operatorname{argmax}}\ b_{k,\mathrm{cvx}}(\mathbf{s}_k) - \sum_{n\in\mathcal{N}} \boldsymbol{\lambda}_n^i s_k^n - c d_k(\mathbf{s}_k)$
  $d\bar{g}_c^{i+1} = \sum_{k\in\mathcal{K}} \mathbf{s}_k^{i+1} - \mathbf{P}^{\mathrm{tot}}$
  $\mathbf{u}^{i+1} = [\frac{d\bar{g}_c^{i+1}}{L_c} + \boldsymbol{\lambda}^i]^+$
  tmp $:=$ tmp $+ \frac{i+1}{2} d\bar{g}_c^{i+1}$
  $\mathbf{v}^{i+1} = [\frac{\mathrm{tmp}}{L_c}]^+$
  $\boldsymbol{\lambda}^{i+1} = \frac{i+1}{i+3}\mathbf{u}^{i+1} + \frac{2}{i+3}\mathbf{v}^{i+1}$
  $i := i + 1$
end for
Build $\hat{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^{i_{\max}+1}$ and $\hat{\mathbf{s}}_k = \sum_{i=0}^{i_{\max}} \frac{2(i+1)}{(i_{\max}+1)(i_{\max}+2)} \mathbf{s}_k^{i+1}$

Proof: Let us define the index sets $\mathcal{I}^- = \{i \in \{1\dots n\} : y_i < 0\}$ and $\mathcal{I}^+ = \{i \in \{1\dots n\} : y_i \geq 0\}$. Then,

$$\mathbf{y}^T\mathbf{z} = \sum_{i\in\mathcal{I}^-} y_i z_i + \sum_{i\in\mathcal{I}^+} y_i z_i \leq \sum_{i\in\mathcal{I}^+} y_i z_i = ([\mathbf{y}]^+)^T\mathbf{z} \leq \|[\mathbf{y}]^+\|\|\mathbf{z}\|.$$

□

The following lemma provides a lower bound for the primal gap, $f^*_{\mathrm{cvx}} - \sum_{k\in\mathcal{K}} b_{k,\mathrm{cvx}}(\hat{\mathbf{s}}_k)$, of (8.11):

Lemma 3 *Let $\boldsymbol{\lambda}^*$ be any optimal Lagrange multiplier, then for any $\hat{\mathbf{s}}_k \in \mathcal{S}_k, k \in \mathcal{K}$, the following lower bound on the primal gap holds:*

$$f^*_{\mathrm{cvx}} - \sum_{k\in\mathcal{K}} b_{k,\mathrm{cvx}}(\hat{\mathbf{s}}_k) \geq -\|\boldsymbol{\lambda}^*\|\|[\sum_{k\in\mathcal{K}} \hat{\mathbf{s}}_k - \mathbf{P}^{\mathrm{tot}}]^+\| \tag{8.24}$$

Proof: From the assumptions of the lemma we have

$$\begin{aligned} f^*_{\mathrm{cvx}} &= \max_{\{\mathbf{s}_k\in\mathcal{S}_k, k\in\mathcal{K}\}} \sum_{k\in\mathcal{K}} b_{k,\mathrm{cvx}}(\mathbf{s}_k) - \boldsymbol{\lambda}^{*T}(\sum_{k\in\mathcal{K}} \mathbf{s}_k - \mathbf{P}^{\mathrm{tot}}) \\ &\geq \sum_{k\in\mathcal{K}} b_{k,\mathrm{cvx}}(\hat{\mathbf{s}}_k) - \boldsymbol{\lambda}^{*T}(\sum_{k\in\mathcal{K}} \hat{\mathbf{s}}_k - \mathbf{P}^{\mathrm{tot}}) \end{aligned} \tag{8.25}$$

and then (8.24) is obtained by applying Lemma 2.

□

From Lemma 3 it follows that if $\|[\sum_{k\in\mathcal{K}} \hat{\mathbf{s}}_k - \mathbf{P}^{\text{tot}}]^+\| \leq \epsilon_c$, then the primal gap is bounded, i.e. for all $\hat{\boldsymbol{\lambda}} \in \mathcal{R}_+^N$

$$-\epsilon_c\|\boldsymbol{\lambda}^*\| \leq f^*_{\text{cvx}} - \sum_{k\in\mathcal{K}} b_{k,\text{cvx}}(\hat{\mathbf{s}}_k) \leq g_{\text{cvx}}(\hat{\boldsymbol{\lambda}}) - \sum_{k\in\mathcal{K}} b_{k,\text{cvx}}(\hat{\mathbf{s}}_k). \tag{8.26}$$

Therefore, if we are able to derive an upper bound ϵ for the dual gap, namely $g_{\text{cvx}}(\hat{\boldsymbol{\lambda}}) - \sum_{k\in\mathcal{K}} b_{k,\text{cvx}}(\hat{\mathbf{s}}_k)$, and an upper bound ϵ_c for the coupling constraints for some given $\hat{\boldsymbol{\lambda}}$ (≥ 0) and $\hat{\mathbf{s}}_k \in \mathcal{S}_k, \forall k$, then we can conclude that $\hat{\mathbf{s}}_k$ is an (ϵ, ϵ_c)-solution for (8.11) (since in this case $-\epsilon_c\|\boldsymbol{\lambda}^*\| \leq f^*_{\text{cvx}} - \sum_{k\in\mathcal{K}} b_{k,\text{cvx}}(\hat{\mathbf{s}}_k) \leq \epsilon$). The next theorem derives these upper bounds for Algorithm 8.1 and provides a concrete value for c.

Theorem 8.3 *Let $\boldsymbol{\lambda}^*$ be an optimal Lagrange multiplier, taking $c = \frac{\epsilon}{\sum_{k\in\mathcal{K}} D_{\mathcal{S}_k}}$ and $i_{\max} + 1 = 2\sqrt{(\sum_k \frac{1}{\sigma_{\mathcal{S}_k}})(\sum_k D_{\mathcal{S}_k})}\frac{1}{\epsilon}$, then after $i_{\max}$ iterations Algorithm 8.1 obtains an approximate solution $\hat{\mathbf{s}}_k, k \in \mathcal{K}$, to the convex approximation (8.11) with a duality gap less than ϵ, i.e.*

$$g_{\text{cvx}}(\hat{\boldsymbol{\lambda}}) - \sum_{k\in\mathcal{K}} b_{k,\text{cvx}}(\hat{\mathbf{s}}_k) \leq \epsilon, \tag{8.27}$$

and the constraints satisfy

$$\|[\sum_k \hat{\mathbf{s}}_k - \mathbf{P}^{\text{tot}}]^+\| \leq \epsilon(\|\boldsymbol{\lambda}^*\| + \sqrt{\|\boldsymbol{\lambda}^*\|^2 + 2}) \tag{8.28}$$

Proof : Using a similar reasoning as in Theorem 3.4 in [21] we can show that for any c the following inequality holds:

$$\bar{g}_{\text{cvx}}(\hat{\boldsymbol{\lambda}}) \leq \min_{\boldsymbol{\lambda}\geq 0} \Big\{ \frac{2L_c}{(i_{\max}+1)^2}\|\boldsymbol{\lambda}\|^2 + \sum_{i=0}^{i_{\max}} \frac{2(i+1)}{(i_{\max}+1)(i_{\max}+2)} [\bar{g}_{\text{cvx}}(\lambda^i) + (\nabla \bar{g}_{\text{cvx}}(\lambda^i))^T(\lambda - \lambda^i)] \Big\}$$

Replacing $\bar{g}_{\text{cvx}}(\lambda^i)$ and $\nabla\bar{g}_{\text{cvx}}(\lambda^i)$ by their expressions given in (8.18) and Theorem 8.2, respectively, and taking into account that the functions $b_{k,\text{cvx}}$ are

concave, we obtain the following inequality:

$$\begin{aligned} g_{\text{cvx}}(\hat{\boldsymbol{\lambda}}) - \sum_{k\in\mathcal{K}} b_{k,\text{cvx}}(\hat{\mathbf{s}}_k) \\ \leq c(\sum_{k\in\mathcal{K}} D_{\mathcal{S}_k}) + \min_{\boldsymbol{\lambda}\geq 0} \big\{ \frac{2L_c}{(i_{\max}+1)^2}\|\boldsymbol{\lambda}\|^2 - \langle \lambda, \sum_k \hat{\mathbf{s}}_k - \mathbf{P}^{\text{tot}} \rangle \big\} \\ = c(\sum_{k\in\mathcal{K}} D_{\mathcal{S}_k}) - \frac{(i_{\max}+1)^2}{8L_c}\|[\sum_k \hat{\mathbf{s}}_k - \mathbf{P}^{\text{tot}}]^+\|^2 \leq c(\sum_{k\in\mathcal{K}} D_{\mathcal{S}_k}). \end{aligned}$$

By taking $c = \frac{\epsilon}{\sum_{k\in\mathcal{K}} D_{\mathcal{S}_k}}$, we obtain (8.27). For the constraints using Lemma 3 and the previous inequality we get that $\|[\sum_k \hat{\mathbf{s}}_k - \mathbf{P}^{\text{tot}}]^+\|$ satisfies the second order inequality in y: $\frac{(i_{\max}+1)^2}{8L_c}y^2 - \|\lambda^*\|y - \epsilon \leq 0$. Therefore, $\|[\sum_k \hat{\mathbf{s}}_k - \mathbf{P}^{\text{tot}}]^+\|$ must be less than the largest root of the corresponding second-order equation, i.e.

$$\|[\sum_k \hat{\mathbf{s}}_k - \mathbf{P}^{\text{tot}}]^+\| \leq \big(\|\lambda^*\| + \sqrt{\|\lambda^*\|^2 + \frac{\epsilon(i_{\max}+1)^2}{2L_c}}\big)\frac{4L_c}{(i_{\max}+1)^2}.$$

By taking $i_{\max} = 2\sqrt{(\sum_k \frac{1}{\sigma_{\mathcal{S}_k}})(\sum_k D_{\mathcal{S}_k})\frac{1}{\epsilon}} - 1$, we obtain (8.28). □

From Theorem 8.3 we can conclude that by taking $c = \frac{\epsilon}{\sum_{k\in\mathcal{K}} D_{\mathcal{S}_k}}$, Algorithm 8.1 converges to a solution with duality gap less than ϵ and the constraints violation satisfy $\|[\sum_k \hat{\mathbf{s}}_k - \mathbf{P}^{\text{tot}}]^+\| \leq \epsilon(\|\boldsymbol{\lambda}^*\| + \sqrt{\|\boldsymbol{\lambda}^*\|^2 + 2})$ after $i_{\max} = 2\sqrt{(\sum_k \frac{1}{\sigma_{\mathcal{S}_k}})(\sum_k D_{\mathcal{S}_k})\frac{1}{\epsilon}} - 1$ iterations, i.e. the convergence speed is of the order $\mathcal{O}(\frac{1}{\epsilon})$.

Note that Algorithm 8.1 provides a fully automatic approach. Once the required application accuracy ϵ and the particular prox-function $d_k(\mathbf{s}_k)$ are defined, all parameters are fixed. The algorithm then automatically updates its stepsize so as to converge fast to the optimal dual value within the specified accuracy. It does not require any stepsize tuning, which otherwise is known to be a very difficult and crucial process. Finally note that combining this algorithm with an outer loop that iteratively updates the convex approximations leads to an overall procedure that converges to a local maximizer of the nonconvex problem cWRS [27][12]. The extension of CA-DSB with the improved dual decomposition approach will be referred to as Improved CA-DSB (I-CA-DSB).

A final remark on Algorithm 8.1 is that the independent convex per-tone problems (line 6 of Algorithm 8.1) are slightly modified with respect to the standard per-tone problems for CA-DSB. This is a consequence of the addition

of the extra prox-function term. One can use state-of-the-art iterative methods (e.g. Newton's method) to solve this convex subproblem with guaranteed convergence. An alternative consists in using an iterative fixed point update approach, which is shown to work well, with very small complexity, and is easily extended to distributed implementation by using a protocol [15][12]. We propose a modified fixed point update formula for the transmit powers s_k^n used by CA-DSB, so as to take the extra prox-term into account. Following the same procedure as explained in [12], consisting of a fixed point reformulation of the corresponding KKT stationarity condition of (8.12), and for the choice of prox-function $d_k(\mathbf{s}_k) = \frac{1}{2}\|\mathbf{s}_k\|^2$, we obtain the following transmit power update formula, that only differs in the presence of the term PROX:

$$s_k^n = \left[\left(\frac{w_n f_s / \log(2)}{\lambda_n + \underbrace{2cs_k^n}_{\text{PROX}} + \sum_{m \neq n} \omega_m f_s a_k^{n,m} - \sum_{m \neq n} \frac{w_m f_s \Gamma |h_k^{m,n}|^2 / \log(2)}{\sum_p |\tilde{h}_k^{m,p}|^2 s_k^p + \Gamma \sigma_k^m}} \right) - \frac{\sum_{m \neq n} \Gamma |h_k^{n,m}|^2 s_k^m + \Gamma \sigma_k^n}{|h_k^{n,n}|^2} \right]_0^{s_k^{n,\text{mask}}}. \tag{8.29}$$

Providing convergence conditions for this type of iterative fixed point updates is outside the scope of this paper. In [17, 12, 18], convergence is proven under certain conditions, and demonstrated for realistic DSL scenarios. This leads to an alternative and fast way of implementing line 6 of Algorithm 8.1, as specified in Algorithm 8.2. The number of iterations in line 2 is typically fixed at 3.

Algorithm 8.2 Iterative fixed point update approach for solving line 6 of Algorithm 8.1

1: **for** $k = 1 \ldots K$ **do**
2: **for** iterations **do**
3: **for** $n = 1 \ldots N$ **do**
4: $s_k^n =$(8.29)
5: **end for**
6: **end for**
7: **end for**

As mentioned, although the improved dual decomposition approach has been elaborated for CA-DSB, it can similarly be applied to other DSM algorithms based on iterative convex approximations, like for instance SCALE, with a similar speed up of convergence. In this case the prox-function can be taken as $d_k(\mathbf{s}_k) = \|\mathbf{s}_k\|^2$, resulting in concrete values for c, $i_{\max}$ and L_c. The extension of SCALE with the improved dual decomposition approach will be referred to as Improved SCALE (I-SCALE).

Finally, we would like to remark that the idea of adding a strictly convex term to a non-strictly convex objective to improve the sensitivity is a known technique that has been proposed before [31, 32, 33, 34]. Algorithm 8.1 however extends this with automatic tuning strategies, concrete convergence orders, and an optimal gradient based method by following the approach of [21], which is here particularly elaborated for DSL DSM.

8.4.2 An Improved Dual Decomposition Approach for Direct DSM Algorithms

In this section we extend the improved dual decomposition approach to direct DSM algorithms such as OSB, ISB, ASB, (MS-)DSB, MIW, etc, corresponding to the structure visualized in Figure 8.1. Using a similar trick as in Section 8.4.1, we define a smoothed dual function $\bar{g}(\boldsymbol{\lambda})$ as follows

$$\bar{g}(\boldsymbol{\lambda}) = \max_{\{\mathbf{s}_k \in \mathcal{S}_k, k \in \mathcal{K}\}} \sum_{k \in \mathcal{K}} b_k(\mathbf{s}_k) - \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} \lambda_n s_k^n + \sum_{n \in \mathcal{N}} \lambda_n P^{n,\text{tot}} - \sum_{k \in \mathcal{K}} c d_k(\mathbf{s}_k) \tag{8.30}$$

where $d_k(\mathbf{s}_k)$ is a prox-function and c is a positive smoothness parameter.

Note that by defining parameter c to a sufficiently small value, the smoothed dual function $\bar{g}(\boldsymbol{\lambda})$ can be brought arbitrarily close to the original dual function $g(\boldsymbol{\lambda})$, i.e. $\bar{g}(\boldsymbol{\lambda}) \approx g(\boldsymbol{\lambda})$. Based on the obtained smoothed dual function $\bar{g}(\boldsymbol{\lambda})$, we propose the *improved dual decomposition approach for direct DSM algorithms* as shown in Algorithm 8.3, where line 6 corresponds to solving the following optimization problem:

$$\begin{aligned} \tilde{\mathbf{s}}_k(\boldsymbol{\lambda}) := \quad & \underset{\mathbf{s}_k}{\text{argmax}} \quad b_k(\mathbf{s}_k) - \sum_{n \in \mathcal{N}} \lambda_n s_k^n - c d_k(\mathbf{s}_k) \\ & \text{s.t.} \qquad 0 \leq s_k^n \leq s_k^{n,\text{mask}}, \quad n \in \mathcal{N}. \end{aligned} \tag{8.31}$$

For the concrete choice of $d_k(\mathbf{s}_k) = \frac{1}{2}\|\mathbf{s}_k\|^2$, the values for c, L_c, D_{S_k} and σ_{S_k} can be computed by the same simple closed-form expressions as given in (8.22).

Algorithm 8.3 uses a similar optimal gradient based scheme on the smoothed dual function as in Algorithm 8.1. Again no stepsize tuning is needed. Besides the improved updating procedure for the Lagrange multipliers (lines 7-11), it involves a slightly different decomposed per-tone problem (8.31) (line 6). This can be solved by using a discrete exhaustive search similar to OSB, a discrete coordinate descent method similar to ISB, or a KKT system approach similar to DSB/MIW/MS-DSB using (8.29), where $a_k^{n,m} = \frac{\Gamma |h_k^{m,n}|^2 / \log(2)}{\sum_{p \neq m} \Gamma |h_k^{m,p}|^2 s_k^p + \Gamma \sigma_k^m}$ [12]. One can also use a virtual reference length approach similar to ASB, ASB2. Note that for ASB, and when using $d_k(\mathbf{s}_k) = \|\mathbf{s}_k\|^2$, this increases the complexity as a polynomial equation of degree 4 is then to be solved instead

Algorithm 8.3 Improved dual decomposition approach for direct DSM algorithms

1: $i := 0$, tmp $:= 0$
2: initialize $\boldsymbol{\lambda}^i$ and ϵ_a (desired accuracy on per-user total powers)
3: initialize required application accuracy ϵ
4: $c := \frac{\epsilon}{\sum_{k\in\mathcal{K}} D_{\mathcal{S}_k}}$, $L_c := \sum_{k\in\mathcal{K}} \frac{1}{c\ \sigma_{\mathcal{S}_k}}$
5: **while** $\exists n : (\text{abs}(\lambda_n^i(\sum_{k\in\mathcal{K}} s_k^n - P^{n,\text{tot}})) \geq \epsilon_a)$ **do**
6: $\quad \forall k : \mathbf{s}_k^{i+1} = \tilde{\mathbf{s}}_k(\boldsymbol{\lambda}^{\boldsymbol{i}})$ obtained by solving (8.31)
7: $\quad dg^{i+1} = \sum_{k\in\mathcal{K}} \mathbf{s}_k^{i+1} - \mathbf{P}^{\text{tot}}$
8: $\quad \mathbf{u}^{i+1} = [\frac{dg^{i+1}}{L_c} + \boldsymbol{\lambda}^i]^+$
9: $\quad$ tmp $:=$ tmp $+ \frac{i+1}{2} dg^{i+1}$
10: $\quad \mathbf{v}^{i+1} = [\frac{\text{tmp}}{L_c}]^+$
11: $\quad \boldsymbol{\lambda}^{i+1} = \frac{i+1}{i+3}\mathbf{u}^{i+1} + \frac{2}{i+3}\mathbf{v}^{i+1}$
12: $\quad i := i + 1$
13: **end while**
14: Build $\hat{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^i$ and $\hat{\mathbf{s}}_k = \mathbf{s}_k^i, \forall k \in \mathcal{K}$

of a cubic equation. Depending on the choice of the algorithm for solving the per-tone problem, there will be a trade-off in complexity versus performance [12]. We will again add the prefix 'I-' to refer to these algorithms using the improved dual decomposition approach, i.e. I-OSB,I-ISB, I-DSB/MIW, I-MS-DSB, I-ASB.

The main difference of Algorithm 8.3 is that line 6 now involves K nonconvex optimization problems, while line 6 of Algorithm 8.1 involves K (strong) convex optimization problems. As a consequence, the smoothed dual function $\bar{g}(\boldsymbol{\lambda})$ is not necessarily differentiable and its gradient is not necessarily Lipschitz continuous. More specifically, this is the case when $\bar{g}(\boldsymbol{\lambda})$ has multiple globally optimal solutions for a given Lagrange multiplier $\boldsymbol{\lambda}$. This non-uniqueness problem becomes a true problem only for a particular type of scenarios, namely symmetric DSL scenarios with large crosstalk, where multiple adjacent tones have multiple globally optimal solutions. This will be analyzed and discussed in more detail in Section 8.5. For these particular scenarios the worst case convergence of order $\mathcal{O}(\frac{1}{\epsilon})$ can not be guaranteed, as in Theorem 8.3, but still we can expect an improved convergence behaviour with respect to the standard subgradient approach. Except for these specific cases, and so for most practical DSL scenarios, the smoothed dual function $\bar{g}(\boldsymbol{\lambda})$ will be differentiable and Lipschitz continuous, and so a worst case convergence speed of $\mathcal{O}(\frac{1}{\epsilon})$ is guaranteed. For instance, in [35] conditions on the channel and noise parameters were given

under which cWRS can be "convexified". For these conditions, differentiability and Lipschitz continuity holds for $\bar{g}(\boldsymbol{\lambda})$ and so application of Algorithm 8.3 will provide a worst case convergence of $\mathcal{O}(\frac{1}{\epsilon})$.

8.5 An Interleaving Procedure for Recovering the Primal Solution from the Dual Solution

The subgradient based dual decomposition approach for solving problem cWRS (8.4) as well as the improved dual decomposition approach presented in Sections 8.4.1 and 8.4.2, converge to the optimal dual variables. However, because of the nonconvex nature of cWRS, extra care must be taken when recovering the optimal primal solution, i.e. optimal transmit powers $\mathbf{s}_k^*, k \in \mathcal{K}$, for (8.4), from the optimal dual variables $\boldsymbol{\lambda}^*$, as was also mentioned in [5][31]. The fact that the objective function of cWRS is not strictly concave, can result in cases where the optimal $\mathbf{s}_k(\boldsymbol{\lambda}^*), k \in \mathcal{K}$, that solves (8.7) is not unique, leading to multiple solutions $\mathbf{s}_k(\boldsymbol{\lambda}^*), k \in \mathcal{K}$, for given optimal dual variables $\boldsymbol{\lambda}^*$. Formally this can be expressed as follows:

$$\{\mathbf{s}_k(\boldsymbol{\lambda}^*), k \in \mathcal{K}\} \in \mathcal{B} = \{(\tilde{\mathbf{s}}_{k,1}, k \in \mathcal{K}), \ldots, (\tilde{\mathbf{s}}_{k,|\mathcal{B}|}, k \in \mathcal{K})\}$$

$$\text{with } \tilde{\mathbf{s}}_{k,m} \in \mathcal{S}_k, k \in \mathcal{K},$$

$$\text{and } \mathcal{L}(\tilde{\mathbf{s}}_{k,m}, k \in \mathcal{K}, \boldsymbol{\lambda}^*) = \max_{\{\mathbf{s}_k \in \mathcal{S}_k, k \in \mathcal{K}\}} \mathcal{L}(\mathbf{s}_k, k \in \mathcal{K}, \boldsymbol{\lambda}^*), \quad m \in \{1, \ldots, |\mathcal{B}|\}, \tag{8.32}$$

where the cardinality of set $\mathcal{B}$ is larger than 1, i.e. $|\mathcal{B}| > 1$. It is important to note that the elements of $\mathcal{B}$ are not necessarily solutions to (8.4), i.e. they do not necessarily satisfy the user total power constraints (8.3). However, there exists at least one element in set $\mathcal{B}$ that does satisfy the total power constraints [5]. In order to obtain convergence to a primal optimal solution for (8.4) in the case that $|\mathcal{B}| > 1$, the dual decomposition approach has to be extended with an extra procedure that chooses an element out of set $\mathcal{B}$ that satisfies the user total power constraints.

A simple example may be given to clarify this issue; suppose we have a DSL scenario consisting of two users ($N = 2$) and two tones ($K = 2$), where the channel matrices (direct and crosstalk components) and noise components for the two tones are the same, i.e. $\mathbf{H}_1 = \mathbf{H}_2$ and $\sigma_1^n = \sigma_2^n,\ n \in \mathcal{N}$, and the weights are also the same $w_1 = w_2$. Furthermore suppose the crosstalk components are very large. In this case, there will be only one user active on each tone [36]. Finally suppose that the optimal dual variables λ_1^*, λ_2^*, where $\lambda_1^* = \lambda_2^*$, are given and the total power constraints are $P^n \leq \text{ON}$, where ON is a fixed power level. For this setup there will be 4 possible solutions to (8.7), namely $\{s_1^1 = \text{ON}, s_2^1 = \text{ON}, s_1^2 = 0, s_2^2 = 0\}, \{s_1^1 = 0, s_2^1 = 0, s_1^2 = \text{ON}, s_2^2 = \text{ON}\}, \{s_1^1 =$

ON, $s_2^1 = 0, s_1^2 = 0, s_2^2 = \text{ON}\}, \{s_1^1 = 0, s_2^1 = \text{ON}, s_1^2 = \text{ON}, s_2^2 = 0\}$. Note that all these solutions correspond to exactly the same objective value but only the last two solutions are primal optimal solutions as they satisfy the user total power constraints. Typical DSM algorithm implementations, however, have a fixed exhaustive search order or iteration order over tones so that one of the two first solutions may be selected and, as a consequence, these algorithms will not provide the primal optimal solutions of (8.4). To obtain convergence to the optimal primal variables of (8.4) an extra procedure should be added to the dual decomposition approach.

Note that the above problem is practically only relevant when the phenomenon of non-unique globally optimal solutions $\mathbf{s}_k(\boldsymbol{\lambda}^*)$ occurs at many tones. This is the case for DSL scenarios that have a subset of strong symmetric crosstalkers with equal line lengths, i.e. lines that generate the same interference to their environment over multiple tones k, with equal weights w_n and user total power constraints $P^{n,\text{tot}}$. Here, we can have many adjacent tones with multiple globally optimal solutions, namely where only one of the subset of strong crosstalkers is active [36]. If no special care is taken when recovering the primal transmit powers, this can lead to extremely slow convergence or even no convergence at all for these scenarios. More specifically, a fixed exhaustive search order or iteration order in typical DSM algorithm implementations will choose the same strong crosstalker over all competing tones, instead of equally dividing the resources over the competing users.

To overcome this problem we propose a very simple, but effective, interleaving procedure, that can be combined with Algorithm 8.3. More specifically this solution consists of alternatingly on a per-tone basis, giving priority to the globally optimal solution that corresponds to a different active strong crosstalker of the symmetric subset. This interleaving procedure replaces line 6 of Algorithm 8.3 with the following:

$$\forall k : \begin{cases} \mathcal{C}_k & = \{\text{all globally optimal solutions } \tilde{\mathbf{s}}_\text{k}(\lambda) \text{ of (8.31) for given } \boldsymbol{\lambda}\}, \\ & = \{\mathcal{C}_k(1), \ldots, \mathcal{C}_k(|\mathcal{C}_k|)\}, \\ \text{index} & = \text{rem}(k, |\mathcal{C}_k|) + 1, \\ \mathbf{s}_k^{i+1} & = \mathcal{C}_k(\text{index}), \end{cases} \tag{8.33}$$

where 'rem$(k, |\mathcal{C}_k|)$' refers to the remainder after dividing k by $|\mathcal{C}_k|$. As the suggested solution requires that all globally optimal solutions in the first step of (8.33) actually be computed, it should be combined with algorithms for the per-tone nonconvex problem that indeed compute all these solutions such as OSB with a fixed order exhaustive search for all tones or a multiple starting point approach such as MS-DSB with a fixed iteration order for all tones.

In the simulation Section 8.6, it will be demonstrated how the usage of (8.33) significantly improves the robustness of the dual decomposition approach for cWRS.

Remark: The above mentioned non-uniqueness also has an impact on the Lipschitz continuity condition of the smoothed gradient. More specifically this condition reduces to [21]:

$$\| \sum_{k \in \mathcal{K}} \tilde{\mathbf{s}}_k(\boldsymbol{\lambda}) - \sum_{k \in \mathcal{K}} \tilde{\mathbf{s}}_k(\boldsymbol{\mu}) \|^2 \leq L_c \|\boldsymbol{\lambda} - \boldsymbol{\mu}\|^2 \text{ with } L_c < \infty \tag{8.34}$$

For the above two-user two-tone symmetric strong crosstalk example, this condition does not hold. This can be shown as follows. Let us compare two cases: (1) optimal dual variables $(\lambda_1^*, \lambda_2^* + \mu)$ corresponding to primal variables $\{s_1^1 = \text{ON}, s_2^1 = \text{ON}, s_1^2 = 0, s_2^2 = 0\}$, (2) optimal dual variables $(\lambda_1^* + \mu, \lambda_2^*)$ corresponding to primal variables $\{s_1^1 = 0, s_2^1 = 0, s_1^2 = \text{ON}, s_2^2 = \text{ON}\}$, where $\boldsymbol{\mu} \geq 0$. For very small μ these two cases have only slightly different dual variables but completely different primal variables. So a small change in Lagrange multipliers can lead to a large change in primal variables. This means that for these specific cases Lipschitz continuity (8.34) is not satisfied and so the convergence speed will be worse than $O(\frac{1}{\epsilon})$. However adding the interleaving trick alleviates this problem, as will be demonstrated in Section 8.6.

Remark: In [37], a randomized LP-based algorithm is proposed for recovering the primal variables from the dual optimum. This algorithm is however designed for the cWRS problem with extra FDMA constraint, which is a simpler problem that can be solved using polynomial time algorithms. Furthermore the algorithm assumes the application of time-sharing.

8.6 Simulation Results

In this section, simulation results are shown that compare the performance of the improved dual decomposition approach with respect to the subgradient based dual decomposition approach. More specifically, in Section 8.6.1 we demonstrate for a DSM algorithm based on iterative convex approximations (CA-DSB) the very fast convergence of the improved dual decomposition approach with respect to the subgradient approach with different stepsize tuning strategies. In Section 8.6.2 we demonstrate how the improved dual decomposition approach in combination with a direct DSM algorithm (MS-DSB) succeeds in providing much faster convergence than with the subgradient based dual decomposition approach. Furthermore the convergence improvement for the interleaving procedure presented in Section 8.5 is demonstrated.

The following parameter settings are used for the simulated DSL scenarios. The twisted pair lines have a diameter of 0.5 mm (24 AWG). The maximum per-user total transmit power is 11.5 dBm for the VDSL scenarios and 20.4 dBm

for the ADSL scenarios. The SNR gap Γ is 12.9 dB, corresponding to a coding gain of 3 dB, a noise margin of 6 dB, and a target symbol error probability of 10^{-7}. The tone spacing Δ_f is 4.3125 kHz. The DMT symbol rate f_s is 4 kHz. Furthermore the prox-function $d_k(\mathbf{s}_k) = \frac{1}{2}\|\mathbf{s}_k\|^2$ (with convexity parameter equal to 1) is used for all simulations, which is an appropriate prox-function for box constraints [23].

8.6.1 Convergence Speed Up for Iterative Convex Approximation based DSM

A first DSL scenario is shown in Figure 8.3. This is a so-called near-far scenario which is known to be challenging, where DSM can make a substantial difference. For this scenario, we compare the convergence behaviour for the improved approach for CA-DSB (Algorithm 8.1) and the standard subgradient based dual decomposition approach for CA-DSB with different stepsize updating strategies, where the stepsize is δ in (8.10). The first stepsize update strategy is one that is guaranteed to converge, namely $\delta = q/i$, where q is the initial stepsize and i is the iteration counter [5]. We will refer to this with 'decreasing step' with some value for q. The second stepsize update strategy is the fixed stepsize, namely $\delta = q$. Note that this scheme is not guaranteed to converge for large values of q. We will refer to this update strategy with 'fixed step'. Furthermore the target accuracy on the dual gap is specified as 0.5%, which has to be seen as an application requirement and not as a tuning parameter. Finally note that we use an iterative fixed point update approach with the same number of inner iterations to solve the per-tone problems for both the improved and the subgradient schemes. The results are shown in Figures 8.4 and 8.5. It can be observed that different initial stepsizes lead to a different convergence behaviour for the subgradient approaches, and this is generally difficult to tune. The subgradient scheme with decreasing stepsize is generally much slower in convergence. The subgradient approach with fixed stepsize is better but can become instable for large values of q as shown in Figure 8.5. Stepsize tuning is crucial for these schemes. In contrary, the improved scheme automatically tunes its stepsize and converges very rapidly in only 40 iterations. Finally note that the curve for the subgradient scheme with fixed step and $q = 2500$ initially has a steeper curve with respect to the improved scheme, but as it approaches the optimal dual value its slope decreases fast and it converges only after 90 iterations.

Finally, we remark that different values for the application accuracy ϵ, i.e. the upper bound on the dual gap, lead to a different number of iterations to converge to that accuracy. For Figure 8.4, we set ϵ to correspond to 0.5% accuracy, which is sufficiently accurate in practice. In Table 8.1, we also show the relation between the specified accuracy (which is actually defined by the application) and the number of iterations so as to converge to that accuracy

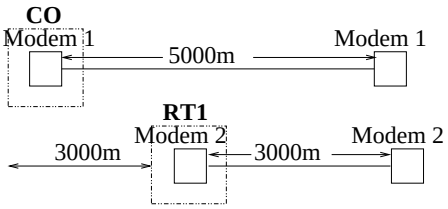


Figure 8.3: 2-user near-far ADSL downstream scenario

for the improved scheme. One can see that for very small specified accuracies the improved scheme requires a very small number of iterations to converge to that accuracy, e.g. in 84 iterations it converges to a specified accuracy of 0.1% of the optimum.

Table 8.1: The relation between the accuracy on the dual gap and the number of iterations so as to converge to that accuracy for the improved dual decomposition approach

Accuracy on dual gap	Number of Iterations
0.60%	22
0.39%	39
0.10%	84
0.05%	110
0.01%	250

8.6.2 Convergence Speed Up for Direct DSM

It was shown in [6] that for direct DSM algorithms the subgradient based dual decomposition approach with a particular stepsize selection procedure works well for ADSL scenarios, i.e. there are typically only 50-100 subgradient iterations needed to converge to the optimal dual variables. However for multi-user VDSL scenarios, which use a much larger frequency range and have to cope with significantly more crosstalk interference, existing subgradient approaches [5][6] are found to have significant convergence problems. We will focus on such VDSL scenarios and demonstrate how the improved approach succeeds in providing much faster convergence.

The different VDSL scenarios are shown in Figures 8.6, 8.7 and 8.9, i.e. four-user VDSL upstream, six-user VDSL upstream, and six-user VDSL upstream scenario with a subset of strong symmetric crosstalkers, respectively. The weights w_n are chosen equal for all users n, namely $w_n = 1/N$. Note that we used the multiple starting point procedure MS-DSB to solve the nonconvex

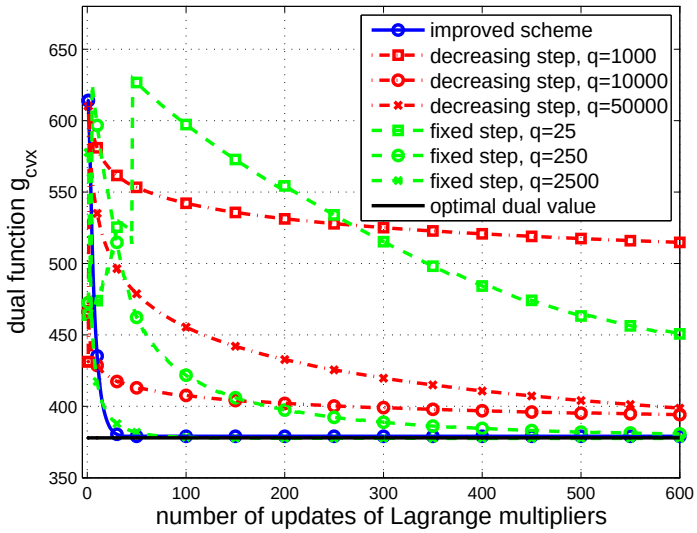


Figure 8.4: Comparison of convergence behaviour between subgradient dual decomposition approach, with different stepsize update strategies, and the improved dual decomposition approach, for CA-DSB

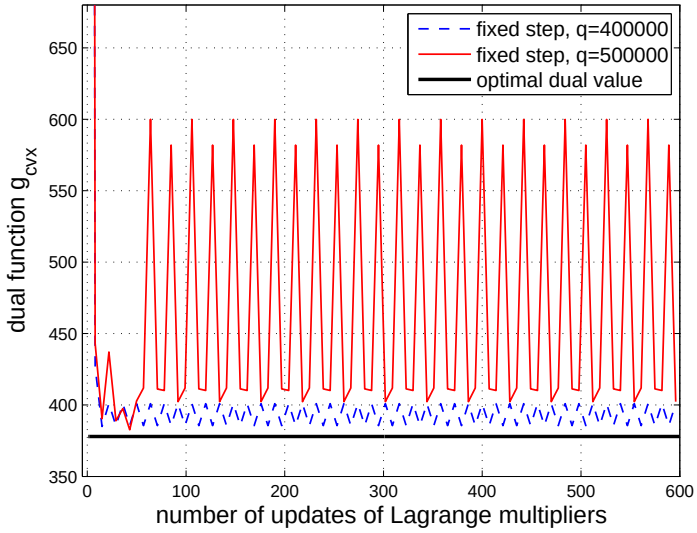


Figure 8.5: Convergence behaviour of subgradient dual decomposition approach, with large fixed stepsizes, for CA-DSB

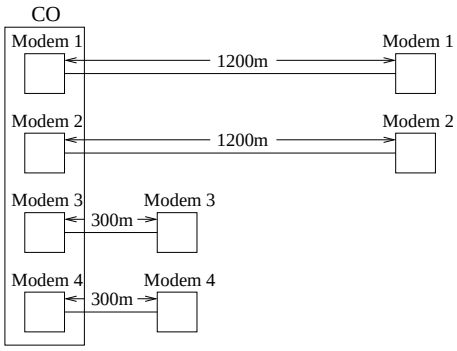


Figure 8.6: 4-user VDSL upstream scenario

per-tone problems for the subgradient based dual decomposition approach as well as the improved dual decomposition approach using (8.29). In [12] it was shown that this procedure provides globally optimal performance for practical ADSL and VDSL scenarios.

The first scenario, shown in Figure 8.6, is a four-user upstream VDSL scenario, consisting of two far-users with line length 1200 m and two near-users with line length 300 m. In the higher frequency range, there is a significant crosstalk coupling. This is a near-far scenario where spectrum management is crucial as to avoid significant performance degradation for the far-end users. Note that the near-end users form a subset of strong symmetric crosstalkers, in the high frequency range. As mentioned in Section 8.5, this can cause significant convergence problems for the dual decomposition approach. In fact, simulations show that the subgradient methods in [6] and [5] fail to converge to the dual variables, i.e. after 20000 iterations the complementarity conditions for some users are far from being satisfied. The main problem is that the stepsize selection procedure, which is a crucial component for fast convergence, is difficult to tune. For decreasing and fixed step sizes as proposed in [5], with different initial stepsizes, the procedure does not converge. For adaptive stepsizes, as proposed in [6], very small stepsizes are selected resulting in a very slow convergence ($>$ 20000 iterations). It is observed that for some users there is a fast convergence to the corresponding complementarity conditions whereas for other users convergence is very slow. The presence of the subset of strong symmetric crosstalkers, can lead to large changes in primal variables for small changes in dual variables, as discussed in Section 8.5, if stepsizes are not tuned carefully. The improved approach of Algorithm 8.3, in contrary, converges very fast to the optimal dual and primal variables. In only 100 iterations convergence is obtained, within an accuracy of 0.05%.

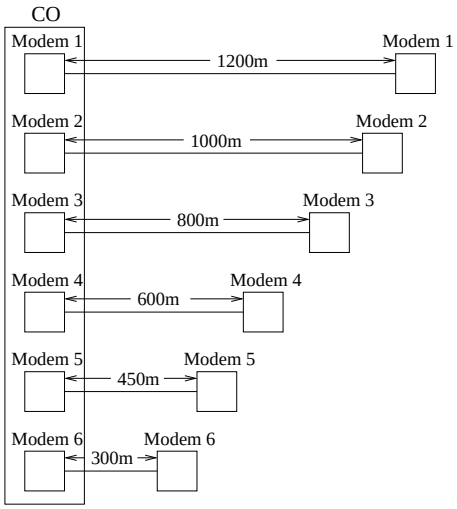


Figure 8.7: 6-user VDSL upstream scenario

The second VDSL upstream scenario, shown in Figure 8.7, consists of six users with different line lengths. Also for this large crosstalk scenario, the standard subgradient approaches [6][5] fail to converge to the optimal dual variables, i.e. after 10000 iterations the complementarity conditions are far from being satisfied. Similarly to the scenario of Figure 8.6, one can observe very different convergence behaviour for the different users to the corresponding complementarity conditions, where typically for a few users convergence is very slow. The improved dual decomposition approach however converges to the optimal dual and primal variables in only 150 iterations, within an accuracy of 0.05%. The optimal transmit powers are shown in Figure 8.8 for illustration.

The VDSL upstream scenario of Figure 8.9 consists of a six-line cable bundle with a subset of three strong symmetric crosstalkers, namely the set of lines with length 300m. The standard subgradient approaches [6] [5] fail to converge to the optimal dual variables. The presence of the strong symmetric crosstalkers significantly slows down the convergence, as it can lead to multiple globally optimal solutions for particular values of the dual variables. Here, stepsize selection is very crucial as a small change in dual variables can lead to a large change in primal variables, as also explained in Section 8.5. The improved dual decomposition approach converges to the optimal dual variables in only 150 iterations, but does not succeed in obtaining the primal optimal variables, because of the existence of multiple globally optimal solutions (i.e. optimal transmit powers) for optimal dual variables that do not satisfy the user total power constraints. More specifically for this scenario, for the obtained optimal dual variables, the obtained transmit powers jump

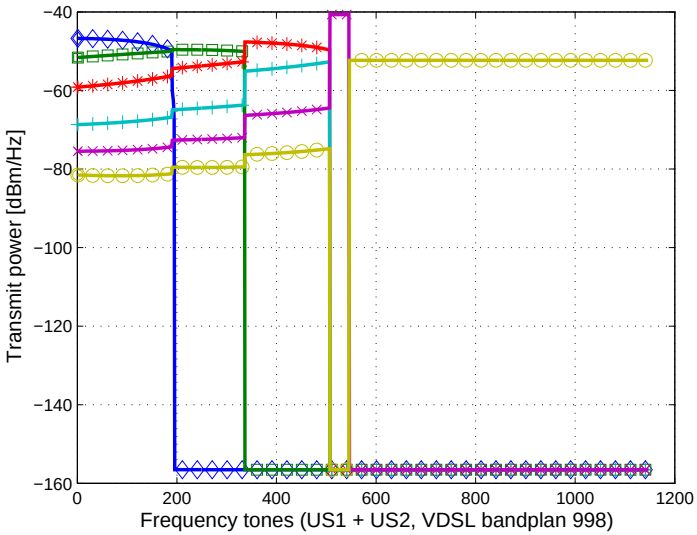


Figure 8.8: Optimal transmit powers for DSL scenario of Fig. 8.7 obtained using the improved dual decomposition approach. Blue diamond, green square, red asterisk, cyan plus, magenta cross and yellow circle curves correspond to transmit powers of users with line length 1200m, 1000m, 800m, 600m, 450m and 300m respectively.

to different solutions, with total powers $\{P^1, P^2, P^3\} = \{P^{1,\text{tot}}, P^{2,\text{tot}}, P^{3,\text{tot}}\}$, and $\{P^4, P^5, P^6\} \in \big\{\{3P^{\text{tot}}, A, A\}, \{A, 3P^{\text{tot}}, A\}, \{A, A, 3P^{\text{tot}}\}\big\}$, with A being very small. These primal solutions are shown in Figures 8.10, 8.11 and 8.12. One can observe that in the low and medium frequency range (used tones 1-727), the users with line lengths 1200 m, 900 m and 600 m are active. In this frequency range the strong crosstalkers will back-off and transmit at small similar transmit powers corresponding to a total power equal to A. However in the high frequency range (used tones 727-1147) where the users with line lengths 1200 m, 900 m and 600 are switched off, the three strong crosstalkers will compete, where only one user can be active in each tone k because of the significant crosstalk interference [36]. As explained in Section 8.5, typical DSM algorithm implementations will select the same active user for each of these tones, namely the user that corresponds to the smallest dual variable, where the dual variable can be seen as a penalty. So instead of dividing the total power over the three users equally, which would lead to a primal solution satisfying the per-user total power constraints, one user gets all power, leading to $P^n = 3P^{n,\text{tot}}$ for user n and $P^m = A$ for users $m \neq n$. Note that this prevents convergence to the optimal primal variables satisfying the per-user total power constraints.

However, when applying the proposed interleaving procedure (8.33), as proposed in Section 8.5, together with the improved dual decomposition approach, we can observe a very fast convergence both in primal and dual variables. Convergence is achieved in only 150 iterations, within an accuracy of 0.05%. The obtained optimal transmit powers are shown in Figure 8.13. In the frequency range between tone 728 and tone 1147, one can observe the interleaving effect. In Figure 8.14 this is zoomed in for tones 970 up to 975.

Remark: In the practical implementation the first step of the interleaving procedure is changed to 'all best solutions that are 99.9% close to each other'. This is to prevent that the procedure is only active when the dual variables are exactly the same. The overall effect of this is a negligible noise on the transmit powers as can be seen in Figure 8.13.

Remark: Note that applying the interleaving procedure combined with the improved dual decomposition approach for the scenarios in Figures 8.6 and 8.7, also leads to a faster convergence in both dual and primal variables.

8.7 Conclusion

Dynamic spectrum management has been recognized as a key technology to significantly improve the performance of DSL broadband access networks by mitigating the impact of crosstalk interference. Many existing DSM algorithms use a standard subgradient based dual decomposition approach to tackle the

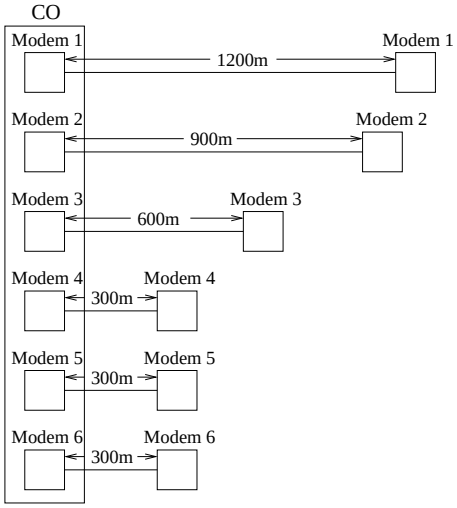


Figure 8.9: 6-user VDSL upstream scenario with subset of strong symmetric crosstalkers

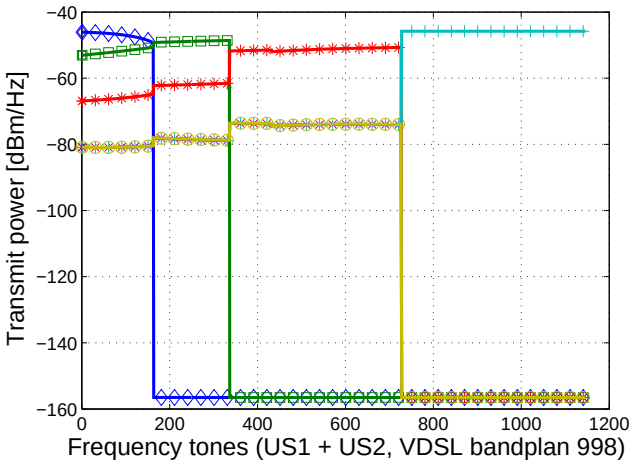


Figure 8.10: Optimal transmit power allocations for DSL scenario of Fig. 8.9 for optimal dual variables $\boldsymbol{\lambda}^*$, where $\{P^1, P^2, P^3, P^4, P^5, P^6\} = \{P^{1,\mathrm{tot}}, P^{2,\mathrm{tot}}, P^{3,\mathrm{tot}}, 3P^{4,\mathrm{tot}}, A, A\}$ with $A << P^{4,\mathrm{tot}}$. Blue diamond, green square, red asterisk, cyan plus, magenta cross, yellow circle curves correspond to transmit powers of users 1,2,3,4,5 and 6 respectively.

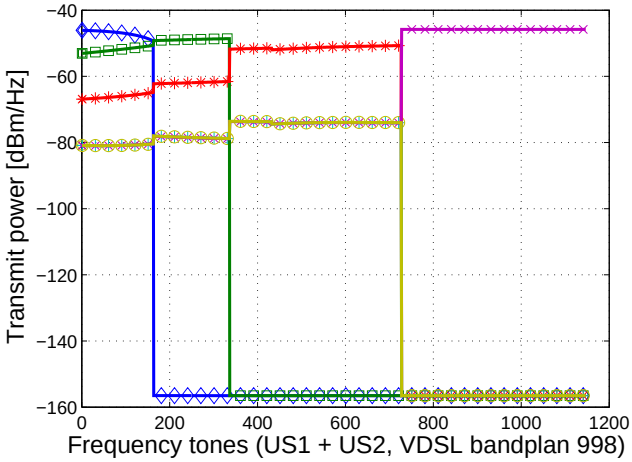


Figure 8.11: Optimal transmit power allocations for DSL scenario of Fig. 8.9 for optimal dual variables $\boldsymbol{\lambda}^*$, where $\{P^1, P^2, P^3, P^4, P^5, P^6\} = \{P^{1,\text{tot}}, P^{2,\text{tot}}, P^{3,\text{tot}}, A, 3P^{5,\text{tot}}, A\}$ with $A << P^{5,\text{tot}}$. Blue diamond, green square, red asterisk, cyan plus, magenta cross, yellow circle curves correspond to transmit powers of users 1,2,3,4,5 and 6 respectively.

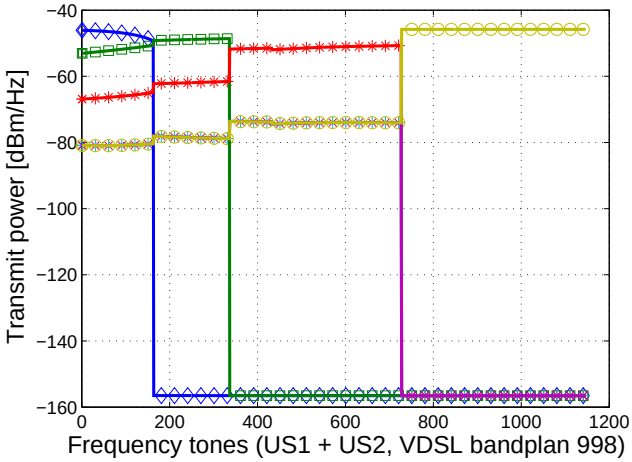


Figure 8.12: Optimal transmit power allocations for DSL scenario of Fig. 8.9 for optimal dual variables $\boldsymbol{\lambda}^*$, where $\{P^1, P^2, P^3, P^4, P^5, P^6\} = \{P^{1,\text{tot}}, P^{2,\text{tot}}, P^{3,\text{tot}}, A, A, 3P^{6,\text{tot}}\}$ with $A << P^{6,\text{tot}}$. Blue diamond, green square, red asterisk, cyan plus, magenta cross, yellow circle curves correspond to transmit powers of users 1,2,3,4,5 and 6 respectively.

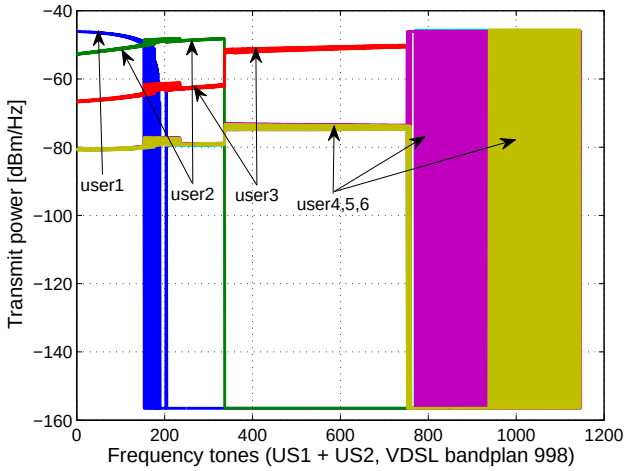


Figure 8.13: Transmit powers for scenario of DSL scenario of Fig. 8.9 obtained using improved dual decomposition approach with the interleaving procedure (8.33). Blue, green, red, cyan, magenta, yellow curves correspond to transmit powers of users 1,2,3,4,5 and 6 respectively.

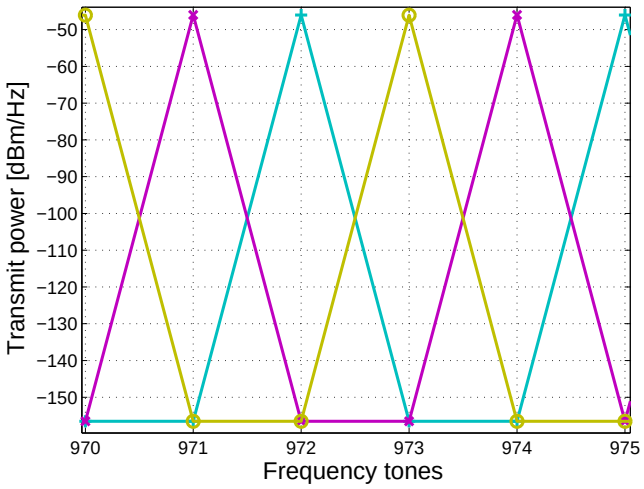


Figure 8.14: Zoom in on tones 970 to 975 of Fig. 8.13.

corresponding nonconvex optimization problems. However, this standard approach is often found to lead to extremely slow convergence or even no convergence at all. Especially for multiuser VDSL scenarios with subsets of strong symmetric crosstalkers significant convergence problems are observed because (1) the stepsize selection procedure of the subgradient updates is very critical, and (2) because special care must be taken when recovering the optimal transmit powers from the optimal dual solution. This paper proposes an improved dual decomposition approach, which consists of an optimal gradient based scheme with an automatic optimal stepsize selection removing the need for a tuning strategy. With this approach it is proven that the convergence of current state-of-the-art DSM algorithms based on iterative convex approximations (CA-DSB, SCALE), is improved by one order of magnitude. The improved dual decomposition approach is also extended for other DSM algorithms (OSB, ISB, ASB, (MS)-DSB, MIW). The addition of an extra interleaving procedure for recovering the optimal transmit powers from the dual optimal solution furthermore improves the convergence of the proposed approach. Simulation results demonstrate that significant convergence speed ups are obtained for practical DSL scenarios. The proposed improved dual decomposition approach makes an important step towards obtaining numerically fast and effective DSM algorithms.

Bibliography

[1] P. Tsiaflakis, I. Necoara, J. A. K. Suykens, and M. Moonen, "An improved dual decomposition approach to DSL dynamic spectrum management," in *Proc. of the 17th European Signal Processing Conference (EUSIPCO)*, Glasgow, Scotland, Aug. 2009.

[2] IE Market Research Corp. (http://www.iemarketresearch.com/), "Global DSL Subscribers and Access Revenues Forecast, 2009-2012: Global DSL subscribers to rise to 331 million in 2012 with DSL access revenues reaching $136.4 billion in 2012," Tech. Rep., Feb 2009, internet:http://www.researchandmarkets.com/reports/686876/.

[3] K. B. Song, S. T. Chung, G. Ginis, and J. M. Cioffi, "Dynamic spectrum management for next-generation DSL systems," *IEEE Communications Magazine*, vol. 40, no. 10, pp. 101–109, Oct. 2002.

[4] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 922–933, May 2006.

[5] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1310–1322, Jul 2006.

[6] P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "A low complexity optimal spectrum balancing algorithm for digital subscriber lines," *Signal Processing*, vol. 87, no. 7, pp. 1735–1753, Jul. 2007.

[7] P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen, "Green DSL: Energy-efficient DSM," in *Proc. of the IEEE International Conference on Communications (ICC)*, Jun. 2009.

[8] M. Wolkerstorfer, D. Statovci, and T. Nordstrom, "Dynamic spectrum management for energy-efficient transmission in DSL," in *Proc. of the 11th IEEE Singapore International Conference on Communication Systems (ICCS)*, Nov. 2008, pp. 1015–1020.

[9] J. M. Cioffi, S. Jagannathan, W. Lee, H. Zou, A. Chowdhery, W. Rhee, G. Ginis, and P. Silverman, "Greener copper with dynamic spectrum management," in *AccessNets, Las Vegas, NV, USA*, Oct. 2008.

[10] Z. Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, Feb. 2008.

[11] P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen, "Throughput and delay of DSL dynamic spectrum management with dynamic arrivals," in *Proc. of the IEEE Global Telecommunications Conference (GLOBECOM)*, Nov. 2008, pp. 1–5.

[12] P. Tsiaflakis, M. Diehl, and M. Moonen, "Distributed spectrum management algorithms for multiuser DSL networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4825–4843, Oct. 2008.

[13] R. Cendrillon and M. Moonen, "Iterative spectrum balancing for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 3, May 2005, pp. 1937–1941.

[14] Y. Xu, T. Le-Ngoc, and S. Panigrahi, "Global concave minimization for optimal spectrum balancing in multi-user DSL networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 2875–2885, Jul. 2008.

[15] J. Papandriopoulos and J. Evans, "SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Transactions on Information Theory*, vol. 55, no. 8, pp. 3711–3724, Aug. 2009.

[16] W. Yu, "Multiuser water-filling in the presence of crosstalk," in *Proc. of the Information Theory and Applications (ITA)*, Feb. 2007.

[17] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spectrum balancing for digital subscriber lines," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4241–4257, Aug. 2007.

[18] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[19] S. Jagannathan and J. M. Cioffi, "Distributed adaptive bit-loading for spectrum optimization in multi-user multicarrier systems," *Elsevier Physical Communication*, vol. 1, no. 1, pp. pp. 40–59, 2008.

[20] ——, "DSM level 2 band-preference using penalty tables," ATIS Contribution NIPP-NAI 2007-062, Minneapolis, Tech. Rep., May 2007, internet: http://jsumanth.googlepages.com/.

[21] I. Necoara and J. A. K. Suykens, "Application of a smoothing technique to decomposition in convex optimization," *IEEE Transactions on Automatic Control*, vol. 11, no. 53, pp. 2674–2679, 2008.

[22] Y. Nesterov, "Smooth minimization of non-smooth functions," *Mathematical Programming (A)*, vol. 103, no. 1, pp. 127–152, May 2005.

[23] ——, *Introductory Lectures on Convex Optimization: A Basic Course*, Kluwer, Ed. Boston, 2004.

[24] T. Starr, J. M. Cioffi, P. J. Silverman, *Understanding digital subscriber lines*. Prentice Hall, 1999.

[25] R. Lui and W. Yu, "Low-complexity near-optimal spectrum balancing for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 3, May 2005, pp. 1947–1951.

[26] W. Lee, Y. Kim, M. H. Brady, and J. M. Cioffi, "Band-preference dynamic spectrum management in a DSL environment," in *Proc. of the IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2006, pp. 1–5.

[27] M. Chiang, C. Tan, D. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.

[28] S. Kontogiorgis, R. D. Leone, and R. Meyer, "Alternating direction splitings for block angular parallel optimization," *Journal of Optimization Theory and Applications*, vol. 90, no. 1, pp. 1–29, 1996.

[29] G. Chen and M. Teboulle, "A proximal-based decomposition method for convex minimization problems," *Mathematical Programming (A)*, vol. 64, pp. 81–101, 1994.

[30] J. E. Spingarn, "Applications of the method of partial inverses to convex programming: decomposition," *Mathematical Programming (A)*, vol. 32, pp. 199–223, 1985.

[31] L. Xiao, M. Johansson, and S. Boyd, "Simultaneous routing and resource allocation via dual decomposition," *IEEE Transactions on Communications*, vol. 52, no. 7, pp. 1136–1144, Jul. 2004.

[32] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Athena Scientific, 1999.

[33] K. C. Kiwiel, "Approximations in bundle methods and decomposition of convex programs," *J. Optimiz. Theory Applicat.*, vol. 84, pp. 529–548, 1995.

[34] R. T. Rockafellar, "Augmented Lagrangians and applications of the proximal point algorithms in convex programming," *Math. Oper. Res.*, vol. 1, pp. 97–116, 1976.

[35] P. Tsiaflakis, C. Tan, Y. Yi, M. Chiang, and M. Moonen, "Optimality certificate of dynamic spectrum management in multi-carrier interference channels," in *Proc. of the IEEE International Symposium on Information Theory (ISIT)*, Jul. 2008, pp. 1298–1302.

[36] S. Hayashi and Z. Q. Luo, "Dynamic spectrum management: When is FDMA sum-rate optimal?" in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 3, Apr. 2007, pp. III–609 – III–612.

[37] Z.-Q. Luo and S. Zhang, "Duality gap estimation and polynomial time approximation for optimal spectrum management," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2675–2689, Jul. 2009.

Chapter 9

Cancellation Resource Optimization for Linear Partial Crosstalk Cancellation

A Dual Decomposition Approach to Partial Crosstalk Cancellation in a Multiuser DMT-xDSL Environment

J. Vangorp, P. Tsiaflakis, M. Moonen, J. Verlinden, and G. Ysebaert

Published in *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Jun. 2007, pp. 1-11., doi:10.1155/2007/37963

Copyright ©2007 Jan Vangorp et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Contributions of second author

- literature study
- co-development of problem statement for multi-user partial crosstalk cancellation
- co-development of dual decomposition based algorithm for partial crosstalk cancellation (single-user and multi-user case)
- co-analysis of computational complexity resource allocation algorithm
- co-analysis of computational complexity dual decomposition based algorithm (single-user and multi-user case)
- co-design of simulation scenarios
- software co-implementation and computer co-simulation
- co-analysis and co-interpretation of simulation results
- text co-redaction and co-editing

Abstract

In modern DSL systems, far-end crosstalk is a major source of performance degradation. Crosstalk cancellation schemes have been proposed to mitigate the effect of crosstalk. However, the complexity of crosstalk cancellation grows with the square of the number of lines in the binder. Fortunately, most of the crosstalk originates from a limited number of lines and, for DMT-based xDSL systems, on a limited number of tones. As a result, a fraction of the complexity of full crosstalk cancellation suffices to cancel most of the crosstalk. The challenge is then to determine which crosstalk to cancel on which tones, given a complexity constraint. This paper presents an algorithm based on a dual decomposition to optimally solve this problem. The proposed algorithm naturally incorporates rate constraints and the complexity of the algorithm compares favourably to a known resource allocation algorithm, where a multi-user extension is made to incorporate the rate constraints.

9.1 Introduction

Far-end crosstalk (FEXT), which is typically 10-15 dB larger than the background noise, is a major source of performance degradation in xDSL systems. One strategy for dealing with this crosstalk is crosstalk cancellation. Several crosstalk cancellation schemes have been proposed. Linear pre- and post filtering [1][2] requires coordination at both the transmitters and receivers. Successive interference cancellation or precompensation [3][4] can be used if there is only coordination available at the receivers or transmitters respectively, e.g. in the case of crosstalk cancellation in an upstream VDSL scenario. For this level of coordination, it is shown in [5][6] that a simple linear zero-forcing canceller or linear precompensator performs near-optimally in an xDSL environment.

Even for these simple linear cancellers, the complexity grows with the square of the number of lines. For example, in a binder of 8 VDSL lines transmitting on 4096 tones at a block rate of 4000 blocks per second, the runtime complexity of crosstalk cancellation exceeds 1 billion multiplications per second.

However, crosstalk exhibits space and tone selectivity [7]. Measurements show that most of the crosstalk originates from a limited number of lines, e.g. those in close proximity. Moreover, crosstalk coupling is heavily dependent on the frequency.

Because most of the crosstalk originates from a limited number of lines on a limited number of tones, a fraction of the complexity of full crosstalk cancellation suffices to cancel most of the crosstalk. This is called partial crosstalk cancellation [7][8].

The challenge in these upstream VDSL scenarios is then to determine for every user which crosstalk to cancel on which tones. In [7], an algorithm based on resource allocation is presented to solve this single-user problem. This paper presents an alternative optimal algorithm, based on a dual decomposition. The complexity of the algorithm is found to be more favourable than the complexity of the resource allocation algorithm, where a multi-user extension is made to incorporate rate constraints.

In Section 9.2, the partial crosstalk cancellation problem is presented and then solved following a dual decomposition approach. A number of observations is made to reduce the complexity without losing the optimality of the solution. In Section 9.3 the complexity of the single-user version of the dual decomposition algorithm is compared to the complexity of the resource allocation algorithm for the single-user case, where each user has an individual complexity constraint. Section 9.4 then extends these results to the multi-user case where all users share a complexity constraint. A search procedure is presented to dynamically distribute the available complexity for crosstalk cancellation according to the rate constraints. Section 9.5 provides some simulation results and finally Section 9.6 concludes the paper.

9.2 Dual Decomposition

9.2.1 System Model

Most current DSL systems use Discrete Multi-Tone (DMT) modulation. The available frequency band is divided in a number of parallel subchannels or tones. Each tone is capable of transmitting data independently from other tones, and so the transmit power and the number of bits can be assigned individually for each tone.

Transmission for a binder of N users can be modelled on each tone k by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k \qquad k = 1 \ldots K.$$

The vector $\mathbf{x}_k = [x_k^1, x_k^2, \ldots, x_k^N]^T$ contains the transmitted signals on tone k for all N users. $[\mathbf{H}_k]_{n,m} = h_k^{n,m}$ is an $N \times N$ matrix containing the channel transfer functions from transmitter m to receiver n. The diagonal elements are the direct channels, the off-diagonal elements are the crosstalk channels. $\mathbf{z}_k$ is the vector of additive noise on tone k, containing thermal noise, alien crosstalk, RFI,.... The vector $\mathbf{y}_k$ contains the received symbols.

The linear zero-forcing crosstalk canceller $\mathbf{W}$ cancels the crosstalk by making a linear combination of the received signals:

$$\hat{\mathbf{x}}_k = \mathbf{W}_k \mathbf{y}_k = \mathbf{W}_k \mathbf{H}_k \mathbf{x}_k + \mathbf{W}_k \mathbf{z}_k \qquad k = 1 \ldots K,$$

where $\mathbf{W}_k$ is chosen based on the zero-forcing criterion such that the equivalent channel $\mathbf{W}_k\mathbf{H}_k$ becomes an identity matrix. In [5][6] it is shown that, due to the characteristics of the xDSL channel, $\mathbf{W}$ exists and does not change the statistics of the noise. In the case of partial crosstalk cancellation $\mathbf{W}_k$ is chosen to be sparse [7], thereby saving on the number of calculations that is required, such that the resulting equivalent channel also becomes sparse.

In this paper, partial crosstalk cancellation is taken into account by introducing an equivalent channel $\tilde{\mathbf{H}}$. This is the same channel as the original channel $\mathbf{H}$, but with off-diagonal elements set to zero where the crosstalk is cancelled. If user n is cancelling crosstalk originating from user m on tone k, then $\tilde{h}_k^{n,m} = 0$.

We denote the transmit power as $s_k^n \triangleq \Delta_f E\{|x_k^n|^2\}$, the noise power as $\sigma_k^n \triangleq \Delta_f E\{|z_k^n|^2\}$. The DMT symbol rate is denoted as f_s, the tone spacing as Δ_f.

It is assumed that each modem treats interference from other modems as noise. When the number of interfering modems is large, the interference is well approximated by a Gaussian distribution. Under this assumption the achievable bit loading of user n on tone k, given the transmit spectra of all modems in the system and the crosstalk cancellation configuration, is

$$b_k^n \triangleq \log_2\left(1 + \frac{1}{\Gamma}\frac{|\tilde{h}_k^{n,n}|^2 s_k^n}{\sum_{m\neq n}|\tilde{h}_k^{n,m}|^2 s_k^m + \sigma_k^n}\right), \tag{9.1}$$

where Γ denotes the SNR-gap to capacity, which is function of the desired BER, the coding gain and noise margin. The data rate for user n is

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

When interference is being cancelled, the assumption of Gaussian noise becomes less valid. Under non-Gaussian noise, equation (9.1) gives a lower bound on the capacity of the channel. However, it remains the best model available for the achievable bitrate.

9.2.2 Partial Crosstalk Cancellation Problem

Because of the runtime complexity of full crosstalk cancellation, only a limited amount of crosstalk can be cancelled. The cancellation of the crosstalk from one user on some tone is done by a cancellation tap. The number of cancellation taps that can be used is constrained by the *cancellation tap constraint* C^{tot} [9]. The partial crosstalk cancellation problem amounts to finding an optimal selection of which crosstalk to cancel, thereby maximizing the capacity of the network.

Secondly, there is a *rate constraint* $R^{n,target}$ for each user. Typically, service providers offer a number of profiles to guarantee a certain Quality of Service. The rate constraint then indicates a minimum data rate required by the user.

The allocation of cancellation taps in partial crosstalk cancellation then results in the following maximization problem:

$$\begin{aligned}
\max_{\mathbf{c}} \quad & \sum_{n=1}^{N} R^n \\
\text{subject to} \quad & C = \sum_{k=1}^{K}\sum_{m=1}^{N}\sum_{n=1}^{N} c_k^{n,m} \leq C^{tot} \\
& R^n \geq R^{n,target} \qquad n = 1 \ldots N \\
\text{with} \quad & [\mathbf{c}_k]_{n,m} = c_k^{n,m} \qquad c_k^{n,m} = \begin{cases} 0 & \Rightarrow \quad \tilde{h}_k^{n,m} = h_k^{n,m} \\ 1 & \Rightarrow \quad \tilde{h}_k^{n,m} = 0 \end{cases}
\end{aligned} \tag{9.2}$$

where $\mathbf{c} = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_K]$. $c_k^{n,m} = 1$ indicates that a cancellation tap is assigned on tone k for cancelling crosstalk on line n originating from line m.

To find the global optimum for this optimization problem, one has to exhaustively search through all possible cancellation tap configurations $\mathbf{c}$. Because the cancellation tap constraint and the rate constraints are coupled over the tones, this results in an exponential complexity in the number of tones. By using a dual decomposition this complexity can be made linear [10][11][12][13][9]. This is done by using Lagrange multipliers to move the constraints coupled over tones to the objective function of the optimization problem [10]:

$$\begin{aligned}
\mathbf{c}^{opt} := \quad & \operatorname*{argmax}_{\mathbf{c}} \quad \sum_{n=1}^{N} \omega_n R^n + \lambda\big(C^{tot} - \sum_{k=1}^{K}\sum_{m=1}^{N}\sum_{n=1}^{N} c_k^{n,m}\big) \\
& \text{subject to} \quad \lambda \geq 0 \\
& \qquad\qquad\quad \omega_n \geq 0, \qquad n = 1 \ldots N,
\end{aligned} \tag{9.3}$$

where λ and ω_n are Lagrange multipliers. For a given set of λ and $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_N]^T$, (9.3) is a maximization of a sum over tones that can be performed by maximizing each tone individually. The optimization problem can then be solved in a per-tone fashion:

$$\begin{aligned}
& \text{for } k = 1 \ldots K, \\
\mathbf{c}_k^{opt} := \quad & \operatorname*{argmax}_{\mathbf{c}_k} \quad \sum_{n=1}^{N} \omega_n f_s b_k^n - \sum_{n=1}^{N}\sum_{m=1}^{N} \lambda c_k^{n,m} \\
& \text{subject to} \quad \lambda \geq 0 \\
& \qquad\qquad\quad \omega_n \geq 0, \qquad n = 1 \ldots N.
\end{aligned} \tag{9.4}$$

Maximization of (9.4) for given Lagrange multipliers can be performed by an exhaustive search. For each tone, all possible combinations for the cancellation

taps of the users should be checked. The combination giving the largest value for this expression is the optimal allocation of canceller taps for this tone.

The constraints can be enforced by choosing appropriate values for the Lagrange multipliers. The λ can be viewed as a cost for crosstalk cancellation taps. Larger values for the Lagrange multiplier result in less cancellation taps being allocated. The data rates of the users are weighted by $\boldsymbol{\omega}$, thereby giving more importance to some users. In this way, all possible trade-offs can be made to enforce the data rate constraints.

To solve (9.2) by (9.4), $\boldsymbol{\omega}$ and λ should be tuned to enforce the constraints. In [10][11], an efficient Lagrange multiplier search procedure is presented for a similar problem. This procedure can be easily adapted for this partial cancellation problem. The basis for this procedure is relation (9.5), which is proven in the Appendix:

$$\left[\begin{array}{cc} -(\Delta\boldsymbol{\omega})^T & \Delta\lambda \end{array} \right] \left[\begin{array}{c} \Delta\mathbf{R} \\ \Delta C \end{array} \right] \leq 0 \tag{9.5}$$

$\mathbf{R} = [R^1, \ldots, R^N]^T$ is a vector with the data rates and C is the number of cancellation taps corresponding to the Lagrange multipliers at hand.

Following [10][11], relation (9.5) leads to the following update formula for the Lagrange multipliers:

$$\begin{aligned} \left[\begin{array}{c} \Delta\boldsymbol{\omega} \\ \Delta\lambda \end{array} \right] &= -\mu \left[\begin{array}{c} \mathbf{R} - \mathbf{R}^{target} \\ C^{tot} - C \end{array} \right] \\ \Rightarrow \left[\begin{array}{c} \boldsymbol{\omega} \\ \lambda \end{array} \right]^{t+1} &= \left(\left[\begin{array}{c} \boldsymbol{\omega} \\ \lambda \end{array} \right]^{t} - \mu \left[\begin{array}{c} \mathbf{R} - \mathbf{R}^{target} \\ C^{tot} - C \end{array} \right] \right)^{+} \end{aligned} \tag{9.6}$$

where $(x)^+$ means $max(0, x)$ and μ is a stepsize parameter. Note that all the Lagrange multipliers are updated in parallel. This update formula is used in Algorithm 9.1, adopted from [10], to converge to the Langrange multipliers that enforce the constraints.

The partial crosstalk cancellation problem (9.2) is a non-convex constrained optimization problem. Without dual decomposition, finding the global optimum requires an exhaustive search over all possible solutions. On a certain tone, a user has to decide which crosstalk of $N-1$ other users has to be cancelled. There are 2^{N-1} possibilities to do this. For N users and K tones, this results in a total complexity of $\mathcal{O}((2^{N-1})^{NK})$.

In [9] it is shown that when using a dual decomposition in multi-carrier systems, the duality gap is zero. Therefore the solution for the dual problem is also the solution for the primal problem.

The dual decomposition decouples the problem over the tones, therefore reducing the exponential complexity in the number of tones K to linear complexity: $\mathcal{O}(K(2^{N-1})^N)$. This amounts to K exhaustive searches of complexity

Algorithm 9.1 Lagrange multiplier search algorithm

```
while distance > tolerance do
  Θ = [ω, λ]^T = best [ω, λ]^T so far
  μ = 1
  while distance ≤ previousDistance do
    previousDistance = distance
    μ = μ × 2
    ΔΘ = [Δω, Δλ]^T = update formula (9.6)
    [R^{Θ+ΔΘ}, C^{Θ+ΔΘ}, c] = exhaustiveSearch(Θ + ΔΘ)
    distance = ‖[R^{Θ+ΔΘ} − R^{target}, C^{tot} − C^{Θ+ΔΘ}]^T‖
  end while
end while
```

$\mathcal{O}((2^{N-1})^N)$. For an 8 user VDSL system, the complexity is reduced from $2^{7\times8\times4096}$ to $4096 \times 2^{7\times8}$. This is an enormous reduction in complexity. Moreover, as shown in the next subsection, the complexity can be even further reduced by observing that many cancellation tap configurations can be eliminated in advance.

9.2.3 Per-tone Search Complexity Reduction

To determine the optimal allocation of crosstalk cancellation taps on a certain tone, all of the $(2^{N-1})^N \approx 2^{N^2}$ possible allocations have to be evaluated. Even for a limited number of users this becomes complex. Fortunately, many of these possibilities can be eliminated based on two observations: user independence and line selection.

- *User Independence:* All users have to decide on a crosstalk cancellation configuration. This leads to an exponential complexity in the number of users N. However, from (9.1) it can be seen that if user n allocates a crosstalk cancellation tap to cancel crosstalk caused by user m (i.e. $\tilde{h}_k^{n,m} = 0$) this only has an influence on the capacity of user n. This corresponds to a per-user decoupling of (9.4), leading to

$$\begin{aligned} &\text{for } k = 1 \dots K, \\ &\text{for } n = 1 \dots N, \\ &\mathbf{c}_k^{n,opt} := \underset{\mathbf{c}_k^n}{\text{argmax}} \quad \omega_n f_s b_k^n - \sum_{m=1}^{N} \lambda c_k^{n,m} \\ &\qquad\qquad \text{subject to} \quad \lambda \geq 0 \\ &\qquad\qquad\qquad\qquad\quad \omega_n \geq 0, \qquad n = 1 \dots N. \end{aligned} \tag{9.7}$$

  As a consequence, the exponential complexity in N is reduced to linear complexity. Instead of one large search over all users, there are N independent searches for the users. This observation results in the following

complexity reduction:

$$(2^{N-1})^N \rightarrow N(2^{N-1}) \tag{9.8}$$

- *Line Selection:* A user has to decide for $N-1$ other users whether or not to cancel the crosstalk originating from these other users. This leads to 2^{N-1} possible crosstalk cancellation configurations. However, from (9.1) it can be seen that to maximize the capacity, one should allocate crosstalk cancellation taps to cancel the users which are causing the largest crosstalk. Therefore, if n crosstalk cancellation taps are available, these should be used to cancel the n largest sources of crosstalk.
  As a consequence, the 2^{N-1} possibilities for crosstalk cancellation are reduced to N possibilities: cancel no crosstalker, cancel the strongest crosstalker, cancel the 2 strongest crosstalkers, ..., cancel all $N-1$ crosstalkers.

$$\begin{array}{l} \text{for } k = 1 \ldots K, \\ \text{for } n = 1 \ldots N, \\ \mathbf{c}_k^{n,opt} := \quad \underset{\mathbf{c}_k^n}{\text{argmax}} \quad \omega_n f_s b_k^n(r) - \lambda r \\ \qquad\qquad \text{subject to} \quad \lambda \geq 0 \\ \qquad\qquad\qquad\qquad\quad \omega_n \geq 0, \qquad n = 1 \ldots N \end{array} \tag{9.9}$$

  where $b(r)$ is the capacity when the r largest crosstalkers are cancelled.

When both observations are combined, N users independently have to choose one of N possible crosstalk cancellation configurations. This results in the following total complexity reduction:

$$(2^{N-1})^N \rightarrow NN \tag{9.10}$$

In a 8-user case, these observations reduce the number of crosstalk cancellation configurations to be evaluated from 2^{56} to 2^6. Note that despite drastic complexity reductions, the solution is still optimal.

9.3 Single-User Algorithms and Complexity Comparison

In this section the complexity of the algorithm based on dual decomposition is analyzed and compared to the complexity of the optimal resource allocation algorithm of [7]. The resource allocation algorithm is a single-user algorithm. Therefore, a single-user formulation of the dual decomposition algorithm is used for the complexity comparison. The results will then be extended to the multi-user case in Section 9.4.

9.3.1 Single-User Resource Allocation Algorithm

The resource allocation algorithm uses the average capacity increase per allocated crosstalk cancellation tap on a certain tone:

$$v_k(r) = \frac{b_k(r) - b_k(0)}{r}, \tag{9.11}$$

with $b_k(r)$ the capacity on tone k when the r largest crosstalkers are cancelled (cfr. Section 9.2.3, line selection). A greedy algorithm then selects the tone k and number of crosstalkers r to cancel by searching the largest value of $v_k(r)$. The average capacity increase per allocated crosstalk cancellation tap should then be recalculated on tone k_s, based on the selected value $v_{k_s}(r_s)$, as follows:

- the average capacity increase for allocating less or equal crosstalk cancellation taps than r_s is set to zero
- the average capacity increase for allocating more crosstalk cancellation taps than r_s is recalculated as $v_k(r) = \frac{(b_k(r)-b_k(r_s))}{(r-r_s)}$, where the increase is now referenced to $b_k(r_s)$.

This is repeated until all available crosstalk cancellation taps are allocated. Note that in each iteration of the algorithm a minimum of 1 and a maximum of $N-1$ crosstalk cancellation taps is allocated. Because of this varying granularity, the crosstalk cancellation tap constraint cannot always be enforced tightly. However, the granularity is small enough to get close to the constraint.

The procedure is presented in Algorithm 9.2. A $K \times (N-1)$ table is initialized containing the average capacity increases per allocated crosstalk cancellation tap. For each of K tones the capacity increase has to be calculated for all $N-1$ crosstalk cancellation configurations. To be able to calculate the capacity increase, the capacity without crosstalk cancellation $b_k(0)$ also has to be calculated for every tone. This results in KN capacity calculations. Another $K(N-1)$ multiplications and additions are required to calculate the average capacity increase per allocated crosstalk cancellation tap. The $N-1$ crosstalk cancellation configurations are based on the line selection observation of Section 9.2.3. This requires a sort over the crosstalkers for each tone. This sort can be accomplished by selecting the crosstalkers one by one and placing them in the correct position of a sorted list. Because the resulting list is sorted at all times, a binary search can be used to find the correct position to place the current crosstalker. This results in a complexity of $\sum_{i=1}^{N-1} \log_2(i)$ comparisons to sort the list.

The table is then sorted to be able to efficiently find the maximum. This can be done analogous to the sorting of the crosstalkers and requires a complexity of $\sum_{i=1}^{K(N-1)} \log_2(i)$ comparisons.

Algorithm 9.2: Single-User Resource allocation algorithm

	capacities	multiplications	additions	comparisons
init: $v_k(r) = \frac{(b_k(r)-b_k(0))}{r}$ $\begin{cases} k=1\ldots K \\ r=1\ldots N-1 \end{cases}$	KN	$K(N-1)$	$K(N-1)$	$K\sum_{i=1}^{N-1}\log_2(i)$
sort $v_k(r)$	0	0	0	$\sum_{i=1}^{K(N-1)}\log_2(i)$
repeat				
$(k_s, r_s) = \text{argmax}_{k,r}\, v_k(r)$	0	0	0	0
$v_{k_s}(r) = 0, \quad \forall r \le r_s$	0	0	0	0
$v_{k_s}(r) = \frac{(b_k(r)-b_k(r_s))}{(r-r_s)}, \quad \forall r > r_s$	$\frac{N-1}{2}+1$	$\frac{N-1}{2}$	$N-1$	0
re-sort $v_k(r)$	0	0	0	$\sum_{i=K(N-1)-(\frac{N-1}{2}-1)}^{K(N-1)}\log_2(i)$
while $\sum_k r_k < C^{tot}$	0	0	1	1

Crosstalk cancellation taps can now be allocated by selecting the element with the maximum average capacity increase of the table, located at the top of the sorted list. On average, $\frac{N-1}{2}$ crosstalk cancellation taps are thereby allocated. $\frac{N-1}{2}$ elements in the table then have to be recalculated to the new reference capacity $b_k(r_s)$. This requires $\frac{N-1}{2}+1$ capacity calculations, $\frac{N-1}{2}$ multiplications and $N-1$ additions.

To keep the list sorted, $\frac{N-1}{2}$ binary searches are performed to find the new positions for the $\frac{N-1}{2}$ updated elements. This requires $\sum_{i=K(N-1)-(\frac{N-1}{2}-1)}^{K(N-1)} \log_2(i)$ comparisons. The number of currently allocated cancellation taps is updated and compared to the cancellation tap constraint C^{tot}.

This is repeated until all available crosstalk cancellation taps are allocated. In [7] it was shown that with a runtime complexity of 30% of full crosstalk cancellation, almost all crosstalk can be cancelled. This means that approximately $\frac{K(N-1)}{3}$ crosstalk cancellation taps have to be allocated. Taking into account that in each iteration of the algorithm $\frac{N-1}{2}$ taps are allocated, there are $\frac{K(N-1)}{3\frac{N-1}{2}}$ iterations required on average.

9.3.2 Single-User Dual Decomposition Algorithm

To be able to compare the algorithm based on dual decomposition to the resource allocation algorithm, a single-user formulation of the partial crosstalk cancellation problem (9.2) is used for user n:

$$\begin{aligned} \max_{\mathbf{c}} \quad & R^n \\ \text{subject to} \quad & C^n = \sum_{k=1}^{K}\sum_{m=1}^{N} c_k^{n,m} \leq C^{n,tot} \\ \text{with} \quad & [\mathbf{c}_k]_{n,m} = c_k^{n,m}, \qquad c_k^{n,m} = \begin{cases} 0 & \Rightarrow \quad \tilde{h}_k^{n,m} = h_k^{n,m} \\ 1 & \Rightarrow \quad \tilde{h}_k^{n,m} = 0 \end{cases} \end{aligned} \tag{9.12}$$

This results in the following dual problem which is decoupled over the tones:

$$\begin{aligned} &\text{for } k = 1 \ldots K, \\ &\mathbf{c}_k^{opt} := \quad \underset{\mathbf{c}_k}{\text{argmax}} \quad b_k^n - \sum_{m=1}^{N} \lambda c_k^{n,m} \\ &\qquad\qquad \text{subject to} \quad \lambda \geq 0. \end{aligned} \tag{9.13}$$

This can be viewed as one optimization of the multi-user problem where all users are allocated a crosstalk cancellation tap budget in advance.

Algorithm 9.3 presents the single-user dual decomposition algorithm. It starts by initializing a $K \times N$ table of capacities for K tones and N possible crosstalk cancellation configurations. To obtain the N possible crosstalk cancellation

configurations, the line selection observation of Section 9.2.3 is used. This requires sorting the crosstalkers which uses $K\sum_{i=1}^{N-1}\log_2(i)$ comparisons.

The algorithm then starts from some initial λ and performs K per-tone exhaustive searches. There are N possible values for λr, which can be calculated in advance. This requires N multiplications. These precalculated values are then subtracted from the corresponding elements of the $K \times N$ table. Finally, K exhaustive searches of N values are performed to obtain the maximum on each tone. This requires $K(N-1)$ comparisons.

The cancellation tap constraint is then checked by summing the number of taps allocated on each tone. If the constraint is not tightly satisfied, the Lagrange multiplier λ is updated and then the per-tone search is repeated. Because there is only one Lagrange multiplier, bisection can be used. This requires typically 10 iterations.

Table 9.1 summarizes the total complexity of the single-user resource allocation algorithm and the dual decomposition algorithm.

Figure 9.1 shows the initialization complexity as a function of the number of users for the single-user resource allocation algorithm and the dual decomposition algorithm for $K = 1000$. It is taken into account that a capacity calculation in an N-user system roughly takes $N + 2$ multiplications and N additions. Assuming the remaining 3 operations (multiplication, addition and comparison) are equally resource consuming, one can see an 18% complexity reduction in the 20-user case.

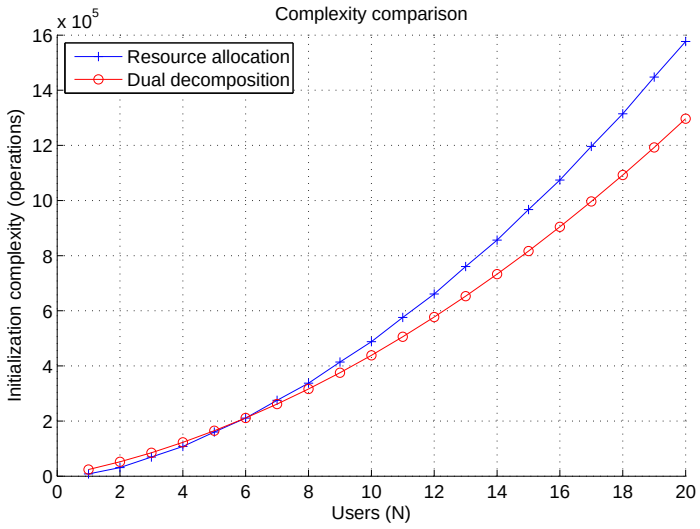


Figure 9.1: Complexity comparison single-user algorithms

Algorithm 9.3: Single-User Dual decomposition algorithm

	capacities	multiplications	additions	comparisons
init: $b_k(r)$ $\begin{cases} k=1\ldots K \\ r=0\ldots N-1 \end{cases}$	KN	0	0	$K\sum_{i=1}^{N-1}\log_2(i)$
repeat				
for $k=1\ldots K$				
$\mathbf{c}_k^{opt} = \text{argmax}_r\, b_k(r) - \lambda r$	0	N	KN	$K(N-1)$
endfor				
update λ based on (9.6)				
while $\sum_k \mathbf{c}_k^{opt} \neq C^{tot}$	0	0	$K-1$	1

Table 9.1: Complexity comparison single-user algorithms

	resource allocation	dual decomposition
capacities	$KN + \frac{K(N-1)}{3^{\frac{N-1}{2}}}\left(\frac{N-1}{2}+1\right)$	KN
multiplications	$K(N-1) + \frac{K(N-1)}{3^{\frac{N-1}{2}}}\frac{N-1}{2}$	$10 \times N$
additions	$K(N-1) + \frac{K(N-1)}{3^{\frac{N-1}{2}}}N$	$10 \times (KN + K - 1)$
comparisons	$K\sum_{i=1}^{N-1}\log_2(i) + \sum_{i=1}^{K(N-1)}\log_2(i) + \frac{K(N-1)}{3^{\frac{N-1}{2}}}\left(1 + \sum_{i=K(N-1)-(\frac{N-1}{2}-1)}^{K(N-1)}\log_2(i)\right)$	$K\sum_{i=1}^{N-1}\log_2(i) + 10 \times \big(K(N-1)+1\big)$

9.4 Multi-User Algorithms and Complexity Comparison

The extension to the multi-user case can be made by dividing the cancellation tap budget over the users in advance. By varying the cancellation tap budget allocated to each user, various trade-offs can be made in the data rates. This reduces the problem to multiple single-user problems. The core complexity of both the resource allocation algorithm and the dual decomposition algorithm is then increased by a factor N. Because of user independence and fixed individual cancellation tap budgets, optimization of the individual users also results in the optimization of the sum rate.

In this section, the single-user algorithms are extended to automatically determine the correct proportions of the cancellation tap budget to be allocated to the users such that the rate constraints are satisfied.

9.4.1 Multi-user Resource Allocation Algorithm

For the resource allocation algorithm in [7], no procedure is available to automatically distribute the cancellation tap budget over the users so that certain data rate constraints are satisfied. However, by introducing weights ω_n, some lines can be emphasized to meet the rate constraints. To achieve a higher data rate for a user, more crosstalk cancellation taps should be allocated to that user. In order to do this, the average benefit of adding a crosstalk cancellation tap for that user is increased by a factor ω_n. A larger weight leads to more crosstalk cancellation taps allocated and thus a higher data rate.

A given set of $\omega_n{}'$s implies a cancellation tap budget for each user (which is known after the optimization is done with these $\omega_n{}'$s). Because of the user independence, this again leads to an optimization of the sum rate. However, the rates are now weighted with $\omega_n{}'$s, thus a weighted rate sum is optimized.

Therefore, the following relation can be derived, analogous to the derivation in the appendix:

$$\Delta\boldsymbol{\omega}\Delta\mathbf{R} \geq 0 \tag{9.14}$$

This is a reduced form of (9.5), which leads to a simplified version of the update formula (9.6):

$$\Delta\boldsymbol{\omega} = -\mu(\mathbf{R} - \mathbf{R}^{target}) \Rightarrow \boldsymbol{\omega}^{t+1} = \left[\boldsymbol{\omega}^{t} - \mu(\mathbf{R} - \mathbf{R}^{target})\right]^{+} \tag{9.15}$$

During I iterations, this update formula can then be used to steer the $\omega_n{}'$s so that the rate constraints are satisfied.

Algorithm 9.4 presents the resulting multi-user resource allocation algorithm with its associated complexities. Note that the table of $KN(N-1)$ average ca-

pacity increases per crosstalk cancellation tap is now globally searched instead of individually per user.

9.4.2 Multi-user Dual Decomposition Algorithm

In the dual decomposition approach, Algorithm 9.1 can be used to find an appropriate distribution of the cancellation tap budget over the users, where the per-tone search is simplified based on the observations in Section 9.2.3. The resulting algorithm and complexities are shown in Algorithm 9.5. Because the updates of the Lagrange multipliers are based on the same update formula as in the resource allocation algorithm, roughly the same number of I iterations is required to enforce the constraints.

In Table 9.2 the total complexities of the multi-user resource allocation algorithm and the multi-user dual decomposition algorithm are compared.

Figure 9.2 shows the initialization complexity as function of the number of users for the resource allocation algorithm and the dual decomposition algorithm for $K = 1000$, under the assumption that $I = 50$ iterations are required to enforce the constraints. It is taken into account that a capacity calculation in an N-user system roughly takes $N + 2$ multiplications and N additions. Assuming the remaining 3 operations (multiplication, addition and comparison) are equally resource consuming, one can see an 88% complexity reduction in the 20-user case.

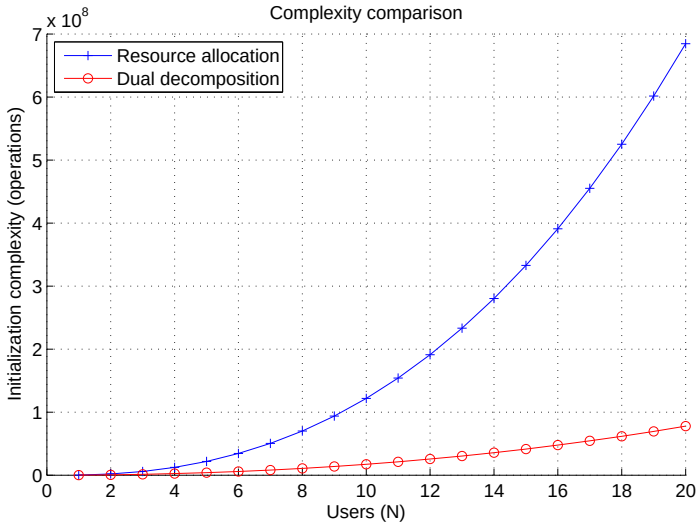


Figure 9.2: Complexity comparison multi-user algorithms

Algorithm 9.4: Multi-user resource allocation algorithm

	capacities	multiplications	additions	comparisons
init: $v_k^n(r) = \frac{(b_k^n(r) - b_k^n(0))}{r}$ $\begin{cases} k{=}1\ldots K \\ r{=}1\ldots N{-}1 \\ n{=}1\ldots N \end{cases}$	KNN	$KN(N-1)$	$KN(N-1)$	$KN\sum_{i=1}^{N-1}\log_2(i)$
repeat				
$v_k^{\boldsymbol{\omega},n}(r) = \omega_n v_k^n(r)$	0	$KN(N-1)$	0	0
sort $v_k^{\boldsymbol{\omega},n}(r)$	0	0	0	$\sum_{i=1}^{KN(N-1)}\log_2(i)$
repeat				
$(k_s, r_s, n_s) = \text{argmax}_{k,r,n}\, v_k^{\boldsymbol{\omega},n}(r)$	0	0	0	0
$v_{k_s}^{\boldsymbol{\omega},n_s}(r) = 0, \quad \forall r \le r_s$	0	0	0	0
$v_{k_s}^{\boldsymbol{\omega},n_s}(r) = \omega_{n_s}\frac{(b_k^{n_s}(r) - b_k^{n_s}(r_s))}{(r-r_s)}, \quad \forall r > r_s$	$\frac{N-1}{2}+1$	$N-1$	$N-1$	0
re-sort $v_k^{\boldsymbol{\omega},n}(r)$	0	0	0	$\sum_{i=KN(N-1)-(\frac{N-1}{2}-1)}^{KN(N-1)}\log_2(i)$
while $\sum_{n=1}^{N}\sum_{k=1}^{K} r_k^n < C^{tot}$	0	0	1	1
update $\boldsymbol{\omega}$ based on (9.6)				
while rate constraints not satisfied				

Algorithm 9.5: Multi-user dual decomposition algorithm

	capacities	multiplications	additions	comparisons
init: $b_k^n(r)$ $\begin{cases} k{=}1\ldots K \\ r{=}0\ldots N{-}1 \\ n{=}1\ldots N \end{cases}$	KNN	0	0	$KN\sum_{i=1}^{N-1}\log_2(i)$
repeat				
for $k = 1 \ldots K$				
for $n = 1 \ldots N$				
$\mathbf{c}_k^{n,opt} = \text{argmax}_r\, \omega_n b_k^n(r) - \lambda r$	0	$N + KNN$	KNN	$KN(N-1)$
endfor				
endfor				
update $\boldsymbol{\omega}, \lambda$ based on (9.6)				
while $\sum_{n=1}^{N}\sum_{k=1}^{K}\mathbf{c}_k^{n,opt} \neq C^{tot}$ and rate constraints not satisfied	0	0	$(N-1)(K-1)$	1

Table 9.2: Complexity comparison multi-user algorithms

	resource allocation	dual decomposition
capacities	$KNN + I \times \frac{KN(N-1)}{3^{\frac{N-1}{2}}}\left(\frac{N-1}{2}+1\right)$	KNN
multiplications	$KN(N-1) + I \times \left(KN(N-1) + \frac{KN(N-1)}{3^{\frac{N-1}{2}}}(N-1)\right)$	$I \times (N + KNN)$
additions	$KN(N-1) + I \times \frac{KN(N-1)}{3^{\frac{N-1}{2}}}N$	$I \times \big(KNN + (N-1)(K-1)\big)$
comparisons	$KN\sum_{i=1}^{N-1}\log_2(i) + I \times \sum_{i=1}^{KN(N-1)}\log_2(i) + \frac{KN(N-1)}{3^{\frac{N-1}{2}}}\Big(1 + \sum_{i=KN(N-1)-(\frac{N-1}{2}-1)}^{KN(N-1)}\log_2(i)\Big)$	$KN\sum_{i=1}^{N-1}\log_2(i) + I \times \big(KN(N-1)+1\big)$

9.5 Simulation Results

In [7] a simplified joint line/tone selection algorithm is also presented. This algorithm has a much lower complexity than the algorithms discussed in this paper and is claimed to be near-optimal. This algorithm can also be extended to the multi-user case by introducing the weights ω. However, this near-optimality largely depends on the scenario. For simple scenarios with only two different line lengths, the simplified joint line/tone selection algorithm indeed performs near-optimal. However, for practical scenarios with lines of varying lengths, this simplified algorithm can be suboptimal depending on the runtime complexity that is allowed.

In figure 9.3 the performance of both the optimal as well as the simplified line/tone selection is presented for different runtime complexities. This is done for an 8-user upstream VDSL scenario, with line lengths varying from 150 m to 1200 m in 150 m intervals. An empirical channel model [14] is used with line diameter of 0.5 mm (24 AWG) that generates both the direct channels and the crosstalk channels. The transmit power is set to -60 dBm on all tones. The SNR gap Γ is set to 12.9 dB, corresponding to a target symbol error probability of 10^{-7}, coding gain of 3 dB and a noise margin of 6 dB. The tone spacing $\Delta_f = 4.3125$ kHz and the DMT symbol rate $f_s = 4$ kHz.

To allow for an easier comparison, cancellation taps are allocated to each line using a single-user algorithm, keeping all other lines at a fixed bitrate with no crosstalk cancellation. Note that for small runtime complexities, the optimal joint line/tone selection algorithm can increase bitrates up to 50% of the performance of the simplified joint line/tone selection algorithm. Especially for the far-end users, which should be protected most from crosstalk, this performance difference is large.

Secondly, note the difference in runtime complexity for different lines to approach the full crosstalk cancellation performance. For long lines, 30% of full crosstalk cancellation is sufficient because only few tones carry a significant amount of bits. As the lines get shorter, up to 50-60% of full crosstalk cancellation is necessary. Therefore, multi-user algorithms are more suitable to solve the partial crosstalk cancellation problem because they can automatically distribute the cancellation tap budget over the users, in contrast to single-user algorithms where the budget has to be distributed in advance, taking into account the different line lenghts.

The simplified joint line/tone selection algorithm requires a high runtime complexity before it starts performing optimal. For low runtime complexities however, the optimal algorithm reaches a much higher perfomance. Thus depending

on the allowed runtime complexity, the optimal joint line/tone algorithm can be preferred over the simplified algorithm, trading of runtime complexity for initialization complexity when the required bitrate is fixed.

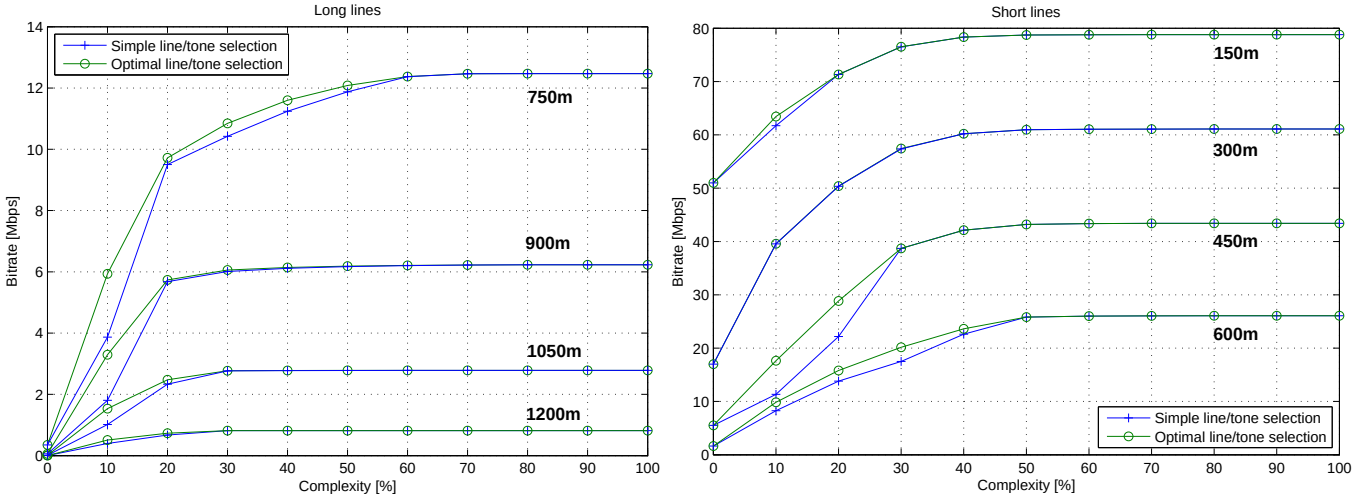
In figure 9.4, rate regions are shown for a symmetric upstream VDSL scenario with two 300m lines. Various crosstalk cancellation complexities are considered when allocating crosstalk cancellation taps optimally. One can see for e.g. a runtime complexity of 25% of the runtime complexity of full crosstalk cancellation that the available cancellation tap budget can be shifted between the users, thereby trading off the performance in terms of bitrate. If full priority is given to one user, only that user will gain the extra capacity due to the crosstalk cancellation. If the priority is divided over the users, both will gain some capacity. For small runtime complexities (almost no crosstalk can be cancelled) and large runtime complexities (all the largest crosstalk components can be cancelled) the trade off that can be made between the users is small.

9.6 Conclusion

In modern DSL systems, crosstalk is a major source of performance degradation. Crosstalk cancellation schemes have been proposed to mitigate the effect of crosstalk. However, the complexity of crosstalk cancellation grows with the square of the number of lines in the binder. Fortunately, most of the crosstalk originates from a limited number of lines on a limited number of tones. As a result, a fraction of the complexity of full crosstalk cancellation suffices to cancel most of the crosstalk, which is exploited by partial crosstalk cancellation. The challenge is then to determine which crosstalk to cancel on which tones, given a certain complexity constraint. In this paper we have presented an algorithm to optimally solve this problem, based on a dual decomposition.

Two cases were considered: single-user and multi-user. In the single-user case, each user has an individual cancellation tap budget to be allocated. It was shown that the dual decomposition algorithm has a favourable complexity compared to the optimal resource allocation algorithm.

In the multi-user case, all users have a common cancellation tap budget. This budget has to be distributed over the users in such a way that rate constraints are satisfied. The dual decomposition approach naturally incorporates these rate constraints. The resource allocation algorithms were extended to this multi-user case to also include these rate constraints. The extension allows for the same search procedure to be used to find the distribution of the cancellation tap budget over the users as used in the dual decomposition algorithm. Also in this multi-user case, the complexity of the dual decomposition algorithm was found to compare favourably with the complexity of the multi-user resource allocation algorithm.



(a) (b)

Figure 9.3: Performance comparison Optimal and Simple line/tone selection algorithms

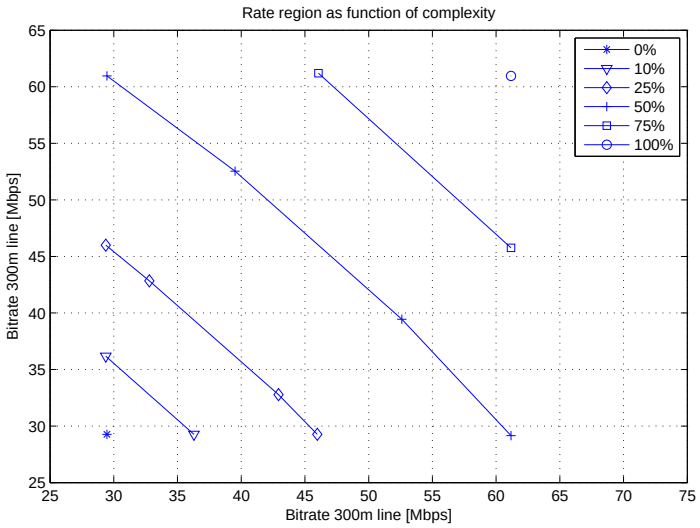


Figure 9.4: Rate regions for various crosstalk cancellation complexities

Appendix : Search Algorithm for the Lagrange Multipliers

The proof presented in [10][11] can be easily adapted for partial crosstalk cancellation. Assume a two-user scenario with signal-level control. Starting from two optimal solutions $(R^{1,\boldsymbol{\omega}_A,\lambda_A}, R^{2,\boldsymbol{\omega}_A,\lambda_A}, C^{\boldsymbol{\omega}_A,\lambda_A})$ and $(R^{1,\boldsymbol{\omega}_B,\lambda_B}, R^{2,\boldsymbol{\omega}_B,\lambda_B}, C^{\boldsymbol{\omega}_B,\lambda_B})$ corresponding to $(\boldsymbol{\omega}_A, \lambda_A)$ and $(\boldsymbol{\omega}_B, \lambda_B)$ respectively, optimality for $(\boldsymbol{\omega}_A, \lambda_A)$ implies

$$\begin{aligned} & \omega_{1,A} R^{1,\boldsymbol{\omega}_B,\lambda_B} + \omega_{2,A} R^{2,\boldsymbol{\omega}_B,\lambda_B} - \lambda_A C^{\boldsymbol{\omega}_B,\lambda_B} \\ \leq\ & \omega_{1,A} R^{1,\boldsymbol{\omega}_A,\lambda_A} + \omega_{2,A} R^{2,\boldsymbol{\omega}_A,\lambda_A} - \lambda_A C^{\boldsymbol{\omega}_A,\lambda_A} \end{aligned} \tag{9.16}$$

Optimality for $(\boldsymbol{\omega}_B, \lambda_B)$ implies

$$\begin{aligned} & \omega_{1,B} R^{1,\boldsymbol{\omega}_A,\lambda_A} + \omega_{2,B} R^{2,\boldsymbol{\omega}_A,\lambda_A} - \lambda_B C^{\boldsymbol{\omega}_A,\lambda_A} \\ \leq\ & \omega_{1,B} R^{1,\boldsymbol{\omega}_B,\lambda_B} + \omega_{2,B} R^{2,\boldsymbol{\omega}_B,\lambda_B} - \lambda_B C^{\boldsymbol{\omega}_B,\lambda_B} \end{aligned} \tag{9.17}$$

Taking the sum of (9.16) and (9.17) results in

$$\begin{aligned} &- \underbrace{\left(\omega_{1,B} - \omega_{1,A}\right)}_{\Delta\omega_1} \underbrace{\left(R^{1,\boldsymbol{\omega}_B,\lambda_B} - R^{1,\boldsymbol{\omega}_A,\lambda_A}\right)}_{\Delta R^1} \\ &- \underbrace{\left(\omega_{2,B} - \omega_{2,A}\right)}_{\Delta\omega_2} \underbrace{\left(R^{2,\boldsymbol{\omega}_B,\lambda_B} - R^{2,\boldsymbol{\omega}_A,\lambda_A}\right)}_{\Delta R^2} \\ &+ \underbrace{\left(\lambda_B - \lambda_A\right)}_{\Delta\lambda} \underbrace{\left(C^{\boldsymbol{\omega}_B,\lambda_B} - C^{\boldsymbol{\omega}_A,\lambda_A}\right)}_{\Delta C} \leq 0 \end{aligned} \tag{9.18}$$

Relation (9.18) is straightforwardly extended to a multi-user scenario:

$$\left[\begin{array}{cc} -(\Delta\boldsymbol{\omega})^T & \Delta\lambda \end{array}\right] \left[\begin{array}{c} \Delta\mathbf{R} \\ \Delta C \end{array}\right] \leq 0 \tag{9.19}$$

$\boldsymbol{\omega} = [\omega_1, \ldots, \omega_N]$ is a vector containing the Lagrange multipliers for the weights for the users, λ is the Lagrange multiplier controlling the number of cancellation taps used. $\mathbf{R} = [R^1, \ldots, R^N]^T$ is a vector with the corresponding data rates and C is the corresponding number of cancellation taps.

Bibliography

[1] G. Taubock and W. Henkel, "MIMO systems in the subscriber-line network," in *Proc. of the 5th International OFDM Workshop*, Hamburg, Germany, Sep. 2000, pp. 18.1–18.3.

[2] R. Cendrillon, M. Moonen, R. Suciu, and G. Ginis, "Simplified power allocation and TX/RX structure for MIMO-DSL," in *IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 4, San Francisco, California, USA, Dec. 2003, pp. 1842–1846.

[3] G. Ginis and J. M. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1085–1104, Jun. 2002.

[4] W. Yu and J. M. Cioffi, "Multiuser detection in vector multiple access channels using generalized decision feedback equalization," in *Proc. of the 5th International Conference on Signal Processing*, Beijing, China, Aug. 2000.

[5] R. Cendrillon, M. Moonen, E. V. den Bogaert, and G. Ginis, "The linear zero-forcing crosstalk canceller is near-optimal in DSL channels," in *Proc. of the IEEE Global Communications Conference (GLOBECOM)*, vol. 4, Dec. 2004, pp. 2334–2338.

[6] R. Cendrillon, G. Ginis, and M. Moonen, "Improved linear crosstalk precompensation for downstream VDSL," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 4, Montreal, May 2004, pp. 1053–1056.

[7] R. Cendrillon, M. Moonen, G. Ginis, K. V. Acker, T. Bostoen, and P. Vandaele, "Partial crosstalk cancellation for upstream VDSL," *EURASIP Journal on Applied Signal Processing*, vol. 2004, no. 10, pp. 1520–1535, Aug. 2004.

[8] ——, "Partial crosstalk precompensation for downstream VDSL," *Signal Processing*, pp. 2005–2019, Nov. 2004.

[9] W. Yu, R. Lui, and R. Cendrillon, "Dual optimization methods for multiuser orthogonal frequency-division multiplex systems," in *IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 1, Dec. 2004, pp. 225–229.

[10] P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "A low complexity optimal spectrum balancing algorithm for digital subscriber lines," *Signal Processing*, vol. 87, no. 7, pp. 1735–1753, Jul. 2007.

[11] P. Tsiaflakis, J. Vangorp, M. Moonen, J. Verlinden, and K. V. Acker, "An efficient search algorithm for the lagrange multipliers of optimal spectrum balancing in multi-user xDSL systems," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 4, May 2006.

[12] R. Cendrillon, M. Moonen, J. Verlinden, T. Bostoen, and W. Yu, "Optimal multi-user spectrum management for digital subscriber lines," in *Proc. of the IEEE International Communications Conference (ICC)*, vol. 1, Jun. 2004, pp. 1–5.

[13] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multi-user spectrum management for digital subscriber lines," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 922–933, May 2006.

[14] T. Starr, J. M. Cioffi, and P. J. Silverman, *Understanding Digital Subscriber Lines*. Prentice Hall, 1999.

Chapter 10

Joint Spectrum and Cancellation Resource Optimization for Linear Partial Crosstalk Cancellation

Joint Spectrum Management and Constrained Partial Crosstalk Cancellation in a Multi-user xDSL Environment

J. Vangorp, P. Tsiaflakis, M. Moonen, and J. Verlinden

Published in *Signal Processing, Special Section: Information Processing and Data Management in Wireless Sensor Networks*, vol. 87, no. 12, Dec. 2007, pp. 3131-3146, doi:10.1016/j.sigpro.2007.06.008

©2007 Elsevier B.V. All rights reserved.

Contributions of second author

- literature study
- co-development of problem statements for joint spectrum coordination and partial crosstalk cancellation
- co-development of optimal dual decomposition based algorithm
- co-analysis of computational complexity
- co-development of complexity reduction techniques
- co-development of ON/OFF power loading procedure
- co-design of simulation scenarios
- software co-implementation and computer co-simulation
- co-analysis and co-interpretation of simulation results
- text co-redaction and co-editing

Abstract

In modern DSL systems, crosstalk is a major source of performance degradation. Crosstalk cancellation techniques have been proposed to mitigate the effect of crosstalk. However, the run-time complexity of these crosstalk cancellation techniques grows with the square of the number of lines. Therefore one has to be selective in cancelling crosstalk to reduce complexity. Secondly, crosstalk cancellation requires signal-level coordination between transmitters or receivers, which is not always available. Because of accessibility constraints, crosstalk between certain lines cannot be cancelled and so has to be mitigated through spectrum management. After a complexity study, this paper presents a solution for the joint spectrum management and constrained partial crosstalk cancellation problem. The complexity of the partial crosstalk cancellation part of the problem is reduced based on a line selection and user independence observation. However, to fully benefit from these observations, power loading has to be applied in the spectrum management part. We therefore also consider ON/OFF power loading, which has a low complexity and shows only a minor performance degradation compared to normal power loading. The resulting algorithm will be compared to currently available algorithms for independent spectrum management and partial crosstalk cancellation.

10.1 Introduction

Current xDSL access networks are evolving into mixtures of various DSL flavours, where for instance traditional ADSL lines provisioning customers over longer distances share binders with VDSL lines deployed from remote terminals. These network topologies suffer from electromagnetic coupling resulting in crosstalk between lines. As xDSL systems currently under development use higher frequencies to meet the demand for high data rates, crosstalk is indeed becoming particularly harmful. Moreover, significant line length variations and mixed deployments from central offices (CO's) and remote terminals (RT's) create a near-far effect in the upstream and downstream direction respectively. This causes crosstalk to sometimes overpower the direct signals. As a result, crosstalk can be 10-15 dB larger than the background noise, and becomes a major limiting factor in the performance of xDSL systems [1].

One strategy for dealing with this crosstalk is crosstalk cancellation. Several crosstalk cancellation techniques have been proposed to remove crosstalk in different scenarios. Linear pre- and post filtering [2][3] requires coordination at both the transmitters and receivers. When this level of coordination is not available, successive interference cancellation or precompensation [4][5] can be used if there is only coordination at the receivers or transmitters respectively. For this level of coordination, it is shown in [6][7] that a simple linear zero-

forcing canceller or linear precompensator performs near optimally in an xDSL environment.

However, even for these simple linear cancellers, the run-time complexity grows with the square of the number of lines. For example, in a binder of 8 VDSL lines transmitting on 4096 tones at a block rate of 4000 blocks per second, the run-time complexity of crosstalk cancellation exceeds 1 billion multiplications per second. Because most of the crosstalk originates from a limited number of lines on a limited number of tones, a fraction of this complexity suffices to cancel most of the crosstalk. This is called partial crosstalk cancellation [8][9].

Crosstalk cancellation requires signal-level coordination at either the transmitter or receiver, i.e. the signals transmitted on interfering lines should be known. Oftentimes, not all interfering lines can be cancelled because they are not accessible. This is the case in a mixed CO-RT deployment where CO and RT reside in different geographical locations. Here partial crosstalk cancellation at the CO side has to be done independent of the partial crosstalk cancellation at the RT side. Secondly, accessibility constraints restrict the number of lines that can have signal-level coordination, even if they are at the same location. For example, crosstalk cancellation may not be possible between lines connected to different line cards.

In such situations, spectrum management [10][11] can be used to mitigate the crosstalk originating from lines that are not accessible. This is a second strategy for dealing with crosstalk. Instead of cancelling the crosstalk after it has occurred, transmit spectra are chosen such that the effect of crosstalk is minimized.

Currently available algorithms independently solve the spectrum management and partial crosstalk cancellation problem. A spectrum management algorithm first selects transmit spectra to avoid crosstalk. As an example, Optimal Spectrum Balancing (OSB) [12][13] can be used to calculate the optimal spectra that minimize the effect of crosstalk. Given these spectra, a partial crosstalk cancellation scheme can be used to cancel the remaining crosstalk. This two-step approach can be very suboptimal. The spectrum management algorithm does not take into account that a certain amount of crosstalk can be canceled afterwards and hence the resulting spectra will be overly conservative.

A better solution can be obtained if the spectrum management and partial crosstalk cancellation problems are solved jointly. In [14], partial crosstalk cancellation based on resource allocation [8] is combined with Iterative Water-filling (IW) spectrum management [15] in an iterative fashion. However, IW tends to be highly suboptimal in near-far scenarios. In this paper, the OSB algorithm is extended to include constrained partial crosstalk cancellation, based on [16].

In Section 10.2 of this paper, the joint spectrum management and constrained partial crosstalk cancellation problem is defined and a solution is proposed based on a dual decomposition. However, the complexity of this solution is found to be too high [16]. Therefore, in Section 10.3, possibilities to reduce this complexity are explored. In a first subsection the complexity of the partial crosstalk cancellation part of the problem is reduced based on a line selection and a user independence observation. These observations do not affect the optimality of the solution. In a second subsection, the complexity of the spectrum management part of the problem is reduced by applying ON/OFF power loading, which has a low complexity and shows only a minor performance degradation compared to the optimal solution. Finally, in Section 10.4, simulation results are presented where the performance of the joint algorithm is compared to the solution obtained by independently solving the spectrum management problem and the partial crosstalk cancellation problem. Finally, Section 10.5 concludes the paper.

10.2 Joint Spectrum Management and Constrained Partial Crosstalk Cancellation

10.2.1 System Model

Most current DSL systems use Discrete Multi-Tone (DMT) modulation. The available frequency band is divided in a number of parallel subchannels or tones. Each tone is capable of transmitting data independently from other tones, and so the transmit power and the number of bits can be assigned individually for each tone. This gives a large flexibility in optimally shaping the transmit spectrum to minimize the effect of crosstalk.

Transmission for a binder of N users can be modeled on each tone k by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k \qquad k = 1 \ldots K.$$

The vector $\mathbf{x}_k = [x_k^1, x_k^2, \ldots, x_k^N]^T$ contains the transmitted signals on tone k for all N users. $[\mathbf{H}_k]_{n,m} = h_k^{n,m}$ is an $N \times N$ matrix containing the channel transfer functions from transmitter m to receiver n. The diagonal elements are the direct channels, the off-diagonal elements are the crosstalk channels. $\mathbf{z}_k$ is the vector of additive noise on tone k, containing thermal noise, alien crosstalk, RFI,... The vector $\mathbf{y}_k$ contains the received symbols.

To take crosstalk cancellation into account, an equivalent channel $\tilde{\mathbf{H}}$ is introduced. This is the same channel as the original channel $\mathbf{H}$, but with off-diagonal elements set to 0 where the crosstalk is cancelled. If user n is cancelling crosstalk originating from user m on tone k, then $\tilde{h}_k^{n,m} = 0$. We refer to [8][9] where procedures are explained for cancelling individual crosstalk channels, based on

particular DSL channel characteristics (row/column-wise diagonal dominance).

We denote the transmit power as $s_k^n \triangleq \Delta_f E\{|x_k^n|^2\}$, $\mathbf{s}_k = [s_k^1, s_k^2, \ldots, s_k^N]^T$, $\mathbf{s} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_K]$ and the noise power as $\sigma_k^n \triangleq \Delta_f E\{|z_k^n|^2\}$. The DMT symbol rate is denoted as f_s, the tone spacing as Δ_f.

It is assumed that each modem treats interference from other modems as noise. When the number of interfering modems is large, the interference is well approximated by a Gaussian distribution. Under this assumption the achievable bit loading of user n on tone k, given the transmit spectra of all modems in the system, is

$$b_k^n \triangleq \log_2\left(1 + \frac{1}{\Gamma}\frac{|\tilde{h}_k^{n,n}|^2 s_k^n}{\sum_{m\neq n}|\tilde{h}_k^{n,m}|^2 s_k^m + \sigma_k^n}\right), \tag{10.1}$$

where Γ denotes the SNR-gap to capacity, which is a function of the desired BER, the coding gain and noise margin. Note that taking the same SNR-gap for all the tones can be suboptimal since it is the mean symbol error probability across the tones that should be constrained.

We also define $\mathbf{b}_k = [b_k^1, b_k^2, \ldots, b_k^N]^T$ and $\mathbf{b} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_K]$. The data rate for user n is then

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

When the bit loading b_k^n of the users is given for a specific tone k, the required transmit power $\mathbf{s}_k$ for the modems in the system can be calculated by [13]

$$\mathbf{s}_k = \left(\mathbf{D}_k - \mathbf{\Lambda}_k\mathbf{A}_k\right)^{-1}\mathbf{\Lambda}_k\sigma_k \tag{10.2}$$

$$\begin{aligned}
\mathbf{D}_k &\triangleq \operatorname{diag}\{|\tilde{h}_k^{1,1}|^2, |\tilde{h}_k^{2,2}|^2, \ldots, |\tilde{h}_k^{N,N}|^2\} \\
\mathbf{\Lambda}_k &\triangleq \operatorname{diag}\{2^{b_k^1}-1, 2^{b_k^2}-1, \ldots, 2^{b_k^N}-1\} \\
[\mathbf{A}_k]_{n,m} &\triangleq a_k^{n,m} \quad \text{with} \quad a_k^{n,m} \triangleq \begin{cases} 0 & n = m \\ \Gamma|\tilde{h}_k^{n,m}|^2 & n \neq m \end{cases} \\
\sigma_k &\triangleq \Gamma[\sigma_k^1, \sigma_k^2, \ldots, \sigma_k^N]^T
\end{aligned}$$

Not all bit loadings lead to feasible power loadings and equation (10.2) can lead to negative powers. To avoid this, a positivity constraint has to be added: $0 < s_k^n$. This positivity constraint allows the detection of infeasible bit loadings.

The total power used by user n is then

$$P^n = \sum_k s_k^n.$$

10.2.2 Problem Statement

The joint spectrum management and constrained partial crosstalk cancellation problem amounts to finding an optimal allocation of transmit power and selection of the crosstalk to cancel, thereby maximizing the capacity of the network. In doing so, there are a number of constraints.

First of all, there is a *total power constraint* $P^{n,tot}$ for each user. This constraint ensures the user's total power does not exceed the maximum allowed total transmit power. On top of this constraint there can be a *spectral mask constraint* $s_k^{n,mask}$ for each tone to guarantee electromagnetic compatibility with other systems.

Secondly, because of the run-time complexity of full crosstalk cancellation, there is a limited amount of resources for crosstalk cancellation. The cancellation of the crosstalk from one crosstalker on a tone is done by one cancellation tap [8][9]. The number of cancellation taps used is constrained by the *cancellation tap constraint* C^{tot} [16]. Furthermore, in a bundle of lines, not all crosstalk can be cancelled. This is the case when receivers are in different geographical locations or when lines are terminating on different line cards. These scenarios can be modeled by multiple cancellation tap constraints $C^{q,tot}$, i.e. one constraint for each subset q of lines with full signal-level access.

Finally, there is a *rate constraint* $R^{n,target}$ for each user. Typically, service providers offer a number of profiles to guarantee a certain Quality-of-Service. The rate constraint then indicates a minimum data rate required by the user.

Joint spectrum management and constrained partial crosstalk cancellation then results in solving the following maximization problem [16]:

$$
\begin{array}{lll}
\max\limits_{\mathbf{s},\mathbf{c}} & \sum\limits_{n=1}^{N} R^n & \\
\text{subject to} & P^n \leq P^{n,tot} & n = 1 \ldots N \\
& 0 \leq s_k^n \leq s_k^{n,mask} & n = 1 \ldots N, k = 1 \ldots K \\
& R^n \geq R^{n,target} & n = 1 \ldots N \\
& C^q = \sum\limits_k \sum\limits_n \sum\limits_m c_k^{n,m} \leq C^{q,tot} & n, m \in i^q, q = 1 \ldots Q \\
\text{with} & [\mathbf{c}_k]_{n,m} = c_k^{n,m} & c_k^{n,m} = \begin{cases} 0 & \Rightarrow \quad \tilde{h}_k^{n,m} = h_k^{n,m} \\ 1 & \Rightarrow \quad \tilde{h}_k^{n,m} = 0 \end{cases} \\
& s_k^n \in \mathcal{S} &
\end{array}
\tag{10.3}
$$

where $\mathbf{c} = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_K]$. $\mathbf{c}_k$ is a matrix containing the crosstalk cancellation configuration for tone k. $c_k^{n,m} = 1$ indicates that a cancellation tap is assigned on tone k for cancelling crosstalk on line n originating from line m. Because of accessibility constraints, n and m are restricted to the subset of line indices

i^q which have full signal-level control. For lines n, m that have no signal-level control, $c_k^{n,m} = 0, \forall k$.

Finally, $\mathcal{S}$ is a set of possible power levels. With (10.3), all possible transmit power levels (together with all possible cancellation tap configurations) are scanned and so this procedure will from now on be referred to as "power loading".

Alternatively, (10.3) can be reformulated as an optimization problem in $\{\mathbf{b}, \mathbf{c}\}$ instead of $\{\mathbf{s}, \mathbf{c}\}$, i.e.

$$
\begin{array}{lll}
\max\limits_{\mathbf{b},\mathbf{c}} & \sum\limits_{n=1}^{N} R^n & \\
\text{subject to} & P^n \leq P^{n,tot} & n = 1 \ldots N \\
& 0 \leq s_k^n \leq s_k^{n,mask} & n = 1 \ldots N, k = 1 \ldots K \\
& R^n \geq R^{n,target} & n = 1 \ldots N \\
& C^q = \sum\limits_k \sum\limits_n \sum\limits_m c_k^{n,m} \leq C^{q,tot} & n, m \in i^q, q = 1 \ldots Q \\
\text{with} & [\mathbf{c}_k]_{n,m} = c_k^{n,m} & c_k^{n,m} = \left\{ \begin{array}{lll} 0 & \Rightarrow & \tilde{h}_k^{n,m} = h_k^{n,m} \\ 1 & \Rightarrow & \tilde{h}_k^{n,m} = 0 \end{array} \right. \\
& b_k^n \in \mathcal{B} &
\end{array}
\tag{10.3bis}
$$

where $\mathcal{B}$ is a set of possible bit loadings. This will be referred to as "bit loading". The cardinalities of $\mathcal{S}$ and $\mathcal{B}$ will both be referred to as B.

10.2.3 Dual Decomposition

Optimization problem (10.3)-(10.3bis) is a non-convex problem. To find the global optimum of (10.3)-(10.3bis), no other solution is known than to exhaustively search through all possible transmit spectra $\mathbf{s}$ (or bit load vectors $\mathbf{b}$) and cancellation tap configurations $\mathbf{c}$. Because some constraints are coupled over the tones, this results in an exponential complexity in the number of tones. In [12][13], the complexity is made linear in K by using a dual decomposition that decouples the problem over the tones. This is done by using Lagrange multipliers to move the constraints coupled over tones into the objective function of the optimization problem. It was shown in [16] that if the number of tones is sufficiently large, the duality gap is zero and therefore the dual decomposition does not affect the solution. Because the cancellation tap constraint is also coupled over the tones, extra Lagrange multipliers are introduced to make the

complexity linear. For the power loading case this results in

$$\begin{aligned}
\mathbf{s}^{opt}, \mathbf{c}^{opt} := \quad & \underset{\mathbf{s},\mathbf{c}}{\text{argmax}} \quad \sum_{n=1}^{N} \omega_n R^n + \sum_{n=1}^{N} \lambda_n \big(P^{n,tot} - \sum_{k=1}^{K} s_k^n\big) \\
& \qquad + \sum_{q=1}^{Q} \nu_q \big(C^{q,tot} - \sum_{k=1}^{K} \sum_{n\in i^q} \sum_{m \in i^q} c_k^{n,m}\big) \\
& \text{subject to} \quad 0 \le s_k^n \le s_k^{n,mask}, \qquad n = 1 \dots N \\
& \qquad \lambda_n \ge 0, \omega_n \ge 0, \qquad n = 1 \dots N \\
& \qquad \nu_q \ge 0, \qquad q = 1 \dots Q \\
& \qquad s_k^n \in \mathcal{S}
\end{aligned} \tag{10.4}$$

where ω_n, λ_n and ν_q are Lagrange multipliers. $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_N]^T$ and $\boldsymbol{\nu} = [\nu_1, \dots, \nu_Q]^T$ can be seen as a cost for power and crosstalk cancellation taps respectively that replace the total power constraint and the cancellation tap constraint. Larger values for these Lagrange multipliers result in less power and allocated cancellation taps. The data rates of the users are weighted by $\boldsymbol{\omega} = [\omega_1, \dots, \omega_N]^T$, thereby giving more importance to some users and replacing the target rate constraints. In this way, all possible trade offs can be made to enforce the data rate constraints.

For a given set of $\boldsymbol{\omega}$, $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ and by collecting the terms per tone and noting that a sum can be maximized by maximizing the individual terms, this optimization problem can then be solved in a per-tone fashion:

for $k = 1 \dots K$:

$$\begin{aligned}
\mathbf{s}_k^{opt}, \mathbf{c}_k^{opt} := \quad & \underset{\mathbf{s}_k,\mathbf{c}_k}{\text{argmax}} \quad \sum_{n=1}^{N} \omega_n f_s b_k^n - \sum_{n=1}^{N} \lambda_n s_k^n - \sum_{q=1}^{Q} \sum_{n\in i^q} \sum_{m \in i^q} \nu_q c_k^{n,m} \\
& \text{subject to} \quad 0 \le s_k^n \le s_k^{n,mask}, \qquad n = 1 \dots N \\
& \qquad \lambda_n \ge 0, \omega_n \ge 0, \qquad n = 1 \dots N \\
& \qquad \nu_q \ge 0, \qquad q = 1 \dots Q \\
& \qquad s_k^n \in \mathcal{S}.
\end{aligned} \tag{10.5}$$

Similar formulas can be given for the bit loading case.

Maximization of (10.5) for given Lagrange multipliers $\boldsymbol{\omega}$, $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ can be performed by an exhaustive search. For each tone, the objective function should be evaluated for all possible combinations of the transmit power levels $\mathbf{s}_k$ and cancellation tap configurations $\mathbf{c}_k$ of the users. The combination giving the largest value for this expression is the optimal allocation of transmit power and cancellation taps for this tone and thus the original constraints can be checked.

To solve (10.3) by (10.5), $\boldsymbol{\omega}, \boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ should be tuned such that after the per-tone exhaustive search the original constraints are satisfied. In [12], an

efficient Lagrange multiplier search procedure for $\boldsymbol{\omega}$ and $\boldsymbol{\lambda}$ is presented. In the Appendix, this procedure is extended to include $\boldsymbol{\nu}$, resulting in an update formula as follows:

$$\begin{bmatrix} \boldsymbol{\omega} \\ \boldsymbol{\lambda} \\ \boldsymbol{\nu} \end{bmatrix}^{t+1} = \left(\begin{bmatrix} \boldsymbol{\omega} \\ \boldsymbol{\lambda} \\ \boldsymbol{\nu} \end{bmatrix}^{t} - \mu \begin{bmatrix} \mathbf{R} - \mathbf{R}^{target} \\ \mathbf{P}^{tot} - \mathbf{P} \\ \mathbf{C}^{tot} - \mathbf{C} \end{bmatrix} \right)^{+} \tag{10.6}$$

where t is the iteration index, $(x)^+$ means $\max(0, x)$ and where $\mathbf{R} = [R^1, \ldots, R^N]^T$, $\mathbf{P} = [P^1, \ldots, P^N]^T$ and $\mathbf{C} = [C^1, \ldots, C^Q]^T$ are vectors with the total powers, data rates and number of cancellation taps corresponding to the Lagrange multipliers at hand. This update formula is used in algorithm 10.1 adopted from [12]. The non-positive inner product of (10.17) ensures that if the stepsize μ is small, update formula (10.6) will bring us closer to the optimum where the distance is expressed as the Euclidean distance to satisfying the constraints with equality. If the current stepsize brings us closer to the optimum, we could have taken a larger step to converge faster, e.g. doubling the stepsize. If we detect that the stepsize is too large (distance to the optimum increases), we take the best Lagrange multipliers found so far and start over. If our initial stepsize is too large, we can start halving it until it is small enough.

Algorithm 10.1 Multi-user Lagrange multiplier search algorithm

```
μ = 1
while distance > tolerance do
  Θ = [ω, λ, ν]^T = best [ω, λ, ν]^T so far
  step = 1
  while distance ≤ previousDistance do
    previousDistance = distance
    ΔΘ = [Δω, Δλ, Δν]^T = update formula (10.6)
    [R^{Θ+ΔΘ}, P^{Θ+ΔΘ}, C^{Θ+ΔΘ}, s, c] = exhaustiveSearch(Θ + ΔΘ)
    distance = ‖[R^{Θ+ΔΘ} − R^{target}, P^{tot} − P^{Θ+ΔΘ}, C^{tot} − C^{Θ+ΔΘ}]^T‖
    if distance > previousDistance & step = 1 then
      μ = μ/2
    else
      μ = μ × 2
      step = 0
    end if
  end while
end while
```

Note that all the Lagrange multipliers are updated in parallel. In [12] it is observed that adding extra Lagrange multipliers does not increase the number of steps required for convergence. The search procedure typically converges in 50 to 150 steps. Therefore, the cancellation tap constraint only adds to the complexity of the per-tone exhaustive search.

10.2.4 Complexity

The joint spectrum management and constrained partial crosstalk cancellation problem (10.3)-(10.3bis) is a non-convex constrained optimization problem. Without the dual decomposition, finding the global optimum requires an exhaustive search over all possible solutions. First, assume there are no accessibility constraints, so that all crosstalk can be cancelled. On a certain tone, a user has to decide which crosstalk from $N-1$ other users has to be cancelled. There are 2^{N-1} possibilities to do this. Together with B possibilities for discrete bit or power loading, this results in a total of $B2^{N-1}$ possibilities for each user on each tone and hence a total complexity of $\mathcal{O}((B2^{N-1})^{KN})$.

The dual decomposition decouples the problem over the tones, therefore reducing the exponential complexity in the number of tones K to linear complexity: $\mathcal{O}(K(B2^{N-1})^N)$. This amounts to K exhaustive searches of complexity $\mathcal{O}((B2^{N-1})^N)$. This is an enormous reduction in complexity. However, this solution is still computationally intractable because of the remaining complexity of the per-tone exhaustive search, which is indeed $(2^{N-1})^N$ times more complex than the per-tone search in the spectrum management problem without crosstalk cancellation. In a 4-user upstream VDSL scenario for example, it takes 20 days to calculate optimal spectra with OSB on a Pentium IV. When partial crosstalk cancellation is added, computing the optimal solution would take about 225 years.

The dual decomposition approach is only feasible if the per-tone exhaustive search can be performed with manageable complexity, as was also concluded in [16]. In the next section, methods are introduced to make this possible.

10.3 Complexity Reduction

The complexity of the per-tone exhaustive search for the joint spectrum management and constrained partial crosstalk cancellation problem is $\mathcal{O}(K(B2^{N-1})^N)$, in the case where all crosstalk can be cancelled (Section 10.2.4). This can be rewritten as $\mathcal{O}(KB^N(2^{N-1})^N)$, clearly showing the complexity due to spectrum management, $\mathcal{O}(B^N)$, and partial crosstalk cancellation, $\mathcal{O}((2^{N-1})^N)$. In this section we focus on reducing the complexity originating from these two individual subproblems.

It will turn out that, unlike in the bit loading case, a significant complexity reduction can be achieved in the power loading case.

10.3.1 Partial Crosstalk Cancellation

This subsection again starts with the assumption that there are no accessibility constraints, i.e. that all crosstalk can be cancelled. Later, observations will be extended to the case with accessibility constraints.

To determine the optimal allocation of crosstalk cancellation taps for a given bit or power loading on a certain tone, all of the $(2^{N-1})^N \approx 2^{N^2}$ possible allocations have to be evaluated. Fortunately, in the power loading case many of these possibilities can be eliminated based on two observations: line selection and user independence.

- *Line Selection:* A user has to decide whether or not it will cancel the crosstalk originating from $N-1$ other users. This leads to 2^{N-1} possible crosstalk cancellation configurations. However, from (10.1) it can be seen that to maximize the capacity for given transmit powers s_k^n (power loading case), one should allocate crosstalk cancellation taps to cancel those users that are causing the largest crosstalk. Therefore, if r crosstalk cancellation taps are available, they should be used to cancel the r largest sources of crosstalk.
  As a consequence, *when power loading is applied*, the 2^{N-1} possibilities for crosstalk cancellation are reduced to N possibilities for each user: no crosstalk cancellation, cancel the strongest crosstalker, cancel the 2 strongest crosstalkers, ..., cancel the $N-1$ strongest crosstalkers. In the N-user case, this observation results in the following complexity reduction:
  $$(2^{N-1})^N \rightarrow N^N \tag{10.7}$$
- *User Independence:* All users have to decide on a crosstalk cancellation configuration. This leads to an exponential complexity in the number of users N. However, from (10.1) it can be seen that if for given transmit powers s_k^n (power loading case) user n allocates a crosstalk cancellation tap to cancel crosstalk caused by user m (i.e. $\tilde{h}_k^{n,m} = 0$) this only has an influence on the capacity of user n. Therefore, *when power loading is applied*, the users are "decoupled", i.e. they can choose a crosstalk cancellation configuration independently.

  The per-tone optimization problem (10.5) (without accessibility constraints) can indeed be rewritten as

$$\begin{aligned} &\max_{\mathbf{s}_k,\mathbf{c}_k} \Big\{ \sum_{n=1}^{N} \omega_n f_s b_k^n - \sum_{n=1}^{N} \lambda_n s_k^n - \sum_{n}^{N} \sum_{m\neq n}^{N} \nu_q c_k^{n,m} \Big\} \\ = &\max_{\mathbf{s}_k} \Big\{ \sum_{n=1}^{N} \max_{c_k^{n,1},c_k^{n,2},\ldots,c_k^{n,N}} (\omega_n f_s b_k^n - \lambda_n s_k^n - \sum_{m\neq n}^{N} \nu_q c_k^{n,m}) \Big\} \end{aligned} \tag{10.8}$$

This means that for each selection of the power load vector $\mathbf{s}_k$, the selection of the corresponding optimal crosstalk cancellation configuration (for a given set of Lagrange multipliers) can be done in a per-user fashion. As a consequence, the exponential complexity in N is reduced to linear complexity. This observation results in the following complexity reduction:

$$(2^{N-1})^N \rightarrow N(2^{N-1}) \tag{10.9}$$

It is noted that in the case of bit loading (10.3bis), (10.2) is used to determine the power necessary to load a certain number of bits on a tone. Cancelling crosstalk for one user now affects all other users. Adding a crosstalk cancellation tap for one user changes the power needed to transmit a certain number of bits for this user, thus also the crosstalk into other users changes. This may again affect the configuration of crosstalk cancellation taps for these other users. Therefore, user independence does not hold when bit loading is applied.

In the power loading case, both observations can be combined. N users then independently have to choose one of N possible crosstalk cancellation configurations. This results in the following total complexity reduction:

$$(2^{N-1})^N \rightarrow NN \tag{10.10}$$

These observations can be easily extended to the case where there are accessibility constraints, reducing the number of crosstalkers that can be cancelled. Assume there are Q line cards. Line card q has access to M_q lines, with $\sum_{q=1}^{Q} M_q = N$. The complexity reduction by line selection and user independence is then summarized in table 10.1.

Table 10.1: Complexity reduction partial crosstalk cancellation

	1 line card	Q line cards
full complexity	$\mathcal{O}((2^{N-1})^N)$	$\mathcal{O}(\prod_{q=1}^{Q}(2^{M_q-1})^{M_q})$
line selection	$\mathcal{O}((N)^N)$	$\mathcal{O}(\prod_{q=1}^{Q}(M_q)^{M_q})$
user independence	$\mathcal{O}(N2^{N-1})$	$\mathcal{O}(\sum_{q=1}^{Q} M_q 2^{M_q-1})$
line selection & user independence	$\mathcal{O}(NN)$	$\mathcal{O}(\sum_{q=1}^{Q} M_q M_q)$

In an 8-user case, the above observations reduce the number of crosstalk cancellation configurations from 2^{56} to 2^6. If there are 2 line cards, each having 4

lines, the number of crosstalk cancellation configurations is reduced from 2^{24} to 2^5. Note that despite these complexity reductions, the computed solution is still globally optimal.

10.3.2 Spectrum Management: ON/OFF Power Loading

In this subsection, the complexity of the spectrum management part of the problem is reduced. Despite the complexity reduction provided by dual decomposition, OSB is still found to be too complex for scenarios with more than 3 users. The reason is the per-tone exhaustive search which still has exponential complexity in the number of users: $\mathcal{O}(B^N)$. In [17][18] an iterative procedure (ISB) is used to make this complexity linear. However, optimality cannot be guaranteed.

In this paper, the complexity is combated by reducing B, the number of possible transmit levels (for power loading). Originally, for OSB, typical values for B are 60 in the case of power loading and 14 in the case of bit loading. Therefore, bit loading would be the most efficient method for OSB. However, as shown in the previous subsection, power loading is necessary to fully benefit from the line selection and user independence observations. By limiting the transmit spectra to ON/OFF power loading, $B = 2$, the complexity is reduced from $\mathcal{O}(B^N)$ to $\mathcal{O}(2^N)$ and the result are simple transmit spectra similar to what is currently used in ADSL systems. This ON/OFF power loading problem equals (10.5) with the spectral mask constraints replaced by

$$s_k^n \in \{0, s_k^{n,ON}\} \text{ with } s^{n,ON} \leq s_k^{n,mask}. \tag{10.11}$$

In order to perform the reduced per-tone exhaustive searches, $s^{n,ON}$ has to be defined. This ON-level should not violate the spectral mask constraint and the resulting spectra should not violate the total power constraint. An ON-level that automatically enforces the total power constraint is

$$s_k^{n,ON} = \min\left(\frac{P^n}{K}, s_k^{n,mask}\right). \tag{10.12}$$

Even when all tones are switched on, the total power constraint is not violated. Therefore, no Lagrange multipliers for the total power constraints have to be searched.

However, with this fixed ON-level, if tones are switched off, not all available power is allocated while it could be redistributed to the tones that are switched on. Because of the lack of this redistribution of power, ON/OFF loading with a fixed ON-level is found to perform worse than Iterative Waterfilling (IW) [15].

The performance of ON/OFF loading can be greatly improved by redistributing the power of tones that are switched off to the other tones. To do this, one can

use an iterative procedure, in each iteration updating the ON-level according to the number of active tones:

$$s_k^{n,ON} = \min\left(\frac{P^n}{(\#\text{ active tones})^{\text{n}}}, s_k^{n,mask}\right). \tag{10.13}$$

However, in some cases, the per-tone exhaustive search does not automatically switch off tones. E.g. in a CO/RT scenario, CO lines do not cause significant crosstalk on the RT lines. Thus these tones are not switched off by the per-tone exhaustive search. As these tones carry (very) few bits, the overall performance would then benefit from a redistribution of power from high frequencies to lower frequencies.

This can be accomplished by setting a threshold on the minimum number of bits that a tone should carry. Tones that do not reach this minimum are then switched off. Tones that are switched off once ($t_k^{n,usable} = 0$) are never activated again, thereby assuring convergence of this iterative procedure. The complete procedure (without crosstalk cancellation) is presented in algorithm 10.2.

Algorithm 10.2 Multi-user ON/OFF loading with adaptive ON-level

$t_k^{n,\text{usable}} = 1, \quad n = 1 \ldots N, k = 1 \ldots K$
repeat
  $s_k^{n,ON} = \min\left(\frac{P^n}{(\#\text{ active tones})^{\text{n}}}, s_k^{n,mask}\right)$
  for $k = 1$ to K **do**
    $s_k^{1,opt}, \ldots, s_k^{n,opt} = \text{argmax}_{s_k^1,\ldots,s_k^n}\, \omega_1 b_k^1 + \ldots + \omega_n b_k^n$
      with $s_k^n \in \{0, s_k^{n,ON} \cdot t_k^{n,\text{usable}}\}$
    $\forall n$: **if** $b_k^n < threshold$ **then** $t_k^{n,\text{usable}} = 0, s_k^{n,opt} = 0$
  end for
until convergence

The procedure typically converges in 4-5 iterations. A good value for the threshold depends on the scenario and the target rates. Low thresholds give good performance for high rates on the RT lines, high thresholds give good performance for high rates on the CO lines. To adaptively select a good threshold for a specific scenario and given target rates, a number of different thresholds can be used. The complexity increase is minor when the thresholds are applied in ascending order. For each new threshold, one can then start from the converged transmit spectra of the previous threshold.

It is noted that while ON/OFF power loading in itself drastically reduced the complexity of the per-tone search, the algorithm can still be combined with the iterative search algorithm (ISB) presented in [17][18]. Instead of performing an exhaustive multi-user search, each user iteratively performs a single-user search

to optimize the weighted rate sum, while other users are kept fixed. Because in this case there are only 2 choices for the power loading in the single-user search, this iterative approach performs as good as an exhaustive search.

An overview of the complexities of various algorithms discussed is shown in table 10.2. U is the number of updates required to find the Lagrange multipliers when there are more than 2 power levels, typically 50-150 [12]. I_1 is the number of iterations needed to converge to an adaptive ON-level, typically 4-5, I_2 is the number of iterations needed to converge in the adaptive threshold case, typically 10-12 for 7 different thresholds. In the ISB case [17][18], I_{ISB} iterations are performed in the per-tone search, typically 2-3.

Table 10.2: Complexity ON/OFF loading

	OSB	ISB
full complexity	$\mathcal{O}(UKB^N)$	$\mathcal{O}(UKI_{ISB}BN)$
fixed on	$\mathcal{O}(K2^N)$	$\mathcal{O}(KI_{ISB}2N)$
adaptive on	$\mathcal{O}(I_1K2^N)$	$\mathcal{O}(KI_1I_{ISB}2N)$
adaptive threshold	$\mathcal{O}(I_2K2^N)$	$\mathcal{O}(KI_2I_{ISB}2N)$

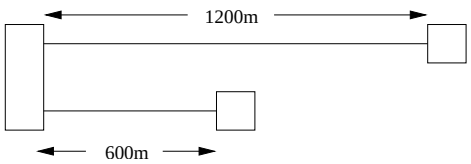
In figure 10.1(b) the rate regions are shown for the scenario of figure 10.1(a), without crosstalk cancellation. Both OSB and ON/OFF loading with adaptive threshold achieve a significant performance gain over Iterative Waterfilling. ON/OFF loading introduces only 10% performance loss compared to OSB.

When combining the line selection and user independence observations with ON/OFF power loading, the optimization of (10.3) for a 4-user upstream scenario can be done in a few minutes instead of the original 225 years (Section 10.2.4).

10.4 Simulation Results

In this section the performance is analyzed when solving the joint spectrum management and constrained partial crosstalk cancellation problem as opposed to independently solving these two problems. An upstream VDSL scenario is considered as shown in figure 10.2, with full signal-level coordination. A line diameter of 0.5 mm (24 AWG) is used and the maximum transmit power is

(a)



(b)

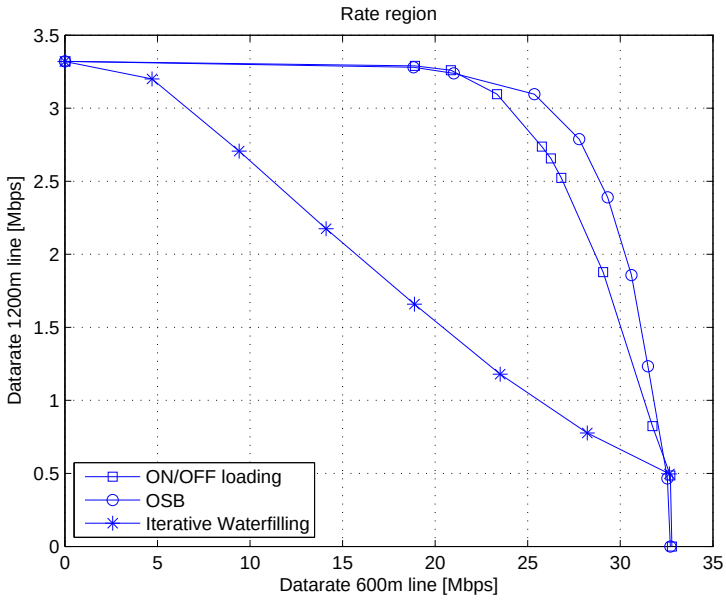


Figure 10.1: 2-user VDSL scenario, adaptive threshold

11.5 dBm. The SNR gap Γ is set to 12.9 dB, corresponding to a target symbol error probability of 10^{-7}, coding gain of 3 dB and a noise margin of 6 dB. The tone spacing is $\Delta_f = 4.3125$ kHz and the DMT symbol rate $f_s = 4$ kHz [19].

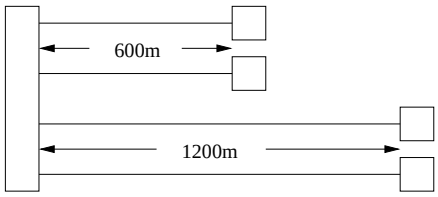


Figure 10.2: 4-user VDSL scenario

To manage the complexity of the problem, ON/OFF loading is used for the spectrum management part of the per-tone search. Furthermore, the ISB procedure is used to search the maximum. In figure 10.3 rate regions are shown to compare the performance of the joint solution and the independent solution obtained by independently solving the spectrum management problem (with ON/OFF loading) and the partial crosstalk cancellation problem.

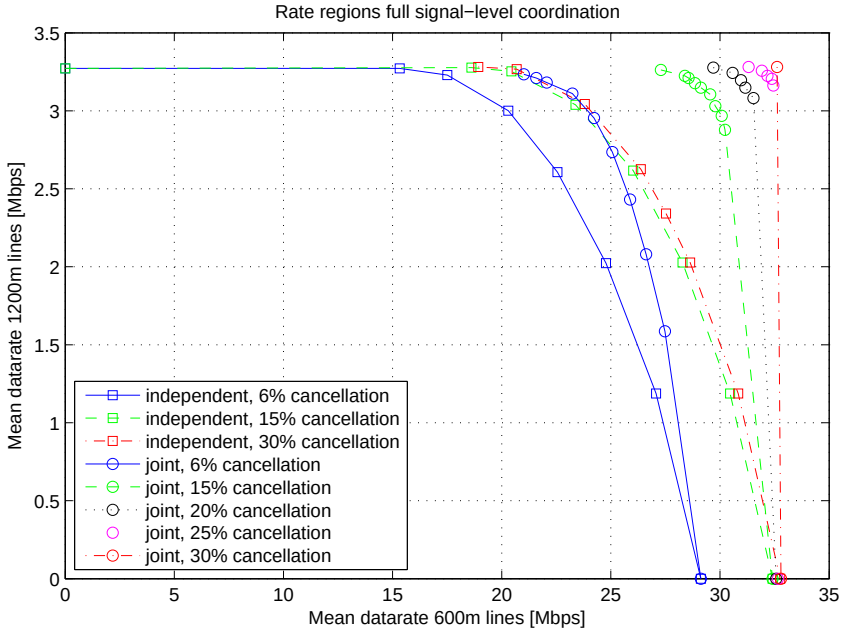


Figure 10.3: Rate regions 4-user VDSL scenario

The rate regions show significant performance gains of the joint solution over the independent solution. Because the independent solution first independently solves the spectrum management problem, the transmit spectra are chosen to avoid crosstalk. This can be seen in figure 10.4(a)(b), where the PSD is shown along with the allocation of cancellation taps for each user and the originating user of the crosstalk that is cancelled. For this strong crosstalk scenario the transmit spectra result in long and short lines occupying different frequency bands. When the partial crosstalk cancellation problem is solved, there is not much crosstalk left to cancel. Therefore, only a limited crosstalk cancellation tap budget can be used effectively. Figure 10.3 shows that no performance is gained by increasing the crosstalk cancellation tap budget beyond 15% of full crosstalk cancellation.

When the spectrum management problem and the partial crosstalk cancellation

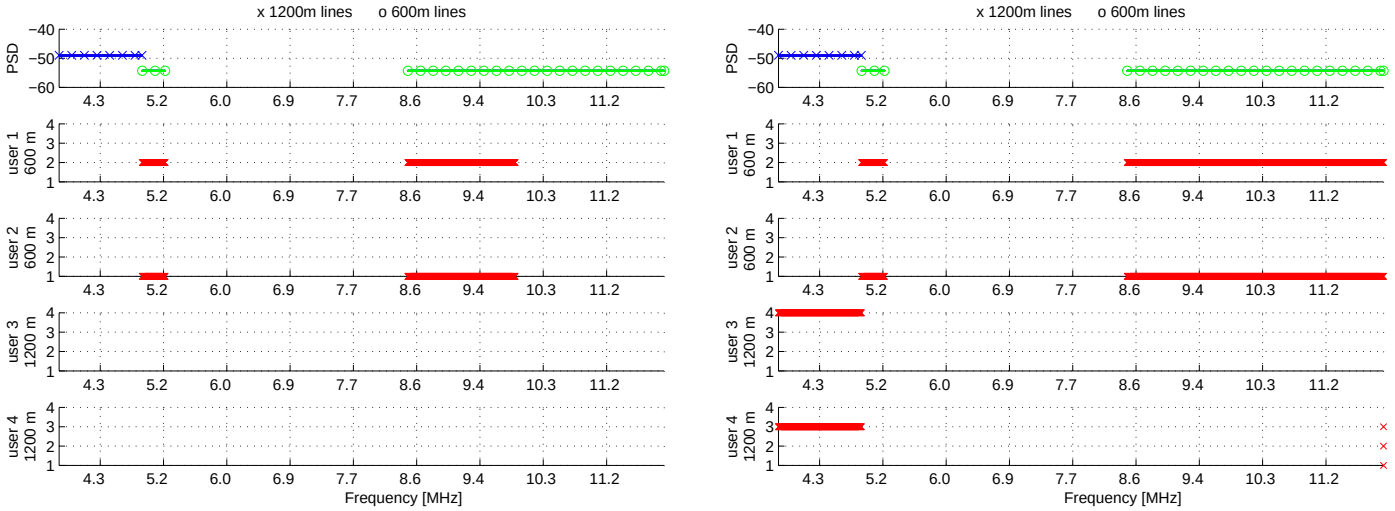
problem are solved jointly, transmit spectra are chosen such that only crosstalk that cannot be cancelled is avoided. This can be seen in figure 10.5(a)(b), where all crosstalk cancellation taps can now be used effectively. Depending on the crosstalk cancellation tap budget, transmit spectra can overlap on frequencies with the highest capacity, resulting in significant performance gains.

When there are restrictions on the signal-level coordination, a choice has to be made as to which lines will be connected to the same line card. In this case, there are two possibilities: connect the lines with similar lengths to the same line card (figure 10.6) or connect lines with different lengths to the same line card (figure 10.10).

Connecting lines with similar lengths to the same line card results in the rate regions of figure 10.7. Only limited performance is gained by increasing the crosstalk cancellation tap budget. This is caused by the fact that the long lines do not have access to the short lines. Therefore, this major source of crosstalk cannot be cancelled. As a result, the spectrum management has to be used to avoid this crosstalk and both groups of lines occupy different frequency bands as shown in figures 10.8 and 10.9. Therefore the joint and independent solutions are similar and only a limited number of crosstalk cancellation taps can be used effectively.

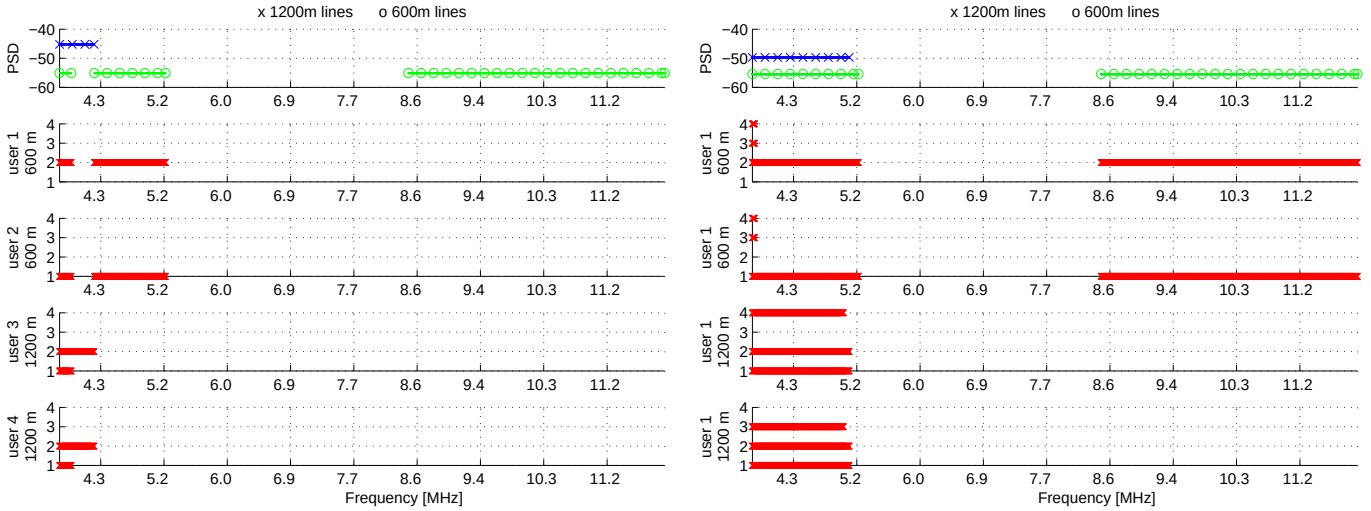
When lines with different lengths are connected to the same line card as shown in figure 10.10, the rate regions of figure 10.11 are obtained. Again, there is no significant difference between the joint and independent solutions. Because the long lines cannot access *all* short lines, there will be severe crosstalk that cannot be cancelled if these lines would use the same frequency band. As a consequence, long and short lines use different frequency bands. Moreover, because lines with similar lengths are on different line cards, no crosstalk cancellation taps can be assigned. This is shown in figures 10.12 and 10.13.

In figure 10.14 the different possibilities for signal-level coordination are summarized for the case where a cancellation tap budget of 30% of full crosstalk cancellation is available. It can be seen that the best performance is obtained when there is full coordination (all lines accessible). Also, in this case there is a significant performance increase when jointly solving the spectrum management and partial crosstalk cancellation problem, reducing the run-time complexity of partial crosstalk cancellation. However, when certain lines are not accessible, the performance of the joint solution is similar to the performance when the spectrum management problem and the partial crosstalk cancellation problem are solved independently. In this case, some performance can be gained by carefully deploying the access network: lines with similar lengths should have signal-level control to maximize performance.



(a) (b)

Figure 10.4: Spectra and cancellation configurations: (a) independent solution, 6% of full cancellation; (b) independent solution, 30% of full cancellation;



(a) (b)

Figure 10.5: Spectra and cancellation configurations: (a) joint solution, 6% of full cancellation; (b) joint solution, 30% of full cancellation;

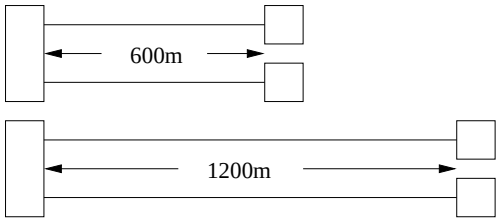


Figure 10.6: 4-user VDSL scenario with constrained signal-level coordination: equal line lengths together

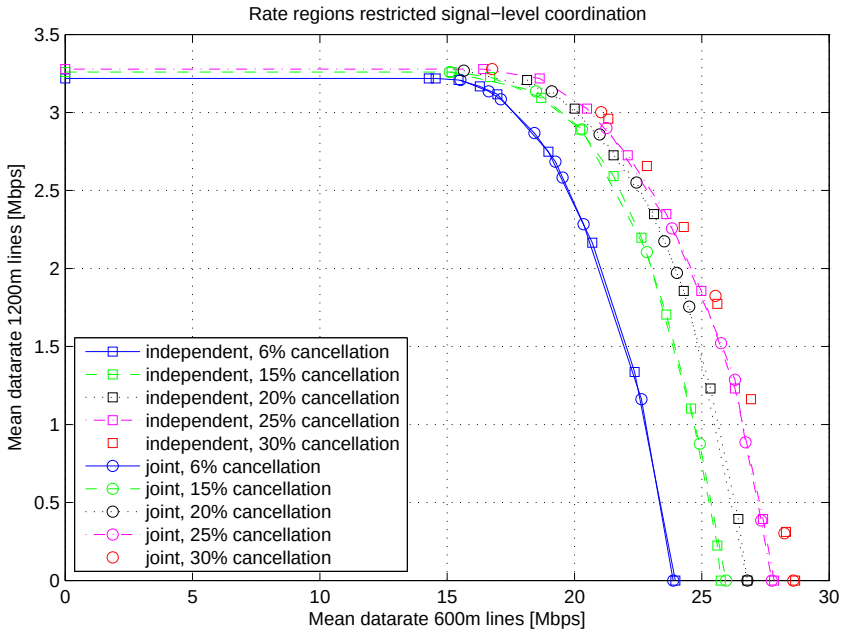
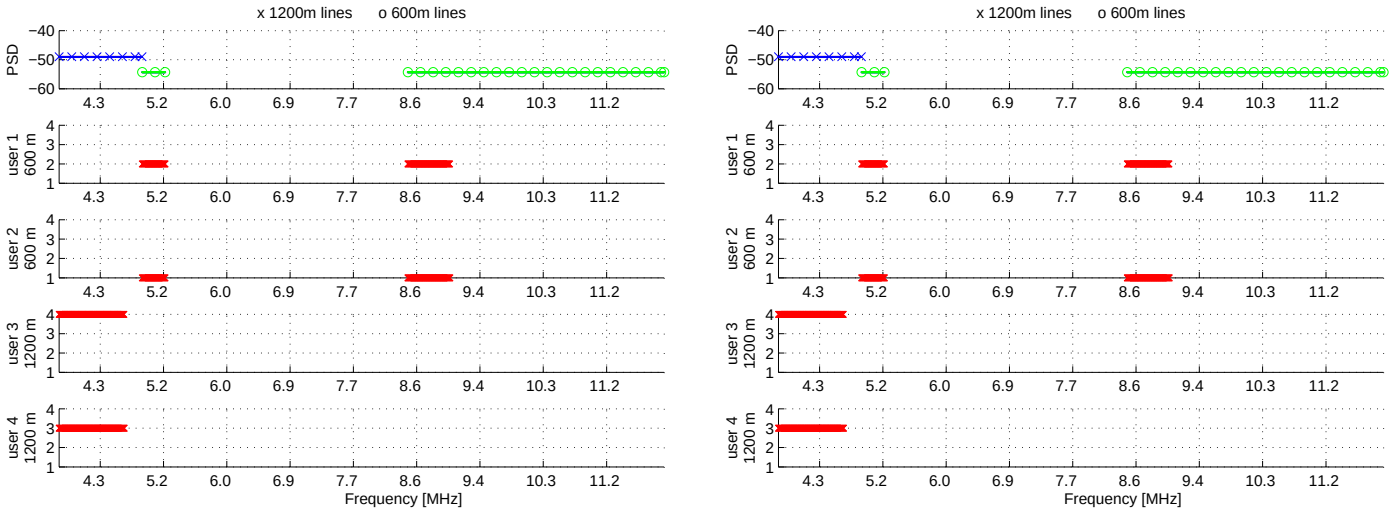
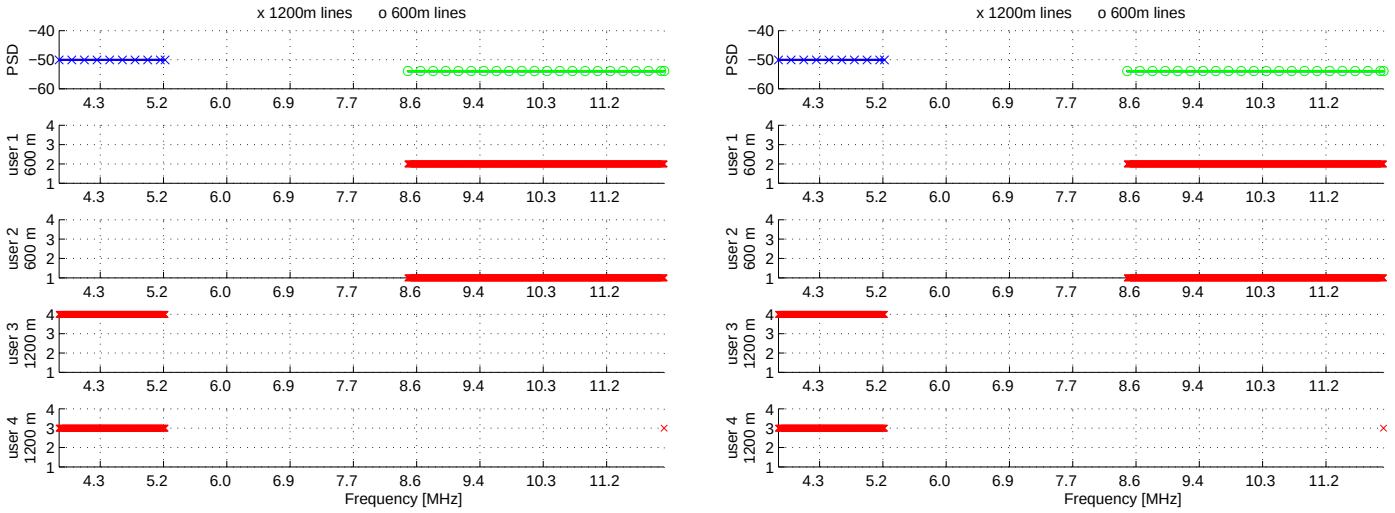


Figure 10.7: Rate regions 4-user VDSL scenario with constrained signal-level coordination: equal line lengths together



(a) (b)

Figure 10.8: Spectra and cancellation configurations with signal-level coordination between lines of equal length: (a) independent solution, 6% of full cancellation; (b) joint solution, 6% of full cancellation;



(a) (b)

Figure 10.9: Spectra and cancellation configurations with signal-level coordination between lines of equal length: (a) independent solution, 30% of full cancellation; (b) joint solution, 30% of full cancellation;

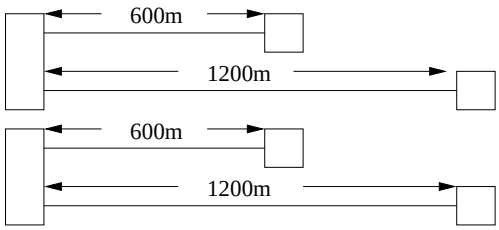


Figure 10.10: 4-user VDSL scenario with constrained signal-level coordination: mixed line lengths

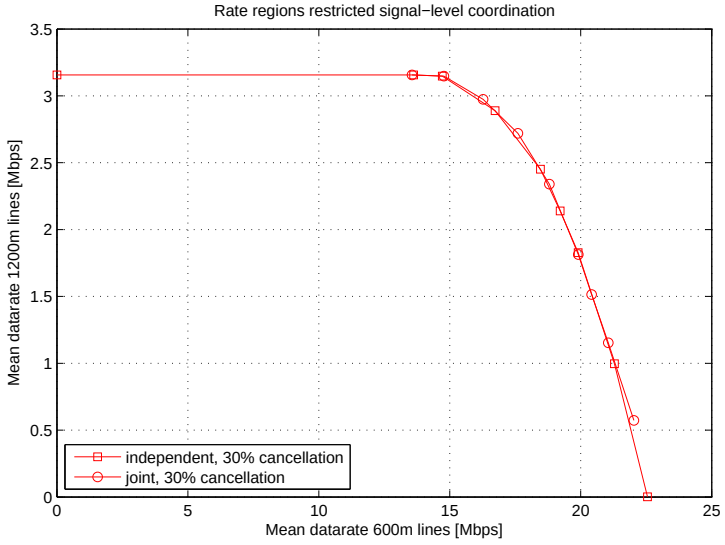
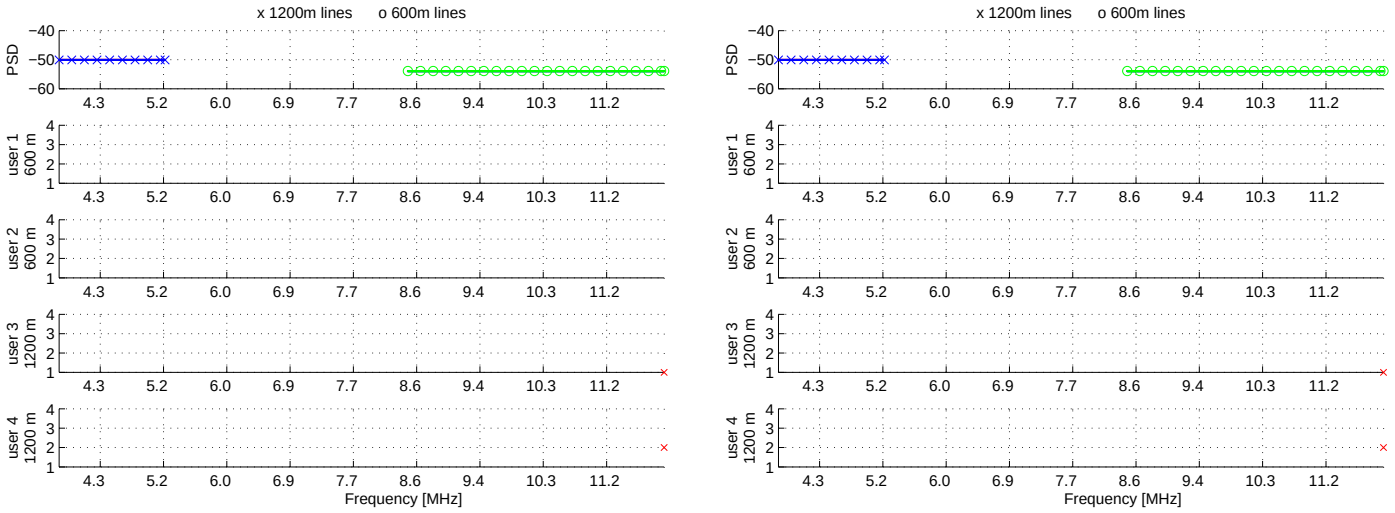
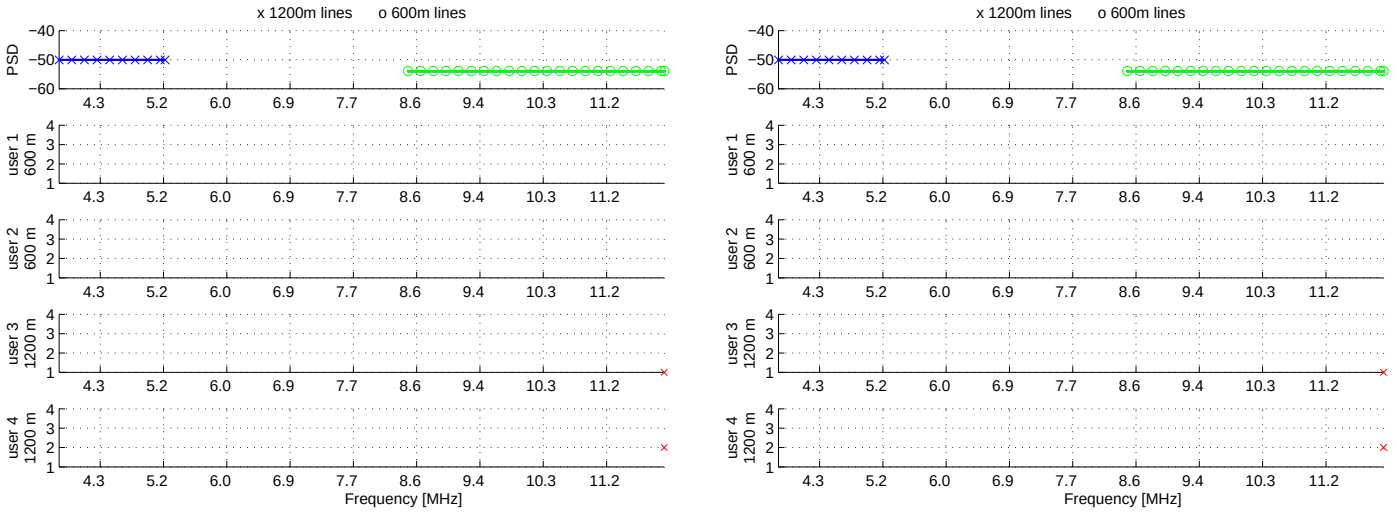


Figure 10.11: Rate regions 4-user VDSL scenario with constrained signal-level coordination: mixed line lengths



(a) (b)

Figure 10.12: Spectra and cancellation configurations with signal-level coordination between lines of different length: (a) independent solution, 6% of full cancellation; (b) joint solution, 6% of full cancellation;



(a) (b)

Figure 10.13: Spectra and cancellation configurations with signal-level coordination between lines of different length: (a) independent solution, 30% of full cancellation; (b) joint solution, 30% of full cancellation;

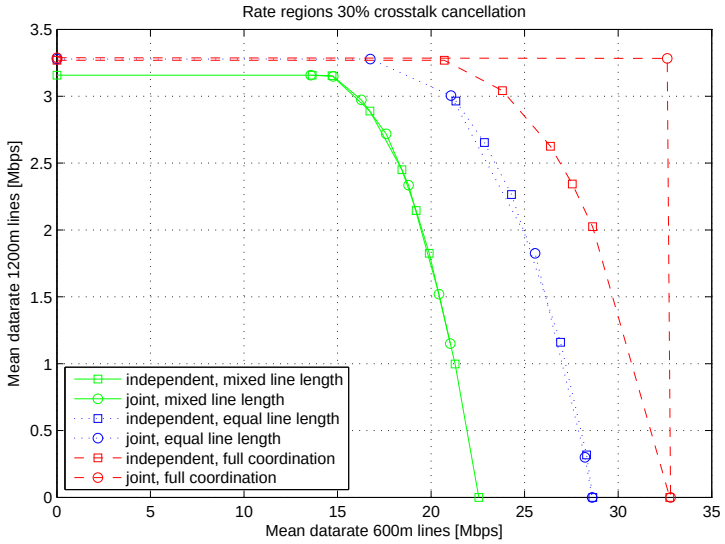


Figure 10.14: Rate regions 4-user VDSL scenario: comparison of constraints on signal-level coordination

10.5 Conclusion

In modern DSL systems, crosstalk is a major source of performance degradation. Crosstalk cancellation techniques have been proposed to mitigate the effect of crosstalk. However, the run-time complexity of these crosstalk cancellation techniques grows with the square of the number of lines. Therefore one has to be selective in cancelling crosstalk to reduce complexity. Secondly, crosstalk cancellation requires signal-level coordination between transmitters or receivers, which is not always available. Because of accessibility constraints, crosstalk between certain lines cannot be cancelled and so has to be mitigated through spectrum management.

In this paper, a solution was presented to jointly solve the spectrum management and constrained partial crosstalk cancellation problem based on a dual decomposition approach. The complexity of the partial crosstalk cancellation part of the solution was reduced based on a line selection and a user independence observation. However, to fully benefit from these observations, power loading has to be applied in the spectrum management part. We have therefore considered ON/OFF power loading, which shows only a minor performance degradation compared to the original power loading.

It was shown that when the spectrum management problem and constrained

partial crosstalk cancellation problem are solved independently, only a limited number of crosstalk cancellation taps can be used effectively because crosstalk is avoided in the first place by the spectrum management. When jointly solving the problems, only crosstalk that cannot be cancelled is avoided, thereby significantly increasing performance. However, this performance increase is only available when there is full signal-level coordination. When only a limited number of lines is accessible, the joint solution provides only little performance increase. It was shown that in this case lines with similar lengths should be made accessible to each other to maximize performance.

Appendix: Search Algorithm for the Lagrange Multipliers

The proof presented in [12] can be easily extended to include $\boldsymbol{\nu}$. First assume a two-user scenario with signal-level control. Starting from two optimal solutions $(R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A}, P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A}, R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A}, P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A}, C^{\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A})$ and $(R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B}, P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B}, R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B}, P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B}, C^{\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B})$ corresponding to $(\boldsymbol{\omega}_A, \boldsymbol{\lambda}_A, \boldsymbol{\nu}_A)$ and $(\boldsymbol{\omega}_B, \boldsymbol{\lambda}_B, \boldsymbol{\nu}_B)$ respectively, optimality for $(\boldsymbol{\omega}_A, \boldsymbol{\lambda}_A, \boldsymbol{\nu}_A)$ implies

$$\begin{aligned} & \omega_{1,A} R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B} + \omega_{2,A} R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B} - \lambda_{1,A} P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B} - \lambda_{2,A} P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B} \\ & -\nu_{1,A} C^{\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B} \\ \leq\ & \omega_{1,A} R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A} + \omega_{2,A} R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A} - \lambda_{1,A} P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A} - \lambda_{2,A} P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A} \\ & -\nu_{1,A} C^{\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A} \end{aligned} \tag{10.14}$$

Optimality for $(\boldsymbol{\omega}_B, \boldsymbol{\lambda}_B, \boldsymbol{\nu}_B)$ implies

$$\begin{aligned} & \omega_{1,B} R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A} + \omega_{2,B} R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A} - \lambda_{1,B} P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A} - \lambda_{2,B} P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A} \\ & -\nu_{1,B} C^{\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A} \\ \leq\ & \omega_{1,B} R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B} + \omega_{2,B} R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B} - \lambda_{1,B} P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B} - \lambda_{2,B} P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B} \\ & -\nu_{1,B} C^{\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B} \end{aligned} \tag{10.15}$$

Taking the sum of (10.14) and (10.15) results in

$$\begin{aligned}
&-\underbrace{\left(\omega_{1,B}-\omega_{1,A}\right)}_{\Delta\omega_1}\underbrace{\left(R^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B}-R^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A}\right)}_{\Delta R^1}\\
&-\underbrace{\left(\omega_{2,B}-\omega_{2,A}\right)}_{\Delta\omega_2}\underbrace{\left(R^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B}-R^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A}\right)}_{\Delta R^2}\\
&+\underbrace{\left(\lambda_{1,B}-\lambda_{1,A}\right)}_{\Delta\lambda_1}\underbrace{\left(P^{1,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B}-P^{1,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A}\right)}_{\Delta P^1}\\
&+\underbrace{\left(\lambda_{2,B}-\lambda_{2,A}\right)}_{\Delta\lambda_2}\underbrace{\left(P^{2,\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B}-P^{2,\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A}\right)}_{\Delta P^2}\\
&+\underbrace{\left(\nu_{1,B}-\nu_{1,A}\right)}_{\Delta\nu_1}\underbrace{\left(C^{\boldsymbol{\omega}_B,\boldsymbol{\lambda}_B,\boldsymbol{\nu}_B}-C^{\boldsymbol{\omega}_A,\boldsymbol{\lambda}_A,\boldsymbol{\nu}_A}\right)}_{\Delta C}\leq 0
\end{aligned} \tag{10.16}$$

Relation (10.16) can be extended straightforwardly to a multi-user scenario with multiple line cards:

$$\begin{bmatrix} -(\Delta\boldsymbol{\omega})^T & (\Delta\boldsymbol{\lambda})^T & (\Delta\boldsymbol{\nu})^T \end{bmatrix}\begin{bmatrix}\Delta\mathbf{R}\\ \Delta\mathbf{P}\\ \Delta\mathbf{C}\end{bmatrix}\leq 0 \tag{10.17}$$

$\boldsymbol{\omega}=[\omega_1,\ldots,\omega_N]$, $\boldsymbol{\lambda}=[\lambda_1,\ldots,\lambda_N]$ and $\boldsymbol{\nu}=[\nu_1,\ldots,\nu_Q]$ are vectors containing the Lagrange multipliers, $\mathbf{R}=[R^1,\ldots,R^N]^T$, $\mathbf{P}=[P^1,\ldots,P^N]^T$ and $\mathbf{C}=[C^1,\ldots,C^Q]^T$ are vectors with the corresponding data rates, total powers and number of cancellation taps.

Bibliography

[1] T. Starr, J. M. Cioffi, P. J. Silverman, *Understanding digital subscriber lines*. Prentice Hall, 1999.

[2] G. Taubock and W. Henkel, "Mimo systems in the subscriber-line network," in *Proc. of the 5th International OFDM Workshop*, Sep 2000, pp. 18.1–18.3.

[3] R. Cendrillon, M. Moonen, R. Suciu, and G. Ginis, "Simplified power allocation and tx/rx structure for MIMO-DSL," in *Proc. of the IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 4, San Francisco, California, USA, Dec 2003, pp. 1842–1846.

[4] G. Ginis and J. M. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE Journal on Selected Areas on Communications*, vol. 20, no. 5, pp. 1085–1104, Jun. 2002.

[5] W. Yu and J. M. Cioffi, "Multiuser detection in vector multiple access channels using generalized decision feedback equalization," in *Proc. of the 5th International Conference on Signal Processing*, Beijing, China, Aug 2000.

[6] R. Cendrillon, G. Ginis, E. V. den Bogaert, and M. Moonen, "A near-optimal linear crosstalk canceler for upstream VDSL," *IEEE Transactions on Signal Processing*, vol. 54, no. 8, pp. 3136–3146, Aug 2006.

[7] R. Cendrillon, G. Ginis, E. V. D. Bogaert, and M. Moonen, "A near-optimal linear crosstalk precoder for downstream VDSL," *IEEE Transactions on Communications*, vol. 55, no. 5, pp. 860–863, May 2007.

[8] R. Cendrillon, M. Moonen, G. Ginis, K. V. Acker, T. Bostoen, and P. Vandaele, "Partial crosstalk cancellation for upstream VDSL," *EURASIP Journal on Applied Signal Processing*, no. 10, pp. 1520–1535, Aug. 2004.

[9] R. Cendrillon, G. Ginis, M. Moonen, and K. V. Acker, "Partial crosstalk precompensation in downstream VDSL," *Signal Processing*, vol. 84, no. 11, pp. 2005–2019, 2004.

[10] T. Starr, M. Sorbara, J. M. Cioffi, P. J. Silverman, *DSL Advances, ch. 11.* Prentice Hall, 2002.

[11] *Spectrum Management for Loop Transmission Systems, ANSI Standard T1.417, Issue 2*, ANSI Std. T1.417, Issue 2, 2003.

[12] P. Tsiaflakis, J. Vangorp, M. Moonen, J. Verlinden, and K. V. Acker, "An efficient search algorithm for the lagrange multipliers of optimal spectrum balancing in multi-user xDSL systems," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 4, May 2006.

[13] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 922–933, May 2006.

[14] H. Bagheri, H. Emami, and M. R. Pakravan, "Iterative joint power control and partial crosstalk cancellation in upstream VDSL," *International Journal of Signal Processing*, vol. 1, no. 3, pp. 195–198, 2004.

[15] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1105–1115, Jun 2002.

[16] W. Yu, R. Lui, and R. Cendrillon, "Dual optimization methods for multiuser orthogonal frequency-division multiplex systems," in *Proc. of the IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 1, Dec 2004, pp. 225–229.

[17] R. Lui and W. Yu, "Low-complexity near-optimal spectrum balancing for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 3, May 2005, pp. 1947–1951.

[18] R. Cendrillon and M. Moonen, "Iterative spectrum balancing for digital subscriber lines," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 3, May 2005, pp. 1937–1941.

[19] *Transmission and Multiplexing (TM); Access transmission systems on metallic access cables; Very high speed Digital Subscriber Line (VDSL); Functional Requirements, ETSI Standard TS 101 270-1 Revision V.1.3.1*, ETSI Std. TS 101 270-1, Rev. V.1.3.1, 2003.

Chapter 11

Joint Spectrum and Receiver Signal Coordination

Multiple Access Channel Optimal Spectrum Balancing for Upstream DSL Transmission

P. Tsiaflakis, J. Vangorp, J. Verlinden, and M. Moonen

Published in *IEEE Communications Letters*, vol. 11, no. 4, Apr. 2007, pp. 298-300.

©2007 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Contributions of first author

- literature study
- co-development of multiple access channel optimal spectrum balancing (MAC-OSB) algorithm
- co-development of multiple access channel iterative spectrum balancing (MAC-ISB) algorithm
- design of simulation scenarios
- software implementation and computer simulation
- co-analysis and co-interpretation of simulation results
- text co-redaction, co-editing and revision

Abstract

Upstream DSL transmission suffers from in-domain crosstalk as well as out-of-domain or alien crosstalk. Here, the use of multi-user receiver signal coordination e.g. generalized decision feedback equalization, can lead to spectacular performance gains. This paper presents a transmission scheme, referred to as MAC-OSB, which focuses on the weighted rate sum capacity by joint receiver signal coordination and transmit spectrum coordination. The proposed scheme incorporates per-user total power constraints, spectral mask constraints and discrete bit or power loading constraints. Furthermore a low-complexity scheme, referred to as MAC-ISB, is presented which performs similar to MAC-OSB. Simulations show large performance gains over existing methods especially for scenarios with significant alien crosstalk.

11.1 Introduction

The increasing demand for high speed services forces DSL systems to use higher frequencies, e.g. up to 30 Mhz for VDSL2. At these frequencies, in-domain crosstalk between different DSL lines in the same cable bundle as well as out-of-domain or alien crosstalk, becomes particularly harmfull and is a dominant source of performance degradation.

By the use of multi-user signal coordination at the transmitter side and/or the receiver side, the crosstalk can be cancelled which leads to large performance gains. In this paper, we are interested in the case where there is only coordination possible at the receiver side. This is referred to as the multiple access channel (MAC) in information theory. The driving application is upstream DSL where multiple transmitters send independent information to one central office (CO) or optical network unit (ONU), which then acts as a joint receiver. Several receiver structures have been proposed to cancel the crosstalk [1][2]. In [3] it is shown that a simple linear zero-forcing receiver performs near-optimally in a DSL environment with only in-domain crosstalk. Unfortunately it will be demonstrated that this receiver performs very sub-optimally in scenarios with significant alien crosstalk. In [4] a procedure is presented for achieving the *unweighted* rate sum capacity in a MAC, consisting of a MMSE-GDFE receiver [2] and per-user iterative waterfilling based transmit power loading (MMSE-GDFE-IWF).

This paper presents a transmission scheme, referred to as *MAC-OSB*, which focuses on the *weighted* rate sum capacity by joint receiver signal coordination and transmit spectrum coordination. The proposed algorithm incorporates per-user total power constraints, spectral mask constraints and discrete bit or power loading constraints. In [5] discrete bit loading was considered for the

single-user case under total power and spectral mask constraints. Furthermore a low-complexity scheme, referred to as MAC-ISB, is presented which performs similar to MAC-OSB. In the simulation section the performance of these schemes is compared with existing schemes.

11.2 System Model

Current DSL systems use Discrete Multi-Tone (DMT) modulation, where the available frequency band is divided in a number of parallel subchannels or tones. The modems in the network are assumed to be synchronized and so each tone is capable of transmitting data independently from other tones. Transmission for a cable bundle of N users can be modelled on each tone k by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k \qquad k = 1 \dots K. \tag{11.1}$$

The vector $\mathbf{x}_k = [x_k^1, x_k^2, \dots, x_k^N]^T$ contains the transmitted signals on tone k for all N users. $\mathbf{z}_k$ is the vector of additive noise on tone k which includes the alien crosstalk. To keep the formulae simple we assume that the noise is pre-whitened $E\{\mathbf{z}_k \mathbf{z}_k^H\} = I_{N\text{x}N}$. $[\mathbf{H}_k]_{n,m} = h_k^{n,m}$ is an $N \times N$ matrix containing the pre-whitened channel transfer functions from transmitter m to receiver n. The diagonal elements are the direct channels, the off-diagonal elements are the in-domain crosstalk channels. The vector $\mathbf{y}_k$ contains the received symbols. We denote the transmit power as $s_k^n \triangleq E\{|x_k^n|^2\}$. The vector containing the transmit power of user n on all tones is $\mathbf{s}^n \triangleq [s_1^n, s_2^n, \dots, s_K^n]^T$.

The upstream DSL channel is a MAC which means that on the one hand a joint decoding can be done at the receiver and on the other hand the transmit spectra can be properly set by each individual transmitter (user). Its sum capacity is given by

$$C_k = \log_2(\det(I + \mathbf{H}_k \mathbf{S}_k \mathbf{H}_k^H)), \tag{11.2}$$

where $\mathbf{S}_k = \text{diag}\{s_k^1, \dots, s_k^N\}$ which is diagonal because transmit signal coordination is indeed not possible.

We will consider total power constraints for every user $\sum_k s_k^n \leq P^{n,tot}, \forall n$, and spectral mask constraints $0 \leq s_k^n \leq s_k^{n,mask}, \forall n, k$.

11.3 Multiple Access Channel Optimal Spectrum Balancing (MAC-OSB)

For conciseness (without loss of generality) we restrict some of the formulae to the 2-user case. The data model (11.1) can then be reformulated as

$$\mathbf{y}_k = \begin{bmatrix} \mathbf{h}_k^1 & \mathbf{h}_k^2 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} + \mathbf{z}_k = \mathbf{h}_k^1 x_k^1 + \mathbf{h}_k^2 x_k^2 + \mathbf{z}_k.$$

Using this reformulation the capacity formula (11.2) can be dissected into the following *unweighted* rate sum

$$\begin{aligned} C_k &= \log_2(\det(\mathbf{I} + \mathbf{h}_k^1 s_k^1 \mathbf{h}_k^{1,H} + \mathbf{h}_k^2 s_k^2 \mathbf{h}_k^{2,H})) \\ &= \log_2(\det(\mathbf{I} + (\mathbf{I} + \mathbf{h}_k^2 s_k^2 \mathbf{h}_k^{2,H})^{-1} \mathbf{h}_k^1 s_k^1 \mathbf{h}_k^{1,H})) \\ &\qquad + \log_2(\det(\mathbf{I} + \mathbf{h}_k^2 s_k^2 \mathbf{h}_k^{2,H})) \qquad (11.3) \\ &= b_k^1 + b_k^2. \end{aligned}$$

The first term of (11.3) represents a bit loading of user 1 when user 1 is detected under crosstalk from user 2. The second term represents a bit loading of user 2 when the detection of user 2 is done after having removed the crosstalk from user 1. This exactly corresponds to the operation of the MMSE-GDFE receiver [2] and thus this receiver is indeed (unweighted rate sum) optimal for given transmit powers. Note that changing the detection order changes per-user rates, but does not change the rate sum. In [4] it is proved that if the MMSE-GDFE receiver is combined with per-user iterative waterfilling (IWF), the *unweighted* rate sum is maximized under per-user transmit power constraints (MMSE-GDFE-IWF).

Our objective will be to optimize the *weighted* rate sum

$$\tilde{C}_k = \omega_1 b_k^1 + \omega_2 b_k^2, \quad \omega_1, \omega_2 \geq 0 \qquad (11.4)$$

where ω_1 and ω_2 are weights given to the different users. Our overall optimization problem is

$$\begin{array}{ll} \max_{(\mathbf{s}^1,\ldots,\mathbf{s}^N,\text{receiver})} & \sum_{k=1}^K \tilde{C}_k \\ \text{subject to} & \sum_k s_k^n \leq P^{n,tot} \quad \forall n \\ & 0 \leq s_k^n \leq s_k^{n,mask} \quad \forall n,k, \end{array}$$

where the optimization is over transmit spectra as well as receiver structures.

Theorem 11.1 *For given transmit spectra, the MMSE-GDFE receiver is the optimal receiver for the weighted rate sum (11.4). The weights define the detection order where users with the smallest weights are detected first and users with the largest weights are detected last.*

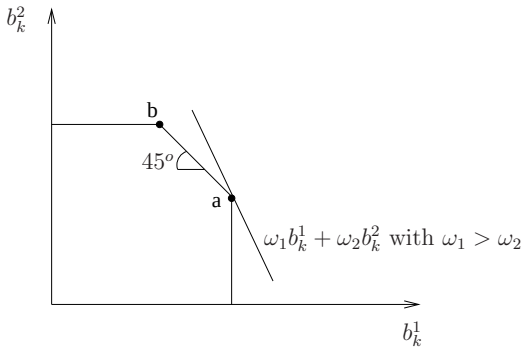


Figure 11.1: Rate region on tone k

Proof: For given transmit spectra, the achievable rate region (for each tone) is known. This is a pentagon with one side making an angle of -45 degrees as shown in Figure 11.1 (2-user case). The two end points a,b of this side are achievable by a MMSE-GDFE receiver. The weighted rate sum optimum is the crossing of the pentagon and the tangent defined by the weights ω_1, ω_2. As the weights are always positive the crossing has to be one of the points a or b. As these points are achievable by the MMSE-GDFE receiver, this receiver is optimal. If $\omega_1 > \omega_2$ the optimum is point a (user 2 detected first), otherwise the optimum is point b (user 1 detected first).

The remaining question is how to find weighted rate sum optimal transmit spectra. This can be done similarly to the Optimal Spectrum Balancing (OSB) dual decomposition procedure [6][7] which finds the optimal power loading for uncoordinated DSL transmission. The basic idea here is to use Lagrange multipliers to move the total power constraints into the cost function and so decouple the problem into independent per-tone optimization problems as follows:

$$\max_{s_k^1,\ldots,s_k^N} \tilde{C}_k - \sum_{i=1}^{N} \lambda_i s_k^i \quad \text{s. t. } 0 \le s_k^n \le s_k^{n,mask} \quad \forall n. \tag{11.5}$$

By tuning the Lagrange multipliers λ_i the total power constraints can be enforced. In [7] an efficient algorithm is presented to find the optimal Lagrange multipliers. For given Lagrange multipliers the optimal power loading for the per-tone optimization problem (11.5) can be found by an exhaustive search over all possible power loadings. The following formulae can be used to calculate the bit loadings corresponding to certain power loadings for $\omega_1 < \omega_2$ (2-user case)

$$b_k^1 \;=\; \log_2(1 + \frac{1}{\Gamma} s_k^1 \mathbf{h}_k^{1,H} (\mathbf{I} + \mathbf{h}_k^2 s_k^2 \mathbf{h}_k^{2,H})^{-1} \mathbf{h}_k^1) \tag{11.6}$$

$$b_k^2 = \log_2(1 + \frac{1}{\Gamma} s_k^2 \mathbf{h}_k^{2,H} \mathbf{h}_k^2) \quad (11.7)$$

Note the addition (compared to (11.3)) of the SNR-gap to capacity Γ which takes into account QAM transmission and is a function of the desired BER, coding gain and noise margin. Note that by including the SNR-gap the theoretical optimality of the scheme is lost, as it is also the case for MMSE-GDFE-IWF [4]. However the proposed scheme provides an achievable rate region which is shown in the simulation section to outperform existing state-of-the-art schemes especially for scenarios with significant alien crosstalk.

This leads to following transmission scheme inspired by OSB to perform optimal transmit spectrum coordination (power loading) under MMSE-GDFE receive coordination, which we will refer to as *multiple access channel optimal spectrum balancing* (MAC-OSB):

Algorithm 11.1 MAC-OSB

pre-whiten noise
initialize weights and Lagrange multipliers
repeat
  update Lagrange multipliers
  per-tone exhaustive search over discrete power loadings
until total power constraints are satisfied

Note that, alternatively, achievable rate regions can be computed for the case of discrete (integer) bit loadings, based on exhaustive searches over all possible bit loadings, and where the corresponding power loadings s_k^1,s_k^2 are computed by inverting (11.6) and (11.7).

11.4 Multiple Access Channel Iterative Spectrum Balancing (MAC-ISB)

The per-tone optimization problem of MAC-OSB is solved by an exhaustive search, with a complexity that is exponential in the number of users, i.e. B^N where B is the number of possible bit or power loadings and N is the number of users. For more than five users this becomes computationally intractable. In order to reduce this complexity one can resort to an iterative approach. This approach iteratively finds the best bit/power loading for user 1 assuming all other bit/power loadings are fixed, then similarly the loading for user 2 is optimized, etc., until user N. Then this is repeated until the loadings do not change anymore. The complexity reduces to a linear complexity LBN where L is the number of iterations which is typically very small (2 or 3). Because of the iterative approach we refer to this method as multiple access

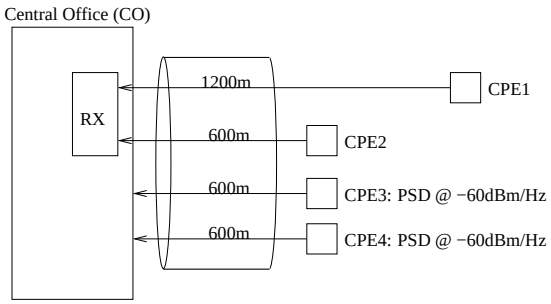


Figure 11.2: VDSL upstream scenario

channel iterative spectrum balancing (MAC-ISB). Simulations show that this algorithm performs similar to MAC-OSB for many scenarios.

11.5 Simulation Results

The VDSL upstream scenario considered here is shown in Figure 11.2. At the CO the lines connected to CPE1 and CPE2 can be coordinated. DSL users CPE3 and CPE4 are two VDSL alien crosstalkers within the same cable bundle transmitting at -60 dBm/Hz. Bandplan FDD 988 is used. The line diameters are 0.4 mm. The maximum transmit power per user is 11.5 dBm for CPE1 and CPE2. The SNR gap Γ is 12.9 dB, corresponding to a coding gain of 3 dB, a noise margin of 6 dB and a target symbol error probability of 10^{-7}. The maximum bit loading is 16 bits. The goal is to maximize the upstream bit rates of CPE1 and CPE2 by choosing optimal transmit spectra with receiver coordination at the CO.

Figure 11.3 shows the rate regions in the absence of the two alien crosstalkers CPE3 and CPE4. In this case the noise is spatially white, and so pre-whitening is not needed, which would otherwise destroy the column-wise diagonal dominant structure (CWDD) of the channel [3]. Therefore the linear zero-forcing receiver in combination with per-user waterfilling has a near-optimal performance [3] which is very close to the performance of MAC-OSB, MAC-ISB and MMSE-GDFE-IWF. The rate region for MAC-OSB and MAC-ISB is obtained by varying the weights $\omega_i,\ i = 1, \ldots, N$. The rate region of MMSE-GDFE-IWF is created by keeping the total power of the users (except one) at the maximum and reducing one of the user's power to zero. Furthermore the rate

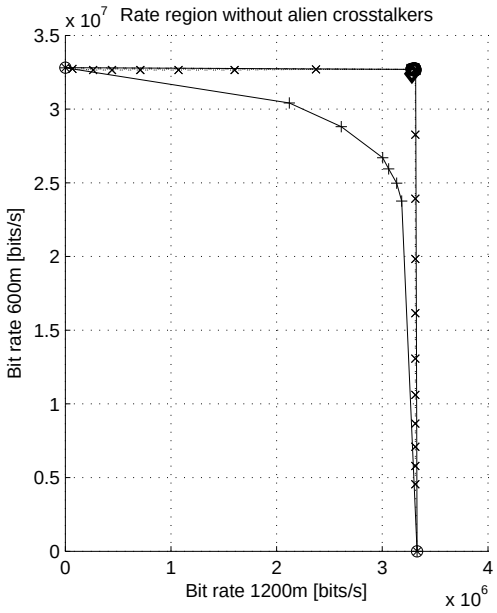


Figure 11.3: Rate region without alien crosstalkers (same legend as Figure 11.4)

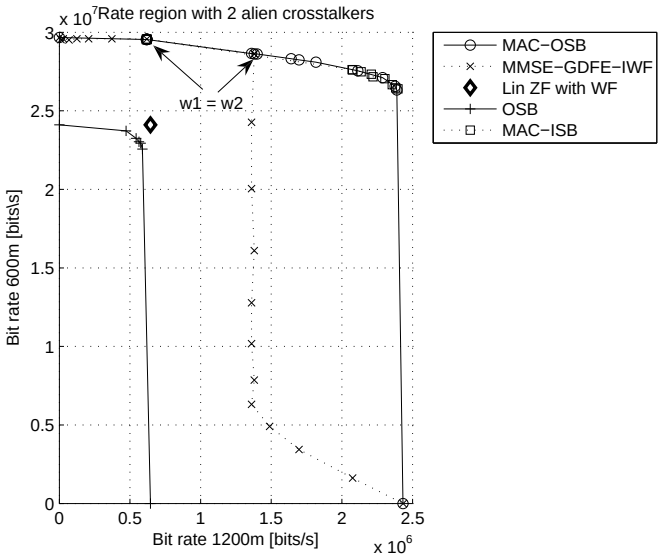


Figure 11.4: Rate region with two alien crosstalkers

region is plotted for OSB which shows the optimal performance for no receiver coordination and optimal transmit spectrum shaping.

Figure 11.4 shows the results of the scenario with the presence of the two alien crosstalkers. It can be seen that the linear zero-forcing receiver performs very sub-optimally. MAC-OSB has a much better performance than MMSE-GDFE-IWF. The MAC-ISB algorithm performs similar to MAC-OSB. Finally it is important to notice that the use of receiver coordination increases the bit rates significantly.

11.6 Conclusion

A transmission scheme, referred to as MAC-OSB, has been presented which focuses on the weighted rate sum capacity in upstream VDSL scenarios by joint receiver signal coordination and transmit spectrum coordination. Furthermore a low-complexity scheme, referred to as MAC-ISB, is presented which performs similar to MAC-OSB. Simulations show that these algorithms outperform existing methods especially for scenarios with significant alien crosstalk.

Bibliography

[1] G. Ginis and J. M. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1085–1104, Jun. 2002.

[2] J. M. Cioffi and G. D. Forney Jr., "Generalized decision feedback equalization for packet transmission with ISI and Gaussian noise," *Communications, Computation, Control and Signal Processing, Kluwer Academic, Boston, Mass, USA*, pp. 79–127, 1997.

[3] R. Cendrillon, G. Ginis, E. V. den Bogaert, and M. Moonen, "A near-optimal linear crosstalk canceler for upstream VDSL," *IEEE Transactions on Signal Processing*, vol. 54, no. 8, Aug. 2006.

[4] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple access channels," *IEEE Transactions on Information Theory*, vol. 50, no. 1, pp. 145–151, Jan. 2004.

[5] E. Baccarelli, A. Fasano, and M. Biagi, "Novel efficient bit-loading algorithms for peak-energy-limited ADSL-type multi-carrier systems," *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1237–1247, May 2002.

[6] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum management for digital subscriber lines," *IEEE Transactions on Communications*, vol. 54, no. 5, May 2006.

[7] P. Tsiaflakis, J. Vangorp, M. Moonen, J. Verlinden, and K. V. Acker, "An efficient lagrange multiplier search algorithm for optimal spectrum balancing in crosstalk dominated xDSL systems," in *Proc. of the IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP)*, May 2006.

Chapter 12

Conclusions and Further Research

12.1 Summary and Conclusions

Crosstalk is *the* dominant source of performance degradation of current DSL broadband access systems. Multi-user coordination techniques such as spectrum and signal coordination are recognized as a key approach for crosstalk mitigation.

This thesis focused on resource management and optimization for such spectrum and signal coordination, where the specific main goals were:

- the design of low-complexity spectrum coordination algorithms with near-optimal performance even for large-scale DSL scenarios,
- the design of distributed spectrum coordination algorithms,
- the extension of the limited physical layer data rate maximization approach of spectrum coordination towards a much wider set of objectives and constraints that are better aligned with the application and network QoS requirements,
- the design of low-complexity resource management and optimization algorithms for joint spectrum and signal coordination.

First, in Section 1.6, an overview was given of state-of-the-art spectrum coordination where the system model, spectrum coordination problems and state-of-the-art spectrum coordination algorithms were described. It was explained that typical spectrum coordination problems correspond to NP-hard non-convex optimization problems, where the transmit powers across all users and tones are

the optimization variables. The algorithm overview mainly focused on dual spectrum coordination algorithms, which tackle the dual spectrum coordination problem formulation and consist of a dual decomposition approach that decomposes the problem into a small-dimension convex master problem and many decomposed small-dimension non-convex slave problems. This lead to two general solution structures as given in Figures 1.17 and 1.18. It is shown that existing dual spectrum coordination algorithms follow one of these two structures, where each algorithm has its particular approaches so as to tackle the corresponding master and slave problems. Then in Section 1.7, a short overview was given of state-of-the-art signal coordination, partial signal coordination and joint spectrum and signal coordination.

In Chapter 4 it was shown that a particular spectrum coordination problem, which was believed to be nonconvex, can actually be transformed into a geometric programming problem, i.e. a convex problem, under specific conditions on the channel environment and transmit power resources. These conditions were characterized using M-matrix theory. Intuitively, the conditions were shown to correspond to DSL scenarios with a small crosstalk level. These conditions provide the first polynomial-time optimality certificate for the particular spectrum coordination problem and for a finite number of tones. A novel spectrum coordination algorithm, referred to as *DSM-GP*, was then proposed. Simulation results demonstrated the usefulness of this optimality certificate for realistic DSL scenarios.

In Chapter 2 the branch-and-bound optimal spectrum balancing (BB-OSB) algorithm was proposed that uses the dual decomposition solution structure of Figure 1.17. It consists of a subgradient update approach with improved stepsize tuning for the master problem. The decomposed slave problems are tackled by a dedicated branch-and-bound approach. BB-OSB is a centralized algorithm that achieves the same globally optimal performance as the existing OSB algorithm [1], but significantly reduces its computational complexity. More specifically, for a 4-user ADSL downstream scenario, BB-OSB succeeds in finding the optimal transmit power allocation in less than 10 minutes, where it takes more than 5 days for OSB. Furthermore BB-OSB allows to compute the optimal transmit power allocations of up to 7-user ADSL downstream scenarios in less than one week of MATLAB simulation time, whereas OSB can only go up to the 4-user case within the same time frame. Although the complexity of BB-OSB remains exponential in the number of users N, it can be used as a benchmark to assess the performance of fast suboptimal spectrum coordination algorithms.

In Chapter 3, a number of distributed spectrum coordination algorithms were proposed with a different level of required message-passing. ASB2 is a fully autonomous algorithm and can be seen as a fixed point update version of the ASB algorithm [2]. CA-DSB and DSB are both distributed algorithms, where the modems update their transmit spectrum based on local measures and in

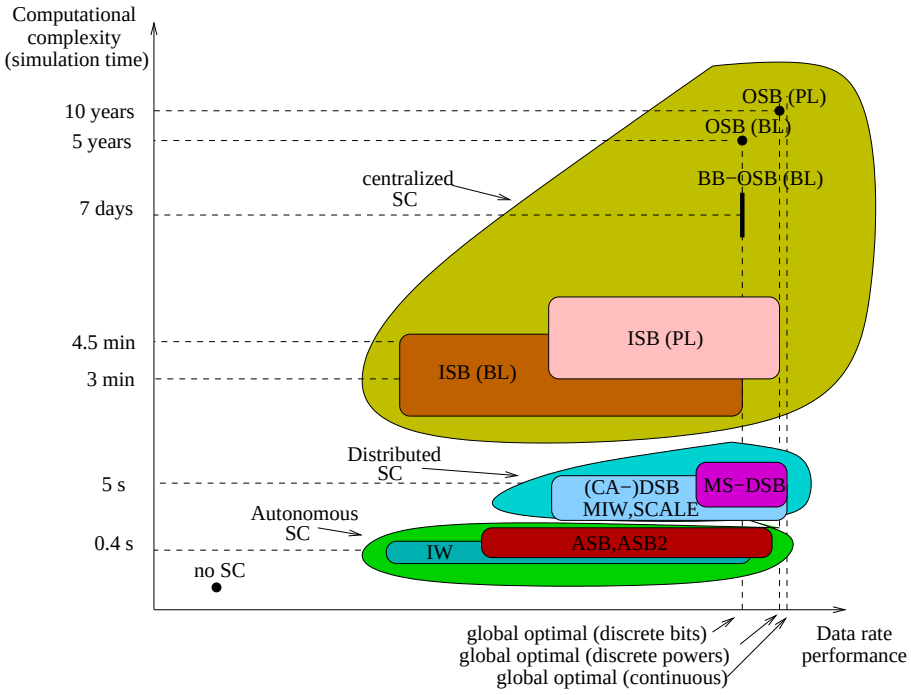


Figure 12.1: Computational complexity versus data rate performance for several SC algorithms. Computational complexity is expressed in MATLAB simulation time required for computing a 7-user ADSL downstream scenario. Data rate performance refers to weighted data rate performance. BL refers to discrete (integer) bit loading, and PL refers to discrete power loading with granularity Δ_s 0.5 dBm/Hz.

addition they are steered by a centralized controller through message-passing. Both algorithms follow a similar iterative convex approximation based design, where CA-DSB follows the solution structure of Figure 1.18 and DSB follows the solution structure of Figure 1.17. CA-DSB and DSB both succeed in finding a locally optimal solution of the corresponding spectrum coordination problem. Note that in [3] a similar algorithm is proposed as the DSB algorithm, which is referred to as modified iterative water-filling (MIW). A fourth proposed algorithm was the MS-DSB algorithm. It consists of a multiple starting point approach to tackle the nonconvex slave problems, where the starting points are carefully chosen along the axes. Extensive simulation results showed that MS-DSB performs globally optimal for practical DSL scenarios, where DSL scenarios up to 8 users can be tackled within a minute of simulation time. For comparison with OSB and BB-OSB, MS-DSB tackles a 4-user ADSL downstream scenario in less than one second.

To give a more concrete view of the relation among the different spectrum coordination algorithms, we have generated the graph of Figure 12.1. It shows the trade-off between the computational complexity and the data rate performance for the spectrum coordination algorithms proposed in this thesis and the state-of-the-art spectrum coordination algorithms as discussed in Section 1.6.3. The computational complexity axis is expressed in terms of MATLAB simulation time, where the concrete simulation times are given so as to simulate a seven-user ADSL downstream scenario. The data rate performance axis indicates the data rate performance in terms of the weighted sum of data rates. The main goal of this graph is to give an idea of the relative complexity and performance among the corresponding spectrum coordination algorithms.

Some of the spectrum coordination algorithms occupy a larger area in the graph than others. This models the variability of the corresponding algorithms as the performance and the complexity can depend on many factors, such as the iteration order, the initial starting points, the particular DSL scenario, the chosen reference line, etc. It can be seen that BB-OSB, which uses discrete bit loading, has the same data rate performance as the OSB (BL) algorithm, but significantly reduces the complexity, i.e. from 5 years to 7 days. The ISB algorithms are much faster than the global optimal algorithms (OSB, BB-OSB) but at the cost of optimality. The distributed algorithms succeed in finding a solution for a 7-user ADSL downstream scenario within a few seconds, whereas the autonomous algorithms require even less than one second. The performance of the autonomous algorithms can however be quite suboptimal and depends for instance on the chosen virtual reference line in the case of ASB(2). The distributed algorithms perform quite well in terms of data rate performance, especially the MS-DSB algorithm, which only requires slightly more complexity than the other distributed algorithms such as CA-DSB, DSB, MIW and SCALE.

In Chapter 8 an improved dual decomposition approach was proposed for spectrum coordination. Many spectrum coordination algorithms are based on a standard (sub)gradient-based dual decomposition approach. For large-scale scenarios, this approach is found to lead to slow convergence or even no convergence at all, one of the reasons being the very difficult tuning of the stepsize parameters. The proposed improved dual decomposition approach is inspired by recent advances in mathematical programming. It uses a smoothing technique for the Lagrangian combined with an optimal gradient based scheme for updating the Lagrange multipliers. The stepsize parameters are furthermore selected optimally removing the need for a tuning strategy. It was shown that this novel approach improves on the convergence of current state-of-the-art spectrum coordination algorithms based on iterative convex approximations (e.g. CA-DSB[4], SCALE[5]) with one order of magnitude without increasing the computational complexity. This resulted in a set of improved algorithms which are referred to as I-CA-DSB and I-SCALE. The improved dual decom-

position approach was then applied to other spectrum coordination algorithms (OSB, ISB, ASB, ...) and further improvements were proposed so as to improve the robustness of the algorithms. The effectiveness of the improved dual decomposition approach was demonstrated for several realistic multi-user DSL scenarios. The proposed improved dual decomposition based optimization approach can be seen as an important step towards obtaining numerically fast and robust spectrum coordination algorithms for large-scale DSL scenarios.

Chapters 5, 6 and 7 then focused on the extension of the traditional physical layer data rate maximization approach of spectrum coordination towards a much wider scope. In Chapter 5 an initial framework was proposed to study upper layer performance by looking at scheduling and spectrum coordination together. This was obtained by extending the system model with a data buffer for each user and with dynamic bit/packet input arrivals. This extension allowed to incorporate upper layer performance metrics such as throughput, stability and delay performance into the spectrum coordination problem formulation. It was then shown how optimal scheduling can be combined with optimal spectrum coordination. Then, somewhat surprisingly, it was shown that suboptimal spectrum coordination algorithms can achieve throughput maximization, when intelligent scheduling algorithms are employed. These were referred to as δ-randomized DSM algorithms, which only require polynomial complexity with a trade-off in complexity and delay performance.

In Chapter 6 the traditional physical layer data rate maximization formulation was extended to a much wider set of energy-aware spectrum coordination problems, which was referred to as "*Green DSL*". A general solution approach was then proposed to tackle these problems. Using this "Green DSL" approach, the energy-rate trade-off was quantified for different practical settings. It was demonstrated that large energy savings can be achieved with only a minor degradation in data rate performance. This result is due to not just the logarithmic dependence of rate on the signal-to-interference-and-noise ratio but also to the characteristics of crosstalk in DSL systems. Furthermore, bounds were proposed on the trade-off between data rate performance and energy saving.

Chapter 7 introduced the concept of *fair greening*. It was first shown that straightforward application of Green DSL can sometimes lead to very unfair data rate and power allocations. Then four formulations of fair greening were proposed to tackle this unfairness problem. These formulations were unified into one general representation, and a unified algorithm, referred to as *General-Green-DSL*, was proposed to solve it. Simulation results were then shown which quantify the intuitions on fairness for the four proposed fair greening formulations, as they offer a range of choices between maintaining a high sum data rate and enforcing various definitions of fairness. The general framework provided in this chapter, together with its general solution approach (General-Green-DSL) can actually be used for modelling a much wider set of spectrum coordination problems that involve different types of utilities on the data rates

and total transmit powers. Furthermore this framework can be extended with the improved dual decomposition approach of Chapter 8.

Finally, Chapters 9, 10 and 11 focused on spectrum and signal coordination. More specifically in Chapter 9 a dual decomposition approach for partial signal coordination was proposed that requires less computational complexity than the existing resource allocation algorithm as proposed in [6]. This was then extended to the multi-user case. In Chapter 10 a low-complexity solution was proposed for joint spectrum coordination and partial signal coordination based on a dual decomposition approach. The spectrum coordination part was limited to a type of ON/OFF spectrum coordination so as to reduce the computational complexity. The resulting algorithm was compared to existing algorithms that consist of independent spectrum coordination and partial signal coordination. Chapter 11 focused on joint spectrum and receiver signal coordination. More specifically, a transmission scheme was presented, referred to as MAC-OSB, which focuses on the weighted rate sum capacity and consists of spectrum coordination in combination with the generalized decision feedback equalizer. The proposed scheme incorporates per-user total power constraints, spectral mask constraints, and discrete bit or power loading constraints. Furthermore a low-complexity scheme, referred to as MAC-ISB, was presented which performs similar to MAC-OSB. Simulation results were shown that demonstrate the large performance gains over existing schemes, especially for scenarios with significant spatially correlated alien crosstalk.

12.2 Further Research

Joint Spectrum and Signal Coordination

The application of signal coordination can cancel the impact of crosstalk, which can result in huge performance gains. However, full signal coordination is not always possible due to (1) complexity and/or power restrictions and (2) physical restrictions, i.e. not all transmitters and/or receivers are colocated.

For these cases, one may combine spectrum and signal coordination, where signal coordination is applied to colocated subsets of transmitters and/or receivers and in addition the transmit spectra are coordinated over all subsets. This can be seen as a combination of multiple access or broadcast channels and interference channels for which the optimal information-theoretic strategy is not known. The optimal transmission scheme for upstream and downstream transmission and for zero or non-zero SNR gap approximation is still an open problem.

The design of low-complexity distributed procedures for such mixed settings can be seen as highly interesting and relevant.

Robust Spectrum and Signal Coordination

Current algorithms for spectrum and signal coordination assume that the channel environment is perfectly known and that the channel does not change during showtime. In practice, however, only an estimate of the channel environment is known. Furthermore, the channel is subject to time-variation due to change in temperature and weather conditions but also due to interfering users turning their modem on and off. This imperfect channel knowledge may lead to suboptimal performance and even instability of the DSL transmission.

Robust spectrum and signal coordination consists of a new approach where the uncertainty of the channel and its time-variation is incorporated into the optimization problem. This requires a number of components that need to be studied carefully. First, good channel estimation algorithms are required so as to obtain good channel estimates. Secondly, it is important to obtain good estimates of the time-variation distributions. These estimates would then have to be incorporated into the design of spectrum and signal coordination techniques so as to obtain procedures that are robust against these uncertainties.

Dynamic Cross-layer Spectrum and Signal Coordination

Spectrum and signal coordination have mainly focused on so-called physical layer resource allocation, where physical layer resources (e.g. transmit spectrum and complexity) are allocated so as to obtain an optimal but static operation point (e.g. data rates/bit error rate), independent of the actual requirements of the applications of the different users.

One interesting direction is to extend the static physical layer optimization approach to a dynamic application-driven cross-layer optimization approach. Chapter 5 already provided an initial step towards this direction by focusing on spectrum coordination and scheduling together. However this can be extended towards a much wider scope and could consist of (1) the design of an extended cross-layer problem formulation that takes into account the specific DSL infrastructure as well as models requirements of new and future applications, and (2) the design of new spectrum and signal coordination algorithms based on the extended cross-layer model.

Spectrum and Signal Coordination Problem Modeling

Chapter 7 proposed a general Green DSL problem formulation for spectrum coordination that can model a wide range of objective functions and constraints. Furthermore, a unifying solution approach, referred to as General-Green-DSL, was proposed to tackle the general problem formulation. This can be extended further with the improved dual decomposition based optimization approach of Chapter 8. However, it is still important to identify which exact objectives

and constraints should be used in this framework. We provide some possible directions on this in the following:

Four energy-fairness objectives were proposed in Chapter 7, that model a different type of fairness. However, the concept of fairness is not well defined. Recent studies are exploring fairness using an axiomatic theory approach [7, 8]. It is important to identify which fairness objective may be relevant to the specific DSL setting.

Secondly, it is important to identify the correct objectives for energy-efficiency. Current energy-aware objectives mainly consist of linear functions of the transmit powers. However, the true relation between the allocated transmit power and the actual power consumption of a modem is still under consideration. The identification of this objective would allow to formulate a Green DSL formulation that minimizes the modem's power consumption directly instead of the transmit power allocation.

Thirdly, within the setting of partial signal coordination it is important to identify the relation between adding a crosstalk canceler and the related power consumption. This would enable a Green DSL framework for joint spectrum and partial signal coordination.

Extension of DSL Spectrum and Signal Coordination to Other Systems

In this thesis, multi-user resource management and optimization techniques for spectrum and signal coordination were designed for the specific area of DSL broadband access. However, these techniques can be extended towards other systems such as power line communications and wireless communication systems. As these systems are characterized by different transmission characteristics, it would be interesting to investigate how existing spectrum and signal coordination techniques should be modified to obtain efficient crosstalk mitigation procedures for these specific systems.

Bibliography

[1] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *Communications, IEEE Transactions on*, vol. 54, no. 5, pp. 922–933, May 2006.

[2] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spectrum balancing for digital subscriber lines," *Signal Processing, IEEE Transactions on*, vol. 55, no. 8, pp. 4241–4257, Aug. 2007.

[3] W. Yu, "Multiuser water-filling in the presence of crosstalk," in *Proc. of the Information Theory and Applications (ITA)*, Feb. 2007.

[4] P. Tsiaflakis, M. Diehl, and M. Moonen, "Distributed spectrum management algorithms for multiuser dsl networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4825–4843, Oct. 2008.

[5] J. Papandriopoulos and J. S. Evans, "Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks," in *Proc. of the IEEE International Conference on Communications (ICC)*, vol. 7, Jun. 2006, pp. 3270–3275.

[6] R. Cendrillon, M. Moonen, G. Ginis, K. V. Acker, T. Bostoen, and P. Vandaele, "Partial crosstalk cancellation for upstream VDSL," *EURASIP Journal on Applied Signal Processing*, no. 10, pp. 1520–1535, Aug. 2004.

[7] T. Lan, D. Kao, and M. Chiang, "An axiomatic theory of fairness," Tech. Rep., May 2009.

[8] H. Boche, S. Naik, and M. Schubert, *Axiomatic framework for resource allocation under fairness and efficiency constraints*, ser. ICT Shaping the World. John Wiley & Sons Ltd, 2009, ch. 9.

Publication List

International Journal Papers

1. P. Tsiaflakis, J. Vangorp, J. Verlinden, and M. Moonen, "Multiple access channel optimal spectrum balancing for upstream DSL transmission," *IEEE Communications Letters*, vol. 11, no. 4, pp. 298–300, Apr. 2007.

2. J. Vangorp, P. Tsiaflakis, M. Moonen, J. Verlinden, and G. Ysebaert, "A dual decomposition approach to partial crosstalk cancellation in a multiuser DMT-xDSL environment," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 37963, pp. 1–11, Jun. 2007.

3. P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "A low complexity optimal spectrum balancing algorithm for digital subscriber lines," *Signal Processing*, vol. 87, no. 7, pp. 1735–1753, Jul. 2007.

4. J. Vangorp, P. Tsiaflakis, M. Moonen, and J. Verlinden, "Joint spectrum management and constrained partial crosstalk cancellation in a multi-user xDSL environment," *Signal Processing, Special Section: Information Proessing and Data Management in Wireless Sensor Networks*, vol. 87, no. 12, pp. 3131–3146, Dec. 2007.

5. P. Tsiaflakis, M. Diehl, and M. Moonen, "Distributed spectrum management for multiuser DSL networks," *IEEE Transactions on Signal Processing*, vol. 87, no. 10, pp. 4825–4843, Oct. 2008.

6. P. Tsiaflakis and M. Moonen, "Distributed spectrum coordination for DSL broadband access," *Journal on Control Engineering and Applied Informatics*, vol. 11, nr. 3, pp. 41–48, Sept. 2009.

7. P. Tsiaflakis, I. Necoara, J.A.K. Suykens, and M. Moonen, "Improved dual decomposition based optimization for DSL dynamic spectrum management," *IEEE Transactions on Signal Processing*, accepted with mandatory minor revisions, Sept. 2009

8. P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen, "Dynamic spectrum management for green DSL," in submission

9. B. Li, P. Tsiaflakis, M. Moonen, J. Maes, and M. Guenach, "Dynamic resource allocation based partial crosstalk cancellation in DSL networks," in submission

International Conference Papers

1. P. Tsiaflakis, J. Vangorp, M. Moonen, J. Verlinden, and K. Van Acker, "An efficient search algorithm for the Lagrange multipliers of Optimal Spectrum Balancing in multi-user xDSL systems," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toulouse, France, May 2006.

2. J. Vangorp, P. Tsiaflakis, M. Moonen, J. Verlinden, and K. Van Acker, "Optimal spectrum balancing in multi-user xDSL systems with ON/OFF power loading," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toulouse, France, May 2006.

3. P. Tsiaflakis, J. Vangorp, M. Moonen, J. Verlinden, and G. Ysebaert, "Partial crosstalk cancellation in a multi-user xDSL environment," in *Proc. of the International Conference on Communications (ICC)*, Istanbul, Turkey, Jun. 2006.

4. J. Vangorp, P. Tsiaflakis, M. Moonen, and J. Verlinden, "Joint spectrum management and constrained partial crosstalk cancellation in a multi-user xDSL environment,", in *Proc. of the European Signal Processing Conference (EUSIPCO)*, Florence, Italy, Sep. 2006.

5. P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "A low complexity branch and bound approach to optimal spectrum balancing for digital subscriber lines," in *Proc. of the IEEE Global Communications Conference (GLOBECOM)*, San Francisco, USA, Nov. 2006.

6. P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "Convex relaxation based low-complexity optimal spectrum balancing for multi-user DSL," in *Proc. of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Honolulu, Hawaii, Apr. 2007.

7. P. Tsiaflakis and M. Moonen, "Low-complexity dynamic spectrum management algorithms for digital subscriber lines,", in *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 2769–2772, Las Vegas, Nevada, Apr. 2008.

8. P. Tsiaflakis, C.W. Tan, Y. Yi, M. Chiang, and M. Moonen, "Optimality certificate of dynamic spectrum management in multi-carrier interference channels," in *Proc. of the IEEE International Symposium on Information Theory (ISIT)*, pp. 1298–1302, Toronto, Canada, Jul. 2008.

9. P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen, "Throughput and Delay of DSL Dynamic Spectrum Management with Dynamic Arrivals," in *Proc. of the IEEE Global Communications Conference (GLOBECOM)*, New Orleans, Louisiana, USA, Nov. 2008.

10. P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen, "Green DSL: Energy-Efficient DSM," in *Proc. of the IEEE International Conference on Communications (ICC)*, pp. 1–5, Dresden, Germany, Jun. 2009.

11. P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen, "Fair greening for DSL broadband access," in *Proc. of the GreenMetrics Workshop in conjunction with ACM Sigmetrics/Performance (ACM/GREENMETRICS)*, pp. 1–5, Seattle, Washington, Jun. 2009.

12. P. Tsiaflakis, I. Necoara, J.A.K. Suykens, and M. Moonen, "An improved dual decomposition approach to DSL dynamic spectrum management," in *Proc. of the 17th European Signal Processing Conference (EUSIPCO)* pp. 1–5, Glasgow, Scotland, Aug. 2009.

International Conference Abstracts

1. P. Tsiaflakis and M. Moonen, "Low-complexity distributed power optimization for DSL networks," in *Proc. of the International Conference on Continuous Optimization (ICCOPT - MOPTA)*, Hamilton, Canada, Aug. 2007.

2. P. Tsiaflakis and M. Moonen, "Low-Complexity dynamic spectrum management for digital subscriber lines," in *Workshop on Optimization and Signal Processing (WOSP)*, Hong Kong, Dec. 2007.

3. M. Chiang, P. Tsiaflakis, Y. Yi, and M. Moonen, "Green DSL: Optimizing Fairly for Energy Efficiency,", in *Proc. of the 20th International Symposium on Mathematical Programming (ISMP)*, Chicago, US, Aug. 2009.

Patent Applications

1. B. Li, P. Tsiaflakis, M. Moonen, J. Maes, and M. Guenach, Applicant: Alcatel-Lucent Bell Labs, 2009[1].

[1]Intellectual property issues prevent further details from being released at this point

Curriculum Vitae



Paschalis Tsiaflakis was born in Genk, Belgium, in 1979. He received the M.Eng. degree (summa cum laude) in electrical engineering from the Katholieke Hogeschool Limburg, Belgium, in 2001 and the M.S. degree (magna cum laude) in electrical engineering from the Katholieke Universiteit Leuven (K.U.Leuven), Belgium, in 2004. Since 2004 he has been pursuing a Ph.D. under the supervision of Prof. M. Moonen at the Department of Electrical Engineering, K.U.Leuven, Belgium. In 2007, he was a Visiting Scholar at the CommNet Group, Princeton University, NJ, USA, under the supervision of Prof. M. Chiang. His research interests include signal processing and optimization of digital communication systems with special emphasis on DSL wired broadband access networks.

Mr. Tsiaflakis was funded by the Research Foundation - Flanders (FWO) from 2004 until 2008. He has been involved in several research projects in cooperation with Alcatel-Lucent, Antwerp, Belgium. He received the Best Multimedia Master Thesis prize awarded by PIMC in 2001, IEEE Travel Grants in 2006 and 2007, an FWO Aspirant scholarship for the period 2004-2008, a three-month FWO scholarship for a Visiting Research Collaboration at Princeton University in 2007, and a one-year PDM-kort postdoc grant.